

procedures. Among companies with 1,000 or more employees, 65.1 percent have mandatory arbitration procedures.

- ✦ Among private-sector nonunion *employees*, 56.2 percent are subject to mandatory employment arbitration procedures. Extrapolating to the overall workforce, this means that 60.1 million American workers no longer have access to the courts to protect their legal employment rights and instead must go to arbitration.
- ✦ Of the employers who require mandatory arbitration, 30.1 percent also include class action waivers in their procedures—meaning that in addition to losing their right to file a lawsuit on their own behalf, employees also lose the right to address widespread rights violations through collective legal action.
- ✦ Large employers are more likely than small employers to include class action waivers, so the share of *employees* affected is significantly higher than the share of employers engaging in this practice: of *employees* subject to mandatory arbitration, 41.1 percent have also waived their right to be part of a class action claim. Overall, this means that 23.1 percent of private-sector nonunion employees, or 24.7 million American workers, no longer have the right to bring a class action claim if their employment rights have been violated.

Introduction

Mandatory arbitration is a controversial practice in which a business requires employees or consumers to agree to arbitrate legal disputes with the business rather than going to court. Although seemingly voluntary in that the employee or consumer can choose whether or not to sign the arbitration agreement, in practice signing the agreement is required if the individual wants to get the job or to obtain the cellphone, credit card, or other consumer product the business is selling. Mandatory arbitration agreements are legally enforceable and effectively bar employees or consumers from going to court, instead diverting legal claims into an arbitration procedure that is established by the agreement drafted by the company and required as a condition of employment or of doing business with it.¹

Much attention has focused on the use of mandatory arbitration agreements in consumer contracts, such as consumer financial contracts, cellphone contracts, and nursing home resident contracts and the implications of such agreements for consumer rights.² There is less awareness of the use of mandatory arbitration agreements in employment contracts, but it is no less of a concern for those workers affected by it. These mandatory employment arbitration agreements bar access to the courts for all types of legal claims, including those based on Title VII of the Civil Rights Act, the Americans with Disabilities Act, the Family and Medical Leave Act, and the Fair Labor Standards Act. If an employment right protected by a federal or state statute has been violated and the affected worker has signed a mandatory arbitration agreement, that worker does not have access to the courts and instead must handle the claim through the arbitration procedure designated in the agreement.

Mandatory employment arbitration is very different from the labor arbitration system used to resolve disputes between unions and management in unionized workplaces. Labor arbitration is a bilateral system jointly run by unions and management, while mandatory employment arbitration procedures are unilaterally developed and forced on employees by employers. Whereas labor arbitration deals with the enforcement of a contract privately negotiated between a union and an employer, mandatory employment arbitration concerns employment laws established in statutes. Research has found that employees are less likely to win arbitration cases and they recover lower damages in mandatory employment arbitration than in the courts. Indeed, employers have a significant advantage in the process given that they are the ones who define the mandatory arbitration procedures and select the arbitration providers.³

Background: The Supreme Court's role in the increased use of mandatory employment arbitration agreements

A crucial 1991 Supreme Court decision, *Gilmer v. Interstate/Johnson Lane*,⁴ upheld the enforceability of mandatory employment arbitration agreements, meaning that such agreements now had the potential to substantially change how the employment rights of American workers are protected. But the practical impact of mandatory employment arbitration depends on whether or not American businesses decide to require that their employees sign these agreements as a term and condition of employment. Research from the 1990s and 2000s found that mandatory employment arbitration was expanding and by the early 2000s nearly one-quarter of the workforce was subject to mandatory arbitration. However there was a lack of subsequent research tracking whether this growth trend had continued beyond the early 2000s and describing the current extent of mandatory employment arbitration (see literature review, next section below).

The lack of basic data on the extent of mandatory arbitration is especially concerning given that recent years have seen a series of court decisions encouraging the expanded use of mandatory arbitration. In two key decisions, *AT&T Mobility LLC v. Concepcion* (2011) and *American Express Co. v. Italian Colors Restaurant* (2013),⁵ the Supreme Court held that class action waivers in mandatory arbitration agreements were broadly enforceable. This meant that businesses could not only use mandatory arbitration agreements to bar access to the courts for individual claims, but they could also shield themselves from class action claims. This gave businesses an additional incentive to include mandatory arbitration agreements in employment and other contracts.

In October 2017, the Supreme Court will hear a consolidated set of cases (*Murphy Oil/Epic Systems/Ernst & Young*) challenging the enforceability of class action waivers in mandatory employment arbitration agreements.⁶ In this set of cases, the central issue is whether requiring this waiver of the ability to use collective action to address employment law violations is a violation of the protections of the right to engage in concerted action contained in Section 7 of the National Labor Relations Act (NLRA). If the Supreme Court

accepts the argument that such waivers are in violation of the NLRA, the Court's decision would effectively put an end to the use of class action waivers in mandatory employment arbitration agreements. However, if the Court sides with the employers' arguments in these cases, this will signal to businesses that the last potential barrier to their ability to opt out of class actions has been removed. This would likely encourage businesses to adopt mandatory employment arbitration and class action waivers even more widely.

Existing research on the extent of mandatory employment arbitration

Despite growing attention to the issue of mandatory employment arbitration, there is a lack of good data on how widespread it has become. A 1992 academic study of conflict resolution procedures used by corporations in nonunion workplaces found that 2.1 percent of the companies surveyed included arbitration in their procedures.⁷ The one major governmental effort to investigate the extent of mandatory arbitration was a 1995 GAO survey, which found that 7.6 percent of establishments had adopted mandatory employment arbitration.⁸

Colvin's 2003 survey of conflict resolution procedures used in the telecommunications industry found that 14.1 percent of establishments in that industry had adopted mandatory arbitration and that these procedures applied to 22.7 percent of the nonunion workforce in the industry (since larger establishments were more likely to have adopted mandatory arbitration).⁹

The overall picture we have is one of mandatory employment arbitration expanding through the 1990s and early 2000s to nearly a quarter of the workforce. This study seeks to determine whether this expansion has continued beyond 2003 and how widespread mandatory employment arbitration is currently.

Findings of this study

To investigate the extent of mandatory employment arbitration, we conducted a national survey of private-sector American business establishments, focusing on the use of mandatory arbitration for nonunion employees. The survey was conducted from March to July 2017 and had a sample size of 627, yielding a margin of error at 95 percent confidence of plus or minus 3.9 percentage points.

More than half of private-sector nonunion workers are subject to mandatory arbitration

On the central question of whether employees were required to sign a mandatory "agreement or provision for arbitration of legal disputes with the company," 50.4 percent

of respondents indicated that employees in their establishment were required to enter into this type of agreement.

Although mandatory employment arbitration is usually established by having employees sign an arbitration agreement, typically at the time of hiring, in some instances businesses adopt arbitration procedures simply by announcing that these procedures have been incorporated into the organization's employment policies. An additional 3.5 percent of establishments had adopted mandatory arbitration using this second mechanism. Combined with the 50.4 percent of employers who require employees to sign an agreement, this means that a total of 53.9 percent of all establishments in the survey had adopted mandatory employment arbitration through one of these two mechanisms.

The establishments that have adopted mandatory arbitration tend to be those with larger workforces. Adjusting for workforce size, overall 56.2 percent of employees in the establishments surveyed were subject to mandatory arbitration procedures. Extrapolating to the overall private-sector nonunion workforce, this corresponds to 60.1 million American workers who are now subject to mandatory employment arbitration procedures and no longer have the right to go to court to challenge violations of their employment rights.¹⁰

Larger companies are more likely to adopt mandatory employment arbitration than smaller companies

As mentioned above, the likelihood that an employer will adopt mandatory employment arbitration varies with the size of the employer. Whereas 53.9 percent of all establishments had mandatory arbitration, among establishments that were part of companies with 1,000 or more employees, 65.1 percent had mandatory arbitration. In general, larger organizations with more sophisticated human resource policies and better legal counsel are more likely to adopt policies like mandatory arbitration that protect them against legal liability.¹¹ They could also become trendsetters over time if smaller employers copy these practices that larger employers have proven to be effective in protecting employers against legal actions.

Mandatory arbitration discourages employees from bringing claims when their rights are violated

Although around 60 million American workers are now subject to mandatory employment arbitration procedures, this does not mean that the number of workers arbitrating workplace disputes has increased correspondingly. It has not. Mandatory arbitration has a tendency to suppress claims. Attorneys who represent employees are less likely to take on clients who are subject to mandatory arbitration,¹² given that arbitration claims are less likely to succeed than claims brought to court and, when damages *are* awarded, they are

likely to be significantly smaller than court-awarded damages.¹³ Attorney reluctance to handle such claims effectively reduces the number of claims that are brought since, in practice, relatively few employees are able to bring employment law claims without the help of an attorney.

In an earlier study, Colvin and Gough (2015) found that an average of 940 mandatory employment arbitration cases per year were being filed with the American Arbitration Association (AAA), the nation's largest employment arbitration service provider.¹⁴ Other research indicates that about 50 percent of mandatory employment arbitration cases are administered by the AAA.¹⁵ This means that there are only about 1,880 mandatory employment arbitration cases filed per year nationally. Given the finding that 60.1 million American workers are now subject to these procedures, this means that only 1 in 32,000 employees subject to these procedures actually files a claim under them each year. These findings indicate that employers adopting mandatory employment arbitration have been successful in coming up with a mechanism that effectively reduces their chance of being subject to any liability for employment law violations to very low levels.

In addition to losing their right to private legal action, nearly 25 million of these workers are also prohibited from participating in class action suits

Although class action waivers are one of the most controversial features of mandatory arbitration procedures, it is important to recognize that mandatory arbitration agreements do not necessarily include class action waivers. Among the survey respondents whose companies had mandatory arbitration procedures, 30.1 percent included class action waivers. These tended to be in establishments with larger workforces, so overall 41.1 percent of employees subject to mandatory arbitration procedures were also subject to class action waivers. Relative to the overall workforce, including both those subject to and those not subject to mandatory arbitration, these estimates indicate that 23.1 percent of all private-sector nonunion employees are subject to class action waivers in mandatory arbitration procedures, corresponding to 24.7 million American workers.

The finding that many employers who have adopted mandatory employment arbitration have not included class action waivers in their procedures stands in contrast to the situation with consumer financial contracts, which the CFPB found almost always include class action waivers along with mandatory arbitration.¹⁶ One explanation for the lower use of class action waivers in the employment setting is the ongoing legal uncertainty about their enforceability given the NLRA issues that the Supreme Court will be deciding in the upcoming *Murphy Oil/Epic Systems/Ernst & Young* cases.

Conclusion: Mandatory arbitration is a growing threat to workers' rights

Mandatory employment arbitration is the subject of fierce legal and policy debates. There is growing evidence that mandatory arbitration produces outcomes different from those of litigation, to the disadvantage of employees, and suffers from due process problems that give the advantage to the employers who impose mandatory arbitration on their workers.¹⁷ What has been less clear is how widespread the impact of mandatory employment arbitration is. In the consumer arena, the CFPB's 2015 study showed that mandatory arbitration clauses are common, being included in a majority of credit card, prepaid card, student loan, and payday loan agreements.¹⁸ By contrast, in the employment arena our knowledge of the extent of mandatory arbitration was limited to a few surveys from the 1990s and early 2000s, the latter of which suggested that nearly a quarter of employees might have been subject to mandatory arbitration by that point in time.

The study described in this report shows that mandatory employment arbitration has continued to grow in extent, and now, in 2017, in over half of American workplaces, employees are subject to mandatory arbitration agreements that take away their right to bring claims against their employer in court. This represents a dramatic and important shift in how the employment rights of American workers are enforced. Rather than having their rights adjudicated through the public courts and decided by juries of their peers, more often now American workers have to bring claims—claims that are based on statutes enacted by Congress or state legislatures—through arbitral forums designated by agreements that their own employers drafted and required them to agree to as a condition of employment.

The employment conditions experienced by the American worker have changed dramatically in recent decades as labor standards and their enforcement have eroded, union representation has declined, and the wage-suppressing effects of globalization have been amplified by an overvalued U.S. dollar and trade agreements that have eroded workers' power. Against this backdrop of increased economic risk and uncertainty for workers and the disruption of traditional protections, laws protecting employment rights such as the minimum wage, the right to equal pay, and the right to a safe workplace free of harassment or discrimination based on race, gender, or religion have become increasingly important as a workplace safety net. However, these protections are at risk of being undermined if there is no effective means of enforcing them.

Mandatory employment arbitration has expanded to the point where it has now surpassed court litigation as the most common process through which the rights of American workers are adjudicated and enforced. It is likely to become an even more widespread practice if the Supreme Court upholds the enforceability of class action waivers in its October 2017 decision. In fact, if the Court rules in favor of the employers in these cases, imposing mandatory arbitration with class action waivers is likely to become the predominant

management practice and workers will find it exponentially more difficult to enforce their rights going forward.

About the author

Alexander J.S. Colvin is the Martin F. Scheinman Professor of Conflict Resolution and Associate Dean for Academic Affairs, Diversity, and Faculty Development at the ILR School, Cornell University. His research and teaching focuses on employment dispute resolution, with a particular emphasis on procedures in nonunion workplaces and the impact of the legal environment on organizations.

Methodological appendix

To measure the current extent of mandatory employment arbitration, we conducted a national-level survey of private-sector employers. The survey was funded by the Economic Policy Institute and administered through telephone- and web-based methods by the Survey Research Institute (SRI) at Cornell University.

The study measured the extent of mandatory employment arbitration by surveying employers rather than by surveying employees because research has found that employees are often unaware or fail to recall that they have signed arbitration agreements and may not understand the content and meaning of these documents.¹⁹ The survey was limited to private-sector employers because public-sector employees typically have their employment regulated by specific public-sector employment laws and employment practices differ substantially between private- and public-sector employers. The survey focused on nonunion employees since unionized employees have their employment governed by collective bargaining agreements, which provide for labor arbitration to resolve disputes. Although both are forms of arbitration, labor arbitration differs in many respects from mandatory employment arbitration and should not be included in the same category.²⁰

The survey population was drawn from Dun & Bradstreet's national marketing database of business establishments. It was stratified by state population to be nationally representative. The survey population was restricted to private-sector business establishments of 50 or more employees, and the analysis was restricted to procedures affecting nonunion employees. The individual respondents were the establishment's human resources manager or whichever individual was responsible for hiring and onboarding employees. The reason for use of this individual as the person to respond to the survey is that mandatory arbitration agreements are typically signed as part of the onboarding paperwork when a new employee is hired. As a result, the manager responsible for this process is the individual most likely to be knowledgeable about the documents the new employee is signing. Typical job titles of individual respondents included human resource director, human resource manager, personnel director, and personnel manager.

Participants were initially contacted by telephone and then given the option of completing phone or web versions of the survey. Follow-up calls were made to encourage participation. Where participants had provided email addresses, a series of emails were also sent to prompt completion of the survey. To encourage participation, respondents were offered the opportunity to win one of ten \$100 Amazon gift cards in a raffle drawing from among participants in the survey.

Data collection started in March 2017 and was completed in July 2017. A total of 1,530 establishments were surveyed, from which 728 responses were obtained, representing an overall response rate of 47.6 percent. Some survey responses had missing data on specific questions; however, 627 respondents provided complete data on the key variables of interest. The response rate and sample size are similar to those obtained in past establishment-level surveys of employment relations and human resource practices. The median establishment size in the sample is 90 employees, and the average size is 226 employees. Most establishments are single-site businesses, while 38.2 percent are part of larger organizations. These larger organizations have an average workforce size of 18,660 employees. Overall, 5.2 percent of establishments in the sample are foreign-owned.

Endnotes

1. For a general discussion of the state of the law and practice around mandatory arbitration, see Stone and Colvin 2015.
2. The Consumer Financial Protection Bureau conducted a study of the widespread use of mandatory arbitration in consumer financial contracts and has proposed a rule limiting the use of class action waivers in these agreements. Mandatory arbitration in nursing home resident contracts was the focus of a proposed rule by the Obama administration banning their use.
3. For an overview of this research, see Stone and Colvin 2015, 18--23.
4. 500 U.S. 20 (1991).
5. *AT&T Mobility LLC v. Concepcion* 563 U.S. 333 (2011); *American Express Co. v. Italian Colors Restaurant* 133 S. Ct. 594 (2013).
6. *NLRB v. Murphy Oil USA, Inc.*, No. 16-307; *Epic Systems Corp. v. Lewis*, No. 16-285; *Ernst & Young LLP v. Morris*, No. 16-300. For more about the *Murphy Oil/Epic Systems/Ernst & Young* cases and the implications of the pending Supreme Court decision, see McNicholas 2017.
7. See Feuille and Chachere 1995, 31.
8. GAO 1995. The GAO's survey initially indicated that 9.9 percent of establishments had mandatory arbitration procedures; however, on follow-up a number of them indicated that they had made mistakes in reporting, such as confusing union labor arbitration procedures with nonunion mandatory employment arbitrations. Adjusting for these erroneous responses, only 7.6 percent of the establishments actually had mandatory employment arbitration.
9. See Colvin 2008.

10. This estimate is based on the Bureau of Labor Statistics report “Union Members – 2016,” released January 26, 2017, which reports an overall private-sector workforce of 115.417 million, among which 8.437 million are union-represented private-sector workers, with the remaining 106.980 million workers being nonunion.
11. See, e.g., Edelman 1992, showing that larger organizations are more likely to adopt organizational policies designed to protect them from the impact of civil rights laws.
12. See Colvin 2014.
13. See Colvin and Gough 2015.
14. See Colvin and Gough 2015 (1027), reporting that 10,335 claims were filed with the AAA over the 11-year period from 2003–2013.
15. See Stone and Colvin 2015, 17.
16. The Consumer Financial Protection Bureau’s Arbitration Study found that over 90 percent of consumer financial contract arbitration clauses that it studied contained class action waivers (CFPB 2015).
17. See Stone and Colvin 2015.
18. CFPB 2015.
19. A study by Zev Eigen (2008) found that a majority of Circuit City employees he interviewed were unaware that they had signed arbitration agreements or of the import of such agreements, even though the company had a longstanding policy of requiring its employees to sign mandatory arbitration agreements and even though Circuit City’s arbitration policy had been the subject of an important case on the enforceability of these agreements that was decided by the Supreme Court in 2001.
20. One of the most important differences is that labor arbitration procedures are jointly established and administered by unions and management, in contrast to mandatory arbitration, which is unilaterally established by the employer. In addition, most labor arbitration procedures do not bar employees from bringing statutory employment claims separately through the courts.

References

- Colvin, Alexander J.S. 2008. “Empirical Research on Employment Arbitration: Clarity Amidst the Sound and Fury?” *Employee Rights and Employment Policy Journal* vol. 11, no. 2, 405–47.
- Colvin, Alexander J.S. 2014. “Mandatory Arbitration and Inequality of Justice in Employment.” *Berkeley Journal of Employment and Labor Law* vol. 35, 71–90.
- Colvin, Alexander J.S., and Mark Gough. 2015. “Individual Employment Rights Arbitration in the United States: Actors and Outcomes.” *ILR Review* vol. 68, no. 5, 1019–42.
- Consumer Financial Protection Bureau (CFPB). 2015. *Arbitration Study: Report to Congress, pursuant to Dodd-Frank Wall Street Reform and Consumer Protection Act § 1028(a)*.
- Edelman, Lauren B. 1992. “Legal Ambiguity and Symbolic Structures: Organizational Mediation of Civil Rights Law.” *American Journal of Sociology* vol. 97, no. 6, 1531–76.

Eigen, Zev. 2008. "The Devil in the Details: The Interrelationship Among Citizenship, Rule of Law and Form-Adhesive Contracts." *Connecticut Law Review* vol. 41, no. 2, 1–50.

Feuille, Peter, and Denise R. Chachere. 1995. "Looking Fair and Being Fair: Remedial Voice Procedures in Nonunion Workplaces." *Journal of Management* vol. 21, 27–36.

General Accounting Office (GAO). 1995. *Employment Discrimination: Most Private Sector Employers Use Alternative Dispute Resolution*. GAO/HEHS 95-150.

McNicholas, Celine. 2017. "Supreme Court Should Uphold Working People's Fundamental Rights in *Murphy Oil*." *Working Economics* (Economic Policy Institute blog), September 27.

Stone, Katherine Van Wezel, and Alexander J.S. Colvin. 2015. *The Arbitration Epidemic*. Economic Policy Institute Briefing Paper no. 414.

Luncheon Address: Reflections on Dwindling Worker Bargaining Power and Monetary Policy

Alan B. Krueger

I commend the Kansas City Federal Reserve Bank for exploring the topic of changing market structures and monetary policy. Many of the product market developments discussed at this year's symposium have important implications for wages, employment, and wage setting. I will focus my remarks on changes in labor market competition and worker bargaining power in the United States, and their implications for central bankers. My theme is that declining competition and worker bargaining power can help explain the puzzle *du jour* of relatively weak wage growth despite historically low unemployment in the United States.

Although economists' go-to model of the labor market is often one with perfect competition—where bargaining power is irrelevant because supply and demand determine the wage, and there is nothing firms can do about it—in many applications I think it is more appropriate to model the labor market as imperfectly competitive, subject to monopsonylike effects, collusive behavior by firms, search frictions and surpluses that are bargained over. As a result of these labor market features, firms should be viewed as wage-setters or wage-negotiators, rather than wage-takers.¹ This perspective can explain many well-documented phenomenon in the labor market, such as the high variability in pay for workers with identical skills in

different industries or firms, the lack of evidence that minimum wage increases reduce employment, and the reluctance of firms to raise wages when vacancies are hard to fill.

I have noticed that many economists are skeptical of the notion that markets are manipulable, that firms or traders have some sway over prices or wages. When I worked at the U.S. Treasury Department in 2009, some of the best finance economists in the world thought it inconceivable that foreign exchange markets or LIBOR could be manipulated. After all, these are the largest and most liquid markets in the world. Only later did we learn that several traders have been convicted of colluding on exchange rates, and that LIBOR was totally rigged.

One economist who thought that labor markets are imperfect and subject to manipulation, however, was Adam Smith. In *The Wealth of Nations* he wrote that employers “are always and everywhere in a sort of tacit, but constant and uniform combination, not to raise the wages of labor above their actual rate. To violate this combination is everywhere a most unpopular action, and a sort of reproach to a master among his neighbors and equals.” And he ridiculed naysayers who doubted that employers colluded “as ignorant of the world as of the subject.” In full conspiracy mode, he added that, “We seldom, indeed, hear of this combination, because it is the usual, and one may say, the natural state of things, which nobody ever hears of.”

Broadly speaking, there are two varieties of economic models that give employers some discretion over wage setting: 1) the first, pioneered by Joan Robinson, is a static monopsony model where a single employer faces the upward sloping market labor supply curve. This could easily be extended to oligopsony, where a small number of employers dominate a market and face upward sloping labor supply curves, or to a Smith-like situation where employers collude to suppress pay below the competitive rate; 2) the second class of models, pioneered by Ken Burdett, Dale Mortensen, Chris Pissarides and Peter Diamond, and extended by Alan Manning (2003), rests on search frictions. It takes time and effort for workers to search for job openings and for firms to search for workers. As a consequence, if a firm pays a little less than the “going wage” it would not lose all of

its workers or find it impossible to hire new ones. In fact, there is no single “going wage” in these models, but a range of plausible offers that firms could make, or bargains that firms and workers can strike.²

As a practical matter, both models are equivalent to assuming that the labor supply curve to a firm is upward sloping, instead of infinitely elastic. Firms operate with costly vacancies in these models, yet resist raising wages because pay would need to be increased for all workers, not just the incremental worker hired. And if employers collude to hold wages to a fixed, below-market rate, or if monopsony power increases over time, then wages could remain stubbornly resistant to upward pressure from increased labor demand in a booming economy.

With this framework as background, I want to make six observations about the labor market that are relevant to the current conundrum of weak wage growth despite low unemployment, and then turn to some reflections on the implications of these observations for monetary policy.

Six Observations on Labor Markets

1. *A high-pressure labor market tends to boost wages and opportunities for low-wage workers.* This was convincingly demonstrated by Arthur Okun in a 1973 Brookings paper, and confirmed by experience in subsequent recoveries, including the current one.³ In a 1999 Brookings paper follow-up to Okun’s work, Larry Katz and I similarly found that the wage Phillips curve relationship is steeper at the lower deciles. In other words, wage growth is more responsive to unemployment for less-skilled and lower-paid workers. Given the tremendous rise in earnings inequality and deterioration in opportunities for workers in the bottom half of the income distribution, the benefits of a high pressure economy cannot be understated. Katz and I also found that the wage Phillips curve moves around over time. For a worker paid the median wage, the unemployment rate threshold required to generate positive real wage growth for the median worker (which we infelicitously called URZERCG, for unemployment rate associated with zero expect real compensation growth), fell from 6.8 percent in the late 1970s and 1980s to 5.4 percent in the 1990s. It appears to have fallen even further in the 2000s (see Bivens 2014), suggesting a tighter labor market is now required to support real wage growth.

This is a reminder that the wage Phillips curve is a useful relationship, but it shifts around from time to time. One reason why it may have shifted in recent decades is because of increased employer monopsony power and declining worker bargaining power.

2. *Average wage growth is weaker than one would expect from conventional relationships.*⁴ (If PowerPoint were allowed over lunch, I could document this assertion beyond a shadow of a doubt for a variety of wage series and specific specifications, but for now you'll have to trust me.) Although nominal wage growth has been creeping up throughout this recovery, over the past 12 months, nominal wage growth has not kept pace with Consumer Price Index inflation. Popular explanations that have been put forth to explain low wage growth in this recovery include: 1) low price inflation; 2) low productivity growth; 3) hidden labor market slack; and 4) demographic changes. These factors likely contribute to slow wage growth to varying extents, but I doubt that they fully explain the wage puzzle.

Based on the wage Phillips curve that I have been estimating for years—which predicts year t 's wage growth less year $t-1$'s inflation as a function of the unemployment rate—annual wage growth is 1 to 1.5 percentage points below what one would expect today. Demographic shifts perhaps shave 0.2 to 0.3 percentage point from wage growth.⁵ Slower productivity growth could explain as much as an additional percentage point of the wage growth puzzle, but less in the last year since productivity growth has picked up yet real wage growth has declined. Moreover, one could argue that productivity growth is endogenously determined as a function of wages. And the role of unmeasured slack is probably minimal because the quit rate (a measure of worker confidence) is back to where it was at the previous business cycle peak, and the prime age employment-to-population rate and labor force participation rate are basically on their long run trend.⁶

3. *There is growing evidence supporting an important role of monopsony power in the job market stemming from both employer concentration and dynamic labor market considerations.* First consider employer concentration. On the one hand, Benmelech, Bergman and Kim (2018) find that the Herfindahl index of establishment-level employment at the county-level for firms classified by four-digit SIC

manufacturing industries grew steadily from 1977 to 2009. They further find that wages are lower in more highly concentrated labor markets, and the connection between wages and employer concentration increased over time. On the other hand, Azar, Marinescu and Steinbaum (2017) find that labor markets—defined as occupational categories within commuting zones—with a higher Herfindahl index of job openings (meaning more employer concentration in terms of hiring) have lower wages. Both studies find surprisingly high degrees of employer concentration, and this is especially for job openings in less populated areas.

Studies of particular professions also find evidence of monopsony power. Perhaps the most studied occupation has been nursing. Sullivan (1989) and Staiger, Spetz and Phibbs (2010), for example, find substantial evidence of monopsony power on the part of hospitals.

Other recent studies have provided evidence of dynamic monopsony power. Using the Longitudinal Employer Household Dynamics (LEHD) data set, for example, Douglas Webber (2015) estimates that the average labor supply elasticity to a firm is 1.08, although there is considerable variability across firms. This is a lot less than infinity! And he finds that firms with more inelastic labor supply pay lower wages, suggesting that they exploit their monopsony power.⁷

4. Monopsony power has probably always existed in labor markets, but the forces that traditionally counterbalanced monopsony power and boosted worker bargaining power have eroded in recent decades. Union membership, for example, has fallen from a quarter of the U.S. workforce in 1980 to only 10.7 percent in 2017. Collective bargaining, which is much less common in the United States today, was an effective counterweight to employer monopsony power. And the effect of this trend on wages is even broader because of what is known as the “union threat effect”; unlike in the past, few employers today preemptively raise pay to head off a possible union drive.

Another counterbalance to monopsony power that is weaker today is the minimum wage. The U.S. federal minimum wage is currently \$7.25 an hour, and has not been raised since July 2009. The real value of the minimum wage is down about 20 percent since 1979.

In that period, by contrast, both the U.K. and Germany enacted national minimum wages that currently stand at about \$10 an hour at current exchange rates. (For the U.K., I use the rate for those ages 25 and older.)

The decline in union representation and the erosion of the real value of the minimum wage have contributed to the significant rise in inequality and polarization of incomes in the United States since the early 1980s. These shifts have also likely contributed to the downward trend in labor's share of national income in the United States since the 1990s, after decades of stability.

One might argue that these changes have made the labor market more competitive, but the fact that employment-to-population rate had trended down, and that regional shocks are now more persistent for wages, employment and labor force, suggests a less competitive labor market with weaker worker bargaining power.⁸

Going forward, worker bargaining power likely will be further eroded by two recent Supreme Court rulings. The *Janus* decision, which bars public sector unions from collecting agency fees from nonmembers, will encourage free riding and further weaken labor unions. And the *Epic* decision allows employers to require employees to pursue disputes in mandatory arbitration instead of filing lawsuits in court and to waive the right to class-action law suits.

5. *There has been a proliferation of practices that enhance monopsony power and weaken worker bargaining power.* Let me highlight five such practices.

First, the reliance on temporary help agencies, staffing firms and outsourcing has increased in the U.S. labor market. One implication of this practice is that firms can wage discriminate, which facilitates the exercise of monopsony power. If a hospital has persistent vacancies for nursing positions, for example, it can reach out to a staffing firm that pays its nurses a higher salary to supply additional nurses without having to raise its wage scale for incumbent nurses.

Second, a quarter of American workers are bound by a noncompete restriction on their current job, or from a previous job. These

restrictions, which may be justified in a limited number of cases to protect returns on specific training or trade secrets, have run amuck. Even Jimmy John's employed the practice for submarine sandwich makers, until they were forced to drop it. Just over one in five workers who earn less than the median wage are bound by a noncompete restriction on their current or a previous job.⁹ Non-compete clauses narrow workers options, and therefore reduce mobility and bargaining power.

Third, a growing fraction of the workforce is covered by occupational licensing restrictions, typically imposed by state and local authorities. Morris Kleiner and I, for example, find that over a quarter of workers are required to obtain a license to perform their job. These restrictions may be justified in some positions that require extraordinary skill or put the public at risk, but they restrict job opportunities and mobility. They restrict mobility because many states do not recognize other states' licenses, so a teacher or a nurse, for example, who is seeking to move to another location would often have to go through the burdensome and costly process of requalifying for a license in another jurisdiction.

Fourth, Orley Ashenfelter and I find that 58 percent of franchise companies have a no-poaching clause that prevents or restricts the ability of one franchisee in a chain from hiring workers employed by other franchisees. This is up from 36 percent in 1996. The practice is particularly common in fast food chains. We find that 80 percent of the 40 largest Quick Service Restaurant franchise chains have a no-poaching requirement. Since the human capital would remain within the chain, there is little business justification for such a clause other than to restrict worker mobility and opportunities.

Fifth, although no-poaching agreements among franchisees within the same chain are an unsettled area of the law, agreements among independent firms to refrain from hiring each other's workers or to set pay or pay increases at a common level are illegal. Nonetheless, as Adam Smith expected, such collusion takes place. There are many colorful recent examples. After Google's co-founder, Sergey Brin, tried to hire a programmer from Apple, for example, Steve Jobs wrote

an email saying, “If you hire a single one of these people that means war.” A class-action civil suit alleging collusion brought on behalf of more than 64,000 software engineers and other employees of Apple, Google, Adobe, Intel, Intuit, Pixar, and Lucasfilm was settled for half a billion dollars in 2015.

Closer to home, the chairs of top U.S. economics departments used to regularly confer at the Annual Meeting of the American Economic Association to jointly agree on pay and course loads for assistant professors until the Justice Department raised concerns about the legality of the practice.

And several suits alleging collusion in hiring or wage setting have been successfully brought on behalf of nurses against hospitals. Eight major hospitals in Detroit, for example, recently reached a \$90 million settlement in a suit alleging that the hospitals colluded to reduce nurses pay. Similar cases are in various stages in Albany, Memphis, San Antonio, and Arizona.

Earlier this week, I spoke with Jeffrey Suhre, a registered nurse and lead plaintiff in the *Detroit Nurses* case, to understand the perspective of an employee who worked at a firm that colluded with other employers to suppress pay. He started working in the emergency room at St. John Providence Hospital in Warren, Michigan, in 1991, and later moved to the critical care unit, tending to patients with open heart surgery and other serious conditions. After 12 or 13 years, Mr. Suhre said he got an inkling that the Human Resource Department at his hospital was coordinating with other hospitals in setting nurses’ pay as a result of some emails that he viewed. He said the nurses were nonunionized, and the hospitals in the area wanted to prevent nurses from jumping from one hospital to another for better pay and working conditions. The executives would often discuss these issues and exchange pay rates at conferences. One indication that the hospitals exploited their monopsony position that he mentioned is that to fill vacancies nurses were often hired from contract agencies at \$38 to \$40 an hour (plus administrative fees), while staff nurse pay at his hospital was only \$30-\$31 an hour. A class action suit was filed on behalf of Mr. Suhre and thousands of other nurses in 2006. He gave a deposition in 2007. He said the hospital “made life hell” for him

after the suit was filed—for example, by increasing his patient load to a level he considered a risk for patients—so he quit in January 2008. Other hospitals were reluctant to hire him. He now works in home health care. The antitrust suit was settled in 2010, but Mr. Suhre did not receive any money until 2012, six years after filing suit. Under the Sherman Antitrust Act damages were limited to four years of employment. The nurses received approximately \$4,000, on average, in damages. He suspects that the collusive practices still continue, but more covertly.

It is also worth emphasizing that collusion is easier when there are fewer companies competing in the labor market. The increase in employer concentration that the United States has experienced thus facilitates collusion.

And collusion doesn't have to be explicit; it could take place because a certain wage, such as the minimum wage, becomes a focal point from which employers are reluctant to deviate. Natalya Shelkova (2014) provides evidence that the large and persistent spike in the wage distribution at the minimum wage is consistent with focal point collusion.

More generally, tacit collusion could come about because employers and workers were shocked by the depth of the Great Recession, making workers fearful of bargaining for higher wages and employers disinclined to offer higher wages despite worker shortages, because they grew accustomed to having a queue of well qualified applicants during the recession and for a long period afterward.

Pressure for collusion to break down increases when the job market becomes really tight, which, I suspect, is part of the reason for the existence of the wage Phillips curve relationship. But in recent years this tendency has been offset by countervailing forces that have strengthened monopsony power and weakened worker bargaining power. This could explain the simultaneous occurrence of record numbers of job openings and only modest wage increases.

When it comes to employer complaints about labor shortages, Minnesota Fed President Neel Kashkari recently said, "If you are not raising wages, then it just sounds like whining." But there is another

possibility. If you are not raising wages and can't find enough workers, you may be colluding. Or resorting to anticompetitive practices.

6. *My final observation is that the occurrence of greater monopsony power would cause lower wages and worker shortages at firms, but not necessarily lower aggregate employment very much.* With lower wages and a small, but positive, labor supply elasticity, there's only a small negative effect on employment. Most estimates in the voluminous literature indicate that aggregate labor supply is fairly inelastic, especially for men (see Killingsworth 1983 and Blundell and MaCurdy 1999). The aggregate labor supply elasticity is probably on the order of only 0.1 or 0.2.

As another indication that aggregate labor supply is fairly inelastic, consider that from 2014:H1 to 2018:H1, the real median weekly earnings of fully time employees increased by 5.4 percent and the economy moved to full employment. Yet the civilian labor force participation rate was essentially unchanged over this period, standing at 62.9 percent in 2014:H1 and 62.8 percent in 2018:H1. Of course, aggregate labor force is being dragged down by an aging workforce, but the increase in participation by prime age workers has not been sufficient to outweigh the downward effects of an aging workforce. (From 2014:H1 to 2018:H2, the participation rate of prime age men and women increased from 81.0 percent to 81.6 percent.)

To be clear, I am not arguing that aggregate labor supply is perfectly inelastic. There is some responsiveness to wages and working conditions, and this is especially the case for the most disadvantaged workers in society.

On balance, however, I would argue that the main effects of the increase in monopsony power and decline in worker bargaining power over the last few decades have been to shrink the slice of the pie going to workers and increase the slice going to employers, not to reduce the size of the pie overall.

This is clearly an important issue, and goes to the heart of the Fed's maximum sustainable employment mandate, so it is a topic that deserves much greater research in the future.

Implications for Monetary Policy

Lastly, I'll turn to the difficult part of my lecture. What does this mean for monetary policy? I readily acknowledge a tremendous amount of uncertainty, so my remarks are mainly intended to start a conversation. In addition, the ongoing structural shift in the labor market toward weaker worker bargaining power is only one factor among *many* that central bankers should consider in setting monetary policy.

And it almost goes without saying that the best tool to address anticompetitive practices in the labor market is antitrust enforcement. On this front, there is reason for a small measure of optimism. Toward the end of the Obama administration, the Department of Justice and Federal Trade Commission issued new guidelines for human resources professionals that clearly stated that, "Agreements among employers not to recruit certain employees or not to compete in terms of compensation are illegal." The Justice Department has said that it will enforce the new guidelines, and the head of the Department's Antitrust Division, Makan Delrahim, recently said, "I've been shocked about how many of these [collusive agreements] there are, but they're real."¹⁰ He has already announced one settlement in a no-poaching case involving two of the largest rail equipment manufacturers.

In addition, in the last month, Washington State's Attorney General Bob Ferguson reached landmark agreements with 15 fast food chains, including McDonald's, Auntie Anne's and Cinnabon, to drop their no-poaching restrictions.

Antitrust policy can only go so far in reversing the erosion of worker bargaining power and offsetting the inefficient aspects of monopsony, however. Is there a role for monetary policy, particularly during a long transition period when monopsony power is rising and worker bargaining power is eroding?

Fifty years ago, in his presidential address to the AEA, Milton Friedman (1968) wrote what was surely the longest and most influential sentence in the history of theory undergirding monetary policy:

The “natural rate of unemployment”, in other words, is the level that would be ground out by the Walrasian system of general equilibrium equations, provided there is embedded in them the actual structural characteristics of the labor and commodity markets, including market imperfections, stochastic variability in demands and supplies, the cost of gathering information about job vacancies and labor availabilities, the costs of mobility and so on.

Somewhere between Adam Smith and Milton Friedman economists’ lexicon evolved from “the natural state of things” to “market imperfections.” Nonetheless, an implication of Friedman’s view is that the *natural rate of employment* falls if monopsony power rises.

How should central banks respond? One view is that they should treat this new development similar to a negative productivity shock. In other words, it is too bad that the Walrasian system has shifted against workers, but that’s embedded in the system and lowers potential output.

This is a reasonable response under Joan Robinson monopsony or an exogenous rise in search frictions. There is probably little a central bank could do to return employment to its previous level in the long run. Moreover, as I have suggested previously, the employment effect is probably small because aggregate labor supply is inelastic.

But if explicit or implicit collusion among employers is an important source of growing monopsony power, allowing the labor market to run hotter than otherwise could possibly cause collusion to break down, because the benefit to an individual firm from raising pay while others are colluding at a fixed wage is greater when demand is greater. If the collusion does whither, wages and employment would rise.

Another consideration concerns the effect of declining worker bargaining power on wages and prices. If weaker nominal wage growth is being passed through in the form of lower prices, then the price stability mandate would call for a more accommodative monetary policy in response to declining worker bargaining power.

The structural labor market shifts that I have emphasized may also have implications for the extent of downward nominal wage rigidity, which is also an important consideration for monetary policy over the business cycle. This is a worthy topic for future research and discussion, but less relevant near the peak of a business cycle.

To conclude, I think it is important for central bankers to be aware of the impact of the growing use of monopsony power and noncompetitive labor market practices on wages, employment and output. What this means for monetary policy, however, is less clear. My tentative advice is that the optimal central bank response depends on: 1) the extent to which weaker wage growth is passed through to prices, or allocated to profits; 2) the elasticity of aggregate labor supply; and 3) the ability of a booming economy to counteract collusive behavior and other anti-competitive labor market forces. These considerations should be part of the conversation along with central banks' other weighty concerns, such as the effect of monetary policy on financial stability, the effect of tariffs and trade wars on inflation and output, and the effects of demographic shifts on potential output.

Endnotes

¹Notice that I don't call these features "imperfections." They are the way the labor market works. The assumption of perfect competition is the deviation from the norm of "imperfection" as far as the labor market is concerned.

²Flinn (2006) explicitly models bargaining between firms and workers over the value that each unique worker-firm match creates. Even absent search frictions, monopsony power would exist if workers have heterogeneous preferences toward working at various companies, such as because of varying commuting costs.

³See Furman (2018) for evidence on the current recovery.

⁴Chairman Powell (2018) has likewise stated, "there is still a bit of a puzzle in that we're hearing about labor shortages now all over the country in many, many different occupations in different geographies. And one would have expected, I would have expected, that wages would move up a little bit more."

⁵I derived this range by estimating a standard cross sectional age-earnings profile, and shifting the age distribution of the workforce back to 1979. Younger workers tend to receive greater annual wage increases, and they represent a shrinking share of the workforce.

⁶On slack, see Furman (2018), Krugman (2018) and Krueger (2017).

⁷Also see Dube, Giuliano and Leonard (2015).

⁸On the persistence of regional shocks, see Dao, Furceri and Loungani (2017) and Charles, Hurst and Schwartz (2018).

⁹See Krueger and Posner (2017) and Starr, Prescott and Bishara (2017).

¹⁰Comment by Assistant Attorney General Makan Delrahim at a conference hosted by the Antitrust Research Foundation at the Antonin Scalia Law School at George Mason University, Jan. 19, 2018.

References

- Azar, José, Ioana Marinescu and Marshall I. Steinbaum. 2017. "Labor Market Concentration," National Bureau of Economic Research, Working Paper w24147.
- Benmelech, Efraim, Nittai Bergman and Hyunseob Kim. 2018. "Strong Employers and Weak Employees: How Does Employer Concentration Affect Wages?" National Bureau of Economic Research, Working Paper w24307.
- Bivens, Josh. 2014. "Working Economics Blog," Sept. 23. Available at <https://www.epi.org/blog/explicit-fighting-inflation-war-ensure-real/>
- Blundell, Richard, and Thomas MaCurdy. 1999. "Labor Supply: A Review of Alternative Approaches," in *Handbook of Labor Economics*, vol. 3, pp. 1559-1695. Elsevier.
- Charles, Kerwin, Erik Hurst and Mariel Schwartz. 2018. "The Transformation Of Manufacturing and the Decline in U.S. Employment," in *NBER Macroeconomics Annual 2018*, vol. 33. University of Chicago Press.
- Dao, Mai, Davide Furceri and Prakash Loungani. 2017. "Regional Labor Market Adjustment in the United States: Trend and Cycle," *Review of Economics and Statistics*, 99, no. 2: 243-257.
- Dube, Arindrajit, Laura M. Giuliano and Jonathan S. Leonard. 2015. "Fairness and Frictions: The Impact of Unequal Raises on Quit Behavior," *American Economic Review*.
- Flinn, Christopher. 2006. "Minimum Wage Effects on Labor Market Outcomes under Search, Bargaining, and Endogenous Contact Rates," *Econometrica*, 74: 1013-1062.
- Friedman, M. 1968. "The Role of Monetary Policy," *American Economic Review*, 58 (1): 1-17.
- Furman, Jason. 2017. "The Real Reason You're Not Getting a Pay Raise," Vox. Aug. 11. Available at <https://www.vox.com/the-big-idea/2018/7/31/17632348/wages-lagging-inequality-income-recovery-recession-wage-puzzle-economics>
- Katz, Lawrence, and Alan Krueger. 1999. "The High-Pressure U.S. Labor Market of the 1990s," *Brookings Papers on Economic Activity*, no. 1, 1-87.
- Killingsworth, Mark R. 1983. *Labor Supply*. Cambridge: Cambridge University Press.
- Kleiner, M.M., and A.B. Krueger. 2013. "Analyzing the Extent and Influence of Occupational Licensing on the Labor Market," *Journal of Labor Economics*, 31(S1), pp. S173-S202.

- Krueger, Alan. 2018. "Where Have All the Workers Gone?: An Inquiry into the Decline of the U.S. Labor Force Participation Rate," *Brookings Papers on Economic Activity*, no. 2, 1- 87.
- _____, and Orley Ashenfelter. 2018. "Theory and Evidence on Employer Collusion in the Franchise Sector," National Bureau of Economic Research Working Paper w24831.
- _____, and Eric Posner. 2018. "A Proposal for Protecting Low-Income Workers from Monopsony and Collusion," The Hamilton Project Policy Proposal 5.
- Krugman, Paul. 2018. "Monopsony, Rigidity, and the Wage Puzzle," May 20. Available at <https://www.nytimes.com/2018/05/20/opinion/monopsony-rigidity-and-the-wage-puzzlewonkish.html>
- Manning, Alan. 2003. *Monopsony in Motion: Imperfect Competition in Labor Markets*, Princeton University Press.
- Okun, Arthur. 1973. "Upward Mobility in a High-Pressure Economy," *Brookings Papers on Economic Activity*, no. 1: 207-261.
- Powell, Jerome. 2018. Interview by Kai Ryssdal, Marketplace, July 12. Transcript available at <https://www.marketplace.org/2018/07/12/economy/powell-transcript>
- Shelkova, Natalya Y. 2014. "Low-Wage Labor Markets and the Power of Suggestion," June 3. Available at SSRN: <https://ssrn.com/abstract=2478219>
- Staiger, D.O., J. Spetz and C.S. Phibbs. 2010. "Is There Monopsony in the Labor Market? Evidence from a Natural Experiment," *Journal of Labor Economics*, 28(2), pp.211-236.
- Starr, Evan P., J.J. Prescott and Norman Bishara. 2017. "Noncompetes in the U.S. Labor Force." Available online at SSRN.
- Sullivan, Daniel. 1989. "Monopsony Power in the Market for Nurses," *The Journal of Law and Economics*, 32, no. 2, Part 2: S135-S178.
- Webber, Douglas A. 2015. "Firm Market Power and the Earnings Distribution," *Labour Economics*, 35: 123-134.

**BEFORE THE
FEDERAL TRADE COMMISSION
Washington, D.C. 20580**

Re: Petition for Rulemaking to Prohibit Worker Non-Compete Clauses

PETITION FOR RULEMAKING

BY

OPEN MARKETS INSTITUTE, AFL-CIO, ARTIST RIGHTS ALLIANCE, CENTER FOR POPULAR DEMOCRACY, COWORKER.ORG, DEMAND PROGRESS EDUCATION FUND, ECONOMIC POLICY INSTITUTE, EIG, INSTITUTE FOR LOCAL SELF-RELIANCE, LAKE RESEARCH PARTNERS, MAKE THE ROAD NEW YORK, NATIONAL EMPLOYMENT LAW PROJECT, ORGANIZATION UNITED FOR RESPECT, PUBLIC CITIZEN, REVOLVING DOOR PROJECT, ROOSEVELT INSTITUTE, SERVICE EMPLOYEES INTERNATIONAL UNION, TOWARDS JUSTICE, UFCW, AND UNITE HERE

AND IN THEIR INDIVIDUAL CAPACITIES

ALAN HYDE, AMY KASTELY, ANN C. MCGINLEY, ARIANA R. LEVINSON, BARBARA BUCHOLTZ, BEN TEMPLIN, CAROL CHOMSKY, CATHERINE FISK, CHARLOTTE GARDEN, CHRIS ODINET, CHRYSTIN ONDERSMA, COLIN P. MARKS, CYNTHIA HO, DALIÉ JIMÉNEZ, EDWARD JANGER, EILEEN APPELBAUM, FRANK PASQUALE, HENRY DRUMMONDS, JANE FLANAGAN, JEFFREY W. STEMPEL, JOAN VOGEL, KAREN CROSS, KATHLEEN ENGEL, LAUREN E. WILLIS, MARTHA T. MCCLUSKEY, MEREDITH A. MUNRO, NANCY MODESITT, NICOLAS CORNELL, ORLY LOBEL, PAMELA FOOHEY, PAUL SECUNDA, RACHEL ARNOW-RICHMAN, RICK BALES, RICHARD CARLSON, ROBERT H. LANDE, RUBEN J. GARCIA, SANJUKTA PAUL, SARA STERNBERG GREENE, SHAUHIN TALESH, SPENCER WEBER WALLER, SUSAN BLOCK-LIEB, TODD RAKOFF, V.B. DUBAL, VIVA R. MOFFAT, WILLIAM S. DODGE, AND YVETTE LIEBESMAN

Table of Contents

Petition for Rulemaking	1
Introduction	1
I. Overview of Non-Compete Clauses	5
<i>A. The Ubiquity of Non-Compete Clauses in Today’s Labor Markets</i>	6
<i>B. Law Governing Non-Compete Clauses</i>	9
II. Employers Can Generally Impose Non-Compete Clauses on Workers	13
<i>A. The Employee-Employer Relationship Is Defined Generally by Unequal Bargaining Power</i>	14
<i>B. Workers Are Unlikely to Bargain Over Non-Compete Clauses</i>	18
<i>C. Given These Power and Behavioral Dynamics in Labor Markets, Non-Compete Clauses Function as Contracts of Adhesion</i>	21
III. Harms to Workers from Non-Compete Clauses	25
<i>A. Impaired Job Mobility</i>	26
<i>B. Effects of Reduced Job Mobility</i>	32
IV. Monopolistic and Oligopolistic Businesses Can Use Non-Compete Clauses to Exclude or Limit Product Market Competition	37
V. The General Justification for Non-Compete Clauses Does Not Stand Up to Scrutiny .	39
<i>A. The Justification for Non-Compete Clauses Rests on a Questionable Premise</i>	40
<i>B. Non-Competes Are a Flawed Means of Protecting Against Free Riding</i>	44
<i>C. Employers Have Less Restrictive Alternatives to Protect Intangible Interests</i>	45
VI. The FTC Should Prohibit Non-Compete Clauses as an Unfair Method of Competition	49
VII. Conclusion	53
Certification	55
Organizational Petitioners	56
Individual Petitioners	59

Petition for Rulemaking

The Open Markets Institute, 19 labor and public interest organizations, and 46 individual advocates and scholars submit this petition pursuant to the Administrative Procedure Act and the Federal Trade Commission Act, 15 U.S.C. § 45 to request the Federal Trade Commission (“FTC”) to initiate a rulemaking to prohibit employers from presenting a non-compete clause to a worker (regardless of whether the worker is classified as an “employee” or an “independent contractor”), conditioning employment or the purchase of a worker’s labor on the worker’s acceptance of a non-compete clause, or enforcing, or threatening to enforce, a non-compete clause against a worker. Under this rule, the FTC could bring enforcement actions against employers and purchasers of labor services (collectively hereafter “employers”) who engage in any of the described conduct.

Introduction

Employers have deprived tens of millions of workers of their freedom to leave their current job to accept employment with another firm or to pursue a business opportunity. They have done so through non-compete clauses (hereafter “non-compete clauses,” “non-competes,” “non-compete conditions,” “non-compete contracts,” and “non-compete requirements” will be used interchangeably). Non-competes require workers, *following separation* from their current employer, to refrain from accepting employment in a similar line of work or establishing a competing business for a specified period in a certain geographic area. For instance, a home health aide’s employer required him to sign a non-compete clause that prohibited him from working for another home health firm or starting his own firm in the county where he lives for

two years after he left his current employer.¹ Approximately 30 million workers,² across a wide range of fields and occupations including camp counselors, engineers, fast food workers, hair stylists, investment managers, and yoga instructors,³ are bound by non-compete clauses.

In labor markets, employers typically have the power to impose non-compete clauses on workers. Workers, who often depend on wages to subsist, are usually at a significant disadvantage in their relationship with employers. In any local or regional labor market in which large numbers of workers are unemployed or underemployed, individual workers are in an especially weak position to negotiate their terms of employment. Only a small fraction of private sector workers belong to a union and can assert collective voice. Compounding these disadvantages of workers, most local labor markets in the United States are highly concentrated on the employer side. As a result, many millions of workers see little or no competition for their services among employers.

Even when employers do compete for their services, workers are likely to focus on bargaining over wages and benefits, *not* contingent terms such as non-compete clauses. Employers generally present non-compete clauses to workers as standard form documents on a take-it-or-leave-it basis. When a worker does notice the clause, he or she may be loath to question or challenge a document that is presented on a take-it-or-leave-it basis, out of fear that the employer might rescind the job offer. Taken together, these factors indicate that non-compete clauses operate as contracts of adhesion.

¹ Sophie Quinton, *These Days, Even Janitors Are Being Required to Sign Non-Compete Clauses*, U.S.A. TODAY, May 27, 2017, <https://www.usatoday.com/story/money/2017/05/27/noncompete-clauses-jobs-workplace/348384001/>.

² U.S. DEP'T OF TREASURY, NON-COMPETE CONTRACTS: ECONOMIC EFFECTS AND POLICY IMPLICATIONS 6 (2016).

³ Rachel Abrams, *'No Poach' Deals for Fast-Food Workers Face Scrutiny by States*, N.Y. TIMES, July 9, 2018, <https://www.nytimes.com/2018/07/09/business/no-poach-fast-food-wages.html>; Steven Greenhouse, *Noncompete Clauses Increasingly Pop Up in Array of Jobs*, N.Y. TIMES, June 8, 2014, <https://www.nytimes.com/2014/06/09/business/noncompete-clauses-increasingly-pop-up-in-array-of-jobs.html>.

Non-compete clauses inflict significant harms on workers. Non-competes bind workers to their current employers and reduce their mobility. Even when employers do not or cannot enforce non-competes, workers subject to a non-compete clause may be deterred from leaving their current job. Because most private-sector workers do not belong to a labor union and lack effective voice on the job, the threat of leaving for another job is often the only source of leverage for many workers. By restricting exit, non-competes rob workers of this power. Research generally finds that non-compete clauses depress wages and wage growth and deter the formation of new firms. The adverse effects on workers extend beyond wages and firm creation rates. Workers subject to non-compete clauses may be compelled to remain in discriminatory, hostile, or unsafe work environments.

Non-competes can also impair product market competition. In a highly concentrated market, monopolists and other powerful firms can use non-compete clauses to deprive rivals of essential workers and thereby impede their ability to compete. Through this strategic use of non-competes, dominant firms can weaken and exclude rivals and maintain market power. Even in the absence of exclusionary intent, non-competes can favor incumbent large firms over small or new entrants. Relative to small firms, larger firms are more likely to be able and willing to take the legal risk of recruiting and hiring workers bound by non-competes.

While they inflict real harms on workers and society, non-competes do not have a credible justification. Employers and their representatives argue that non-competes allow employers to protect intangible investments, such as trade secrets and employee training, from “free riding” by rival firms. If employers are unable to protect against this type of free riding, they will underinvest in intangibles, according to this story. But this rationale does not stand up to scrutiny. First, it assumes that the broad dissemination of information and knowledge is

generally bad for society. Second, even if used to prevent workers from sharing information, non-compete clauses are a flawed method of defending against free riding. They are too broad because they deprive workers of the freedom to use their full experience, knowledge, and skills. At the same time, they are also too narrow because they provide only porous protection to intangibles. Third, employers, insofar as they need to protect their investment in intangibles, have several less restrictive alternatives, such as trade secret law, improved employee retention policies, and employment contracts.

The FTC should initiate a rulemaking to prohibit employers from presenting non-compete clauses as a condition of employment or other work agreement or enforcing or threatening to enforce a non-compete against workers. The FTC has expansive authority to interpret the FTC Act's prohibition on "unfair methods of competition"⁴ and has the power to write competition regulations under the Administrative Procedure Act.⁵ Given that non-compete clauses inflict real harms on workers and competition and offer no credible offsetting benefits to society, they arguably should be presumptively illegal under the Sherman Act.⁶ Using its broader legal authority under Section 5 of the FTC Act, the FTC should hold these clauses to be an unfair method of competition. Even if employers do not or cannot enforce non-competes, these clauses

⁴ 15 U.S.C. § 45. *See* *FTC v. Ind. Fed. of Dentists*, 476 U.S. 447, 454 (1986) (internal citations omitted) ("The standard of 'unfairness' under the FTC Act is, by necessity, an elusive one, encompassing not only practices that violate the Sherman Act and the other antitrust laws, but also practices that the Commission determines are against public policy for other reasons[.]"); *FTC v. Sperry & Hutchinson Co.*, 405 U.S. 233, 244 (1972) (emphasis added) ("[L]egislative and judicial authorities alike convince us that the Federal Trade Commission does *not* arrogate excessive power to itself if, in measuring a practice against the elusive, but congressionally mandated standard of fairness, it, like a court of equity, considers public values beyond simply those enshrined in the letter or encompassed in the spirit of the antitrust laws.").

⁵ The Magnuson-Moss Warranty Act imposes special procedures on FTC rules on unfair or deceptive acts and practices but it does not apply to rules on unfair methods of competition. 15 U.S.C. § 57a(a)(2). As a result, FTC rules on unfair methods of competition are governed by the Administrative Procedure Act's general notice-and-comment requirements for rulemakings. *Nat'l Petroleum Refiners Ass'n v. FTC*, 482 F.2d 672 (D.C. Cir. 1973).

⁶ *In re Polygram Holding, Inc.*, 136 FTC 310, 344 (2003), *pet'n denied Polygram Holding, Inc. v. FTC*, 416 F.3d 29 (D.C. Cir. 2005) ("A plaintiff may avoid full rule of reason analysis, including the pleading and proof of market power, if it demonstrates that the conduct at issue is inherently suspect owing to its likely tendency to suppress competition.").

still chill worker mobility and bind employees to their current employer. Accordingly, the FTC should write a rule holding that the use or enforcement of non-compete clauses is a per se violation of Section 5 of the FTC Act and that employers presenting, enforcing, or otherwise using non-competes with workers are subject to FTC enforcement action.

I. Overview of Non-Compete Clauses

Non-compete clauses restrict a worker from competing, or working for a firm competing, with his or her former employer for a specified time.⁷ In concrete terms, they restrict a worker from accepting employment or starting a business (1) in a line of work or industry, (2) in a geographic area, and (3) for a period of time following his or her departure or other separation from his or her current employer. Employers, in the first instance, draft the occupational (or industrial), geographic, and temporal scope of the non-compete contract. For example, an accountant bound by a non-compete may be prohibited from working as an accountant for another firm, starting her own accounting firm for a year following separation from her current firm, or even working in a *non-accounting* role at another accounting firm. When a worker leaves a current job in violation of a non-compete clause or expresses an interest in doing so, the employer can bring, or threaten to bring, legal action against the employee for breach of contract.

State law is the principal authority governing non-compete clauses today. States differ greatly in their treatment of non-competes. A few states, notably California, prohibit the enforcement of non-compete clauses against all workers. Other states bar the enforcement of non-compete clauses against certain classes of workers. Most states generally permit employers to enforce non-compete clauses against workers. In these states, courts apply a “reasonableness” test to decide the enforceability of non-competes.

⁷ Non-competes that are used as part of a sale or transfer of a business are distinguishable from worker non-competes and are not addressed in this petition.

A. The Ubiquity of Non-Compete Clauses in Today's Labor Markets

Millions of workers are subject to non-compete requirements. While traditionally associated with highly paid corporate executives, non-competes today are found across a wide range of fields and types of work. A recent study estimated that nearly 30 million American workers, or approximately one-in-every-five, are bound by non-competes with their present employers.⁸ Almost 60 million workers, or approximately two out of five, have been bound by a non-compete at *some* point in their careers.⁹ In general, workers with higher earnings and more education are more likely to be subject to a non-compete clause. Among architects and engineers in the manufacturing sector, approximately 50% of workers are bound by a non-compete.¹⁰ Another study found that 70% of the reviewed firms used non-compete contracts with their top executives.¹¹

But other workers are by no means exempt from non-competes. Among workers without a bachelor's degree and earning less than \$40,000 per year, approximately 12% are working under a non-compete clause.¹² To illustrate how ubiquitous non-compete clauses are today: they have been imposed on, among others, arborists, fast food workers, drilling rig operators, factory managers, journalists, and teachers.¹³ Firms have also imposed non-competes on independent contractors whose services they have retained.¹⁴

⁸ Evan Starr, J.J. Prescott & Norman Bishara, *Noncompetes in the U.S. Labor Force* 14 (2018).

⁹ *Id.*

¹⁰ *Id.* at 48.

¹¹ Mark J. Garmaise, *Ties That Truly Bind: Noncompetition Agreements, Executive Compensation, and Firm Investment*, 27 J.L. ECON. & ORG. 376, 396 (2011).

¹² Starr et al., *supra* note 8, at 2-3.

¹³ Abrams, *supra* note 3; Conor Dougherty, *How Noncompete Clauses Keep Workers Locked In*, N.Y. TIMES, May 13, 2017, <https://www.nytimes.com/2017/05/13/business/noncompete-clauses.html>; Aruna Viswanatha, *Noncompete Agreements Hobble Junior Employees*, WALL ST. J., Feb. 2, 2016, <https://www.wsj.com/articles/noncompete-agreements-hobble-junior-employees-1454441651>; Rachel Cohen, *Fining Teachers for Switching Schools*, AM. PROSPECT, Nov. 3, 2016, <http://prospect.org/article/fining-teachers-switching-schools>.

¹⁴ See, e.g., *Ag Spectrum Co. v. Elder*, 865 F.3d 1088, 1093 (8th Cir. 2017) (refusing to enforce an agricultural supply firm's non-compete against a salesperson classified as an independent contractor); John Howley, *Justice for Janitors: The Challenge of Organizing in Contract Services*, 15 LAB. RES. REV. 60, 71 (1990) (janitors classified as

A few examples illustrate the ubiquity of non-compete clauses and the expansive scope of non-competes. Until through at least March 2015,¹⁵ Amazon required temporary warehouse workers to assent to very broad non-compete clauses.¹⁶ After leaving Amazon, workers subject to this clause could not work for a period of 18 months for another company that “engage[s] in or support[s] the development, manufacture, marketing, or sale of any product or service that competes with any product or service sold, offered, or otherwise provided by Amazon, or intended to be sold, offered, or otherwise provided by Amazon in the future.”¹⁷ Since Amazon bills itself as “the everything store” and accordingly sells a wide variety of products, this non-compete appeared to prohibit workers from a very broad set of employment options across the United States (and potentially around the world as well). The company conditioned severance pay on an employee affirming his or her commitment to honoring the non-compete clause.¹⁸

The sandwich chain Jimmy John’s has included¹⁹ a broad non-compete clause in the hiring packet given to store employees.²⁰ Under these clauses, Jimmy John’s workers could not work for a competing restaurant, defined as “any business which derives more than ten percent (10%) of its revenue from selling submarine, hero-type, deli-style, pita, and/or wrapped or rolled

“independent contractors were made to non-competes). For a history of the legal development of the “independent contractor” worker identity, see Veena B. Dubal, *Wage Slave or Entrepreneur: Contesting the Dualism of Legal Worker Identities*, 105 CALIF. L. REV. 65 (2017).

¹⁵ Lauren C. Williams, *Amazon Gets Rid of Strict Non-Compete Clause for Contract and Temporary Employees*, THINKPROGRESS, Mar. 30, 2015, <https://thinkprogress.org/amazon-gets-rid-of-strict-non-compete-clause-for-contract-and-temporary-employees-f7b12b94cfa9/>.

¹⁶ Spencer Woodman, *Exclusive: Amazon Makes Even Temporary Warehouse Workers Sign 18-Month Non-Competes*, VERGE, Mar. 26, 2015, <https://www.theverge.com/2015/3/26/8280309/amazon-warehouse-jobs-exclusive-noncompete-contracts>.

¹⁷ *Id.*

¹⁸ *Id.*

¹⁹ Jimmy John’s entered into settlements with the Illinois and New York attorneys general and committed to drop and not enforce non-competes against hourly employees in these states. It is unclear whether Jimmy John’s continues to use non-competes in other states. Samantha Bomkamp, *Jimmy John’s Agrees to Pay \$100,000 to Illinois AG Over Noncompete Contracts*, CHI. TRIB., Dec. 7, 2016, <http://www.chicagotribune.com/business/ct-jimmy-johns-settlement-1208-biz-20161207-story.html>.

²⁰ Aruna Viswanatha, *Sandwich Chain Jimmy John’s to Drop Noncompete Clauses from Hiring Packets*, WALL ST. J., June 21, 2016, <https://www.wsj.com/articles/sandwich-chain-jimmy-johns-to-drop-noncompete-clauses-from-hiring-packets-1466557202>.

sandwiches and which is located within three (3) miles of” *any* Jimmy John’s location.²¹ A lawsuit brought by the Illinois Attorney General against Jimmy John’s alleged that this restriction had a term for two years following an employee leaving Jimmy John’s.²² Because Jimmy John’s is a national chain with nearly 2,000 locations, the non-compete clauses potentially prohibited Jimmy John’s employees from working at a wide range of fast food restaurants across much of the United States.²³

The non-compete clause of payday lender Check Into Cash is another example of the expansive scope of non-competes. The Illinois Attorney General alleged that Check Into Cash required all storefront employees, including low-wage employees, to accept a non-compete clause that prohibited them from working for a wide range of “rivals” within 15 miles of any Check Into Cash location for one after leaving.²⁴ According to the Attorney General, Check Into Cash’s non-compete prohibited employment at not only payday lenders, title lenders, and pawn shops but possibly also “retail stores or auto dealerships that extend credit on an incidental basis or entities like Western Union or the U.S. postal service that transmit money.”²⁵

²¹ Dave Jamieson, *Jimmy John’s Makes Low-Wage Workers Sign ‘Oppressive’ Noncompete Agreements*, HUFFINGTON POST, Oct. 13, 2014, https://www.huffingtonpost.com/2014/10/13/jimmy-johns-non-compete_n_5978180.html.

²² Daniel Wiessner, *Jimmy John’s Settles Illinois Lawsuit Over Non-Compete Agreements*, REUTERS, Dec. 7, 2016, <https://www.reuters.com/article/us-jimmyjohns-settlement/jimmy-johns-settles-illinois-lawsuit-over-non-compete-agreements-idUSKBN13W2JA>.

²³ Viswanatha, *supra* note 20. For Jimmy John’s workers, non-compete contracts were not the only barrier to labor market mobility. Previously, franchisees, under their agreement with Jimmy John’s, were prohibited from hiring workers who worked for other Jimmy John’s franchisees. See Rachel Abrams, *7 Fast-Food Chains to End ‘No Poach’ Deals That Lock Down Low-Wage Workers*, N.Y. TIMES, July 12, 2018, <https://www.nytimes.com/2018/07/12/business/fast-food-wages-no-poach-deal.html>.

²⁴ Press Release, Ill. Att’y Gen., Attorney General Madigan Reaches Settlement with National Payday Lender for Imposing Unlawful Non-Compete Agreements (Jan. 7, 2019), http://illinoisattorneygeneral.gov/pressroom/2019_01/20190107b.html.

²⁵ *Id.*

B. Law Governing Non-Compete Clauses

Non-compete clauses are principally governed by state law today.²⁶ State law treatment of these clauses varies greatly and is complicated and muddled. An Ohio court described this patchwork of state law as “a sea – vast and vacillating, overlapping and bewildering. One can fish out of it any kind of strange support for anything, if he lives so long.”²⁷ Despite this welter of state law, states can be categorized as not enforcing non-competes for all workers, not enforcing non-competes for certain classes of workers, or permitting enforcement of non-compete clauses to varying degrees.²⁸ Only a few states categorically do *not* enforce non-compete clauses, while most states generally enforce them. Notwithstanding the law of the state where an employee resides, some employers have attempted to use “choice of law” provisions to import state law more favorable to the use and enforcement of non-compete contracts.²⁹

In California,³⁰ North Dakota,³¹ and Oklahoma,³² state law prohibits judicial enforcement of non-compete clauses. Employers in these states do not necessarily violate the law by conditioning employment on employee assenting to a non-compete clause. They, however, cannot enforce non-competes in court.³³ If an employee departs for another job or starts a

²⁶ The case law on non-competes dates back centuries and includes some of the earliest decisions on restraints of trade. *Mitchel v. Reynolds*, 24 Eng. Rep. 347 (Q.B. 1711). At the federal level, employees can challenge non-compete clauses as an unreasonable non-price vertical restraint under the Sherman Act. *E.g.*, *Butler v. Jimmy John’s Franchise, LLC*, 2018 WL 3631577 (S.D. Ill. 2018).

²⁷ *Arthur Murray Dance Studios of Cleveland v. Witter*, 105 N.E.2d 685, 687 (Ohio Ct. C.P. 1952).

²⁸ One business scholar has developed an index to capture how strongly or weakly a state enforces non-compete clauses. Norman D. Bishara, *Fifty Ways to Leave Your Employer: Relative Enforcement of Covenants Not to Compete, Trends and Implications for Employee Mobility Policy*, 13 U. PA. J. BUS. L. 751 (2011).

²⁹ Viva R. Moffat, *Making Non-Competes Unenforceable*, 54 ARIZ. L. REV. 939, 952-53 (2012).

³⁰ Cal. Bus. & Prof. Code § 16600.

³¹ N.D. Cent. Code § 9-08-06.

³² OK Stat. § 15-219A. Oklahoma law prohibits the enforcement of non-compete clauses so long as employees do not poach customers from their former employer.

³³ For example, California law holds that “[e]xcept as provided in this chapter, every contract by which anyone is restrained from engaging in a lawful profession, trade, or business of any kind is to that extent void.” Cal. Bus. & Prof. Code § 16600.

business in violation of a non-compete clause, employers in these states generally cannot go to court and seek to enforce this contract and obtain remedies.³⁴

Other states bar enforcement of non-compete clauses for certain classes of workers. For example, Illinois law prohibits employers from requiring workers making less than \$13 per hour to assent to non-compete clauses.³⁵ In 2015, Hawaii enacted a law that restricts the enforcement of non-competes against information technology professionals.³⁶ Several states have barred the enforcement of non-compete clauses against physicians.³⁷

Most states, including states that have limited prohibitions or restrictions, allow the enforcement of non-compete clauses. Courts in these states determine the legality of non-compete clauses using a “reasonableness” framework. While the precise formulation of this reasonableness standard varies, many states have adopted the analytical test in the Restatement (Second) of Contracts.³⁸ Under the Restatement’s test, a non-compete clause ancillary to a valid agreement is “*unreasonably* in restraint of trade if (1) the restraint is greater than is needed to protect the business and goodwill of the employer; or (2) the promisee’s need is outweighed by the hardship to the promisor and the likely injury to the public.”³⁹

The test in the Second Restatement raises several legal issues for courts to decide.⁴⁰ First, courts decide whether a valid agreement, to which the non-compete clause is ancillary, exists. If

³⁴ See, e.g., *Edwards v. Arthur Andersen LLP*, 44 Cal.4th 937, 955 (2008) (“Noncompetition agreements are invalid under section 16600 in California, even if narrowly drawn, unless they fall within the applicable statutory exceptions of sections 16601, 16602, or 16602.5.”).

³⁵ 820 ILCS 90.

³⁶ Haw. Rev. Stat. § 480-4.

³⁷ E.g., RSA 239:31-a (New Hampshire); N.M.S.A. §§ 24-11-1-5; Tex. Bus. & Com. Code §§ 15.50-52.

³⁸ *Griffin Toronjo Pivateau, Putting the Blue Pencil Down: An Argument for Specificity in Noncompete Agreements*, 86 NEB. L. REV. 672, 677-678 (2008). E.g., *Summits 7, Inc. v. Kelly*, 178 Vt. 396, 399 (2005); *BDO Seidman v. Hirshberg*, 93 N.Y.2d 382, 388-89 (1999).

³⁹ RESTATEMENT (SECOND) OF CONTRACTS § 188(1) (1981) (emphasis added).

⁴⁰ One court has stated that determining the enforceability of a non-compete clause requires answering 41 distinct questions. *Arthur Murray*, 105 N.E.2d at 695-99.

a non-compete clause is not ancillary to a valid agreement, it is per se illegal.⁴¹ Courts have taken two general approaches to what constitutes a “valid agreement” in the employment relationship.⁴² Some state courts hold that continued at-will employment is a valid agreement, to which a non-compete clause can be ancillary.⁴³ In an at-will employment arrangement, either the employer or the employee can end the relationship for practically any or no reason at all.⁴⁴ Courts in other states hold that the promise of continued at-will is not a valid agreement because the employer has not bound itself in any way and retains the option to terminate the employee immediately without cause.⁴⁵ In these states, an employer must bind itself in a tangible way in exchange for the employee assenting to a non-compete clause. Under this stricter approach, employers who agree to terminate workers only for “just cause” have likely created a valid agreement to which a non-compete clause can be ancillary.⁴⁶

Second, for non-compete clauses that are ancillary to a valid agreement, courts determine whether the scope of the non-compete is appropriately tailored to protect the employer’s legitimate interests. State law varies on what qualifies as a legitimate interest. A business’s legitimate interests can include customer relationships, trade secrets, and other intangibles.⁴⁷

⁴¹ *E.g.*, *Freiburger v. J-U-B Engineers, Inc.*, 141 Idaho 415, 419-20 (2005).

⁴² Tracy L. Staidl, *The Enforceability of Noncompetition Agreements When Employment Is At-Will: Reformulating the Analysis*, 2 EMPLOYEE RTS. & EMP. POL’Y J. 95, 104-05 (1998).

⁴³ *See, e.g.*, *Summits* 7, 178 Vt. at 404 (“A noncompetition agreement presented to an employee at any time during the employment relationship is ancillary to that relationship and thus requires no additional consideration other than continued employment.”).

⁴⁴ Jay M. Feinman, *The Development of the Employment at Will Rule*, 20 AM. J. LEG. HIST. 118 (1976).

⁴⁵ *E.g.*, *Rogers v. Runfola & Assoc., Inc.*, 565 N.E.2d 540, 542 (Ohio 1991).

⁴⁶ For instance, state law in Texas previously held that “[a]n ‘employment-at-will’ relationship is not binding upon either the employee or the employer. Either may terminate the relationship at any time. Thus, an employment-at-will relationship, although valid, is not an otherwise enforceable agreement.” *Travel Masters, Inc. v. Star Tours, Inc.*, 827 S.W.2d 830, 832-33 (Tex. 1991), superseded by statute as stated in *Alex Sheshunoff Mgmt Servs. L.P. v. Johnson*, 209 S.W.3d 644, 653 (Tex. 2006).

⁴⁷ Cynthia L. Estlund, *Between Rights and Contract: Arbitration Agreements and Non-Compete Covenants as a Hybrid Form of Employment Law*, 155 U. PA. L. REV. 379, 393 (2006). For an analysis of legitimate interests in the context of media employees, see Cathy Packer & Johanna Cleary, *Rediscovering the Public Interest: An Analysis of the Common Law Governing Post-Employment Non-Compete Contracts for Media Employees*, 24 CARDOZO ARTS & ENT. L.J. 1073, 1095-1105 (2006).

Notably, restraining competition is *not* a legitimate business interest.⁴⁸ When deciding whether a non-compete clause's scope is acceptable, courts analyze whether the covered activity, space, or time of the non-compete is overbroad and beyond what is necessary to protect the employer's legitimate interests.⁴⁹

Third, courts examine the offsetting harms to the bound employees and the public. They examine whether the employee has reasonable alternative employment options that are not proscribed by the non-compete clause. On occasion, courts have deemed jobs with significantly lower wages or located at a great distance from the employee's residence not to be reasonable alternatives.⁵⁰ Courts also sometimes examine the effect of enforcing the non-compete clause on broader market competition. Specifically, they consider whether enforcement of the non-compete would deprive rivals of skilled labor and impair their ability to compete.⁵¹

States that permit the enforcement of non-compete clauses take one of three approaches to non-competes that violate their reasonableness test.⁵² A few states follow the so-called "red pencil" doctrine. Under this doctrine, courts deem a noncompete clause with any overbroad provision to be unenforceable in its entirety.⁵³ Other states apply the "blue pencil" doctrine. In blue pencil states, courts strike any overbroad provisions and, if the remaining non-compete is

⁴⁸ See, e.g., *Marsh USA Inc. v. Cook*, 354 S.W.3d 764, 770 (Tex. 2011) (internal citations omitted) ("Where the object of both parties in making such a contract is merely to restrain competition, and enhance or maintain prices, there is no primary and lawful purpose of the relationship to justify or excuse the restraint.").

⁴⁹ E.g., *Golden Rd. Motor Inn Inc. v. Islam*, 376 P.3d 151, 155-56 (Nev. 2016). See also *Pivateau*, *supra* note 38, at 677-81.

⁵⁰ E.g., *Marinelli v. Medco Health Solutions, Inc.*, 951 F.Supp.2d 303, 320-21 (D. Conn. 2013); *Bennett v. Storz Broadcasting Co.*, 270 Minn. 525, 537 (1965).

⁵¹ E.g., *Weissman v. Transcon. Printing U.S.A., Inc.* 205 F.Supp.2d 415, 426-27 (E.D. Pa. 2002). Some courts have also stated that the public interest analysis includes determining whether enforcing the non-compete clause would compel the unemployed worker to seek public assistance. E.g., *Basicomputer Corp. v. Scott*, 791 F.Supp. 1280, 1291 (N.D. Ohio 1991).

⁵² WHITE HOUSE, NON-COMPETE AGREEMENTS: ANALYSIS OF THE USAGE, POTENTIAL ISSUES, AND STATE RESPONSES 11 (2016).

⁵³ E.g., *Star Direct, Inc. v. Dal Pra*, 319 Wis.2d 274, 313 (2009); *Ward v. Process Control Corp.*, 247 Ga. 583, 584 (1981).

still valid, enforce this revised version of the non-compete clause.⁵⁴ Finally, in another group of states (“reformation states”), courts rewrite overbroad non-compete clauses and enforce the “reformed” clause.⁵⁵ The red pencil doctrine promotes careful drafting by employers because overbroad non-competes will be held as categorically unenforceable. In contrast, the blue pencil doctrine and reformation permit, and even may encourage, employers to draft overbroad non-compete clauses.⁵⁶

II. Employers Can Generally Impose Non-Compete Clauses on Workers

Non-compete clauses function as contracts of adhesion instead of products of free bargaining between employees and employers. In general, the employee-employer relationship is defined by inequality. Tens of millions of Americans generally have nothing to sell but their labor and skills and so must work to subsist. Even in this nominally full employment economy, millions of Americans are unemployed or underemployed so competition among workers is intense in many labor markets. Since less than one-in-ten private sector workers belongs to a union, most workers have no means of asserting collective voice against their employer and must engage with their employer on an individual basis. To compound the weak position of workers, the employer side of most local labor markets is highly concentrated. This limits employment options for millions of workers and restricts their ability to bargain for better terms of employment.

Competition for workers, to the extent it exists, cannot be counted on to police the terms of employment. In situations in which employers do compete to hire employees, bargaining and

⁵⁴ *E.g.*, *Valley Med. Specialists v. Farber*, 982 P.2d 1277, 1286 (Ariz. 1999); *Reddy v. Cmty. Health Found. of Man*, 298 S.E.2d 906, 915 (W. Va. 1982).

⁵⁵ *E.g.*, *Hillard v. Medtronic. Inc.*, 910 F. Supp. 173, 177 (M.D. Pa. 1995).

⁵⁶ Pivateau, *supra* note 38, at 689-91. *See also* White House, *supra* note 52, at 11 (“[S]ome states [under the red pencil doctrine] provide disincentives for employers to write non-compete contracts that are unenforceable by refusing to enforce and making void a non-compete contract that contains any unenforceable provisions.”).

competition are likely to focus on salient features such as wages and benefits. In contrast, competition and bargaining are not likely to occur over terms contingent on future events such as non-compete clauses.

A. The Employee-Employer Relationship Is Defined Generally by Unequal Bargaining Power

Individual workers generally enter employment relationships in a highly unequal position relative to their employers. Most workers do not have significant sources of non-wage income and must work to meet their basic needs. Competition among workers is generally intense, with millions of workers on standby due to unemployment or underemployment. Other features of American labor markets strengthen employers and weaken workers. The default rule of at-will employment, low rates of unionization, and employer-side concentration in many local labor markets further disempower workers.

A large fraction of workers have only their labor to sell and have no other significant source of income.⁵⁷ According to a survey conducted by the Federal Reserve Board of Governors, only 19% of individuals between 30 and 39 and 26% of individuals between 40 and 49 receive *any* income from interest, dividends, or rental property.⁵⁸ Furthermore, many Americans do not have the savings to meet a modest emergency expense. In 2017, 41% of Americans stated that they couldn't afford a \$400 emergency expense without borrowing or selling assets (down from 50% in 2013).⁵⁹ A 2014 survey found that more than 60% of Americans did not have the savings to pay a \$1,000 emergency medical expense.⁶⁰ And even in

⁵⁷ See *Arthur Murray*, 105 N.E.2d at 704 (“The average, individual employee has little but his labor to sell or to use to make a living. He is often in urgent need of selling it and in no position to object to boiler plate restrictive covenants placed before him to sign. To him, the right to work and support his family is the most important right he possesses. His individual bargaining power is seldom equal to that of his employer.”).

⁵⁸ BD. OF GOV. OF THE FED. RES. SYS., REPORT ON THE ECONOMIC WELL-BEING OF U.S. HOUSEHOLDS IN 2017 12 (2018).

⁵⁹ *Id.* at 21.

⁶⁰ Neal Gabler, *The Secret Shame of Middle-Class Americans*, ATLANTIC (2016).

the absence of an emergency expense, more than 20% of Americans are unable to pay in full their monthly obligations on, among other items, housing, utilities, or credit cards.⁶¹

In the United States, the default of at-will employment further tilts the balance of power in favor of employers. Employers can dismiss workers for any reason, or no reason at all, under the prevailing at-will employment regime.⁶² Statutes such as the Civil Rights Act and National Labor Relations Act have qualified the rights of employers to terminate workers. For example, employers cannot dismiss workers based on race or gender or for labor organizing activities.⁶³ Nonetheless, under at-will employment, employers have broad discretion to dismiss workers or modify the terms of employment. Given the at-will employment rule, employers have the freedom to condition new or continued employment on the acceptance of their terms.⁶⁴

While the corollary to at-will *termination* for employers is at-will *resignation* for workers, this does not necessarily create functional equality between workers and employers. Except arguably for the very smallest firms, an individual worker is much more dependent on the employment relationship than the employer. The employer has other employees and can hire new employees or use other employees to make up for one worker's rejection of a job offer or resignation. Many workers, however, depend on one job for most, or all, of their income. In the words of Samuel Issacharoff, "the hiring stage is most like a first date between a polygamist [the employer] and a monogamist [the employee]."⁶⁵

The structure of labor markets weakens the position of workers and strengthens the position of employers. In most labor markets, workers compete aggressively against each other.

⁶¹ BD. OF GOVERNORS, *supra* note 58, at 22.

⁶² See generally Clyde W. Summers, *Employment at Will in the United States: The Divine Right of Employers*, 3 U. PA. J. LAB. & EMP. L. 65 (2000).

⁶³ 42 U.S.C.A. § 2000a; 29 U.S.C.A. § 158.

⁶⁴ Rachel Arnow-Richman, *Modifying At-Will Employment Contracts*, 57 B.C. L. REV. 427, 431 (2016).

⁶⁵ Samuel Issacharoff, *Contracting for Employment: The Limited Return of the Common Law*, 74 TEX. L. REV. 1783, 1795 (1996).

Even in an economy considered to be at full employment, millions of Americans are involuntarily unemployed or underemployed.⁶⁶ With the rise of gig platforms, some employers and other purchasers of labor participate in a global labor pool and can hire from anywhere in the world.⁶⁷ Geographic markets are, in effect, worldwide for *purchasers* of certain types of labor. This vigorous competition among workers can lead to an erosion in wages and labor market standards. While certain employers can hire from anywhere in the world, the labor market's geographic scope for workers is typically narrower, with workers generally looking for employment near their place of residence.⁶⁸

Low union density also limits the bargaining power of workers. Through unionization, workers band together and exercise collective power in negotiating with employers. Whereas an individual worker is typically dispensable and wields little power against an employer, a group of workers can exercise significant power, including by threatening to strike and disrupt the employer's business. Through unions, workers can obtain, for instance, higher wages⁶⁹ and greater job security.⁷⁰ While unions once represented nearly a third of American workers, only 10.5% of all workers belonged to a union in 2018.⁷¹ In the private sector, fewer than 7% of

⁶⁶ Broader unemployment measures and labor force participation rates indicate that the U.S. economy is still far from true full employment. *See* Bureau of Labor Stats., Table A-15. Alternative Measures of Labor Underutilization, <https://www.bls.gov/news.release/empsit.t15.htm>; Bureau of Labor Stats., Labor Force Statistics from the Current Population Survey, <https://data.bls.gov/timeseries/LNS12300000>.

⁶⁷ Alana Semuels, *The Online Gig Economy's 'Race to the Bottom'*, ATLANTIC, Aug. 31, 2018, <https://www.theatlantic.com/technology/archive/2018/08/fiverr-online-gig-economy/569083/>.

⁶⁸ *See* Ioana Marinescu & Roland Rathelot, *Mismatch Unemployment and the Geography of Job Search 2* (2014), <http://www.sole-jole.org/15260.pdf> (“Documenting the geography of job search, we find that job seekers are more likely to apply to jobs closer to home: for example, a job seeker is 50% as likely to apply to a vacancy that is 25 miles away relative to a vacancy that is in the job seekers’ own zip code. Still, we find that 16% of applications are to out-of-state vacancies.”). Notably, on online job platforms, workers display a great deal of inertia in switching between employers. *See* Arindrajit Dube et al., *Monopsony in Online Labor Markets* 14-15 (Nat’l Bureau of Econ. Res. Working Paper No. 24,416, Mar. 2018), <https://www.nber.org/papers/w24416> (finding low elasticity of labor supply among workers on Amazon’s MTurk platform).

⁶⁹ *E.g.*, Henry Farber et al., *Unions and Inequality Over the Twentieth Century: New Evidence from Survey Data* 43 (2018), http://tuvalu.santafe.edu/~snaidu/papers/union_sub3.pdf.

⁷⁰ Summers, *supra* note 62, at 77.

⁷¹ Bureau of Labor Stats., Union Members Summary, <https://www.bls.gov/news.release/union2.nr0.htm>.

workers are members of a union today.⁷² Furthermore, an important segment of the labor force— independent contractors⁷³—is not protected by federal labor law and may face antitrust investigations and lawsuits if they attempt to bargain collectively and engage in other concerted activity.⁷⁴ Because most workers are not represented by a union, they meet employers on an individual, not a collective, basis.

On the employer side, most local labor markets are highly concentrated, as defined by the Department of Justice and Federal Trade Commission’s Horizontal Merger Guidelines.⁷⁵ Labor markets in smaller cities and rural areas are most likely to be concentrated.⁷⁶ In a significant number of labor markets, workers have only one actual or prospective employer—a true monopsony.⁷⁷ Because of this concentration among employers, millions of American workers have only one or a few employment options.⁷⁸

Long-term and recent wage trends are consistent with the general powerlessness of American workers and the structural advantage enjoyed by employers. Median worker productivity has risen greatly since the late 1970s. During the same time, however, median wages have stagnated. Between 1973 and 2017, median worker productivity increased by 77.4%,

⁷² *Id.*

⁷³ Lauren Weber, *The Second-Class Office Workers*, WALL ST. J., Sep. 14, 2017, <https://www.wsj.com/articles/the-contractors-life-overlooked-ground-down-and-stuck-1505400087>.

⁷⁴ See, e.g., *FTC v. Superior Court Trial Lawyers Assn.*, 493 U.S. 411, 436 (1990) (holding that lawyers’ boycott of Washington, D.C. public defender service is per se illegal). See generally Sanjukta M. Paul, *The Enduring Ambiguities of Antitrust Liability for Worker Collective Action*, 47 LOY. U. CHI. L.J. 969 (2016).

⁷⁵ José Azar, Ioana Marinescu, Marshall Steinbaum & Bledi Taska, *Concentration in U.S. Labor Markets: Evidence from Online Vacancy Data 1* (2018), <https://www.econstor.eu/bitstream/10419/177183/1/dp11379.pdf>. Other research has found that employer-side concentration in local labor markets has increased since the 1970s. Efraim Benmelech, Nittai Bergman & Hyunseob Kim, *Strong Employers and Weak Employees: How Does Employer Concentration Affect Wages?* 3 (2018), https://www.kellogg.northwestern.edu/faculty/benmelech/html/BenmelechPapers/BBK_2018_January_31.pdf.

⁷⁶ Azar et al, *supra* note 75, at 12.

⁷⁷ José Azar, Ioana E. Marinescu & Marshall Steinbaum, *Labor Market Concentration A.2* (2017), <https://www.econstor.eu/bitstream/10419/177058/1/dp11254.pdf>. This employer-side concentration depresses wages. See *id.* at 2 (“Going from the 25th to the 75th level of concentration decreases posted wages by 17% in the baseline IV specification, and by 5% in the baseline OLS specifications.”).

⁷⁸ See Azar et al, *supra* note 75, at 2 (“When we weight markets by [Bureau of Labor Statistics’] total employment, we find that 17 percent of workers work in highly concentrated labor markets[.]”).

while hourly pay of the typical worker rose by 12.4%.⁷⁹ In other words, workers have received just a small share of the rewards from their increased productivity. Even in this current period of nominally full employment, median wages have only grown modestly.⁸⁰

Against this background of weak employees and strong employers, workers are unlikely to be able to avoid non-compete clauses. Due to their economic situation and the structure of labor markets in the United States, most workers have little leverage in the hiring context. Consequently, they typically must accept the terms of employment presented to them by employers. Indeed, workers' initial and continued employment generally are contingent on their acceptance of the employer's terms. Under these circumstances, most workers must acquiesce to an employer's insistence on a non-compete clause.

B. Workers Are Unlikely to Bargain Over Non-Compete Clauses

Even for individual workers who have some power, competition and bargaining are not likely to discipline employers' use of non-compete clauses. Insofar as workers can and do negotiate individually over terms, this bargaining is likely to center on the immediate terms of employment such as hours, wages, and benefits. In contrast, however, workers are much less likely to be aware, or take notice, of other employment terms, especially those contingent on a future event such as resignation or termination from a job. Given these biases, any competition among employers for workers likely focuses on the most salient dimensions of employment, such as wages and benefits, and not on less salient terms such as non-compete clauses.

Behavioral research has examined how individuals make decisions in general and in the face of uncertainty in particular. Whether as consumers or employees, individuals have limited

⁷⁹ Econ. Pol'y Inst., *The Productivity-Pay Gap* (2018), <https://www.epi.org/productivity-pay-gap/>.

⁸⁰ Ernie Tedeschi, *Unemployment Looks Like 2000 Again. But Wage Growth Doesn't.*, N.Y. TIMES UPSHOT, Oct. 22, 2018, <https://www.nytimes.com/2018/10/22/upshot/mystery-slow-wage-growth-economy.html>.

time and interest to evaluate all the terms in an economic relationship or transaction. Studies have found that individuals when shopping or bargaining for a product or service focus on certain terms and neglect other terms.⁸¹ Due to this “bounded rationality,” in which individuals do not consider the full universe of terms, bargaining and competition are likely to center on the salient terms and be weak or non-existent for the non-salient terms.⁸²

Empirical research has documented bounded rationality in several areas. In a modern economy, individuals engage in many economic transactions, each with its own set of price and non-price attributes. Given this informational overload, individuals are likely to simplify their decision-making by focusing on a few salient terms in each transaction. In the context of most consumer contracts:

[T]he close reading and comparison needed to make an intelligent choice among alternative forms seems grossly arduous. Moreover, many of the terms concern risks that in any individual transaction are unlikely to eventuate. It is notoriously difficult for most people, who lack legal advice and broad experience concerning the particular transaction type, to appraise these sorts of contingencies.⁸³

For example, a survey by the Consumer Financial Protection Bureau found that customers looking for a credit card showed most interest in a card’s fees, interest rate, and rewards and the reputation of the issuer.⁸⁴ In contrast, they were generally either uninterested in

⁸¹ Russell Korobkin, *Bounded Rationality, Standard Form Contracts, and Unconscionability*, 70 U. CHI. L. REV. 1203 (2003); Robert A. Hillman & Jeffrey J. Rachlinski, *Standard-Form Contracting in the Electronic Age*, 77 N.Y.U. L. REV. 429 (2002).

⁸² For example, considering just one common term, one study estimated that “reading privacy policies carries cost in time of approximately 201 hours a year, worth about \$3,534 annually per American Internet user.” Aleecia M. McDonald & Lorrie Faith Cranor, *The Cost of Reading Privacy Policies*, 4 ISJLP 543, 565 (2008).

⁸³ Todd D. Rakoff, *Contracts of Adhesion: An Essay in Reconstruction*, 96 HARV. L. REV. 1173, 1226 (1983).

⁸⁴ CONSUMER FIN. PROTECTION BUREAU, ARBITRATION STUDY § 3.4.1 (2015), https://files.consumerfinance.gov/f/201503_cfpb_arbitration-study-report-to-congress-2015.pdf.

or unaware of whether account-related disputes with the card issuer would be subject to mandatory arbitration.⁸⁵

The FTC has implicitly recognized the effects of bounded rationality in its consumer protection rulemaking. In the Credit Practices Rule, the FTC prohibited certain default remedies in credit contracts, including confessions of judgment and taking a security interest in a borrower's existing household goods, as an "unfair act or practice."⁸⁶ The FTC reasoned that competition cannot be expected to restrain creditors' use of these remedies and so consumers could not reasonably avoid these remedies. The FTC stated that "[b]ecause remedies are relevant only in the event of default, and default is relatively infrequent, consumers reasonably concentrate their search on such factors as interest rates and payment terms."⁸⁷ The FTC concluded that in the market for consumer credit competition disciplined lenders on certain loan terms, but not on others such as a creditor's remedies in the event of a borrower's delinquency or default.

Given the boundaries on human rationality, workers likely focus on immediate terms of employment and devote little, if any, attention to contingent terms. As Cynthia Estlund has written, "Arbitration and non-compete agreements constrain employees only in a fairly remote and uncertain future event; and we may expect employees to overdiscount the likelihood of these events or the importance of the rights at stake."⁸⁸

⁸⁵ *Id.* Research indicates that competition in the credit card market is heavily driven by behavioral biases. Despite a very large number of credit card issuers, interest rates on balances are "sticky" and insensitive to underlying changes in interest rates. The failure of competition to lower interest rates suggests that, when shopping for cards, "many consumers are insensitive to interest-rate differentials because they believe they will pay within the grace period (although they repeatedly fail to do so)." Lawrence M. Ausubel, *The Failure of Competition in the Credit Card Market*, 81 AM. ECON. REV. 50, 75-76 (1991).

⁸⁶ 16 C.F.R. § 444.2.

⁸⁷ Credit Practices, 49 FED. REG. 7740, 7744 (Mar. 1, 1984).

⁸⁸ Estlund, *supra* note 47, at 413.

Considering these behavioral limitations, bargaining and competition cannot be expected to discipline employers' use of non-compete clauses with their workers. In the context of non-competes, the Treasury Department described the potential effect of behavioral biases in a 2016 report:

[W]orkers do not pay attention to non-compete contracts and do not realize how much bargaining power and future employment flexibility they are foregoing. Only later, when workers consider exiting a firm, do they become aware of the existence and/or implications of the non-compete agreement.⁸⁹

To the degree workers can and do bargain with employers and “shop around” for the best job, they are unlikely to focus on the existence of a non-compete clause. Even when they have some power at the hiring stage, workers are still likely vulnerable to employers using non-compete clauses to restrain their future mobility in labor markets.

C. Given These Power and Behavioral Dynamics in Labor Markets, Non-Compete Clauses Function as Contracts of Adhesion

Non-compete clauses function as contracts of adhesion. Todd Rakoff, in his seminal article on the topic, defined contracts of adhesion as documents that are drafted as standard forms by one party and presented to the other party on a take it-or-leave it basis.⁹⁰ Through contracts of

⁸⁹ U.S. DEP'T OF TREASURY, *supra* note 2, at 9.

⁹⁰ Rakoff, *supra* note 83, at 1177. He also provides a more precise seven-part definition. He defines a contract of adhesion as follows: “(1) The document whose legal validity is at issue is a printed form that contains many terms and clearly purports to be a contract. (2) The form has been drafted by, or on behalf of, one party to the transaction. (3) The drafting party participates in numerous transactions of the type represented by the form and enters into these transactions as a matter of routine. (4) The form is presented to the adhering party with the representation that, except perhaps for a few identified items (such as the price term), the drafting party will enter into the transaction only on the terms contained in the document. This representation may be explicit or may be implicit in the situation, but it is understood by the adherent. (5) After the parties have dickered over whatever terms are open to bargaining, the document is signed by the adherent. (6) The adhering party enters into few transactions of the type represented by the form - few, at least, in comparison with the drafting party. (7) The principal obligation of the adhering party in the transaction considered as a whole is the payment of money.” *Id.*

adhesion, corporations can reallocate legal rights and duties and structure legal environments to their own advantage. Due to power dynamics in labor markets and behavioral biases among workers, market competition is unlikely to discipline employers' use of non-compete requirements. Scholars studying non-compete clauses have written that "a reasonable interpretation of the data is that non-competes are typically a take-it-or-leave-it proposition."⁹¹ Non-competes can be included in an employment agreement or tucked inside an employee handbook (along with many other documents) whose unqualified acceptance is a condition of employment.⁹² In practice, non-compete clauses bear a close resemblance to mandatory arbitration clauses in that they deprive workers of future rights contingent on certain events. Instead of "memorializ[ing] a negotiated set of terms," employers use non-competes "to extract waivers of rights, thus realigning statutory and default rules to better reflect employers' interests."⁹³

Employees generally lack bargaining power and so they have little choice to accept terms of employment, such as non-compete clauses. Most workers are at a systematic power disadvantage relative to employers. First, they depend on labor income to subsist and often have limited resources. Second, most workers do not belong to a union and cannot exercise collective voice in negotiating with employers. Third, many workers today are in highly concentrated local labor markets and so have only a small set of potential employers.

Some employers further tilt the power imbalance in their favor by delaying presentation of the non-compete clause to workers. They withhold the non-compete until after a worker has

⁹¹ Starr et al., *supra* note 8, at 20.

⁹² Kenneth G. Dau-Schmidt & Timothy A. Haley, *Governance of the Workplace: The Contemporary Regime of Individual Contract*, 28 COMP. LAB. L. & POL'Y J. 313, 318, 345 (2007).

⁹³ Rachel Arnow-Richman, *Cubewrap Contracts: The Rise of Delayed Term, Standard Form Employment Agreements*, 49 ARIZ. L. REV. 637, 639 (2007).

accepted an offer of employment or commenced employment. At this stage, workers have rejected potential alternative employment opportunities and their continued employment (in a world of at-will employment) is often contingent on their acceptance of the non-compete.⁹⁴ One survey found that 70% of respondents were “asked to sign the non-compete *after* accepting the offer.”⁹⁵ Like shrink wrap standard form contracts whose terms the consumer can only learn of after purchasing and opening a product, workers learn of these non-competes only after they have committed to a job, or even begun employment, with a company.⁹⁶

Market competition cannot be expected to discipline employers’ use of non-compete clauses. Behavioral biases ensure that competition among employers in recruiting workers, when and where it does exist, is not likely to constrain the use of non-compete clauses. To the extent workers face competition for their services and can negotiate terms of employment, their bargaining is likely to focus on wages, benefits, hours, and other non-contingent aspects of the employment relationship. In contrast, they are likely to discount or ignore contingent terms such as non-compete clauses that may affect them in the future. Due to these biases, competition among employers is likely to focus on wages and benefits and not on contingent terms such as non-compete clauses.

Because non-compete clauses function as contracts of adhesion in an environment characterized by power disparities and behavioral biases, employers have broad power to impose these restrictive agreements on workers. The conditions of contractual formation are very different from the textbook theory of contract in which bargaining and negotiation are

⁹⁴ Matt Marx & Lee Fleming, *Non-Compete Agreements: Barriers to Entry . . . and Exit?*, 12 INNOV. POL’Y & ECON. 39, 49 (2012).

⁹⁵ *Id.* (emphasis added).

⁹⁶ Rachel Arnow-Richman, *Cubewrap Contracts and Worker Mobility: The Dilution of Employee Bargaining Power Via Standard Form Noncompetes*, 2006 MICH. ST. L. REV. 963, 977-80.

preconditions of reaching agreement.⁹⁷ Contracts of adhesion permit standardization and obviate the need for costly and time-consuming individual bargaining. They, however, create significant risks for consumers and workers. Under a legal system that generally enforces these terms, the drafting party has great power to reassign legal duties and rights in its own favor and against the receiving party.⁹⁸ Employers can use non-compete clauses to restrict the mobility of their workers, often without assuming any legal duty or responsibility in return.

Empirical research supports treating non-compete clauses as contracts of adhesion. Most workers read and sign the non-compete condition without attempting to bargain or negotiate. In a representative survey of workers, only approximately 10% of individuals presented with a non-compete sought to modify the terms of the non-compete or request benefits in exchange for signing it.⁹⁹ Fewer than one in five consulted a lawyer over the non-compete clause, and this consultation of a lawyer is strongly correlated with efforts to negotiate around the non-compete clause.¹⁰⁰ Most workers believed that their hiring was contingent on their signing the non-compete.¹⁰¹ Only 11% of workers thought they would still be hired if they refused to sign the non-compete clause.¹⁰²

Other evidence suggests that even highly educated, highly-paid workers do not bargain over, or attempt to resist, employers' non-compete clauses. A study of automated speech recognition experts found that, across age and tenure groups, at least 85% of study participants agreed to a non-compete clause, with the figure at or above 94% for the youngest and least

⁹⁷ MARGARET JANE RADIN, *BOILERPLATE: THE FINE PRINT, VANISHING RIGHTS, AND THE RULE OF LAW* 3-18 (2013).

⁹⁸ *Id.* at 33-34.

⁹⁹ Starr et al., *supra* note 8, at 18-19.

¹⁰⁰ *Id.* at 19.

¹⁰¹ *Id.*

¹⁰² *Id.*

experienced workers.¹⁰³ In other words, regardless of age or experience, most study participants accepted non-compete clauses. When presented with a non-compete clause before starting employment, fewer than one in six in the sample group asked a lawyer to review the non-compete.¹⁰⁴ This figure fell to fewer than one in twenty when employers presented the non-compete on the first day of work.¹⁰⁵ And not a single worker in the study group bargained over a non-compete.¹⁰⁶

Non-compete clauses are analogous to arbitration clauses in the employee-employer relationship. Like arbitration clauses, non-competes are contingent on the occurrence of a future event (resignation or termination in the case of non-competes and a legal dispute in the case of arbitration agreements). Employees are unlikely to be able to resist non-compete and arbitration clauses during the hiring process. And to the extent certain workers do have bargaining power and face competition for their services, bargaining and competition are likely to center on wages, benefits, and hours, as opposed to contingent terms such as non-compete and arbitration clauses.

III. Harms to Workers from Non-Compete Clauses

Non-compete clauses deprive workers of labor market mobility. Employers can file suit to enforce non-competes against workers who accept new employment and seek damages and injunctive relief. Most workers bound by a non-compete face three options: 1) stay with their current employer, 2) find employment in a line of work or a geographic area that is outside the scope of the non-compete, or 3) accept unemployment until the non-compete expires.¹⁰⁷ Workers subject to non-compete clauses may be forced to choose between potentially three unattractive

¹⁰³ Matt Marx, *The Firms Strikes Back: Non-compete Agreements and the Mobility of Technical Professionals*, 76 AM. SOC. REV. 695, 708 (2011).

¹⁰⁴ *Id.* at 706.

¹⁰⁵ *Id.*

¹⁰⁶ Marx & Fleming, *supra* note 94, at 51.

¹⁰⁷ *Id.* at 47-50.

alternatives. Even when employers do not, or cannot, enforce non-compete clauses, these restraints can discourage workers from seeking new employment and pursuing business opportunities. In practice, non-competes amount to workers “bartering away [their] personal freedom”¹⁰⁸ for present employment.

By restricting workers’ freedom to leave, non-compete clauses have material effects on the welfare of workers. Because they limit competition for workers among employers, non-competes have a negative effect on wages. Furthermore, they close off potential entrepreneurial opportunities and reduce new business creation. Importantly, the negative effects of non-competes extend beyond reduced wages and business formation. Non-competes can also compel workers to stay in a job where they are subject to gender or racial discrimination, sexual harassment, or other forms of mistreatment on the job or exposed to threats to their health and safety.

A. Impaired Job Mobility

Workers subject to a non-compete clause face substantially reduced labor market mobility. Under a non-compete clause, workers confront the choice of staying with their current employer, pursuing work outside the scope of the non-compete clause, or choosing unemployment. Workers who have independent sources of income (for example, an inheritance or sizable business investments) may be able to withstand a significant period of unemployment and “wait out” a non-compete clause.¹⁰⁹ Most workers, however, do not have significant non-wage sources of income.¹¹⁰ When bound by a non-compete clause, they typically have to stay

¹⁰⁸ Harlan M. Blake, *Employee Agreements Not to Compete*, 73 HARV. L. REV. 625, 650 (1960).

¹⁰⁹ For example, apparently to comply with his non-compete agreement with Microsoft, tech executive Vic Gundotra took a year off before moving to Google. He used the sabbatical for “philanthropic pursuits.” Ben Romano, *Microsoft Loses Another to Google*, SEATTLE TIMES, June 29, 2006, https://blogs.seattletimes.com/microsoftpri0/2006/06/29/microsoft_loses_another_to_google/.

¹¹⁰ BD. OF GOVERNORS, *supra* note 58, at 12.

with their current employer or enter a new line of work or industry or move to a place outside the scope of the non-compete. Women may experience the restrictive effects of non-competes in an especially acute way because they often have less geographic and occupational flexibility than men.¹¹¹

A worker bound by a non-compete clause runs serious risks in pursuing an employment or entrepreneurial opportunity in violation of the non-compete. Her current employer may threaten to file suit to enforce the non-compete and seek damages for breach of contract and an injunction prohibiting the worker from pursuing employment or business in violation of the non-compete.¹¹² In some states, employers can enforce non-competes even against workers who have been terminated.¹¹³ In addition to suing the worker, her current employer may threaten her new employer for tortious interference of employment.¹¹⁴ When informed of the existence of a non-compete clause, a prospective employer may terminate the bound worker, or withdraw the offer of employment, to mitigate its own legal risk.¹¹⁵

Apart from being enforced in courts, the mere existence of a non-compete can deter workers from pursuing alternative opportunities. Employers appear to use non-competes to

¹¹¹ See Orly Lobel, Opinion, *Companies Compete but Won't Let Their Workers Do the Same*, N.Y. TIMES, May 4, 2017, <https://www.nytimes.com/2017/05/04/opinion/noncompete-agreements-workers.html> (“[W]hile noncompete restrictions impose hardships on every worker, for women these restrictions tend to be compounded with other mobility constraints, including the need to coordinate dual careers, family geographical ties and job market re-entry after family leave.”).

¹¹² See, e.g., *C&W Facility Servs., Inc. v. Mercado* 2018 WL 4854630, *4 (D. Mass. Oct. 5, 2018) (granting preliminary injunction to janitorial services company seeking to enforce non-compete contract against former employee).

¹¹³ Kenneth J. Vanko, “*You’re Fired! And Don’t Forget Your Non-Compete*”: *The Enforceability of Restrictive Covenants in Involuntary Discharge Cases*, 1 DEPAUL BUS. & COM. L.J. 1 (2002). Florida law takes a particularly absolutist approach to the enforcement of non-competes clauses against workers. See *Twenty Four Collection, Inc. v. Keller*, 389 So.2d 1062, 1063 (Fla Dist. Ct. App. 1980) (“[I]t is established law that a court is not empowered to refuse to give effect to such a contract on the basis of a finding, as was the case below, that enforcement of its terms would produce an ‘unjust result’ in the form of an overly burdensome effect upon the employee.”).

¹¹⁴ E.g., *Automated Concepts Inc. v. Weaver*, 2000 WL 1134541 at *5-7 (N.D. Ill. Aug. 9, 2000).

¹¹⁵ See, e.g., Viswanathan, *supra* note 13 (“Within weeks, . . . Thomson Reuters asked Ms. Russell-Kraft to leave. Her previous employer, Law360, also in New York, had sent a letter to the company citing a noncompete agreement she signed when she started at the newswire as a 25-year-old, her first full-time job in journalism.”).

discourage workers from departing in general and do not necessarily file suit to enforce the restriction against many (or in some cases *any*) employees. Harlan Blake wrote, “For every covenant that finds its way to court, there are thousands which exercise an *in terrorem* effect on employees who respect their contractual obligations and on competitors who fear legal complications if they employ a covenantor, or who are anxious to maintain gentlemanly relations with their competitors.”¹¹⁶

Even in states where non-compete clauses are unenforceable, workers are subject to non-competes at roughly the same rate as workers in states where non-competes are enforceable. Although they cannot enforce non-competes in court in these states, many employers still condition employment on worker acceptance of non-compete clauses. Indeed, the relevant law appears to have minimal effects on employers’ use of non-compete restrictions. In California and North Dakota (two of the three states where non-competes are unenforceable), approximately 19% of workers are under a non-compete requirement—the same figure as for workers in states where non-competes are most likely to be enforced.¹¹⁷ When looking at only single-state firms, which are most likely to tailor their practices to conform to relevant state law, “the incidence of noncompetes in [California and North Dakota] is 14%, only slightly less than the 16.5% incidence level . . . in the highest enforcing states.”¹¹⁸

The chilling effect of non-competes on worker mobility appears to be very real. Interviews with automated speech recognition professionals found that these workers generally complied with non-competes even without actual litigation by their employer.¹¹⁹ Of the workers

¹¹⁶ Blake, *supra* note 108, at 682.

¹¹⁷ Starr et al., *supra* note 8, at 16.

¹¹⁸ *Id.*

¹¹⁹ See Marx, *supra* note 103, at 707 (“[F]irms strategically manage the process of obtaining signatures, waiting to present the non-compete until an employee’s bargaining power is minimized. Firms appear to accomplish these outcomes with minimal expenditure. Only one informant reported being formally sued and taken to court by an employer; for the others, merely the threat of litigation sufficed to exert a chilling effect on their career plans.”).

who temporarily left their field of expertise due to a non-compete, *none* did so because they were sued by their employer.¹²⁰ They undertook this dramatic occupational change “on the expectation of what might happen if they refused to act in accordance with the employment agreement they had signed.”¹²¹

Some employers use non-compete contracts regardless of whether they can or will enforce them through legal action. For instance, although it insisted on a broad non-compete clause with storefront employees, Jimmy John’s asserted that it had *never* enforced a non-compete clause against an hourly worker.¹²² According to a legal services attorney who has represented workers bound by non-competes, “most workers obey initial threats [over non-competes] rather than going to court over them . . . [and] ‘there are people who have been affected by this, and it doesn’t even occur to them to get a lawyer.’”¹²³ Taken together, this evidence suggests that employers use non-compete clauses to discourage workers from seeking, or even exploring, alternative work and business opportunities. This chilling effect on worker mobility is not necessarily captured in litigation and may be the most significant consequence of non-compete clauses.¹²⁴

Empirical evidence shows that non-competes do reduce labor market mobility. The latest research finds that workers “bound by non-competes have . . . 11% longer [job] tenures.”¹²⁵ This

¹²⁰ Marx & Fleming, *supra* note 94, at 49.

¹²¹ *Id.*

¹²² Wiessner, *supra* note 22.

¹²³ Quinton, *supra* note 1.

¹²⁴ Arnow-Richman, *supra* note 96, at 981-82. One commentator has described “the pervasive use of noncompetes in Massachusetts is part of the dark matter of the legal landscape in the state. You know it’s there, exerting some gravitational force, but you can’t see it or measure it. You never really know how many employees didn’t move to another job, didn’t start their own companies, and didn’t take the risk of challenging their noncompete agreements in court.” Lee Gesmer, *Why Has Silicon Valley Outperform Boston/Route 128 as a High Tech Hub?*, MASS LAW BLOG, Dec. 6, 2007, <https://masslawblog.com/noncompete-agreements/why-has-silicon-valley-outperformed-boston-route-128-as-a-high-tech-hub/>.

¹²⁵ Evan Starr, J.J. Prescott & Norman Bishara, *Noncompetes and Employee Mobility* 3 (2019).

negative effect on job mobility is largely insensitive to the degree of non-compete enforceability in a state.¹²⁶ Notably, even workers not bound by non-competes appear to be adversely affected the existence of non-competes. Given the lack of transparency over who is bound by non-competes in a labor market, non-competes can increase employers' costs of recruiting and hiring and impede job market mobility for workers not subject to non-competes.¹²⁷ In labor markets with a higher incidence of non-competes and in states that favor enforcement of non-competes, *all* workers receive fewer job offers and have longer job tenures.¹²⁸

Other research finds that workers do stay with their current employer longer when they work in the shadow of a non-compete condition and that job tenure is longer in states where non-competes are more likely to be enforced. After Michigan rewrote its antitrust law and –made non-competes inadvertently easier to enforce in court – inventors (excluding inventors in the then-turbulent auto industry) experienced an 8.1% reduction in mobility relative to peers in other states,¹²⁹ while inventors with firm-specific knowledge and skills saw a 15.4% decline in mobility.¹³⁰ Another study found that workers in the computer industry in California, where non-competes are unenforceable, had greater job mobility than their peers in other states, where non-competes are (to varying degrees) enforceable.¹³¹ This finding was supported in another study on tech sector workers.¹³²

¹²⁶ *Id.*

¹²⁷ Evan Starr, Justin Frake & Rajshree Agarwal, *Mobility Constraint Externalities* 6-7 (2018).

¹²⁸ *Id.* at 26.

¹²⁹ Matt Marx, Deborah Strumsky & Lee Fleming, *Mobility, Skills, and the Michigan Non-Compete Experiment*, 55 *MGMT. SCI.* 875, 887 (2009).

¹³⁰ *Id.*

¹³¹ Bruce Fallick, Charles A. Fleischman & James B. Rebitzer, *Job-Hopping in Silicon Valley: Some Evidence Concerning the Microfoundations of a High-Technology Cluster*, 88 *REV. ECON. & STATS.* 472, 481 (2006).

¹³² See Natarajan Balasubramanian et al., *Locked In? The Enforceability of Covenants Not to Compete and the Careers of High-Tech Workers* 30 (Ross School of Business Working Paper No. 1339, Jan. 2017), (“We find that stricter [non-compete] enforceability is associated with longer job spells[.]”).

Impairment of job mobility extends to workers at the very top of the income distribution. Chief executive officers have longer tenures and are less likely to switch firms in states where non-compete clauses are more likely to be enforceable in court.¹³³

Alternatively, workers subject to non-compete clauses sometimes escape them through a temporary but costly “career detour”¹³⁴ or relocation. First, they can find a job in another line of work in which they cannot necessarily put their skills to full use. One study found that workers bound by non-competes reported more than a 50% higher probability of leaving for a firm in a different industry, relative to workers who were not subject to non-competes.¹³⁵ A survey of automated speech recognition professionals found that nearly one-third of workers subject to a non-compete left their field of expertise until their non-compete expired.¹³⁶ In one arguably extreme case, a survey respondent, despite having a PhD and 20 years of experience, accepted a data entry job to comply with the prohibitions in the non-compete clause.¹³⁷ This “detour” can be costly: workers who leave their field of specialization can experience “reduced compensation, atrophy of their skills, and estrangement from their professional networks.”¹³⁸

Second, some bound workers can relocate to another state or even country where they can work without fear of violating the non-compete clause. Due to the change in Michigan law under which non-competes became enforceable, the state saw an exodus of knowledge workers, an effect that was especially pronounced among those likely to engage in collaborative

¹³³ Mark J. Garmaise, *Ties That Truly Bind: Noncompetition Agreements, Executive Compensation, and Firm Investment*, 27 J. L. ECON. & ORG. 376, 400 (2011).

¹³⁴ Marx, *supra* note 103, at 696.

¹³⁵ Starr, Prescott & Bishara, *supra* note 125, at 3.

¹³⁶ Marx, *supra* note 103, at 705.

¹³⁷ Marx & Fleming, *supra* note 94, at 49.

¹³⁸ *Id.* at 48.

projects.¹³⁹ Relocation entails its own costs and challenges for workers, including uprooting one's family, severing existing familial and social ties, and starting afresh in a new place.

The loss of exit opportunities is particularly harmful in present-day labor markets. Due to workers' general lack of power in the workplace, their freedom to leave for another employer or pursue a business opportunity is critical today. Most workers are not members of a union. In the private sector, less than 7% of workers are represented by a union.¹⁴⁰ Furthermore, the Supreme Court has granted employers the right to use mandatory arbitration clauses to block employees' class action lawsuits alleging violation of employment, labor, and other laws governing the workplace.¹⁴¹ Because of the low rate of unionization in the private sector and the neutering of collective litigation methods, workers have little ability to act collectively to address shared grievances in the workplace. In general, they can only make claims as individual employees, which carries serious risk of retaliation and even termination. Exit provides a degree of leverage for otherwise disempowered workers, providing *some* freedom to leave unsatisfactory jobs and find better employment.¹⁴² Given the limited role of voice in the contemporary workplace, employer deprivation of workers' freedom to exit through non-compete clauses is especially harmful.

B. Effects of Reduced Job Mobility

By impairing labor market mobility and binding workers to their present employer, non-compete clauses inflict material harms on workers. Empirical studies have found that stronger enforcement of non-compete clauses is associated with lower wages and lower wage growth over

¹³⁹ Matt Marx, Jasjit Singh & Lee Fleming, *Regional Disadvantage? Employee Non-Compete Agreements and Brain Drain*, 44 RES. POL'Y 394, 403 (2015). See also Balasubramanian et al., *supra* note 132, at 30 ("We find that stricter [non-compete] enforceability is associated with . . . a greater likelihood of leaving the state[.]").

¹⁴⁰ Bureau of Labor Stats., Union Members Summary, <https://www.bls.gov/news.release/union2.nr0.htm>.

¹⁴¹ *Epic Sys. Corp. v. Lewis*, 138 S. Ct. 1612 (2018).

¹⁴² ALBERT O. HIRSCHMAN, *EXIT, VOICE, AND LOYALTY: RESPONSES TO DECLINE IN FIRMS, ORGANIZATIONS, AND STATES* (1970).

time. Research has also found that stronger enforcement depresses the formation of new businesses. The adverse effects of non-competes likely extend beyond quantifiable measures such as wage levels, wage growth, and new firm creation. Given the ubiquity of harassment and discrimination at the workplace and unsafe job conditions, non-competes may also force workers to stay in hostile or dangerous work environments.

In a 2016 report, the Treasury Department concluded that increased enforcement of non-competes is associated with both lower wages and lower wage growth over time. By extrapolating from existing research on non-competes, it estimated that “maximal enforcement” states slow wage growth over time, especially older workers’ wage growth, relative to “minimal enforcement states.” The difference in wages between the two categories of states was “5 percent at age 25 and 10 percent at age 50.”¹⁴³

Another study found that “wages are 4% lower in an average enforcing state relative to a non-enforcing state.”¹⁴⁴ This wage depression was stronger among workers without a graduate-level degree and workers with long tenures at their present job.¹⁴⁵ Looking exclusively at tech workers: In states that are more likely to enforce non-compete clauses, initial wages and wage growth over time were found to be lower, relative to states that are less likely to enforce non-competes.¹⁴⁶ When widely used in a labor market, non-competes introduce costs and uncertainty into hiring and can depress job market mobility—and thereby wages—for all workers, not only those directly bound by them. Two factors—a higher incidence of non-competes and state law

¹⁴³ U.S. DEP’T OF TREASURY, *supra* note 2, at 20.

¹⁴⁴ Evan Starr, *Consider This: Training, Wages, and the Enforceability of Covenants Not to Compete* 17 (2018).

¹⁴⁵ *Id.* at 27.

¹⁴⁶ Balasubramanian et al., *supra* note 132, at 30.

favoring the enforcement of non-competes—appear to depress wages for *all* workers, not only those bound by non-competes.¹⁴⁷

Even among elite workers who have power in the labor market and could reasonably bargain over terms of employment, non-compete contracts may depress incomes. An analysis that looked exclusively at the effects of non-competes on chief executive officers found that stronger state-level enforcement of non-competes reduces the growth rate of compensation by 8.2%.¹⁴⁸

While empirical studies generally find that non-competes depress wages, the research does not categorically find this relationship. Two studies have found evidence of a compensating wage premium for workers who accept non-compete clauses. In one study, physicians in group practices that used non-compete clauses had higher incomes and income growth over time than their peers in groups that did not use non-competes.¹⁴⁹ Looking at all workers, another analysis found a 10% wage premium for workers who receive a non-compete *before* they accept a job offer.¹⁵⁰

Non-compete clauses also appear to discourage workers from pursuing business opportunities and establishing new businesses. Workers subject to non-competes are often restricted from starting a business in competition with, or in the same line of business, as their current employer. One study of the biotechnology industry examined the rate of new firm creation after firms were acquired or made an initial public offering.¹⁵¹ (After either event, employees usually receive a significant amount of money that they can use to start their own

¹⁴⁷ Starr, Frake & Agarwal, *supra* note 128, at 26.

¹⁴⁸ Garmaise, *supra* note 11, at 402.

¹⁴⁹ Kurt Lavetti, Carol Simon & William D. White, *Buying Loyalty: Theory and Evidence from Physicians* 33-34 (2012).

¹⁵⁰ Starr et al., *supra* note 8, at 3.

¹⁵¹ Toby E. Stuart & Olav Sorenson, *Liquidity Events and the Geographic Distribution of Entrepreneurial Activity*, 48 ADMIN. SCI. Q. 175 (2003).

firm.) Examining startup activity after acquisitions or initial public offerings of biotechnology firms, the study found that the rate of new biotech startups is higher in states that do not enforce non-competes relative to states that do enforce non-compete contracts.¹⁵²

Strong enforcement of non-competes also frustrates the ability of venture capital to support startup businesses. An increase in the supply of venture capital funding increases the number of new firms most in states that are less likely to enforce, or do not enforce, non-compete clauses.¹⁵³

Non-compete clauses, by deterring and preventing workers from switching jobs, can bind workers to discriminatory and hostile work environments. Discrimination in the workplace is pervasive. Discrimination at work can include denials of promotions and raises, inferior assignments, and reduced visibility because of an individual's race or gender. For example, 57% of African Americans have reported being denied equal treatment on pay or promotions due to their race.¹⁵⁴ Studies have found a similar incidence of discrimination toward women in the workplace. Forty-two percent of women have reported being denied promotions, paid less for the same work that male coworkers perform, and given less rewarding assignments due to their gender.¹⁵⁵ Furthermore, hostile work environments are also common. For instance, 35% of

¹⁵² *Id.* at 197.

¹⁵³ Sampsa Samila & Olav Sorenson, *Noncompete Covenants: Incentives to Innovate or Impediments to Growth* 57 MGMT. SCI. 425 (2011). Strong enforcement of non-compete clauses also can encourage prospective entrepreneurs to relocate to other states. As discussed earlier, in 1985, the Michigan legislature revised the state's antitrust law and unwittingly made non-compete clauses easier to enforce in court. Examining this legal change, a study found that the more employer-friendly approach to non-compete encouraged skilled technical workers, who may be especially capable of and interested in starting businesses, to leave the state. Marx, Singh & Fleming, *supra* note 139, at 403.

¹⁵⁴ NPR ET AL., DISCRIMINATION IN AMERICA: EXPERIENCES AND VIEWS OF AFRICAN AMERICANS 6 (2017), <https://www.npr.org/assets/img/2017/10/23/discriminationpoll-african-americans.pdf>.

¹⁵⁵ Kim Parker & Cary Funk, *Gender Discrimination Comes in Many Forms for Today's Working Women*, PEW RESEARCH CTR., Dec. 14, 2017, <http://www.pewresearch.org/fact-tank/2017/12/14/gender-discrimination-comes-in-many-forms-for-todays-working-women/>.

women have experienced sexual harassment in the workplace.¹⁵⁶ By restricting job switching, non-compete clauses can lock workers into discriminatory or hostile work environments.

The freedom to leave is especially critical for victims of discrimination, harassment, and other mistreatment who are lower on the firm hierarchy. These workers generally have few legal protections on the job and no feasible path for seeking redress for the mistreatment they suffer.¹⁵⁷ For example, many women who are victims of sexual harassment may believe they have no option but to leave their current employer.¹⁵⁸ In one study, “[t]argets of sexual harassment were 6.5 times as likely as nontargets to change jobs” during the study period.¹⁵⁹

Non-competes can also force workers to remain at jobs with unsafe or dangerous working conditions. In 2017 in the United States, 5,147 workers died from traumatic injuries on the job, which means that fatal job injuries occurred at a rate of 3.5 per 100,000 workers.¹⁶⁰ In 2017, private sector employers reported 2.8 million cases of workplace injury or illness, or a rate 2.8 cases per 100 full-time workers.¹⁶¹ And unhealthy workplaces can inflict lasting harms on workers: exposure to toxins on the job is “responsible for more than 50,000 deaths and 190,000 illnesses each year, including cancers and other lung, kidney, skin, heart, stomach, brain, nerve

¹⁵⁶ 11/22: *More than One in Three Women Report Sexual Harassment in the Workplace*, MARIST POLL, Nov. 22, 2017, <http://maristpoll.marist.edu/1122-more-than-one-in-three-women-report-sexual-harassment-in-the-workplace/#sthash.5qBGru84.dpbs>.

¹⁵⁷ Rachel Arnow-Richman, *Of Power and Process: Handling Harassers in an At-Will World*, 128 YALE L.J. FORUM 85, 89-90 (2018).

¹⁵⁸ Heather McLaughlin, Christopher Uggen & Amy Blackstone, *The Cost of Sexual Harassment*, GENDER & SOC., June 7, 2017, <https://gendersociety.wordpress.com/2017/06/07/the-cost-of-sexual-harassment/>.

¹⁵⁹ Heather McLaughlin, Christopher Uggen & Amy Blackstone, *The Economic and Career Effects of Sexual Harassment on Working Women*, 31 GENDER & SOC. 333, 344 (2017).

¹⁶⁰ Press Release, Bur. Lab. Stats., *Census of Fatal Occupational Injuries Summary* (Dec. 18, 2018), <https://www.bls.gov/news.release/cfoi.nr0.htm>.

¹⁶¹ Press Release, Bur. Lab. Stats., *Employer-Reported Workplace Injury and Illnesses, 2017* (Nov. 8, 2018), <https://www.bls.gov/news.release/osh.nr0.htm>.

and reproductive disease.”¹⁶² These figures likely undercount the incidence of death, injury, and illness on the job.¹⁶³ The threat of dangerous, unsafe, and unhealthy working conditions is real.¹⁶⁴ When workers want to leave for comparatively safer or healthier work environments, non-competes can bar them from finding jobs with employers that provide safer work environments.

IV. Monopolistic and Oligopolistic Businesses Can Use Non-Compete Clauses to Exclude or Limit Product Market Competition

On top of the direct harms to workers, non-compete clauses can protect dominant incumbents against competition in product and labor markets. Incumbents can use non-compete clauses to tie up scarce labor and thereby deprive current and would-be rivals of essential workers. Furthermore, non-competes can favor large incumbents over small rivals. Workers subject to non-competes may be more likely to move to large firms relative to small firms because larger firms are more able and willing to tolerate the risk of, and defend against, lawsuits. Conversely, workers restricted by non-competes may be *less* likely to move to small firms over bigger ones for the same reason. In short, non-competes can serve as an entry barrier, which helps maintain concentrated product and labor markets, and hinder the creation and growth of small businesses.

Dominant corporations can use non-compete clauses with workers to marginalize and exclude existing and potential rivals. In markets in which firms are dependent on highly specialized or otherwise scarce labor, a monopolist can deprive rivals of essential workers and hobble their ability to compete. For instance, a monopolistic hospital can use non-competes to

¹⁶² AFL-CIO, DEATH ON THE JOB: THE TOLL OF NEGLIGENCE 40 (2018).

¹⁶³ *Id.* at 12-13.

¹⁶⁴ *Id.* at 8.

deprive a rival hospital of physicians.¹⁶⁵ Under these circumstances, the non-compete functions as a de facto exclusive dealing contract with physicians and starves rival hospitals of an essential “input” (physicians), which could violate the Sherman or Clayton Acts.¹⁶⁶ In a field such as medicine in which long-term relationships between providers and patients are important, the use of non-competes by powerful incumbents can be particularly damaging to rivals and patients.¹⁶⁷

Firms with market power can use also non-competes to block an important source of potential competition: their own employees. In many fields, employees have the skills and experience to start firms and compete with their current employers. For example, an entrepreneurial doctor may be well positioned to leave a hospital or group practice and establish her own independent practice.¹⁶⁸ Through non-compete clauses, incumbent firms can and do “frequently discourage” employees from breaking out and starting rival firms.¹⁶⁹

The use and enforcement of non-competes, in general, can favor larger incumbent firms over smaller, emerging firms. Workers bound by non-competes may believe large firms are a safer destination than smaller firms. Relative to small firms, larger firms may be better equipped to buy out non-compete clauses and defend themselves and the affected employees against

¹⁶⁵ See, e.g., *BRFHH Shreveport LLC v. Willis Knighton Med. Ctr.*, 176 F.Supp.3d 606, 625 (W.D. La. 2016) (“Vantage also asserts that its allegations of Willis-Knighton's non-compete agreements with its physicians and its control of physician referrals are anticompetitive under section 2 of the Sherman Act.”). Warren Greenberg, *Marshfield Clinic, Physician Networks, and the Exercise of Monopoly Power*, 33 HSR: HEALTH SERVS. RES. 1461, 1470 (1998) (“The Marshfield Clinic also enforced a non-compete clause with physicians who were formerly employed by Marshfield. Such physicians could not practice within 30 miles of Marshfield for three years after termination from the Clinic, resulting in less competition to the Marshfield Clinic.”). See generally Steven C. Salop, *The Raising Rivals’ Cost Foreclosure Paradigm, Conditional Pricing Practices, and the Flawed Incremental Price-Cost Test*, 81 ANTITRUST L.J. 371 (2017); Jonathan B. Baker, *Exclusion as a Core Competition Concern*, 78 ANTITRUST L.J. 527 (2013).

¹⁶⁶ See, e.g., *McWane, Inc. v. FTC*, 783 F.3d 814, 842 (11th Cir. 2015) (affirming FTC’s decision that McWane engaged in illegal exclusive dealing). For antitrust framework on exclusive dealing arrangements, see *Tampa Elec. Co. v. Nashville Coal Co.*, 365 U.S. 320 (1961).

¹⁶⁷ Michelle Andrews, *Did Your Doctor Disappear Without a Word? A Noncompete Clause Could Be the Reason*, N.Y. TIMES, Mar. 15, 2019, <https://www.nytimes.com/2019/03/15/business/physician-non-compete-clause.html>.

¹⁶⁸ *Id.*

¹⁶⁹ Steven Klepper & Peter Thompson, *Disagreements and Intra-industry Spinoffs*, 28 INT’L J. INDUS. ORG. 526, 531 (2010).

lawsuits seeking to enforce non-competes.¹⁷⁰ In the course of her research on non-competes, Orly Lobel found that employees bound by non-competes often preferred to minimize the legal risks of departure “by going to an established competitor that has the resources to protect and indemnify them in the case of legal liability.”¹⁷¹ After Michigan permitted the enforcement of non-competes in the mid-1980s, inventors in that state who changed jobs “were considerably more likely to join larger firms.”¹⁷²

The combined effect of non-compete clauses can be to protect monopolistic and oligopolistic market structures. Dominant and near-dominant firms can employ non-competes as exclusive dealing clauses that deny key workers to actual and would-be rivals. Furthermore, the use of non-competes can favor large firms over small firms, who may be less able or willing to litigate potential lawsuits arising from these contractual restrictions.

V. The General Justification for Non-Compete Clauses Does Not Stand Up to Scrutiny

The general justification for non-compete clauses is unpersuasive on closer examination and should be treated with skepticism. The case for non-compete clauses presumes a need for employers to protect their investment in intangibles through a quasi-property right. These intangibles include trade secrets, customer lists, and employee training. In the absence of non-competes, proponents argue, the free movement of workers from one employer to another allows rival companies to “free ride” on the investment in intangibles made by the first employer.¹⁷³

¹⁷⁰ Marx, *supra* note 103, at 709. As an illustration of how the litigation process favors larger, established firms over new entrants and small firms, smaller biotechnology firms with higher relative costs of litigation appear less likely to patent a new invention in the same subclass as rivals. Josh Lerner, *Patenting in the Shadow of Competitors*, 38 J. L. & ECON. 463, 489-91 (1995). *See id.* at 472 (“In general, small firms believed that their patents were infringed more frequently, but were considerably less likely to litigate these infringements.”).

¹⁷¹ ORLY LOBEL, *TALENT WANTS TO BE FREE: WHY WE SHOULD LEARN TO LOVE LEAKS, RAIDS, AND FREE RIDING* 202 (2013).

¹⁷² Marx & Fleming, *supra* note 94, at 52.

¹⁷³ *See, e.g.*, Edmund W. Kitch, *The Law and Economics of Rights in Valuable Information*, 9 J. LEGAL STUD. 683, 691 (1980) (“It is often impossible to determine whether a former employee who has to work for a competitor has taken trade secret information, and whether he has disclosed that information to his new employer. The former

This justification mistakenly ignores how the dissemination and sharing of information can often benefit society. For example, copyright and patent law, while creating a property right over intangibles, recognize this logic and accordingly limit the scope of protection. Through non-competes, employers effectively expand intellectual property protection and disrupt the balance that Congress has attempted to strike.

Insofar as employers do need to protect their investment in intangibles, non-compete clauses are an overbroad and yet also ineffective tool. In the name of protecting an employer's discrete investment, non-compete clauses restrict an individual's job mobility and freedom to use his or her full experience, knowledge, and skills. Yet, non-competes do not protect the intangible itself and do not restrict employees from sharing it with rivals and other third parties. In lieu of non-compete contracts, employers have less restrictive methods of protecting their intangibles. These include trade secret law, improved employee retention policies, and employment contracts. Employers can use these legal tools to protect their intangibles against free riding without imposing a one-sided labor market restraint on workers.

A. The Justification for Non-Compete Clauses Rests on a Questionable Premise

While the justification for non-competes presumes that free riding is categorically bad, this story is, at best, incomplete and, at worst, specious. The sharing of information among individuals and firms is often desirable for society and should not be indiscriminately restricted through restraints such as non-competes. Moreover, the neoclassical economic theory that provides the justification for non-competes offers only qualified support for the use of non-competes.

employee may give the information to the new firm without disclosing its confidentiality and represent the information as his own to impress the new firm value. A restrictive covenant keeps the ex-employee away from the competitor.”).

Even accepting that firms principally use non-competes to protect their intangibles, information sharing is not a categorical “evil” that state action should police at any cost. What is disparaged as free riding often is the broad dissemination of knowledge that contributes to economic growth and innovation.¹⁷⁴ This sharing can contribute to the growth of new firms and new industries as workers are free to combine their knowledge with knowledge possessed by other workers and firms.¹⁷⁵ Excess protection for knowledge directly through intellectual property or indirectly through contractual restraints such as non-compete clauses can frustrate this iterative dynamic.¹⁷⁶

Non-competes appear to discourage the socially desirable dissemination and sharing of knowledge and other intangible assets. In a widely-cited article, Ronald Gilson, building on the work of urban planning scholar AnnaLee Saxenian,¹⁷⁷ examined the decline of the Route 128 corridor in Boston as a tech hub and the parallel rise of Silicon Valley in California.¹⁷⁸ He suggests that the differential treatment of non-compete clauses in the two states is a powerful explanatory factor. Until recently, Massachusetts law generally supported the enforcement of non-competes,¹⁷⁹ whereas California law has long held that non-competes are unenforceable in

¹⁷⁴ Brett M. Frischmann & Mark A. Lemley, *Spillovers*, 100 COLUM. L. REV. 101, 111-14 (2006). For a summary of the economic and legal literature casting doubt on the free riding theory, see Alan Hyde, *Intellectual Property Justifications for Restricting Employee Mobility: A Critical Appraisal in Light of the Economic Evidence*, IN RESEARCH HANDBOOK ON THE ECONOMICS OF LABOR AND EMPLOYMENT LAW 357 (Cynthia L. Estlund & Michael L. Wachter eds., 2012).

¹⁷⁵ See generally LOBEL, *supra* note 171.

¹⁷⁶ *Id.* at 76-97. Intellectual property protections, such as copyrights and patents, are likely already overprotective. James Boyle has written that recent expansions of intellectual property protection have not had a sound empirical basis and described policymaking in this area as “an evidence-free zone.” JAMES BOYLE, THE PUBLIC DOMAIN: ENCLOSING THE COMMONS OF THE MIND 205-29 (2008).

¹⁷⁷ ANNALEE SAXENIAN, REGIONAL ADVANTAGE: CULTURE AND COMPETITION IN SILICON VALLEY AND ROUTE 128 (1996).

¹⁷⁸ Ronald J. Gilson, *The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants not to Compete*, 74 N.Y.U. L. REV. 575 (1999).

¹⁷⁹ Mass. Gen. Laws. C. 149 § 186. Bob Salsberg, ‘Garden’ Clause in New Law Requires Pay During Noncompete, WBUR, Sep. 30, 2018, <http://www.wbur.org/news/2018/09/30/garden-clause-new-law-requires-pay-during-noncompete>.

court.¹⁸⁰ Gilson contends that California law’s hostility toward non-compete clauses promoted the sharing of information and knowledge and fostered the creation of new businesses.¹⁸¹ He argues that, in contrast, the relatively pro-enforcement orientation of Massachusetts law hindered the dissemination of knowledge and frustrated the creation of new firms in the technology sector.¹⁸²

Gilson’s theory is supported by other research. Relatively easy labor mobility across firms promotes innovation and invention.¹⁸³ Even for firms themselves, too little employee turnover can create an insular ethos resistant to new ideas and thereby become an impediment to innovative activity.¹⁸⁴

The neoclassical economic theory on which proponents of non-competes rely offers only qualified support for the free riding arguments. This theory holds that employers should be allowed to recover just enough on their investment to ensure that they invest in the future.¹⁸⁵ Permitting an employer to recover a *sufficient* amount is very different from allowing the employer to recover the *maximal* amount from their investment. If an employer can recoup its investment in job training within six months of a worker joining, society has no interest in ensuring that the worker stays with the employer indefinitely so that the employer can extract a maximal return on the training. Indeed, if the employer has recouped its investment in training or

¹⁸⁰ Cal. Bus. & Prof. Code § 16600.

¹⁸¹ See generally Gilson, *supra* note 178.

¹⁸² *Id.* at 602-13.

¹⁸³ See, e.g., LOBEL, *supra* note 175, at 40 (“[L]ocalities with dense connections between innovators, knowledge flows, and human capital enjoy dramatically more innovation than smaller, protective, and more isolated settings.”); Marx, Singh & Fleming, *supra* note 139, at 403 (finding that strong enforcement of non-competes encouraged out-migration of “workers who are more collaborative and whose work is more impactful, stripping enforcing states of some of their most valuable knowledge workers.”); Lee Fleming & Koen Frenken, *The Evolution of Inventor Networks in the Silicon Valley and Boston Regions*, 10 ADV. COMPLEX SYS. 53 (2007).

¹⁸⁴ See LOBEL, *supra* note 175, at 129 (“Pathologies of groupthink—whereby cohesive groups overlook important alternatives because of their desire for consensus and conformity—and [not invented here] mentalities are exaggerated when companies are overly stable.”).

¹⁸⁵ Mark A. Lemley, *Property, Intellectual Property, and Free Riding*, 83 TEX. L. REV. 1031, 1046-50 (2005)

other intangible during the worker's tenure, it has no basis for restraining the worker for a single day *after* he or she leaves.

Importantly, non-competes can *discourage* certain forms of investment in intangibles. While non-competes can stimulate employer investment in the training of workers, these contracts can simultaneously discourage self-training by bound workers.¹⁸⁶ Deprived of the freedom to leave, these workers may be less likely to invest in training and other self-improvement because they have less power to obtain higher salaries and wages at their current firm or elsewhere.¹⁸⁷

The laws governing the protection of intangibles already reflect the risk of overprotection of intangibles. For instance, state law holds that non-competes with an unlimited duration are generally unenforceable.¹⁸⁸ Moreover, intellectual property law in the United States incorporates this consideration too. It aims to provide adequate incentive to creators and inventors, not maximal incentive. Accordingly, copyright and patent laws include a term limit on protection and include several important exceptions to the scope of protection. For instance, copyright protects the expression of an idea, but does not protect the idea itself.¹⁸⁹

Non-competes enable employers to circumvent the boundaries of intellectual property law. Copyright, patent, and trade secret law have important limitations on their respective scopes of protection. By imposing non-competes on workers, employers can do “an end run around

¹⁸⁶ LOBEL, *supra* note 175, at 178.

¹⁸⁷ *Id.*

¹⁸⁸ Piveteau, *supra* note 38, at 680. Harlan Blake wrote, “Every postemployment restraint, for whatever reason imposed, has inevitable effects which in some degree oppose commonly shared community values. In view of our feeling that a man should not be able to barter away his personal freedom, even this small degree of servitude is distasteful.” Blake, *supra* note 108, at 650. A non-compete in perpetuity could, in effect, compel the worker to remain with her present employer indefinitely and would raise serious 13th Amendment concerns. *See* U.S. Const. amend. XIII, § 1 (“Neither slavery nor involuntary servitude, except as a punishment for crime whereof the party shall have been duly convicted, shall exist within the United States, or any place subject to their jurisdiction.”).

¹⁸⁹ Harper & Row Publishers, Inc. v. Nation Enter., 471 U.S. 539, 556 (1985).

these [intellectual property] regimes.”¹⁹⁰ Employers can overprotect intangible investments and upset the balance that the federal government and the states have tried to strike in copyright, patent, and trade secret laws. Through non-competes, employers can prevent what, from their perspective, is free riding detrimental to their profits, but, for the public, amounts to beneficial cooperation and sharing of information and knowledge.

B. Non-Competes Are a Flawed Means of Protecting Against Free Riding

Even for the purpose of preventing free riding, non-competes are an ineffective tool. They are overbroad means of protecting against improper appropriation. They are unlike copyright and patent, which protect a specific creative work, invention, or process. Non-competes do not directly control, or protect, intangible assets. Instead, non-competes restrict the possessor of the intangible—the worker. As one scholar of intellectual property has written, “noncompetes regulate the *inputs* to creation and invention, whereas IP rights regulate the inventive or creative *outputs*.”¹⁹¹ Non-competes restrain workers’ labor market freedom without providing effective protection to the intangible at issue. Accordingly, they are too broad and too narrow.¹⁹²

Although some employers may use non-competes to protect trade secrets or customer lists, a fundamental mismatch still exists between that justification and the operation of non-compete clauses. To protect, for example, a trade secret or customer list, employers restrain the labor market freedom of workers and their ability to earn a living. Non-competes prevent workers from using their *full* set of experience, knowledge, and skills at another employer in

¹⁹⁰ Viva R. Moffat, *The Wrong Tool for the Job: The IP Problem with Noncompetition Agreements*, 52 WM. & MARY L. REV. 873, 879 (2010).

¹⁹¹ *Id.* at 914 (emphasis in original).

¹⁹² *Id.*

order to protect an employer's *discrete* intangible.¹⁹³ To prevent a highly specific threat (rival free-riding on the employers' investment in an intangible), non-competes lock workers into their current job, inflicting harms on the bound workers themselves as well as broader society. While human beings cannot be separated from their experience, knowledge, and skills, non-competes strip workers of the freedom to *use* their talents, many of which they acquired or developed on their own and not from employers or any single employer. Through non-competes, employers "convert general training into firm-specific human capital by denying workers the opportunity to apply those skills outside the firm."¹⁹⁴

At the same time, non-compete clauses are not effective at protecting firms' investments in intangibles. Copyright and patent law operate against the entire world and grant the owner a cause of action against any and all parties who infringe their intellectual property. In contrast, non-compete clauses are a creation of contract and operate only against the bound employee, and potentially an employer who recruits this worker.¹⁹⁵ Accordingly, non-competes bind an employee who possess an intangible to his or her employer but do not protect the intangible itself. Even when bound by a non-compete clause, an employee can still covertly disclose a customer list or trade secret to a rival company or another third party. A non-compete clause does little or nothing to address this threat.

C. Employers Have Less Restrictive Alternatives to Protect Intangible Interests

¹⁹³ See LOBEL, *supra* note 175, at 37 ("At their most dangerous, human capital controls such as noncompete agreements temporarily prevent workers who have trained and labored in a specific field with a specific set of knowledge from using their expertise in pursuing their passions and perhaps also from earning a living."); Viva R. Moffat, *Human Capital as Intellectual Property? Non-Competes and the Limits of IP Protection*, 50 AKRON L. REV. 903, 928 (2016) ("By limiting the employment possibilities for employees, non-competes seek to control not only the output of human ingenuity and creativity, but also the source of it—the human capital itself.").

¹⁹⁴ Marx, *supra* note 103, at 698.

¹⁹⁵ See Viva R. Moffat, *Making Non-Competes Unenforceable*, 54 ARIZ. L. REV. 939, 980 (2012) ("[Non-competes] are too narrow because they operate only between an employer and an employee and thus do not protect the intellectual property against the world.").

To the extent they need to protect intangibles against socially harmful free riding, employers have a range of less restrictive means of achieving this objective. Indeed, employers have a vast array of tools in their legal arsenal through which to assert quasi-property rights over intangibles. Even apart from non-competes, Orly Lobel has argued that “human capital controls have wildly expanded and are widespread in almost every industry.”¹⁹⁶ Employers can protect their intangibles through trade secret law and non-disclosure agreements that prevent employees and former employees from sharing or publicizing protected information. In addition to these legal protections for intangibles, employers concerned about the loss of valuable intangibles due to employee departure can improve their retention policies or offer workers employment contracts. These methods allow employers to protect intangibles without imposing a broad one-sided restraint on workers’ mobility.

Trade secret law gives businesses the power to protect intangibles that may not be eligible for copyright or patent protection. Employers can use trade secret law to prohibit both current and former employees from divulging valuable information to competitors and the public. Trade secrets are defined as information that derives its value from being unknown or unascertainable to the public and that is subject to reasonable efforts to maintain its secrecy.¹⁹⁷ Misappropriation of trade secrets, whether through espionage or breach of a contractual or common law duty with an employer, is actionable.¹⁹⁸ The Economic Espionage Act (amended and strengthened several times in the past five years) makes misappropriation of trade secrets a federal crime.¹⁹⁹ In some states, courts can enjoin workers possessing trade secrets from working

¹⁹⁶ Orly Lobel, *The New Cognitive Property: Human Capital Law and the Reach of Intellectual Property*, 93 TEX. L. REV. 789, 797 (2015).

¹⁹⁷ UNIF. TRADE SECRETS ACT § 1(4) (amended 1985). Nearly all states have adopted the Uniform Trade Secrets Act or some modification of it. H.R. REP. NO. 114-529, at 4 (2016).

¹⁹⁸ *Id.* at §§ 1(1) & (2).

¹⁹⁹ 18 U.S.C. § 1839.

for a rival or starting a competing business when they determine that the workers will inevitably disclose the trade secret.²⁰⁰ Under this “inevitable disclosure” doctrine, courts can impose an extraordinary “non-compete remedy” on a worker to protect his or her employer’s intangibles.²⁰¹ Trade secret law grants employers a potent tool with which to protect valuable intangibles.

Employers can also condition employment on employees’ signing non-disclosure agreements. Through a non-disclosure requirement, employers can convert statutory duties under state trade secret law into contractual duties. Employees bound by a non-disclosure condition are prohibited from divulging or publicizing information listed in the contract.²⁰² Today, non-disclosure clauses often prohibit the sharing of a long list of information including “any other information not generally known to the public which, if misused or disclosed, could reasonably be expected to adversely affect Company’s business.”²⁰³ Non-disclosure agreements do restrict employee’s freedom to communicate to new employers and colleagues but are less restrictive than non-competes.

For employers who find these legal tools are inadequate for protecting their intangible investments and believe preventing employee departure is still essential, they have multiple options that are superior to non-compete clauses. Employers can retain workers, and ensure their commitment and loyalty, through higher wages and salaries, more generous benefits, and fair treatment. This dynamic is an important pro-worker element of the more fluid employee-employer relationship in the United States today.²⁰⁴ Indeed, this threat of exit is an important

²⁰⁰ Eleanore R. Godfrey, *Inevitable Disclosure of Trade Secrets: Employee Mobility v. Employer’s Rights*, 3 J. HIGH TECH. L. 161 (2004).

²⁰¹ *E.g.*, *PepsiCo, Inc. v. Redmond*, 54 F.3d 1262, 1269 (7th Cir. 1995).

²⁰² Orly Lobel: *Enforceability TBD: From Status to Contract in Intellectual Property Law*, 96 B.U. L. REV. 869, 874 (2016).

²⁰³ *Id.* at 874-76.

²⁰⁴ Katherine V.W. Stone, *Knowledge at Work: Disputes Over the Ownership of Human Capital in the Changing Workplace*, 34 CONN. L. REV. 721, 733-37 (2002).

source of power for at least some workers in the contemporary workplace.²⁰⁵ Employers can retain workers, who may otherwise leave and accept employment with another company, through raises and promotions.

In lieu of non-compete clauses, employers can also offer employment contracts to workers entrusted with important information. Employment contracts depart from the default rule of at-will employment and offer guaranteed employment to workers for a term.²⁰⁶ They bind both the employer and the employee to commit to maintaining the relationship for a fixed period and limit the ability of both sides to end the contract before the completion of the term. And unlike non-compete clauses, employment contracts do not restrain a worker from finding employment *after* he or she has departed or otherwise left.

Employment contracts also serve an important channeling function. Because non-competes are often one-sided obligations, employers can use them as a matter of course without considering whether they are necessary to protect intangibles.²⁰⁷ In contrast to non-compete clauses often under which only the worker is bound, both the employer and the employee make a binding commitment in an employment contract. By retaining workers through term contracts, employers are required to deliberate on the importance of the intangibles involved. If employers seeking to protect intangibles cannot rely on non-competes and have to offer employment contracts and bind themselves, they are compelled to identify and evaluate the significance of the intangibles at stake and determine an employment term that is just long enough to recoup their investment in the intangibles.

²⁰⁵ Arnow-Richman, *supra* note 96, at 983-84.

²⁰⁶ *See, e.g., Luteran v. Loral Fairchild Corp.*, 455 Pa. Super. 364, 370, 688 A.2d 211 (1997). (“In order to rebut the presumption of at-will employment, a party must establish one of the following: (1) an agreement for a definite duration; (2) an agreement specifying that the employee will be discharged for just cause only; (3) sufficient additional consideration or (4) an applicable recognized public policy exception.”).

²⁰⁷ Staidl, *supra* note 42, at 119.

VI. The FTC Should Prohibit Non-Compete Clauses as an Unfair Method of Competition

Through a rulemaking, the FTC should declare worker non-compete clauses to be an unfair method of competition and classify them as per se illegal under the FTC Act. Non-competes, in general, function as contracts of adhesion that impair labor market mobility. Even if an employer does not intend to, or cannot enforce, them in state court, non-competes can deter workers from leaving a job and chill labor market mobility. By binding workers to their current employer, non-competes reduce wages, depress business formation, and lock workers into discriminatory, hostile, or unsafe workplaces. On top of these harms to workers, dominant firms and other powerful incumbents can also use non-competes to deprive rivals and new entrants of specialized workers and exclude these competitors from the market. In contrast to these real harms, the business justifications for non-compete clauses are fallacious. Businesses can protect their investment in intangibles in more effective ways that are also less restrictive for workers, including trade secret law, non-disclosure agreements, and employment contracts.

The FTC has broad authority to interpret “unfair methods of competition”²⁰⁸ and should use this authority to prohibit non-competes. To deter employers’ use of non-competes, the FTC should prohibit them as an unfair method of competition and not merely hold them to be unenforceable in court. Under the requested rule, an employer who presents, enforcers, or otherwise uses worker non-competes would be liable under the FTC Act.

Tens of millions of workers are subject to non-compete clauses in the workplace today. Due to disparities in bargaining power and behavioral biases among workers, employers can generally condition employment on an employee’s acceptance of a non-compete clause. Non-

²⁰⁸ 15 U.S.C. § 45.

competes in the workplace function as contracts of adhesion. Even in states where they are unenforceable, non-competes clauses discourage worker mobility and lock workers into their current jobs. For workers, non-competes depress wages, frustrate their ability to start new businesses, and compel them to remain in discriminatory and hostile work environments. In addition to harming workers, firms with market power can use non-competes to exclude rivals. Through worker non-competes, incumbent firms can control the supply of labor and deprive rivals and new entrants of the workers that they need to grow and compete. Furthermore, non-competes can divert workers to larger firms and hinder the growth and entry of small and new firms.

The business justifications for employer use of non-compete clauses do not stand up to scrutiny. The underlying assumption is that employers need to guard their investment in intangibles (for example, customer lists, trade secrets, and training) against free riding by rivals and others. What is disparaged as free riding is often beneficial sharing of information and knowledge among workers and across firms. While firms have a motive to defend against perceived free riding by competitors, their private incentive to protect intangibles can conflict with the public interest in allowing the free dissemination and sharing of information and knowledge. Furthermore, non-competes are a flawed means of protecting against free riding. They are too broad and too narrow—binding a worker to his or her current employer and preventing full freedom to use his or her skills and yet failing to truly protect the relevant intangible. Employers have alternative and less restrictive methods of protecting their intangibles. They can use trade secret law and non-disclosure agreements to protect customer lists and trade secrets, and training, offering promotions, raises, and job contracts to retain employees.

The FTC has expansive authority to interpret the FTC Act’s prohibition on unfair methods of competition.²⁰⁹ Congress intended the FTC to be an expert policymaker on antitrust and delegated authority to the Commission to identify and restrict unfair methods of competition over time.²¹⁰ The Supreme Court has affirmed this congressional grant of policymaking power to the FTC, declaring the Commission, in defining “the congressionally-mandated standard of fairness,” can “like a court of equity, consider[] public values beyond simply those enshrined in the letter or encompassed in the spirit of the antitrust laws.”²¹¹ Furthermore, modern administrative law grants agencies, such as the FTC, the authority to interpret broad, open-ended statutes such as the FTC Act.²¹²

The FTC Act authorizes the FTC to outlaw practices beyond those that are reasonably certain to reduce competition or create a monopoly. Congress enacted the FTC Act and created the Commission to stop anticompetitive practices in their “incipiency.”²¹³ The Supreme Court has stated that “Congress enacted § 5 of the Federal Trade Commission Act to combat in their

²⁰⁹ See generally Sandeep Vaheesan, *Resurrecting “A Comprehensive Charter of Economic Liberty”: The Latent Power of the Federal Trade Commission*, 19 U. PA. J. BUS. L. 645 (2017).

²¹⁰ See Neil W. Averitt, *The Meaning of “Unfair Methods of Competition” in Section 5 of the Federal Trade Commission Act*, 21 B.C. L. REV. 227, 237 (1980) (“The judicial decisions which have reviewed [the FTC Act’s] legislative history confirm that the Commission has, as it must have, considerable flexibility in determining which particular acts or practices will constitute ‘unfair methods of competition.’”).

²¹¹ *Sperry & Hutchinson*, 405 U.S. at 244. See also *Ind. Fed. Of Dentists*, 476 U.S. at 454 (“The standard of ‘unfairness’ under the FTC Act is, by necessity, an elusive one, encompassing not only practices that violate the Sherman Act and the other antitrust laws, but also practices that the Commission determines are against public policy for other reasons[.]”).

²¹² *Chevron, USA, Inc. v. Natural Resources Defense Council, Inc.*, 467 U.S. 837, 842-43 (1984). See Justin (Gus) Hurwitz, *Chevron and the Limits of Administrative Antitrust*, 76 U. PITT. L. REV. 209, 263-64 (2014) (“[F]irst, . . . the breadth of [Section 5] constructions likely to be considered permissible is very large; and second, . . . the proper forum in which to challenge such interpretations is not before the Article III courts. Given the breadth of the statute, once the matter has reached that point, there is great weight in favor of the FTC’s position receiving *Chevron* deference.”).

²¹³ See generally A. Everette MacIntyre & Joachim J. Volhard, *The Federal Trade Commission and Incipient Unfairness*, 41 GEO. WASH. L. REV. 407 (1973). See also Marc Winerman, *The Origins of the FTC: Concentration, Cooperation, Control, and Competition*, 71 ANTITRUST L.J. 1, 74 (2003) (“[The principal Congressional supporters] of the FTC Act wanted a new agency that would prosecute if the Department faltered, enforcing a flexible new standard that could reach where the Sherman Act might not.”).

inefficiency trade practices that exhibit a strong potential for stifling competition.”²¹⁴ The principal Senate sponsor of the FTC Act wanted the FTC to become the “social machinery which will protect the individual from oppression and wrong.”²¹⁵ A Senate colleague echoed this statement, stating “no one here—I can speak with confidence for the entire Senate—would put one obstacle in the way of punishing dishonesty, of preventing oppression, of prohibiting exactions.”²¹⁶

Considering the documented harms and unconvincing business justifications for non-competes, the FTC should hold worker non-compete clauses to be an unfair method of competition and categorize them as per se illegal. For the reasons presented, non-competes are competitively suspect. Under existing Sherman Act precedent, non-competes arguably should trigger a strong presumption of illegality because they hurt competition for workers and can impair product market competition and rest on dubious justifications.²¹⁷ Section 5’s unfair methods of competition prong is expressly broader than the Sherman and Clayton Acts.²¹⁸ Accordingly, the FTC has the authority to go further and classify non-competes as per se illegal.

In exercising its Section 5 authority, the FTC’s competition rule should make the presentation or the enforcement of non-competes illegal. Prohibiting non-compete clauses, as

²¹⁴ *FTC v. Texaco Inc.*, 393 U.S. 223, 225 (1968).

²¹⁵ 51 CONG. REC. 11,109 (1914) (statement of Sen. Newlands).

²¹⁶ *Id.* at 14,782 (statement of Sen. Burton).

²¹⁷ *In re Polygram Holding, Inc.*, 136 FTC 310, 344 (2003), *pet’n denied Polygram Holding, Inc. v. FTC*, 416 F.3d 29 (D.C. Cir. 2005) (“A plaintiff may avoid full rule of reason analysis, including the pleading and proof of market power, if it demonstrates that the conduct at issue is inherently suspect owing to its likely tendency to suppress competition.”).

²¹⁸ *See Ind. Fed. of Dentists*, 476 U.S. at 454 (“The standard of ‘unfairness’ under the FTC Act is, by necessity, an elusive one, encompassing not only practices that violate the Sherman Act and the other antitrust laws, but also practices that the Commission determines are against public policy for other reasons[.]”). For example, the FTC has held that Section 5 reaches invitations to collude, which may fall outside the Sherman Act’s prohibition on “restraints of trade.” *See, e.g., In re Drug Testing Compliance Group, LLC*, 2015 WL 9254882, *8 (FTC 2015) (“Mr. Crossett's communication to Competitor A is an attempt to arrange a customer allocation agreement between the two companies. The invitation, if accepted, would be a per se violation of the Sherman Act. The Commission has long held that invitations to collude violate Section 5 of the FTC Act[.]”).

opposed to making them only unenforceable in court, is essential. Regardless of whether they are enforceable in state court today, many employers condition employment on workers acquiescing to a non-compete provision. For instance, California law bars the enforcement of non-competes.²¹⁹ Nonetheless, nearly one-in-five workers in California is subject to a non-compete clause.²²⁰ The mere existence of non-compete contracts, even when legally not binding, still inflicts real harms on workers. Under the requested rule, employers who use non-competes would violate the FTC Act and be subject to FTC enforcement actions.

VII. Conclusion

Through non-compete clauses, employers have deprived tens of millions of workers of the freedom to leave their current job to accept a new job or start a business. Non-competes prohibit workers, following separation from an employer, from seeking employment in a similar line of work or industry or establishing a competing business for a specified period in a geographic area. Approximately 30 million workers, across a wide range of fields and occupations including accountants, engineers, and fast food workers, are bound by non-compete clauses.

In labor markets, employers generally have the power to impose non-compete clauses on workers to the detriment of workers. Due to workers' dependence on wages and lack of union representation and concentration among employers in local labor markets, the employee-employer relationship is defined by inequality. Even when they have employers competing for their services, workers are likely to bargain about wages and benefits, not contingent terms such as non-compete clauses. These factors taken together indicate that non-compete clauses function

²¹⁹ Cal. Bus. & Prof. Code § 16600.

²²⁰ Starr et al., *supra* note 8, at 16.

as contracts of adhesion. Employers present non-competes clauses to workers as a standard form contract on a take it-or-leave it basis.

By restricting labor market mobility, non-compete clauses inflict significant harms on workers and the broader public. Non-compete clauses bind workers to their current employers and thereby depress wages and wage growth and deter the formation of new firms. The effects extend beyond wages and firm creation rates. Due to non-competes, many workers may, in effect, be compelled to remain in discriminatory, hostile, or dangerous work environments.

Non-compete clauses also can impair product market competition and help protect monopolistic and oligopolistic market structures. In concentrated markets, dominant incumbent firms can use non-compete clauses as a way of depriving rivals and new entrants of essential workers and limit the growth of actual and would-be competitors. Even in the absence of any exclusionary intent, non-competes can direct workers toward larger existing firms and away from smaller firms and new entrants.

While the harms from non-compete clauses are real, their justifications are unpersuasive. Employers and their representatives justify worker non-compete clauses as a method of protecting their intangible investments, such as trade secrets and employee training, from free riding by rival firms. If they were unable to protect against this type of free riding, they would underinvest in intangibles, according to this theory. This rationale does not stand up to scrutiny and depends on a series of questionable, if not false assumptions. Insofar as employers should be permitted to protect intangibles, they have several less restrictive and more appropriately tailored alternatives to non-compete clauses.

The FTC should initiate a rulemaking to prohibit employers from presenting non-compete clauses as a condition of employment. The FTC has expansive authority to interpret the

FTC Act’s prohibition on “unfair methods of competition.” Given that non-compete clauses inflict real harms on workers and competition and rest on unpersuasive theoretical justifications, the FTC should hold these clauses to be an unfair method of competition. Accordingly, the FTC should hold these clauses to be a per se violation of Section 5 of the FTC Act. Under this rule, employers who use non-competes with their workers would violate federal law and face legal liability under the FTC Act.

Certification

The undersigned certifies, that, to the best knowledge and belief of the undersigned, this petition includes all information and views on which the petition relies, and that it includes representative data and information known to the petitioner, including information that is unfavorable to the petition.

Organizational Petitioners

The **Open Markets Institute (OMI)** is a non-profit organization dedicated to promoting fair and competitive markets. It does not accept any funding or donations from for-profit corporations. Its mission is to safeguard our political economy from concentrations of private power that undermine competition and threaten liberty, democracy, and prosperity. OMI regularly provides expertise on antitrust law and competition policy to Congress, federal agencies, courts, journalists, and members of the public.

The **AFL-CIO** is a democratically governed federation of 55 unions representing 12.5 million working people.

The **Artist Rights Alliance (ARA)** is an artist-run non-profit advocacy group representing creators in the digital landscape. ARA's work is significant to anyone who creates and makes a living from their creations. ARA's objectives are two-fold: First, economic justice for musicians and music creators in the digital domain. Second, ensuring that the current and future generations of creators retain the rights needed to create and benefit from the use of their work and efforts. ARA has grown into a national organization based on representation, advocacy, and mobilization for sustainable careers in the digital age.

The **Center for Popular Democracy (CPD)** works to create equity, opportunity and a dynamic democracy in partnership with high-impact base-building organizations, organizing alliances, and progressive unions. CPD strengthens our collective capacity to envision and win an innovative pro-worker, pro-immigrant, racial and economic justice agenda.

coworker.org is a digital lab dedicated to supporting worker voice through our platform, trainings, and building networks of workers to improve their jobs.

The **Demand Progress Education Fund** educates its two million members and the general public about matters pertaining to the democratic nature of our nation's communications infrastructure and governance structures, and the impacts of corporate power over our economy and democracy.

The **Economic Policy Institute (EPI)** is a nonprofit, nonpartisan think tank created in 1986 to include the needs of low- and middle-income workers in economic policy discussions. EPI's mission is to inform and empower individuals to seek solutions that ensure broadly shared prosperity and opportunity. EPI believes every working person deserves a good job with fair pay, affordable health care, and retirement security. To achieve this goal, EPI conducts research and analysis on the economic status of working America. EPI proposes public policies that protect and improve the economic conditions of low- and middle-income workers and assesses policies with respect to how they affect those workers.

EIG is a bipartisan public policy organization, combining innovative research and data-driven advocacy to address America's most pressing economic challenges and advance solutions that empower entrepreneurs and investors to forge a more dynamic economy.

The **Institute for Local Self-Reliance (ILSR)** is a national research and advocacy organization that challenges concentrated economic and political power, and instead champions an approach in which ownership is broadly distributed, institutions are humanly scaled, and decision-making is accountable to communities.

Lake Research Partners is a national leader in public opinion research and strategy for Democratic and progressive candidates, causes, and campaigns.

Make the Road New York is a democratic, community-based membership organization representing more than 23,000 working class and immigrant families throughout New York City, Long Island, and Westchester, New York.

The **National Employment Law Project (NELP)** is a nonprofit organization with more than 45 years of experience advocating for the employment and labor rights of low wage and unemployed workers. NELP seeks to ensure that all employees receive the full protection of employment and labor laws, and that employers are not rewarded for skirting those basic rights. NELP promotes policies at the federal, state, and local level to protect workers' rights and has litigated and participated as amicus in numerous cases in the federal appellate courts and the U.S. Supreme Court.

Organization United for Respect (OUR) is a national organization working to reshape the economy so that all can live free, full lives with their families and loved ones.

Public Citizen was founded 46 years ago and serves as the people's voice in the nation's capital. It identifies excessive corporate power as the most serious threat to the values and policy objectives we most treasure: justice, health and safety, ecological sustainability, a functioning democracy, freedom, and equality. Public Citizen has delved into an array of areas, but its work on each issue shares an overarching goal: To ensure that all citizens are represented in the halls of power. Public Citizen carries out an advocacy agenda through divisions with specialized and extraordinary expertise. Its Health Research Group is recognized as the leading campaigner for pharmaceutical safety. Its Litigation Group operates the preeminent public interest Supreme Court practice. Its Global Trade Watch is recognized as a leading force for fair trade. Its Energy Project combines consumer and environmental advocacy for a sustainable future. Its Congress Watch project runs cutting-edge advocacy campaigns on a diverse array of issues, from worker safety to clean government. Moreover, Public Citizen's Austin office has helped turn Texas into a world-leading wind energy producer.

The **Revolving Door Project** is a nonpartisan effort to educate civil society in order to counteract the advantage that Wall Street and corporate America have in how the executive branch writes the rules of the economy. It does this by alerting and educating the media and activists when hard working people are being taken advantage of and by whom. If the executive branch is to write rules that structure the economy away from rent extraction and in the direction of greater economic equality, public-interest minded people must hold key executive branch positions.

The **Roosevelt Institute**, a New York-based think tank, promotes bold policy reforms that would redefine the American economy and democracy. With a focus on curbing corporate power and reclaiming public power, Roosevelt is helping people understand that the economy is shaped by choices—via institutions and the rules that structure markets—while also exploring the economics of race and gender and the changing 21st-century economy. Roosevelt is armed with a transformative vision for the future, working to move the country toward a new economic and political system: one built by many for the good of all.

Service Employees International Union (SEIU) unites 2 million diverse members in the United States, Canada, and Puerto Rico. SEIU members working in the healthcare industry, in the public sector, and in property services believe in the power of joining together on the job to win higher wages and benefits and to create better communities while fighting for a more just society and an economy that works for all, not just corporations and the wealthy.

Towards Justice is a Denver-based non-profit law firm that represents workers in attacking systemic abuses in the labor market through impact litigation, strategic policy advocacy, and capacity building. Towards Justice is particularly interested in attacking anti-competitive practices in the labor market that undermine worker power.

The **UFCW** is the largest private sector union in the United States, representing 1.3 million professionals and their families in grocery stores, meatpacking, food processing, retail shops and other industries. Its members help put food on our nation's tables and serve customers in all 50 states, Canada and Puerto Rico

UNITE HERE is a labor union that represents 270,000 working people across Canada and the United States. Its members work in the hotel, gaming, food service, manufacturing, textile, distribution, laundry, transportation, and airport industries. Its membership is diverse. Its members are predominantly women and people of color, and hail from all corners of the planet. Together, members are building a movement to enable people of all backgrounds to achieve greater equality and opportunity.

Individual Petitioners

Individual petitioners' institutional and other organizational affiliations are provided solely for identification purposes.

Alan Hyde is Distinguished Professor and Sidney Reitman Scholar at Rutgers Law School. He has been a visiting professor at Yale, Columbia, NYU, Toronto, Michigan, Cornell, Fordham, Cardozo, and Brooklyn law schools. He is the author of *Working in Silicon Valley: Economic and Legal Analysis of a High-Velocity Labor Market* (2003), an early analysis of the short job tenures typical of high technology in California and their benefit for startup firms, knowledge diffusion, and technical innovation. He lectures and publishes frequently on intellectual property issues in employment, in the US and around the world. He is a member of the American Law Institute and helped shape the chapter on non-competes in the new Restatement of Employment Law.

Amy Kastely is a senior professor of law at St. Mary's University Law School and a member of the bar in Texas and New Mexico. She is a nationally recognized authority on contract law, having co-authored a widely-known text entitled *Contracting Law*. She also has written numerous articles exploring how law is shaped by narratives of race, gender, class, and other systems of subordination. She served as lead counsel in *Esperanza et al. v. City of San Antonio*, the first case recognizing the importance of cultural rights in public arts funding. In addition, she has represented the Esperanza and numerous community coalitions in litigation and organizing projects involving a broad spectrum of important issues, including protection of the Edward's Aquifer; the right of communities to use public streets, sidewalks, and parks for cultural events and political expression; racial bias in San Antonio's historic preservation practices; and the public's right to witness government deliberations and to hold government officials accountable to democratic values.

Ann C. McGinley is William S. Boyd Professor of Law at the University of Nevada, Las Vegas, Boyd School of Law. The co-director of the Workplace Law Program, McGinley has published three books and more than sixty law review articles and book chapters. Her most recent book is *Masculinity at Work: Employment Discrimination Through a Different Lens* (NYU Press, 2016). McGinley has published articles about gender effects on lawyers' workplace conditions and continues to research this topic; she is the editor of the upcoming *Feminist Judgments: Rewritten Opinions in Employment Discrimination Law* (Cambridge University Press, forthcoming 2020). Professor McGinley has lectured at many universities in the United States and abroad and is a Visiting Foreign Professor at Universidad Adolfo Ibañez in Santiago, Chile where she lectures annually in Spanish about U.S. sexual harassment law. She currently serves on Nevada's Task Force on Sexual Harassment and Discrimination Law and Policy.

Ariana R. Levinson is a professor of law at the University of Louisville Brandeis School of Law who has taught employment law for the past ten years of her fourteen-year academic career. She has authored many law review articles, including some on the topic of technology and the

workplace. Before entering academia, she worked as an attorney representing working people, including spending a year as the fellow in the AFL-CIO Legal Department.

Barbara Bucholtz is professor of law at the University of Tulsa College of Law. She received her J.D. from Valparaiso University of Law and her LL.M. in environmental law from George Washington University (with highest honors). Prior to joining the Tulsa law faculty, she practiced corporate law, business litigation, class action litigation, and estate planning in the Chicago area and in Tulsa. She is a former law clerk for the Northern District of Oklahoma. She serves on the boards of several organizations in the business and the nonprofit sectors. Her teaching interests include contracts, sales, international private law, comparative corporate law, corporate law, securities law, international trade, American legal history and jurisprudence, the law of nonprofit organizations, and legal analysis and writing. She writes and speaks on topics that include international trade agreements, environmental law, and nonprofit associations.

Ben Templin is professor of law at the Thomas Jefferson School of Law. Prior to joining the faculty in 2003, he was a corporate attorney at Wilson Sonsini Goodrich & Rosati where his practice focused on general corporate law for early stage technology companies. Following graduation from the University of California, Boalt Hall School of Law, he taught legal methods to undergraduates at the University of California, Berkeley. His web site provides instruction in how to study and prepare for law school exams. Before going to law school, he was an editor in computer magazine publishing. He has published a series of law review articles on Social Security reform and government investment in private enterprise.

Carol Chomsky is professor of law and former Associate Dean for Academic Affairs at the University of Minnesota Law School, where she has been on the faculty since 1985. She teaches contracts, sales, and legislation and regulation, and she teaches and supervises field placements in the Judicial Externship class. She is co-author of *Contracts: A Contemporary Approach* (3d ed. 2018) and *Learning Sales Law* (2016), both published by West Academic. Her scholarship addresses topics in history, contracts law, and pedagogy. She was co-President of the Society of American Law Teachers in 2000-2002 and served as President of Minnesota Women Lawyers in 1993-1994. Before entering academia, she clerked for Judge Spottswood W. Robinson on the D.C. Circuit and practiced law in Washington, D.C.

Catherine Fisk is the Barbara Nachtrieb Armstrong Professor of Law at the University of California, Berkeley, where she teaches courses on employment and labor law, civil procedure, and the legal profession. She is the author of *Working Knowledge: Employee Innovation and the Rise of Corporate Intellectual Property*, a multiple prize-winning book examining the history of employer-employee disputes over workplace knowledge and intellectual property. She has written articles on the law of noncompetition agreements and trade secrets, as well as books and articles on numerous other topics related to the law of the workplace. Prior to joining the Berkeley Law School faculty, she held chaired professorships at Duke University and the University of California, Irvine. She is a graduate of Princeton University and Berkeley.

Charlotte Garden is an expert in labor & employment law. She is an associate professor (with tenure) at the Seattle University School of Law, where she teaches labor law, employment law,

constitutional law, appellate litigation, and legislation & regulation. Her scholarship focuses on the intersection of work/labor/technology and the constitution. Her articles have appeared in the Emory Law Journal, Boston University Law Review, George Washington Law Review, Fordham Law Review, William & Mary Law Review, the University of Chicago Legal Forum, and the Harvard Civil Rights-Civil Liberties Law Review. She is a co-author of two leading labor & employment law casebooks: Modern Labor Law in the Private and Public Sectors; and the forthcoming ninth edition of Employment Law Cases and Materials. In 2019, Cambridge University Press will publish her co-edited volume, The Cambridge Handbook of U.S. Labor Law.

Chris Odinet is an associate professor of law and affiliate associate professor in entrepreneurship at the University of Oklahoma. His primary teaching and research interests focus on the intersection of law, credit, consumer protection, and technology. He is also active in law reform efforts, serving on committees of the Uniform Law Commission, the European Law Institute, and the Association of American Law Schools. He is also the co-editor of the Annual Survey of Consumer Finance Law, which appears in the American Bar Association's The Business Lawyer publication.

Chrystin Ondersma is professor of law and Judge Morris Stern Scholar at the Rutgers Law School (Newark). She joined the faculty in Spring 2010. Her scholarship focuses on Bankruptcy and Commercial Law. She received her J.D. magna cum laude in 2007 from Harvard Law School where she was the recipient of a Goldsmith Academic Fellowship and an executive editor of the Harvard Civil Rights-Civil Liberties Law Review. From 2007 to 2008 she clerked for the Honorable Michael Daly Hawkins of the U.S. Court of Appeals for the Ninth Circuit. Prior to joining the Rutgers faculty, she was an associate in the Business Finance and Restructuring Department at Weil, Gotshal & Manges LLP in New York.

Colin P. Marks graduated magna cum laude from the University of Houston Law Center in 2001 where he served as an associate editor on the Houston Law Review. After law school, he clerked for the Honorable Harold R. DeMoss Jr. on the United States Fifth Court of Appeals for two years. In the fall of 2003, he joined the law firm of Baker Botts, L.L.P., in Houston, Texas where he was an associate in the trial department. At Baker Botts, his practice concentrated on commercial litigation, as well as some pro bono criminal work. He left Baker Botts in the summer of 2006 to join St. Mary's University School of Law. He is a member of the American Law Institute (ALI), Co-Chair of the ABA's UCC Annual Survey Subcommittee, and member of the executive committee for the AALS Sections on Agency, Partnerships, LLCs and Unincorporated Associations, and Section on Contracts. In 2018 he was elected President of the South Eastern Association of Law Schools and will begin his one year term in 2019.

Cynthia Ho is the Clifford E. Vickrey Research Professor at Loyola University of Chicago School of Law, where she is also the Director of the Intellectual Property Program. Prior to joining the faculty at Loyola, she litigated intellectual property cases and also was involved in evaluating IP assets and liabilities. She holds a B.A. from Boston University and a J.D. from Duke Law. She is a member of the New York Bar.

Dalié Jiménez is a professor of law at the University of California, Irvine School of Law. She teaches and writes in the areas of contract, bankruptcy, and consumer protection law.

Edward Janger is the David M. Barse Professor of Law, Associate Dean for Research and Scholarship, and Co-Director of the Brooklyn Law School Center for the Study of Business Law and Regulation. He teaches and writes in the areas of bankruptcy, secured credit and consumer finance.

Eileen Appelbaum is Co-Director of the Center for Economic and Policy Research, Washington, D.C., Fellow at Rutgers University Center for Women and Work, and Visiting Professor at the University of Leicester, UK. Prior to joining CEPR, she held positions as Distinguished Professor and Director of the Center for Women and Work at Rutgers University and as professor of economics at Temple University. She is past president of the Labor and Employment Relations Association. Her research focuses on organizational restructuring and outcomes for firms and workers; private equity and financialization; and work-family policies. *Private Equity at Work: When Wall Street Manages Main Street*, coauthored with Rosemary Batt, was selected by the Academy of Management as one of the four best books of 2014 and 2015, and was a finalist for the 2016 George R. Terry Award. *Unfinished Business*, coauthored with Ruth Milkman, examines the effects of paid family leave in California on employers and employees. It has been widely cited in discussions of national paid family and medical leave policy. Her current research examines the implications of consolidation of hospitals and decentralization of health services to outpatient care centers for employees in these two segments of the healthcare industry.

Frank Pasquale has researched and written extensively on law and political economy. He edited a special issue of *Critical Analysis of Law*, entitled *New Economic Analysis of Law*. He is the author of *The Black Box Society* (Harvard University Press, 2015), which develops a social theory of reputation, search, and finance, and has been translated into Chinese, Korean, French, and Serbian. The book offered critical legal commentary on algorithmic approaches to profiling, and recommended law & policy to improve the information economy. He has served on the NSF-sponsored Council on Big Data, Ethics, & Society, and has co-authored a casebook on administrative law and co-authored and authored over 50 scholarly articles.

Henry Drummonds is professor of law at Lewis and Clark Law School in Portland, Oregon.

Jane Flanagan is a Leadership in Government fellow with the Open Society Foundations and a Visiting Scholar at IIT Chicago-Kent School of Law. She is the former chief of the Workplace Rights Bureau within the Illinois Attorney General's Office, a bureau she founded and led from November 2015 through December 2018. Under her leadership, the bureau brought national attention to the increasing use of noncompete agreements for low-wage workers and negotiated settlements to release thousands of such workers nationwide from those agreements. Previously, she was an assistant attorney general in Maryland and counsel to Maryland's Division of Labor and Industry. She began her career in private practice litigating wage and hour collective action cases on behalf of employees.

Jeffrey W. Stempel is the Doris S. & Theodore B. Lee Professor of Law at the William S. Boyd School of Law, University of Nevada, Las Vegas where he teaches insurance law, civil procedure, contracts, professional responsibility, and evidence. Before joining the UNLV faculty in 1999, he was the Fonvielle & Hinkle Professor of Litigation at Florida State University College of Law and professor of law at Brooklyn Law School. Prior to becoming a law teacher, he was a civil litigator. He is a member of the American Law Institute, the European Law Institute, The Association of the Bar of the City of New York, the American Bar Association, the American Bar Foundation, the American Judicature Society, and the Law & Society Association. He received his B.A. degree from the University of Minnesota in 1977 and his J.D. degree from Yale Law School in 1981 and is admitted in Nevada and Minnesota.

Joan Vogel is professor of law at Vermont Law School. She specializes in employment law, anthropology of law, consumer law, medical malpractice, and tort reform. The courses she has taught at Vermont Law School include commercial law, employment discrimination law, employment law, law and anthropology, and torts. She received a BA degree from George Washington University in 1973. She earned an MA degree in anthropology in 1975 and a JD degree in 1981, both from the University of California at Los Angeles. She served as a teaching associate in UCLA's Anthropology Department from 1975 to 1976 and performed research in African law. In 1979 and 1980, she clerked with two Los Angeles law firms, Levy and Goldman, and Schwartz, Steinsapir, Dohrmann and Krepack. She was a law clerk for Judge Alfred T. Goodwin of the U.S. Court of Appeals, Ninth Circuit, in Oregon from 1981 to 1982. From 1982 to 1989, she served as assistant professor at the University of Pittsburgh School of Law, an associate visiting professor at Albany Law School, and an associate professor at Oklahoma City University Law School. She joined the faculty at VLS in 1989. She testified before the Vermont Senate Judiciary Committee on modification of the state's Human Rights Commission in 1995 and was a consultant for the drafting of Vermont's employment law bill in 1997. In 1998, she helped to write the tobacco reimbursement statute that passed the Vermont Legislature. She has presented widely on topics of legal pluralism, new teaching methods in labor law, tort reform, and on the "Lemon Laws." She has served as chair of the Law and Anthropology and the Labor and Employment Law sections of the Association of American Law Schools.

Karen Cross, as Associate Dean for Administration at John Marshall, develops law school policies and assists with the law school's pending acquisition by the University of Illinois at Chicago. She recently returned to the law school after taking leave to work in central administration on Northwestern University's Evanston campus. As a law student, Karen served as editor of the Harvard Civil Rights-Civil Liberties Law Review. After law school, she conducted research as a Fulbright scholar in former Yugoslavia and worked as an associate with Cleary, Gottlieb, Steen & Hamilton in New York. Since joining the John Marshall faculty, Professor Cross has taught at the Catholic University of Portugal in Lisbon, the University of San Diego's summer program in Moscow, the Central European University in Budapest, and the MBA program for Executives and International Managers at UIC. Her teaching and scholarship focuses on contract law, international economic law, arbitration, and higher education law and policy. She is a contributor to Kluwer Arbitration Blog and ASIL Insights, and her scholarship has appeared in the Journal of Collective Bargaining in the Academy (refereed), Journal of

International Economic Law (refereed), Ohio State Law Journal, and Michigan Journal of International Law, among other publications. She was awarded a Fulbright Distinguished Chair to Portugal in 2008 and has since served on many peer review committees for the Fulbright program. She teaches contracts, education law and policy, and arbitration-related courses. She received an M.S. in Higher Education Administration and Policy from Northwestern, a J.D., cum laude, from Harvard Law School, and a B.A., with distinction, from the University of Wisconsin-Madison.

Kathleen Engel is a research professor of Law at Suffolk University Law School. For several decades, she taught employment law, including the law of non-competes. Prior to her teaching career, she practiced employment law at Burnham, Hines and Dilday.

Lauren E. Willis is professor of law and William M. Rains Fellow at Loyola Law School, Los Angeles, teaches contracts and studies the effects of contractual fine print, among other topics.

Martha T. McCluskey is professor of law and William J. Magavern Scholar, State University of New York at Buffalo. Her scholarship focuses on economic policy and inequality. Her teaching and writing have included constitutional law, employment law, low-wage labor, gender equity and economic policy, and regulated industries. She is a member scholar of the Center for Progressive Reform, where she has worked on consumer and worker protection regulatory issues, and President of the Association for Promotion of Political Economy and the Law (APPEAL).

Meredith A. Munro is a partner at King & Greisen LLP, an all-women owned law firm specializing in employment law and civil rights. She has practiced law for over twenty-five years, in both private practice and the government. Her focus is on serving individuals, including employees and independent contractors, in ensuring that their federal and state law rights are protected. She most recently spoke on the topic of employee mobility and restrictive covenants before the Colorado Bar Association's Labor and Employment Section in January 2019 and is a member of the NELA and PELA, active advocates for legislation on behalf of workers.

Nancy Modesitt is a professor of law at the University of Baltimore School of Law. She teaches employment law, employment discrimination law, torts, and introduction to lawyering skills. She has written extensively on a variety of employment topics, including employment discrimination issues and whistleblowing claims. Before becoming a law professor, she was an attorney at the United States Department of Justice and in private practice at several large law firms. When she was in private practice, she negotiated and drafted noncompetition agreements on behalf of companies.

Nicolas Cornell is assistant professor of law at the University of Michigan Law School. He teaches and writes in the areas of contract law, moral philosophy, and private law theory. Prior to joining the faculty at Michigan, he was an assistant professor of Legal Studies and Business Ethics at the Wharton School, University of Pennsylvania. He holds a J.D. from Harvard Law School, a Ph.D. in philosophy from Harvard University. He clerked for the Vermont Supreme Court, and he is a member of the Pennsylvania Bar.

Orly Lobel is the Don Weckstein Professor of Law and founding member of the Center for Intellectual Property Law and Markets at the University of San Diego. A graduate of Harvard Law School and Tel-Aviv University, her interdisciplinary research is published widely in top journals in law, economics, and psychology and has recently been named among the most cited public law scholars in the nation. She has received numerous fellowships, grants and awards for her research and her books *You Don't Own Me: How Mattel v. MGA Entertainment Exposed Barbie's Dark Side* (Norton 2018) and *Talent Wants to Be Free: Why We Should Learn to Love Leaks, Raids and Free Riding* (Yale University Press 2013) are the winners of several prestigious awards.

Pamela Foohey is an associate professor at the Indiana University Maurer School of Law since 2014. Her research centers on bankruptcy, commercial law, and consumer law. Her work primarily involves empirical studies of bankruptcy and related parts of the legal system, combining quantitative and qualitative, interview-based research. Her work in business bankruptcy focuses on non-profit entities. She also is a co-investigator on the Consumer Bankruptcy Project, a long-term research project studying persons who file bankruptcy which has been the leading empirical study of consumer bankruptcy for the past 35 years. The results of her research have been featured in top media outlets, including The New York Times, NPR, Bloomberg, and the Washington Post.

Paul Secunda is professor of law at Marquette University Law School. He teaches employee benefits law, labor law, employment discrimination law, employment law, education law, civil procedure, and trusts and estates. He is also the founder and former faculty advisor of the Marquette Benefits and Social Welfare Law Review, which began publication in 2015. Professor Secunda is the faculty advisor for the student-run Marquette University Labor and Employment Law Society (LELS).

Rachel Arnow-Richman is the Chauncey Wilson Memorial Research Professor of Law at the University of Denver Sturm College of Law where she directs the Workplace Law Program. Prof. Arnow-Richman teaches and publishes in the areas of contracts, employment law and employment discrimination. She is co-author of two textbooks on Employment Law and has written numerous articles on topics including employment at will, covenants-not-to-compete, gender discrimination, and work family/conflict. She is a frequent media contributor on topics related to wrongful termination, employee competition, and equal employment opportunity, including sexual harassment. She currently serves on the Executive Committee of the American Association of Law Schools' Contracts & Commercial Law Section and is a past Chair of the Committee on Labor & Employment Law. Prior to joining the Denver Law School, Prof. Arnow-Richman was an Associate Professor at Texas A&M Law School, and she has held visiting appointments Colorado Law School, Fordham Law School, and Temple Law School. Prior to entering law teaching, she practiced employment law and commercial litigation at Drinker, Biddle & Reath in Philadelphia and clerked for the New Jersey Supreme Court.

Rick Bales is a faculty member at of Ohio Northern University Pettit College of Law (though he will be visiting at Akron 2018-19). He has published more than eighty scholarly articles and authored or co-authored five books on employment arbitration, alternative dispute resolution in

the workplace, employment law, labor law, and arbitration law. He has spoken widely on topics pertaining to innovative ways of teaching employment and labor law courses. Recent presentations have been in Berlin, Brisbane, Cartagena, Dhaka, Hanoi, Hong Kong, Manila, Milan, Moscow, Paris, Phnom Penh, Prato (Italy), Saigon, Siem Reap, and Yangon. He has twice been selected as a Fulbright Specialist, once each to Kuala Lumpur and Jakarta. He was Dean of ONU Law from 2013-16. Before coming to ONU, he was a faculty member at NKU Chase College of Law, where he was the Director of the Advocacy Center and he was Associate Dean of Faculty Development. Before that, he taught at the University of Montana Law School, Southern Methodist University Law School, and University of Houston Law School, and he litigated employment cases for Baker Botts and Baker Hostetler. He received his law degree from Cornell Law School and his B.A. from Trinity University.

Richard Carlson is a professor at South Texas College of Law-Houston and has taught employment law and other subjects since 1985. He earned his B.A. degree at Wake Forest University and his J.D. at the University of Georgia. He has written a number of articles on employment law and is co-author with Scott Moss of *Employment Law*, a casebook and text on employment law.

Robert H. Lande is the Venable Professor of Law at the University of Baltimore School of Law. Before this he worked at Jones Day and at the Federal Trade Commission. He has authored or co-authored more than 75 U.S. and 18 foreign legal publications. Eleven of his articles have been re-published in books or collections of articles. He has been quoted in the media hundreds of times about antitrust issues and has appeared on TV in the United States, France, the United Kingdom, and China. He has spoken at national events sponsored by many organizations, including the American Bar Association, Association of American Law Schools, National Association of Attorneys General, American Antitrust Institute, and the American Economic Association. He has testified before the US House of Representatives Judiciary Committee the US Senate Commerce Committee, the Antitrust Modernization Commission, and the federal antitrust enforcement agencies. He has given competition advice to enforcement officials from several foreign nations and has given legal talks in Italy, Spain, Japan, England, Belgium, Venezuela, and Peru. He is a past chair of the AALS Antitrust Section and has held many positions in the ABA Antitrust Section. He is an elected member of the American Law Institute and a member of the District of Columbia Bar. He received his JD and Masters in Public Policy from Harvard University and his BA from Northwestern University.

Ruben J. Garcia is Associate Dean for Faculty Development and Research, Co-Director of the UNLV Workplace Law Program, and professor of law at the University of Nevada, Las Vegas. He has held academic appointments at the University of California, Davis, California Western School of Law, the University of California San Diego, and the University of Wisconsin Law School, where he was a William H. Hastie Fellow. The courses that he teaches include constitutional law, employment law, labor law and professional responsibility.

Sanjukta Paul is assistant professor of law at Wayne State University. She is a graduate of Yale Law School and former judicial clerk for the U.S. Court of Appeals for the Ninth Circuit. Her current work involves the intersection of antitrust law and labor policy, and has appeared or is

forthcoming in the *UCLA Law Review*, the *Duke Journal of Law & Contemporary Problems*; the *Berkeley Journal of Employment & Labor Law*; and the *Cambridge Handbook of U.S. Labor Law*. Her forthcoming book *Solidarity in the Shadow of Antitrust: Labor & the Legal Idea of Competition* will be published by Cambridge University Press. Her paper *The Enduring Ambiguities of Antitrust Liability for Worker Collective Action* was recognized with the Jerry S. Cohen Memorial Fund's award for the best antitrust scholarship of 2016 (category prize). She previously served as David J. Epstein Fellow in Public Interest Law & Policy at UCLA Law School and was a litigator representing workers and civil rights plaintiffs for several years.

Sara Sternberg Greene studies consumer law, bankruptcy, poverty law, access to justice and tax. Broadly concerned in her scholarship with the relationship between law and inequality, she has recently focused on the role of the law in perpetuating and exacerbating poverty and inequality. Her work has been published or is forthcoming in the *New York University Law Review*, the *Duke Law Journal*, the *Minnesota Law Review*, and the *American Bankruptcy Law Journal*, among others. She has presented her Access to Justice research at the Department of Justice to the White House Legal Aid Interagency Roundtable, and along with others, integrated her research on the Earned Income Tax Credit into a federal policy proposal, *The Rainy Day EITC: A Reform to Boost Financial Security by Helping Low Wage Workers Build Emergency Savings*. Senators Cory Booker (D-NJ) and Jerry Moran (R-KS) adopted the proposal and are co-sponsors of a bipartisan bill proposing the Refund to Rainy Day Savings Act. She received her BA in 2002 from Yale University (magna cum laude and with distinction). She received her JD in 2005 from Yale Law School, where she received the Stephen J. Massey Prize for excellence in advocacy and served as notes editor for the *Yale Law Journal* and articles editor for the *Yale Law and Policy Review*. After clerking for Judge Richard Cudahy on the United States Court of Appeals for the Seventh Circuit, she focused on housing law and tax credit matters at the law firm Klein Hornig in Boston before beginning a Ph.D. program. She received her Ph.D. in social policy and sociology from Harvard University in 2014.

Shahin Talesh is professor of law and holds an affiliate appointments in the department of Sociology and department of Criminology, Law & Society at the University of California, Irvine. He also is the Faculty Director for the Civil Justice Research Initiative and directs the Law & Graduate Studies program. He is an interdisciplinary scholar whose work spans law, sociology, and political science. His research interests include the empirical study of law and business organizations, dispute resolution, consumer protection, insurance, and the relationship between law and social inequality. His most recent empirical study addresses the intersection between organizations, risk, and consumer protection laws, focusing on private organizations' responses to and constructions of laws designed to regulate them, consumers' mobilization of their legal rights and the legal cultures of private organizations. His scholarship has appeared in multiple law and peer-reviewed social science journals including *Law and Society Review* and has won multiple awards in Sociology, Political Science and Law & Society.

Spencer Weber Waller is the John Paul Stevens Chair in Competition Law and the Director of the Institute for Consumer Antitrust Studies at the Loyola University of Chicago School of Law.

Susan Block-Lieb is a professor at Fordham Law School, where she holds the Cooper Family Chair in Urban Legal Studies. She specializes in bankruptcy, consumer debt adjustment, corporate reorganization and restructuring, commercial and consumer finance, including consumer financial protection, and commercial laws more generally, including international and transnational law and lawmaking in this context. She has written extensively on unification, harmonization and modernization of international private law by international organizations, on global governance by the G20, IMF, World Bank and UN, on sovereign debt restructuring and more generally on transnational legal orders. Together with Terence C. Halliday, she is author of *Global Lawmakers: International Organizations in the Crafting of World Markets* (Cambridge Univ. Press 2017), an intensive study of global lawmaking by the UN Commission on International Trade Law. She has consulted for the International Monetary Fund, served as a reporter and observer at World Bank task forces on insolvency and secured transactions law reform initiatives, and has been a delegate from the American Bar Association to UNCITRAL's Insolvency Working Group for more than 15 years.

Todd Rakoff is Byrne Professor of Administrative Law at Harvard Law School, and has also served as Dean of the J.D. Program there. Among other publications, he has for over thirty years been one of the editors of Gellhorn and Byse's *Administrative Law*, currently in its twelfth edition.

V.B. Dubal is an associate professor of Law at the University of California, Hastings College of the Law. Her research focuses on the intersection of law and social change in the work context. She joined the Hastings Faculty in 2015, after a post-doctoral fellowship at Stanford University (also her undergraduate alma mater). Prior to that, she received her J.D. and Ph.D. from UC Berkeley, where she used historical and ethnographic methodologies to study workers and worker collectivities in the San Francisco taxi industry. Her work on taxi workers, Uber drivers, and Silicon Valley tech workers has been featured in top-ranked law reviews and featured in the local and national media.

Viva R. Moffat is a professor of law at the University of Denver's Sturm College of Law. She teaches and writes in the areas of contract law and intellectual property law, and she recently stepped down after five years as associate dean of academic affairs. Her scholarly work has been published in the *William & Mary Law Review*, the *Harvard Journal of Law & Technology*, the *UC Davis Law Review*, and the *Berkeley Technology Law Journal*, among others. Prior to joining the University of Denver faculty, she practiced law in Denver and in San Francisco, and she clerked for Judge Robert R. Beezer of the U.S. Court of Appeals for the Ninth Circuit in Seattle.

William S. Dodge is Martin Luther King, Jr. Professor of Law at the University of California, Davis, School of Law, where he teaches contracts and international business transactions. He is the co-author of *Transnational Business Problems* (5th ed. 2014). He earned his B.A. summa cum laude from Yale University and his J.D. from Yale Law School. He served as a law clerk for Judge William A. Norris on the U.S. Court of Appeals for the Ninth Circuit and for Justice Harry A. Blackmun on the U.S. Supreme Court. He served as co-reporter for the American Law

Institute's Restatement (Fourth) of Foreign Relations Law, and he currently serves as an adviser for its Restatement (Third) of Conflicts

Yvette Liebesman is a professor of law at Saint Louis University School of Law, where she is also advisor for the Intellectual Property Law concentration. Her research interests include examining restraints on an individual's Right of Publicity. Her article *When Selling Your Personal Name Mark Extends to Selling Your Soul* was published in the *Temple Law Review*.



United States Department of Agriculture
Economic Research Service

 An official website of the United States government [Here's how you know](#)



United States Department of Agriculture
Economic Research Service



(/) Economic Research Service (/)

MENU

Commuting Zones and Labor Market Areas

Note: Updates to this data product are discontinued.

The ERS Commuting Zones (CZs) and Labor Market Areas (LMAs) were first developed in the 1980s as ways to better delineate local economies. County boundaries are not always adequate confines for a local economy and often reflect political boundaries rather than an area's local economy. CZs and LMAs are geographic units of analysis intended to more closely reflect the local economy where people live and work. Beginning in 1980 and continuing through 2000, hierarchical cluster analysis was used along with the Census Bureau's journey to work data to group counties into these areas. In 2000, there were 709 CZs delineated for the U.S., 741 in 1990, and 768 in 1980. LMAs are similar to CZs except that they had to have a minimum population of 100,000 persons. LMAs were only estimated in 1980 and 1990. This was done in order for the Census Bureau to create microdata samples using decennial census data (1980 PUMS-D, 1990 PUMS-L) that avoided disclosure. The LMAs were discontinued in 2000 because researchers found them to be too large and not as useful as the CZs. The identical methodology was used to develop CZs for all three decades.

Using a consistent methodology, more recent (2010) CZs were developed and are available on the Penn State website, see Labor-sheds for Regional Analysis (<https://sites.psu.edu/psucz/data/>).

Data Set	Last Updated	Next Update
2000 commuting zones (/webdocs/DataFiles/48457/cz00_eqv_v1.xls?v=3378.5)	2/22/2012	
1980 and 1990 commuting zones and labor market areas (/webdocs/DataFiles/48457/czlma903.xls?v=3378.5)	2/22/2012	

Commuting Zones and Labor Market Areas

Overview (</data-products/commuting-zones-and-labor-market-areas/>)

Documentation (</data-products/commuting-zones-and-labor-market-areas/documentation/>)

Last updated: Tuesday, March 26, 2019

For more information, contact: Website Administrator
(<mailto:webadmin@usda.gov>)

[Return to top](#)

ERS Links

[ERS Home \(/\)](#)

[Accessibility \(/help/accessibility/\)](#)

[Careers \(/about-ers/careers-at-ers/\)](#)

[Contact Us \(/contact-us/\)](#)

[E-Mail Updates \(https://visitor.r20.constantcontact.com/disp?llr=fnrkbelab&p=oi&m=1111458399653\)](#)

[Help \(/help/\)](#)

[Information Quality \(/about-ers/policies-and-standards/information-quality/\)](#)

[Site Map \(/sitemap/\)](#)

[Privacy Policy & Non-Discrimination Statement \(/privacy/\)](#)

USDA Links

[FOIA \(/freedom-of-information-act/\)](#)

[Plain Writing \(https://www.usda.gov/plain-writing\)](#)

[Report Fraud \(https://www.usda.gov/oig/forms/contractor-fraud\)](#)

[USA.gov \(http://www.usa.gov\)](#)

[White House.gov \(http://www.WhiteHouse.gov\)](#)

Sign Up for Updates

[Subscribe \(http://visitor.r20.constantcontact.com/disp?llr=fnrkbelab&p=oi&m=1111458399653\)](#)



Economic Research Service (/)

U.S. DEPARTMENT OF AGRICULTURE (HTTPS://WWW.USDA.GOV/)



<http://t.co/11458399653> (<https://www.usda.gov/economic-research-service/>) <https://www.usda.gov/ers/blog> (<https://www.usda.gov/ers/blog>) <https://www.usda.gov/ers/blog> (<https://www.usda.gov/ers/blog>) <https://www.usda.gov/ers/blog> (<https://www.usda.gov/ers/blog>)

Kauffman Foundation State Entrepreneurship Rankings

By The Editors For Ewing Marion Kauffman Foundation

Feature Charticle

Kauffman Early-State Entrepreneurship Index Rankings

Rank	State	<u>Overall Score</u>	Share of adults becoming entrepreneurs in a given month (%)	Share of entrepreneurs driven by opportunity (%)	Job start
1	California	2.86	0.4	87.0	
2	Missouri	1.86	0.3	88.5	
3	Florida	1.75	0.4	85.4	
4	Wyoming	1.68	0.5	85.6	
5	Texas	1.67	0.4	83.8	
6	Maine	1.63	0.3	78.4	
7	Montana	1.54	0.4	86.2	
8	Arizona	1.31	0.4	89.6	
9	Oklahoma	1.17	0.4	87.7	
10	Minnesota	1.14	0.3	88.0	
11	New York	1.12	0.3	87.0	
12	Washington	1.09	0.3	92.0	
13	Idaho	1.08	0.3	83.9	
14	Nevada	1.06	0.4	86.8	
15	North Dakota	1.05	0.4	88.3	
16	Massachusetts	0.89	0.3	88.6	
17	New Jersey	0.75	0.3	85.5	
18	South Dakota	0.66	0.3	92.7	
19	Colorado	0.54	0.3	83.6	
20	Alaska	0.45	0.4	79.3	

21	Georgia	0.43	0.4	87.3
22	Utah	0.39	0.3	90.0
23	Mississippi	0.31	0.4	82.5
24	Tennessee	0.29	0.3	89.4
25	Vermont	0.17	0.4	76.6
26	Nebraska	0.01	0.3	94.0
27	North Carolina	-0.02	0.3	87.2
28	Michigan	-0.03	0.2	87.6
29	Oregon	-0.05	0.3	80.0
30	Louisiana	-0.22	0.3	75.6
31	Iowa	-0.24	0.2	94.0
32	New Mexico	-0.31	0.4	76.0
33	Arkansas	-0.33	0.3	84.2
34	District of Columbia	-0.51	0.2	74.5
35	West Virginia	-0.54	0.3	89.0
36	Hawaii	-0.62	0.3	81.4
37	South Carolina	-0.70	0.3	77.7
38	Maryland	-0.78	0.3	75.9
39	Kentucky	-0.82	0.2	85.9
40	Delaware	-0.86	0.2	93.4
41	Alabama	-0.88	0.2	78.9
42	Connecticut	-0.94	0.2	81.9
43	Kansas	-1.15	0.3	80.2
44	Virginia	-1.23	0.2	80.0
45	Indiana	-1.26	0.2	83.4
46	Wisconsin	-1.29	0.3	73.3
47	Pennsylvania	-1.30	0.2	84.7
48	Illinois	-1.36	0.2	74.1
49	New Hampshire	-1.40	0.2	82.8
50	Ohio	-1.85	0.2	82.6
51	Rhode Island	-2.93	0.2	68.7

Read the full report [here](#).

Findings:

- California is the most entrepreneurial state, with a Kauffman Early-Stage Entrepreneurship (KESE) score of 2.86.
- Missouri, Florida, Wyoming, and Texas round out the top 5, with Rhode Island coming in dead last with a KESE score of -2.93.

- The KESE score is an aggregate of four entrepreneurship indicators:
 - Share of adults becoming entrepreneurs in a given month.
 - Share of entrepreneurs driven by opportunity rather than necessity.
 - Jobs created by startup per 1,000 people.
 - Share of firms surviving one year after founding.

Noncompete agreements

Ubiquitous, harmful to wages and to competition, and part of a growing trend of employers requiring workers to sign away their rights

Report • By Alexander J.S. Colvin and Heidi Shierholz • December 10, 2019

Executive summary

In recent decades, the U.S. labor market has been marked by rising inequality and largely stagnant wages among all but the highest-paid workers. At the same time, job mobility and other measures of labor market fluidity have declined substantially. There are many factors underlying these trends, but growing empirical evidence suggests that one among the vast set of factors is the rise of the use of noncompete agreements.

Noncompete agreements are employment provisions that ban workers at one company from going to work for, or starting, a competing business within a certain period of time after leaving a job. It is not difficult to see that noncompetes may be contributing to weak wage growth, given that changing jobs is how workers often get a raise. And given that noncompetes limit the ability of individuals to start businesses or take other jobs, it also is not difficult to see that noncompetes may be contributing to the declines in dynamism in the U.S. labor market. But how common are they? This report uses data from a national survey of private-sector American business establishments to investigate the extent of noncompete usage. We find:

- ✦ Roughly half, 49.4%, of responding establishments indicated that at least some employees in their establishment were required to enter into a noncompete agreement. Nearly a third, 31.8%, of responding establishments indicated that *all* employees in their establishment were required to enter into a noncompete agreement, regardless of pay or job duties.

SECTIONS

1. Executive summary • 1
2. Introduction • 2
3. Existing research on the extent of noncompete agreements in use • 3
4. Findings of this study • 4
5. Policy solutions and conclusion • 12

About the authors • 13
Methodological appendix • 13
Endnotes • 14

- ❖ The survey data do not allow us to determine the precise share of workers nationwide that are subject to noncompete agreements. However, we can calculate a range, and we find that somewhere between 27.8% and 46.5% of private-sector workers are subject to noncompetes. Applying this share to today's private-sector workforce of 129.3 million means that somewhere between 36 million and 60 million private-sector workers are subject to noncompete agreements.
- ❖ The extent of noncompete use identified in this survey is substantially greater than what a high-quality 2014 survey found: 18.1% of workers covered by noncompete agreements. The difference likely is attributable to the fact that the surveys were three years apart, suggesting that the use of noncompetes is growing. It also likely is attributable to the fact that ours was a survey of business establishments, while the earlier instrument was a survey of workers in the private sector or in a public health care system. While businesses know whether their workers are subject to noncompete agreements, workers may not know or remember they are covered by a noncompete, and thus may underreport being subject to them.
- ❖ While establishments with high pay or high levels of education are generally more likely to use noncompete agreements, noncompetes also are common in workplaces with low pay and where workers have low education credentials.
- ❖ Noncompete agreements are common across the country, including in California, despite noncompetes being unenforceable under California law. Even though these agreements would not stand up if challenged in California courts, businesses still can use them to pressure employees into not going to work for competitors.
- ❖ The use of noncompetes is part of a broader trend of employers requiring their workers to sign a variety of restrictive contracts as a condition of employment. In addition to noncompetes, another common restrictive contract is mandatory arbitration, in which businesses require employees to agree to arbitrate any legal disputes with the business. We find that employers who use mandatory arbitration also are significantly more likely to use noncompetes.

Given the ubiquity of noncompetes, the real harm they inflict on workers and competition, and the fact that they are part of a growing trend of employers requiring their workers to sign away their rights as a condition of employment, noncompetes can and should be prohibited either through legislation or through regulation.

Introduction

In recent decades, the U.S. labor market has been marked by rising inequality and largely stagnant wages among all but the highest-paid workers. At the same time, job mobility and other measures of labor market fluidity have declined substantially.¹ There are many things underlying these trends, but growing empirical evidence suggests that one among the vast set of factors is the rise of the use of noncompete agreements.

Noncompete agreements are employment provisions that ban workers at one company from going to work for, or starting, a competing business within a certain period of time

after leaving a job. It is not difficult to see that noncompetes may be contributing to weak wage growth, given that changing jobs is how workers often get a raise. One study, for example, finds that workers in states that enforce noncompetes earn less than similar workers in states that do not enforce noncompetes.² And given that noncompetes limit the ability of individuals to start businesses or take other jobs, it also is not difficult to see that noncompetes may be contributing to the declines in dynamism in the U.S. labor market. One study found that being bound by a noncompete is associated with an 11% increase in the length of time in a job,³ and another found that greater enforceability of noncompetes reduces the formation of new firms by 12%.⁴

But how common are noncompete agreements?

Existing research on the extent of noncompete agreements in use

A high-quality study on the extent of noncompete agreements nationally involving a survey of 11,500 workers in 2014 found that 18.1% of workers in the private sector or in a public health care system said they were subject to a noncompete agreement.⁵ A key methodological aspect of this survey is that workers themselves were asked whether they were subject to a noncompete agreement. One potential downside to this approach is that it could lead to an underestimate of the share of workers who are subject to noncompetes if workers do not know or remember that they are subject to them. The survey's findings suggest that signing a noncompete may not always be a memorable occasion—for example, it found that when asked to sign a noncompete, 88% of workers simply sign it rather than negotiate over the terms. It also found that more than 30% of workers who are asked to sign noncompetes are asked *after* they already have accepted the job, often on the first day of work, which is a time when new employees often are signing many forms and may not pay a great deal of attention to each form.⁶ Noncompetes also can be tucked inside a larger employee handbook, the provisions of which employees are required to unconditionally agree to as a condition of employment.⁷ In light of these factors, there appears to be meaningful potential for underestimation when asking workers whether they are subject to noncompete agreements.

One way around this problem is an establishment survey—namely, a survey in which business establishments are asked whether their workers are subject to noncompete agreements, rather than asking workers themselves. The establishment surveys that have been conducted to date on this topic, however, have been done on narrow sectors of the labor market and/or have asked firms whether they use noncompetes, but not what share of workers within the firms are subject to them. As a result, these surveys cannot provide additional information on the total share of workers economywide who are subject to noncompete agreements.⁸

Findings of this study

To help shed further light on the extent of noncompete agreements, we used data from a national survey of private-sector American business establishments with 50 or more employees. The survey used a random sample and was conducted from March 2017 to July 2017. It had a sample size of 634, yielding a 95% confidence interval for top-line estimates of plus or minus 3.9 percentage points. The individual respondents were the establishment's human resources manager or whichever individual was responsible for hiring and onboarding employees. The reason for use of this individual as the person to respond to the survey is that noncompete agreements often are signed as part of the onboarding paperwork when a new employee is hired. As a result, the manager responsible for this process is the individual most likely to be knowledgeable about the documents the new employee is signing.

This survey allows us to estimate the share of businesses in which all employees are subject to noncompete agreements, and the share of businesses in which at least some employees are subject to noncompetes. In what follows, we report these estimates for the private sector as a whole and by establishment size, state, industry, average wage level, and typical education level. We then calculate a range for the number of *workers* subject to noncompete agreements.

Roughly half of businesses use noncompete agreements

Roughly half, 49.4%, of responding establishments indicated that at least some employees in their establishments were required to enter into a noncompete agreement. Employers who reported using noncompetes for some but not all employees did not provide information on the proportion of employees who are subject to noncompetes. Some employers in this group did, however, report which employees were subject to noncompetes, with many reporting it was either managers or sales workers. Some employers in this group mentioned other specific occupations—for example, doctors being subject to noncompetes in the case of a medical employer, and on-air talent being subject to noncompetes in the case of a media company. Nearly a third, 31.8%, of responding establishments indicated that *all* employees in their establishment were required to enter into a noncompete agreement, regardless of pay or job duties.

Noncompete agreements by size of employer

Table 1 shows, by the size of the employer, the share of employers that use noncompete agreements (i.e., the share of workplaces where **any** employees are subject to noncompetes) and the share that impose them on all employees. As the third column in the table shows, smaller establishments—those with 50–100 employees—are less likely than larger establishments to use noncompete agreements. Larger organizations with more sophisticated human resources policies and legal counsel may be more likely to

Table 1

Noncompete agreements in U.S. workplaces, by size of employer

Employer workforce size	Sample size	Share of workplaces where all employees are subject to noncompete agreements	Share of workplaces where any employees are subject to noncompete agreements
50–99 employees	254	30.3%	43.7%**
100–499 employees	203	36.4%*	54.2%*
500–999 employees	29	31.0%	48.3%
1,000–4,999 employees	54	22.2%	51.8%
5,000 or more employees	94	30.8%	53.2%

Notes: Percentages indicate the share of workplaces in each row category where either all employees are subject to noncompete agreements or at least some employees are subject to noncompete agreements. The symbols *, **, and *** indicate that the use of noncompete agreements is significantly different from the other categories in the table combined at the 0.10 level, 0.05 level, and 0.01 level, respectively.

Source: Original data from national survey of private-sector workplaces (see the methodological appendix).

Economic Policy Institute

adopt policies that make workers less able to leave to take another job. However, it is worth noting that those findings are shifted when focusing on only those establishments where all workers are subject to noncompetes. Mid-sized establishments (100–499 employees) are more likely than larger and smaller establishments to have all employees signing noncompetes.

Noncompete agreements by state

The incidence of noncompete agreements varies across the country. **Table 2** shows the percentage of establishments that use noncompetes in each of the 12 largest states by population.⁹ A striking result is that noncompetes are widely used nationwide, with more than 40% of establishments in each of the 12 largest states having at least some employees covered by noncompetes. This includes 45.1% of establishments in California, *despite noncompete agreements being unenforceable under California state law*.¹⁰ Even though these agreements would not stand up if ever challenged in court in California, businesses still can use them to pressure employees into not going to work for competitors. Most noncompete agreements never make it to court: workers assume they are valid, or workers can't afford to take on the risk and expense of possible litigation. A typical employee who is reminded that they have signed a noncompete or receiving an intimidating letter from the employer's legal counsel simply may accept that working for a competitor is not an option rather than taking the risk of being sued. This results in a chilling effect, as workers stay in their jobs regardless of the actual enforceability of their

Table 2

Noncompete agreements in U.S. workplaces, by state

State (in order of population size)	Sample size	Share of workplaces where all employees are subject to noncompete agreements	Share of workplaces where any employees are subject to noncompete agreements
<i>California</i>	82	29.3%	45.1%
<i>Texas</i>	28	50.0%**	60.7%
<i>Florida</i>	28	39.3%	46.4%
<i>New York</i>	43	23.3%	44.2%
<i>Illinois</i>	28	14.3%**	50.0%
<i>Pennsylvania</i>	45	31.1%	42.2%
<i>Ohio</i>	27	44.3%	66.7%*
<i>Georgia</i>	35	34.3%	51.4%
<i>North Carolina</i>	31	29.0%	51.6%
<i>Michigan</i>	29	37.9%	55.2%
<i>New Jersey</i>	43	25.6%	48.8%
<i>Virginia</i>	28	46.4%*	64.3%

Notes: Percentages indicate the share of workplaces in each row category where either all employees are subject to noncompete agreements or at least some employees are subject to noncompete agreements. The symbols *, **, and *** indicate that the use of noncompete agreements is significantly different from the other categories in the table combined at the 0.10 level, 0.05 level, and 0.01 level, respectively.

Source: Original data from national survey of private-sector workplaces (see the methodological appendix).

Economic Policy Institute

noncompete agreements.¹¹

Noncompete agreements by industry

Rates of usage of noncompete agreements vary widely across industry. **Table 3** shows use of noncompetes within major industries (based on North American Industry Classification System (NAICS) codes). Noncompetes are used by approximately 70% of establishments in business services and in wholesale trade, but used much less in transportation, in education and health services, and in leisure and hospitality. However, it is striking that even within leisure and hospitality, a *quarter* of establishments use noncompetes, and one in seven responding establishments in leisure and hospitality use noncompetes for *all* their workers.

Noncompete agreements by pay level

To further investigate the interaction between workforce characteristics and the use of

Table 3

Noncompete agreements in U.S. workplaces, by industry

Industry	Sample size	Share of workplaces where all employees are subject to noncompete agreements	Share of workplaces where any employees are subject to noncompete agreements
<i>Construction</i>	65	30.7%	47.7%
<i>Manufacturing</i>	135	34.8%	54.1%
<i>Wholesale trade</i>	34	32.3%	67.6%**
<i>Retail trade</i>	55	25.4%	41.8%
<i>Transportation</i>	38	21.0%	36.8%*
<i>Information</i>	24	25.0%	54.2%
<i>Finance, insurance, and real estate</i>	31	35.5%	58.1%
<i>Business services</i>	75	52.0%***	70.7%***
<i>Education and health</i>	94	28.7%	39.4%**
<i>Leisure and hospitality</i>	28	14.3%**	25.0%***
<i>Other Services</i>	35	31.4%	42.9%

Notes: Percentages indicate the share of workplaces in each row category where either all employees are subject to noncompete agreements or at least some employees are subject to noncompete agreements. The symbols *, **, and *** indicate that the use of noncompete agreements is significantly different from the other categories in the table combined at the 0.10 level, 0.05 level, and 0.01 level, respectively.

Source: Original data from national survey of private-sector workplaces (see the methodological appendix).

Economic Policy Institute

noncompete agreements, we can look at the relationship between noncompetes and pay levels. The survey included a question about the average pay level of workers in the establishment. **Table 4** reports the percentage of workplaces with noncompetes by the average pay level of workers. Average pay levels among the survey respondents are divided into quartiles and annual salaries are converted to equivalent hourly wages for ease of comparison. The use of noncompetes tends to be higher for higher-wage workplaces than lower-wage workplaces. However, it is striking that more than a quarter—29.0%—of responding establishments where the average wage is less than \$13.00 use noncompetes for all their workers.

Noncompete agreements by employee education level

Another workforce characteristic the survey asked about is the education level of the workforce. In **Table 5**, we look at the use of noncompete agreements by the most

Table 4

Noncompete agreements in U.S. workplaces, by average employee pay level

Average hourly wage level	Sample size	Share of workplaces where all employees are subject to noncompete agreements	Share of workplaces where any employees are subject to noncompete agreements
<i>Less than \$13.00</i>	124	29.0%	37.9%***
\$13.00-\$16.99	139	30.9%	56.8%**
\$17.00-\$22.49	131	32.8%	46.6%
\$22.50 and greater	148	36.5%	55.4%*

Notes: Percentages indicate the share of workplaces in each row category where either all employees are subject to noncompete agreements or at least some employees are subject to noncompete agreements. The symbols *, **, and *** indicate that the use of noncompetes is significantly different from the other categories in the table combined at the 0.10 level, 0.05 level, and 0.01 level, respectively.

Source: Original data from national survey of private-sector workplaces (see the methodological appendix).

Economic Policy Institute

common education level of employees in the establishment. The results show much higher use of noncompetes for employees with higher education levels, with significantly higher levels in establishments where workers typically have a four-year college degree or more education; about 45% of responding establishments where the typical education level is a college degree or higher used noncompetes for all their employees. The extent of the use of noncompetes in workplaces with workers that have lower education credentials is striking, however. For example, noncompetes are used for all workers in more than a quarter of workplaces where the typical worker has only a high school diploma.

Noncompete agreements and mandatory arbitration

The use of noncompete agreements is part of a broader trend of employers requiring their workers to sign a variety of restrictive contracts as a condition of employment. In addition to noncompetes, another common restrictive contract is mandatory arbitration, a controversial practice in which businesses require employees to agree to arbitrate any legal disputes with the business. Mandatory arbitration agreements effectively bar employees from going to court, instead forcing workers to resolve workplace disputes in an individual arbitration process that overwhelmingly favors the employer.¹² The survey data used in this study finds that more than half (53.9%) of responding establishments have mandatory arbitration procedures.¹³

In **Table 6**, we look at the use of noncompetes by whether mandatory arbitration is used in the establishment. These results indicate that employers who use mandatory arbitration also are significantly more likely to use noncompetes for some or all of their workers. This

Table 5

Noncompete agreements in U.S. workplaces, by employee education level

Typical employee education level	Sample size	Share of workplaces where all employees are subject to noncompete agreements	Share of workplaces where any employees are subject to noncompete agreements
<i>Some high school</i>	25	20.0%	32.0%*
<i>High school diploma</i>	262	27.1%**	43.9%**
<i>Some college</i>	170	27.6%	48.8%
<i>College degree or more</i>	175	44.8%***	61.6%***

Notes: Percentages indicate the share of workplaces in each row category where either all employees are subject to noncompete agreements or at least some employees are subject to noncompete agreements. The symbols *, **, and *** indicate that the use of noncompetes is significantly different from the other categories in the table combined at the 0.10 level, 0.05 level, and 0.01 level, respectively.

Source: Original data from national survey of private-sector workplaces (see the methodological appendix).

Economic Policy Institute

Table 6

Noncompete agreements and mandatory arbitration in U.S. workplaces

Mandatory arbitration	Sample size	Share of workplaces where all employees are subject to noncompete agreements	Share of workplaces where any employees are subject to noncompete agreements
<i>Uses mandatory arbitration</i>	284	42.6%***	53.7%**
<i>Does not use mandatory arbitration</i>	326	28.9%***	43.3%**

Notes: Percentages indicate the share of workplaces in each row category where either all employees are subject to noncompete agreements or at least some employees are subject to noncompete agreements. The symbols *, **, and *** indicate that the use of noncompetes is significantly different from the other categories in the table combined at the 0.10 level, 0.05 level, and 0.01 level, respectively.

Source: Original data from national survey of private-sector workplaces (see the methodological appendix).

Economic Policy Institute

is perhaps not surprising, given that in practice, noncompete agreements bear a close resemblance to mandatory arbitration agreements in that they deprive workers of future rights contingent on certain events. These results suggest that employers who require their workers to enter into one type of restrictive contract are more likely to require their workers to sign additional restrictive contracts.

Table 7

Workplaces and workers in private sector subject to noncompete agreements

By share	
<i>Workplaces where all employees are subject to noncompete agreements</i>	31.8%
<i>Workplaces where any employees are subject to noncompete agreements</i>	49.4%
<i>Private-sector workers covered by noncompete agreements</i>	27.8%–46.5% (low- to high-end estimate)
By number	
<i>Private-sector workers covered by noncompete agreements</i>	36–60 million (low- to high-end estimate)

Source: Original data from national survey of private-sector workplaces (see report text on estimating lower- and upper-bound estimates and the report's methodological appendix).

Economic Policy Institute

The share of workers subject to noncompete agreements

As mentioned earlier, 49.4% of responding establishments indicated that *at least* some employees in their establishment were required to enter into a noncompete agreement, and 31.8% of responding establishments indicated that *all* employees in their establishment were required to enter into a noncompete agreement. Unfortunately, the 17.6% of employers who reported using noncompetes for *only* some employees did not provide information on the proportion of employees subject to noncompetes. Because of this, we are unable to determine the precise share of workers nationwide that are subject to noncompetes. However, we can provide a range. In the next two sections, we show that somewhere between 27.8% and 46.5% of private-sector workers are subject to noncompetes. Applying these shares to today's private-sector workforce of 129.3 million means that somewhere between 36 million and 60 million private-sector workers are subject to noncompete agreements.¹⁴ These shares and numbers are presented in summary **Table 7**.

The extent of noncompete use that we find in this survey is substantially above what a study of workers in 2014 found: 18.1% of workers. The difference is likely attributable to the fact that the surveys were three years apart, suggesting that the use of noncompetes is growing. It also is likely attributable to the fact our survey was a survey of business establishments, while the earlier survey was a survey of workers in the private sector or in a public health care system. While businesses know whether their workers are subject to noncompetes, workers may not know or remember they are covered by a noncompete, and thus may underreport being subject to them.

Estimating a lower bound on the number of workers subject to noncompete agreements

To calculate the lower bound takes three steps, but the basic idea is to simply ignore firms where not everyone signed a noncompete agreement. First, we include workers in businesses where all employees are subject to noncompetes. Adjusting for establishment size, the 31.7% of businesses where all employees in the establishment are subject to a noncompete agreement translates into 25.0% of the private-sector workforce (recall from Table 1 that establishments with all employees signing noncompetes are more likely to be relatively small). In addition to businesses where all employees have signed noncompetes, there is a second set of businesses where we know the share of workers with noncompetes. If an establishment is unionized, we know from the survey what share of workers is in the union and what share of workers is not in the union; there were some unionized businesses that reported that *all* of their nonunion workers signed noncompetes. Adding in these workers, we arrive at a lower bound of 30.0% of employees being covered by noncompetes.

To arrive at our final lower-bound estimate, we make one additional correction. As mentioned above, the survey was restricted to private-sector businesses of 50 or more employees. Data from the Bureau of Labor Statistics show that in 2017 (the year the survey was collected), 27.9% of private-sector employment was in firms with fewer than 50 employees.¹⁵ In order to adjust our lower-bound estimate to account for firms with fewer than 50 employees, we need a lower bound on the share of workers in small firms who are subject to noncompetes. Recall that Table 1 shows that smaller firms tend to be *more* likely to have all their workers sign noncompetes than larger firms. This means that it is likely safe to assume that the share of firms with fewer than 50 workers that have all their workers sign noncompetes is not substantially smaller than it is for larger firms. However, because we are calculating a lower bound, we prefer to be very conservative. As such, we assign the smallest share of firms that have all their workers sign noncompetes from Table 1, 22.2%, to firms with fewer than 50 employees. Making an adjustment to account for small firms—namely, assuming a lower bound of 22.2% of employees being covered by noncompetes in the 27.9% of firms that have fewer than 50 employees and a lower bound of 30.0% of employees being covered by noncompetes in the remaining 72.1% of firms—yields an overall lower bound of 27.8% of private-sector employees being covered by noncompetes.¹⁶

Estimating an upper bound on the number of workers subject to noncompete agreements

To calculate the upper bound takes two steps. Adjusting for establishment size, the 49.4% of businesses where at least some employees in the establishment are subject to a noncompete agreement translates into 51.7% of the private-sector workforce. We know the actual noncompete share cannot be higher than this level. But to arrive at our final upper-bound estimate, we make an additional correction to account for firms with fewer than 50 employees. Recall that Table 1 shows that smaller firms tend to be less likely to use noncompetes than larger firms. There is a 10.5 percentage-point difference in the share of

firms that use noncompetes in the two lowest-size categories in Table 1. For simplicity, we simply apply that difference to the lowest category to get an estimate, 33.2%, of the share of firms that use noncompetes that have fewer than 50 employees. Making an adjustment to account for small firms—namely, assuming an upper bound of 33.2% of employees being covered by noncompetes in the 27.9% of firms that have fewer than 50 employees and an upper bound of 51.7% of employees being covered by noncompetes in the remaining 72.1% of firms—yields an overall upper bound of 46.5% of private-sector employees being covered by noncompetes.¹⁷

Policy solutions and conclusion

Sens. Todd Young (R-Ind.) and Chris Murphy (D-Conn.) have introduced the Workforce Mobility Act of 2019, which prohibits the use of noncompete agreements in almost all situations, with minimal, sensible exceptions—for example, for owners and senior executives in the sale of a business.¹⁸ The bill explicitly permits employers to protect trade secrets by requiring workers to sign agreements not to disclose such secrets. The bill also provides for civil fines of \$5,000 per week of violation and creates a private right of action with damages and attorneys’ fees available for successful lawsuits. Further, the bill contains outreach and public education provisions, requiring employers to post a notice and requiring the secretary of labor to conduct outreach. If passed, this bipartisan bill effectively would stop the abuse of noncompete agreements nationwide.

However, given that this bill may be unlikely to pass at the federal level in a reasonable time frame, states can act to limit the abuses of noncompete agreements. In recent years, many states have passed laws limiting employers’ ability to impose noncompete agreements on their employees.¹⁹ Noncompetes also could be prohibited by regulation. The Federal Trade Commission is reviewing a petition seeking a rule prohibiting noncompete agreements as an unfair method of competition.²⁰ A group of senators also urged the FTC to conduct rulemaking to bring an end to the abusive use of noncompete clauses in employment contracts²¹, as did 19 state attorneys general.²²

Our survey results show that somewhere between 27.8% and 46.5% of the private-sector workforce—between 36 million and 60 million workers—are subject to noncompete agreements. Similar to surveys using household data,²³ our data show that while establishments with high pay or high levels of education are more likely to use noncompetes, noncompetes also are common in workplaces with low pay and where workers have low education credentials. Given the ubiquity of noncompetes, the real harm they inflict on workers and competition, and the fact they are part of a growing trend of employers requiring their workers to sign a variety of contracts that take away their rights, noncompetes can and should be prohibited.

About the authors

Alexander J.S. Colvin is the Kenneth F. Kahn dean and the Martin F. Scheinman Professor of Conflict Resolution at the ILR School, Cornell University. His research and teaching focus on employment dispute resolution, with a particular emphasis on procedures in nonunion workplaces and the impact of the legal environment on organizations. **Heidi Shierholz** is a senior economist and the director of policy at EPI. She previously served as chief economist at the U.S. Department of Labor.

Methodological appendix

To measure the current extent of noncompete agreement usage we conducted a national-level survey of private-sector employers. The survey was funded by the Economic Policy Institute and administered through telephone- and web-based methods by the Survey Research Institute (SRI) at Cornell University.

The study measured the extent of noncompete usage by surveying employers rather than by surveying employees, to sidestep the possibility that some employees may be unaware or fail to recall that they have signed noncompete agreements and may not understand the content and meaning of these documents. The survey was limited to private-sector employers because public-sector employees typically have their employment regulated by specific public-sector employment laws, and employment practices differ substantially between private- and public-sector employers. The survey focused on nonunion employees. In particular, if workplaces had unionized employees, questions were asked only about nonunion employees. Thus, when tabulating the share of businesses where all employees sign noncompetes, we only counted firms with no union members who said all employees signed noncompetes, since we do not have information on whether the union members signed noncompetes, and anecdotal evidence indicates that it is very rare for unions to agree to include noncompete clauses in the collective bargaining agreements they negotiate.

The survey population was drawn from Dun & Bradstreet's national marketing database of business establishments. It was stratified by state population to be nationally representative. The survey population was restricted to private-sector business establishments of 50 or more employees. The individual respondents were the establishment's human resources manager or whichever individual was responsible for hiring and onboarding employees. The reason for use of this individual as the person to respond to the survey is that noncompete agreements often are signed as part of the onboarding paperwork when a new employee is hired. As a result, the manager responsible for this process is the individual most likely to be knowledgeable about the documents the new employee is signing. Typical job titles of individual respondents included human resource director, human resource manager, personnel director, and personnel manager.

Randomly selected participants were contacted initially by telephone and then given the option of completing phone or web versions of the survey. Follow-up calls were made to encourage participation. Where participants had provided email addresses, a series of emails also was sent to prompt completion of the survey. To encourage participation, respondents were offered the opportunity to win one of 10 \$100 Amazon gift cards in a raffle drawing from among participants in the survey.

Data collection started in March 2017 and was completed in July 2017. A total of 1,530 establishments were surveyed, from which 728 responses were obtained, representing an overall response rate of 47.6%. Some survey responses had missing data on specific questions; however, 634 respondents provided complete data on the key variables of interest. The response rate and sample size are similar to those obtained in past establishment-level surveys of employment relations and human resource practices. The median establishment size in the sample is 90 employees, and the average size is 226 employees. Most establishments are single-site businesses, while 38.2% are part of larger organizations. These larger organizations have an average workforce size of 18,660 employees. Overall, 5.2% of establishments in the sample are foreign-owned.

Endnotes

1. For details on declining labor market fluidity, see Raven Molloy, Christopher L. Smith, Riccardo Trezzi, and Abigail Wozniak, “Understanding Declining Fluidity in the U.S. Labor Market,” Brookings Papers on Economic Activity, Spring 2016.
2. See Evan Starr, “Consider This: Training, Wages and the Enforceability of Covenants Not to Compete,” *ILR Review*, 72, no. 4 (August 2019): 783–817.
3. See Evan Starr, J.J. Prescott, and Norman D. Bishara, *Noncompetes in the U.S. Labor Force*, University of Michigan Law & Econ Research Paper no. 18–013, August 2019.
4. See Jessica Jeffers, *The Impact of Restricting Labor Mobility on Corporate Investment and Entrepreneurship*, Social Science Research Network (SSRN), July 5, 2019.
5. See Evan Starr, J.J. Prescott, and Norman D. Bishara, *Noncompetes in the U.S. Labor Force*, University of Michigan Law & Econ Research Paper no. 18–013, August 2019. A similar share, 15.5%, was found in a smaller survey conducted in 2017; see Alan B. Krueger and Eric Posner, *A Proposal for Protecting Low-Income Workers from Monopsony and Collusion*, The Hamilton Project, February 2018.
6. See Evan Starr, J.J. Prescott, and Norman D. Bishara, *Noncompetes in the U.S. Labor Force*, University of Michigan Law & Econ Research Paper no. 18–013, August 2019.
7. See Kenneth G. Dau-Schmidt and Timothy A. Haley, “Governance of the Workplace: The Contemporary Regime of Individual Contract,” *Articles by Maurer Faculty*, Paper 168 (2007).
8. For a discussion of surveys on noncompetes that use establishment data, see page 520 of Norman D. Bishara and Evan Starr, “The Incomplete Noncompete Picture,” *Lewis & Clark Law Review*, Vol. 20, no. 2 (June 2016): 497–546.
9. We only report the noncompete rate for the 12 largest population states to ensure we have a

sufficient number of observations per state to provide reliable estimates: each of these states had at least 25 observations in the sample. Although the survey is national in coverage, smaller states had fewer observations per state.

10. California adopted its ban in 1872. See R.J. Gilson, “The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not to Compete,” *New York University Law Review* Vol. 74, no. 3 (June 1999), 575.
11. See also Evan Starr, J.J. Prescott, and Norman D. Bishara, *The in Terrorem Effects of (Unenforceable) Contracts*, University of Michigan Law & Econ Research Paper no. 16–032, October 2016.
12. Katherine V.W. Stone and Alexander J.S. Colvin, *The Arbitration Epidemic*, Economic Policy Institute, December 2015.
13. Alexander J.S. Colvin, *The Growing Use of Mandatory Arbitration*, Economic Policy Institute, April 2018.
14. This estimate is based on the fact that Current Employment Statistics data from the U.S. Bureau of Labor Statistics show there were 129.3 million private-sector employees in the United States in October 2019. See Bureau of Labor Statistics, Current Employment Statistics (BLS-CES), “Table B-1. Employees on Nonfarm Payrolls by Industry Sector and Selected Industry Detail,” accessed November 27, 2019.
15. Bureau of Labor Statistics, Business Employment Dynamics (BLS-BDM), “Table F. Distribution of Private Sector Employment by Firm Size Class: 1993/Q1 through 2019/Q1, Not Seasonally Adjusted,” *Historical Series*, accessed Nov. 18, 2019.
16. Note that if we simply had assumed that *no* workers in establishments with fewer than 50 employees signed noncompetes, we would have found an overall lower bound of 21.6%.
17. Note that if we simply had assumed that *no* workers in establishments with fewer than 50 employees signed noncompetes, we would have found an overall upper bound of 37.3%.
18. Workforce Mobility Act of 2019, S.2614, 116th Cong. (2019).
19. For more on state action on noncompetes, see Jane Flanagan and Terri Gerstein, “Welcome Developments on Limiting Noncompete Agreements,” *Working Economics* (Economic Policy Institute blog), Nov. 7, 2019.
20. Open Markets Institute, et al., “Petition for Rulemaking to Prohibit Worker Non-Compete Clauses.” Federal Trade Commission, Washington, D.C., 2019.
21. Richard Blumenthal et al, Letter to Federal Trade Commission Chairman Joseph Simons, March 20, 2019.
22. Keith Ellison et al., Letter to Federal Trade Commission Chairman Joseph Simons, November 15, 2019.
23. Evan Starr, J.J. Prescott, and Norman D. Bishara, “Noncompetes in the U.S. Labor Force,” University of Michigan Law & Econ Research Paper no. 18–013, August 2019.

PROFITING FROM THE POOR

Thousands of Poor Patients Face Lawsuits From Nonprofit Hospitals That Trap Them in Debt

Across the country, low-income patients are overcoming stigmas surrounding poverty to speak out about nonprofit hospitals that sue them. Federal officials are noticing. Help us keep the pressure on.

by Maya Miller and Beena Raghavendran, Sept. 13, 2019, 5 a.m. EDT



Social worker Raquel Nelson was sued for \$2,200 by Methodist Le Bonheur Healthcare, a Memphis, Tennessee-based nonprofit that brought more than 8,300 lawsuits against patients for unpaid medical bills over five years. (Andrea Morales for MLK50)

This article was produced in partnership with MLK50, which is a member of the ProPublica Local Reporting Network.

ProPublica is a nonprofit newsroom that investigates abuses of power. Sign up to receive our biggest stories as soon as they're published.

Over the past few months, several hospitals have announced major changes to their financial assistance policies, including curtailing the number of lawsuits they file against low-income patients unable to pay their medical bills.

Investigative reports have spurred the moves, and they prompted criticism from a top federal official.

“We are learning the lengths to which certain not-for-profit hospitals go to collect the full list price from uninsured patients,” Seema Verma, the administrator of the Centers for Medicare and Medicaid Services, told board members of the American Hospital Association on Tuesday, according to published remarks. “This is unacceptable. Hospitals must be paid for their work, but it’s actions like these that have led to calls for a complete Washington takeover of the entire health care system.”

In June, ProPublica published a story with MLK50 on the Memphis, Tennessee-based nonprofit hospital system Methodist Le Bonheur Healthcare. It brought more than 8,300 lawsuits against patients, including dozens against its own employees, for unpaid medical bills over five years. In thousands of cases, the hospital attempted to garnish defendants’ paychecks to collect the debt.

After our investigation, the hospital temporarily suspended its legal actions and announced a review. That resulted in the hospital raising its workers’ wages, expanding its financial assistance policy and announcing that it would not sue its lowest-income patients. “We were humbled,” the hospital’s CEO, Michael Ugwueke, told reporters.

The same month, NPR reported that Virginia’s nonprofit Mary Washington Hospital was suing more patients for unpaid medical bills than any hospital in the state. Dr. Marty Makary, a surgeon at Johns Hopkins University, and fellow researchers had documented 20,000 lawsuits filed by Virginia hospitals in 2017 alone. The research team found that nonprofit hospitals more frequently garnished wages than their public and for-profit peers.

In mid-August, The Oklahoman reported that dozens of hospitals across the state had filed more than 22,250 suits against former patients since 2016. Saint Francis Health System, a nonprofit that includes eight hospitals, filed the most lawsuits in the three-year span.

In the first week of September, The New York Times reported that Carlsbad Medical Center in New Mexico had sued 3,000 of its patients since 2015. That report was also based on findings from Makary, who just published the book “The Price We Pay: What Broke American Health Care — and How to Fix It.”

And this week, Kaiser Health News and The Washington Post chronicled how Virginia’s state-run University of Virginia Health System sued patients more than 36,000 times over a six-year span.

There is no federal law mandating that nonprofit hospitals provide a specific amount of charity care, nor is there readily accessible data measuring how aggressively each hospital pursues patients for unpaid

bills. But consumer advocates say the revelations in recent coverage on hospitals' litigation practices are troubling.

"It's dismaying to see how common it is," said Jenifer Bosco, an attorney with the National Consumer Law Center who helped craft a Model Medical Debt Protection Act.

Nearly half of the nation's 6,200 hospitals are nonprofits, meaning they are exempt from paying most local, state and federal taxes in return for providing community benefits.

But the issue of nonprofit hospitals engaging in aggressive debt collection practices that push the very communities they are designed to assist into poverty isn't new.

In 2014, ProPublica reported on a small Missouri hospital that filed 11,000 lawsuits over a five-year span. In response, Sen. Chuck Grassley, R-Iowa, opened an investigation, and the hospital forgave the debts owed by thousands of former patients.

In 2003, The Wall Street Journal detailed how Yale-New Haven Hospital in Connecticut had pursued a patient's widow to pay off his late wife's 20-year-old medical bills. The hospital canceled the debt following the article.

"Some of these things are really outrageous," said Jessica Curtis, a policy expert with Community Catalyst who helped draft billing protections for patients in the Affordable Care Act. "There are really aggressive tactics being used and little consideration or understanding for how those tactics actually impact people."

Grassley, chairman of the Senate Finance Committee, sent a letter to the commissioner of the Internal Revenue Service in February to renew his inquiries into whether nonprofit hospitals provide sufficient community benefits to qualify for tax breaks.

Since publishing our story on Methodist hospital in Memphis, we've continued to work with communities in the city to better understand the toll these lawsuits are taking.

We've learned from our reporting that, because of the stigma around owing money, people who've been sued sometimes don't want to discuss it with a reporter. So we've tried to reach people in several ways, including letters sent in the mail, flyers posted in spots they might frequent and graphics we're sharing on Facebook. We're learning a bit more every day about what resonates with the community, and we hope to report back on that soon.

In the meantime — and we tell this to every person we can — these stories are stronger and more accurate when people who've been sued share their

experiences with us. **Hearing from more people who have been sued can help us hold more institutions accountable.**

If you've been sued by a nonprofit hospital or physician group, we want to hear from you. If you work or have worked for an organization that takes unusually aggressive legal action against people unable to pay, we'd also like to hear from you.

This questionnaire is no longer receiving responses.

Filed under: [Health Care](#), [Regulation](#), [Nonprofits](#)



Maya Miller

Maya Miller is an engagement reporter at ProPublica working on community-sourced investigations.

✉ Maya.Miller@propublica.org



Beena Raghavendran

Beena Raghavendran is an engagement reporter focused on local reporting at ProPublica.



WAGE AND HOUR DIVISION
UNITED STATES DEPARTMENT OF LABOR

Fact Sheet #17A: Exemption for Executive, Administrative, Professional, Computer & Outside Sales Employees Under the Fair Labor Standards Act (FLSA)

Revised September 2019

NOTICE: On August 30, 2023, the Department of Labor (Department) announced issuance of a [Notice of Proposed Rulemaking \(NPRM\)](#), *Defining and Delimiting the Exemptions for Executive, Administrative, Professional, Outside Sales, and Computer Employees*. The NPRM proposes to update and revise the regulations issued under section 13(a)(1) of the Fair Labor Standards Act implementing the exemption from minimum wage and overtime pay requirements for executive, administrative, and professional employees. Proposed revisions include increasing the standard salary level and the highly compensated employee total annual compensation threshold, as well as providing an automatic updating mechanism that would allow for the timely and efficient updating of all the thresholds to reflect current earnings data.

This fact sheet provides general information on the exemption from minimum wage and overtime pay provided by Section 13(a)(1) of the FLSA as defined by Regulations, [29 C.F.R. Part 541](#).

The [FLSA](#) requires that most employees in the United States be paid at least the [federal minimum wage](#) for all hours worked and [overtime pay](#) at not less than time and one-half the regular rate of pay for all hours worked over 40 hours in a workweek.

However, Section 13(a)(1) of the FLSA provides an exemption from both [minimum wage](#) and [overtime pay](#) for employees employed as bona fide executive, administrative, professional and outside sales employees. Section 13(a)(1) and Section 13(a)(17) also exempt certain computer employees. To qualify for exemption, employees generally must meet certain tests regarding their job duties and be paid on a salary basis at not less than \$684* per week. Employers may use nondiscretionary bonuses and incentive payments (including commissions) paid on an annual or more frequent basis, to satisfy up to 10 percent of the standard salary level. Job titles do not determine exempt status. In order for an exemption to apply, an employee's specific job duties and salary must meet all the requirements of the Department's regulations.

See other fact sheets in this series for more information on the exemptions for [executive](#), [administrative](#), [professional](#), [computer](#) and [outside sales](#) employees, and for more information on the [salary basis](#) requirement.

Executive Exemption

To qualify for the executive employee exemption, all of the following tests must be met:

- The employee must be compensated on a [salary basis](#) (as defined in the regulations) at a rate not less than \$684* per week;
- The employee's primary duty must be managing the enterprise, or managing a customarily recognized department or subdivision of the enterprise;
- The employee must customarily and regularly direct the work of at least two or more other full-time employees or their equivalent; and

- The employee must have the authority to hire or fire other employees, or the employee's suggestions and recommendations as to the hiring, firing, advancement, promotion or any other change of status of other employees must be given particular weight.

Administrative Exemptions

To qualify for the administrative employee exemption, all of the following tests must be met:

- The employee must be compensated on a salary or fee basis (as defined in the regulations) at a rate not less than \$684^{*} per week;
- The employee's primary duty must be the performance of office or non-manual work directly related to the management or general business operations of the employer or the employer's customers; and
- The employee's primary duty includes the exercise of discretion and independent judgment with respect to matters of significance.

Professional Exemption

To qualify for the **learned professional** employee exemption, all of the following tests must be met:

- The employee must be compensated on a salary or fee basis (as defined in the regulations) at a rate not less than \$684^{*} per week;
- The employee's primary duty must be the performance of work requiring advanced knowledge, defined as work which is predominantly intellectual in character and which includes work requiring the consistent exercise of discretion and judgment;
- The advanced knowledge must be in a field of science or learning; and
- The advanced knowledge must be customarily acquired by a prolonged course of specialized intellectual instruction.

To qualify for the **creative professional** employee exemption, all of the following tests must be met:

- The employee must be compensated on a salary or fee basis (as defined in the regulations) at a rate not less than \$684^{*} per week;
- The employee's primary duty must be the performance of work requiring invention, imagination, originality or talent in a recognized field of artistic or creative endeavor.

Computer Employee Exemption

To qualify for the computer employee exemption, the following tests must be met:

- The employee must be compensated **either** on a salary or fee basis (as defined in the regulations) at a rate not less than \$684^{*} per week **or**, if compensated on an hourly basis, at a rate not less than \$27.63 an hour;
- The employee must be employed as a computer systems analyst, computer programmer, software engineer or other similarly skilled worker in the computer field performing the duties described below;
- The employee's primary duty must consist of:
 1. The application of systems analysis techniques and procedures, including consulting with users, to determine hardware, software or system functional specifications;
 2. The design, development, documentation, analysis, creation, testing or modification of computer systems or programs, including prototypes, based on and related to user or system design specifications;
 3. The design, documentation, testing, creation or modification of computer programs related to machine operating systems; or
 4. A combination of the aforementioned duties, the performance of which requires the same level of skills.

Outside Sales Exemption

To qualify for the outside sales employee exemption, all of the following tests must be met:

- The employee's primary duty must be making sales (as defined in the FLSA), or obtaining orders or contracts for services or for the use of facilities for which a consideration will be paid by the client or customer; and
- The employee must be customarily and regularly engaged away from the employer's place or places of business.

Highly Compensated Employees

Highly compensated employees performing office or non-manual work and paid total annual compensation of \$107,432 or more (which must include at least \$684² per week paid on a salary or fee basis) are exempt from the FLSA if they customarily and regularly perform at least one of the duties of an exempt executive, administrative or professional employee identified in the standard tests for exemption.

Blue-Collar Workers

The exemptions provided by FLSA Section 13(a)(1) apply only to "white-collar" employees who meet the salary and duties tests set forth in the Part 541 regulations. The exemptions do not apply to manual laborers or other "blue-collar" workers who perform work involving repetitive operations with their hands, physical skill and energy. FLSA-covered, non-management employees in production, maintenance, construction and similar occupations such as carpenters, electricians, mechanics, plumbers, iron workers, craftsmen, operating engineers, longshoremen, construction workers and laborers are entitled to minimum wage and overtime premium pay under the FLSA, and are not exempt under the Part 541 regulations no matter how highly paid they might be.

Police, Fire Fighters, Paramedics & Other First Responders

The exemptions also do not apply to police officers, detectives, deputy sheriffs, state troopers, highway patrol officers, investigators, inspectors, correctional officers, parole or probation officers, park rangers, fire fighters, paramedics, emergency medical technicians, ambulance personnel, rescue workers, hazardous materials workers and similar employees, regardless of rank or pay level, who perform work such as preventing, controlling or extinguishing fires of any type; rescuing fire, crime or accident victims; preventing or detecting crimes; conducting investigations or inspections for violations of law; performing surveillance; pursuing, restraining and apprehending suspects; detaining or supervising suspected and convicted criminals, including those on probation or parole; interviewing witnesses; interrogating and fingerprinting suspects; preparing investigative reports; or other similar work.

Other Laws & Collective Bargaining Agreements

The FLSA provides minimum standards that may be exceeded, but cannot be waived or reduced. Employers must comply, for example, with any Federal, State or municipal laws, regulations or ordinances establishing a higher minimum wage or lower maximum workweek than those established under the FLSA. Similarly, employers may, on their own initiative or under a collective bargaining agreement, provide a higher wage, shorter workweek, or higher overtime premium than provided under the FLSA. While collective bargaining agreements cannot waive or reduce FLSA protections, nothing in the FLSA or the Part 541 regulation relieves employers from their contractual obligations under such bargaining agreements.

Where to Obtain Additional Information

For additional information, visit our Wage and Hour Division Website: <http://www.dol.gov/agencies/whd> and/or call our toll-free information and helpline, available 8 a.m. to 5 p.m. in your time zone, 1-866-4USWAGE (1-866-487-9243).

This publication is for general information and is not to be considered in the same light as official statements of position contained in the regulations.



The contents of this document do not have the force and effect of law and are not meant to bind the public in any way. This document is intended only to provide clarity to the public regarding existing requirements under the law or agency policies.

8/20/2016 10:00 AM

**Public Comments of 19 State Attorneys General
in Response to the Federal Trade Commission’s January 9, 2020 Workshop on
Non-Compete Clauses in the Workplace**

We, the undersigned Attorneys General, submit this Comment in response to the Federal Trade Commission’s (FTC) request for public comments in connection with the FTC’s January 9, 2020 workshop on Non-Competes in the Workplace.¹ In this Comment, we offer our perspective on the use of non-compete covenants under antitrust law, and the Attorneys General’s specific interest in, and ability to address, these issues. This Comment addresses some of the discussion during the workshop, particularly as it relates to low-income workers and others negatively impacted by non-compete agreements, and responds to certain of the FTC’s questions on which it invited public comments. Specifically, this Comment provides the Attorneys General’s perspective on:

- 1) the impact of non-compete clauses on labor market participants;
- 2) the traditional business justifications for non-compete clauses;
- 3) the competitive harm from non-competes that support identifying them as unfair methods of competition; and
- 4) recommendations for further FTC action regarding non-competes, including rulemaking, enforcement, economic research and study, and public education.

I. Introduction

We, as State Attorneys General, have a strong interest in the competitiveness of our markets, including labor markets. We care about our residents as workers and consumers, and we want to ensure that companies and organizations compete fairly for the labor of workers through wages and other benefits. We are interested in ensuring that our economies prosper in an environment free of anticompetitive restraints. One mechanism that is increasingly responsible for anticompetitive effects in labor markets and beyond are covenants not to compete (CNCs).

CNCs have a long history in the United States. They have been regulated to varying degrees by the states, who take different approaches to legislating and enforcing CNCs. California, for example, has long banned the enforcement of CNCs, while most other states have not.² This is the type of experimentation and variation that our system of government is designed to promote.³ As State Attorneys General, we support federal rulemaking that is consistent with our ability to pursue enforcement and legislative priorities to the benefit of workers and consumers.

¹ FED. TRADE COMM’N, NON-COMPETES IN THE WORKPLACE: EXAMINING ANTITRUST AND CONSUMER PROTECTION ISSUES, <https://www.ftc.gov/news-events/events-calendar/non-competes-workplace-examining-antitrust-consumer-protection-issues> (last visited March 1, 2020).

² Cal. Bus. & Prof. Code § 16600; *e.g.* *Muggill v. Reuben H. Donnelley Corp.*, 398 P.2d 147, 149 (CA 1965) (“This section invalidates provisions in employment contracts prohibiting an employee from working for a competitor after completion of his employment or imposing a penalty if he does so . . .”).

³ *New State Ice Co. v. Liebmann*, 285 U.S. 262, 311 (1932) (Brandeis, J., dissenting) (“It is one of the happy incidents of the federal system that a single courageous state may, if its citizens choose, serve as a laboratory; and try novel social and economic experiments without risk to the rest of the country.”).

Current conditions in labor markets and a deeper understanding of the competitive harms workers may be experiencing have led to new scrutiny of CNCs.⁴ Some states have recently limited the enforceability of CNCs.⁵ Nineteen State Attorneys General submitted a comment to the FTC in July 2019 advocating greater study of antitrust and labor issues and increased state-federal collaboration in the area. This area has become a priority because of new evidence that shows the impacts of CNCs, particularly how abusive CNCs harm low-wage workers through reduced wages. Harms from CNCs are not confined to low-wage workers, however, and can reach any number of industries, from technology to health care. Also, new research shows that CNCs are prevalent, even in low-wage workers' employment agreements. Moreover, the traditional justifications that underlie the general tolerance of CNCs are facing increased skepticism.

The FTC's workshop was an excellent gathering of experts in this area that highlighted what we now understand. The participants also identified the things we need to further understand and suggested new policy approaches, and how those policy changes might be effectuated. We appreciate the opportunity to respond to the FTC's workshop, and recommend ways forward to enhance protections for our workers and consumers through both enforcement work and a targeted FTC rulemaking that would provide a floor, not a ceiling or constraint, on states' ability to implement state-specific remedies.

This Comment proceeds by first describing the competitive harms caused by CNCs, including harms within labor markets and harms beyond labor markets. It then argues that traditional rationales supporting the enforceability of CNCs are not persuasive. This Comment then concludes by offering recommendations on how the FTC and the States can address the harms CNCs can cause to laborers, employers and consumers.

II. CNCs Can Harm Competition

Ample research shows CNCs present a danger to competition. Harm from CNCs is not limited to the employees bound by them. Negative externalities from CNCs harm other workers in the same labor market, regardless of state boundaries. These harms present competition issues as they not only directly impact workers but also impact competition by hindering competitive entry and dynamism in industries altogether. These harms consequently may harm consumers through higher prices and lower quality because the consumers are losing the benefits from innovation and entry in the industry that would otherwise occur. Below, we describe each harm in turn, starting

⁴ Transcript of Fed. Trade Comm'n Workshop, Non-Competes in the Workplace: Examining Antitrust and Consumer Protection (Jan. 9, 2020), Comm'n Phillips at 220:19-21 [hereinafter FTC Trans.], *available at* https://www.ftc.gov/system/files/documents/public_events/1556256/non-compete-workshop-transcript-full.pdf

⁵ Recently, for example, Hawaii banned the enforcement of CNCs for technology workers, and Colorado now bans them for doctors. *See* Haw. Rev. Stat. Ann. § 480-4(d) (“[I]t shall be prohibited to include a noncompete clause or a nonsolicit clause in any employment contract relating to an employee of a technology business. The clause shall be void and of no force and effect.”); Colo. Rev. Stat. Ann. § 8-2-113 (3)(a) (“Any covenant not to compete provision of an employment, partnership, or corporate agreement between physicians that restricts the right of a physician to practice medicine . . . upon termination of the agreement is void.”). For a discussion on non-competes in medicine, see also Michelle Andrews, *Did Your Doctor Disappear Without a Word? A Noncompete Clause Could Be the Reason*, N.Y. TIMES (March 15, 2019) <https://www.nytimes.com/2019/03/15/business/physician-non-compete-clause.html>; Lavetti, et al., *The Impacts of Restricting Mobility of Skilled Service Workers: Evidence from Physicians*, J. OF HUMAN RESOURCES (2019).

with harm to workers and how harms can cross states lines, then tracing how that harm reaches consumers.

a. Harm within Labor Markets

i. Reduced Wages

CNCs harm a labor market by reducing or even eliminating competition for workers' labor by denying workers the ability to change firms.⁶ This reduction in competition for labor can happen even when firms independently choose to implement CNCs. If enough firms in a market include CNCs in their employment contracts, the resulting harms are similar, if not worse, than if the market suffered from a monopsonist employer: reduced wages, job mobility, and entry. To be sure, a monopsonist using CNCs risks a Section 2 violation. But our concerns with CNCs go beyond that scenario, as the economist Alan Krueger observed, “[n]ew practices have emerged to facilitate employer collusion, such as noncompete clauses and no-raid pacts, but the basic insights are the same: employers often implicitly, and sometimes explicitly, act to prevent the forces of competition from enabling workers to earn what a competitive market would dictate, and from working where they would prefer to work.”⁷

Given the prevalence of CNCs, the above scenario is not hypothetical. Studies show that the use of CNCs is common, and the states see this firsthand. As an example, one state AG office has received nearly 45 individual complaints about 45 different businesses regarding potentially abusive CNCs over the last six months. Nationally, CNCs restrict the mobility of around 25% of the country's workforce.⁸ Of the employees who are subject to CNC, 53% are paid hourly.⁹ With CNCs this prevalent, the benefits supposedly produced by these agreements should by now have been conclusively established across the spectrum of workers. On the contrary, the evidence points decidedly in the other direction for low-wage workers, and casts doubt as to other categories of workers as well.

A lack of employee leverage and employer accountability has made CNCs particularly problematic and contributes to their prevalence. Employers face little downside from including CNCs in their

⁶ See Workshop Comment to FTC from Sen. Rubio 1 (“[T]he proliferation of noncompete agreements has caused great harm to American workers and their families.”); Workshop Comment to FTC from Sen. Murphy, et al. (“At their core, non-competes inherently manipulate competitive labor market forces by narrowing the available employment options for workers.”).

⁷ Alan B. Krueger, *The Rigged Labor Market*, MILKEN INST. REV. (April 28, 2017), <http://www.milkenreview.org/articles/the-rigged-labor-market>.

⁸ Evan Starr, J.J. Prescott & Norman Bishara, *Noncompetes in the US Labor Force*, 17 (“We find that noncompetes are a regular part of the employment relationship: 38.1% of the sample report agreeing to a noncompete at some point in their lives, while 18.%, or roughly 28 million individuals, report currently working under one.”); Alexander Colvin & Heidi Shierholz, *Noncompete agreements: Ubiquitous, Harmful to Wages and to Competition, and Part of a Growing Trend of Employers Requiring Workers to Sign Away Their Rights* 10 (Econ. Policy Inst. Working Paper, 2019) (“[W]e are unable to determine the precise share of workers nationwide that are subject to noncompetes. However, we can provide a range. In the next two sections, we show that somewhere between 27.8% and 46.5% of private-sector workers are subject to noncompetes.”).

⁹ Michael Lipsitz & Evan Starr, *Low Wage Workers and the Enforceability of Noncompete Agreements* 6 (Working Paper, 2019), <https://ssrn.com/abstract=3452240>.

employment agreements.¹⁰ They understand that CNCs are rarely truly negotiated,¹¹ especially in low-wage situations. They understand that there is almost never a consequence to an employer for using abusive CNCs or for defining particular terms of the CNC – the labor market, the time period, or the geographic area – in a way that exceeds any legitimate business purpose of the CNC.¹² This has caused increasing harms to competition for labor as the resulting problem of firms acting in parallel through CNCs make it difficult for a worker to change positions without relocating outside the range of the CNC or working in a different industry.¹³ That such a situation would result in suppressed wages follows from the fact that changing jobs, or the credible threat of changing jobs, is one of the most effective ways to increase compensation.¹⁴

Moreover, an overbroad CNC can cause harm without an employer seeking to enforce it. It can harm an employee without any action from an employer beyond the contract itself. And this is simply because an employee may believe she's bound by an overbroad CNC and never bothers to seek a better position. Or, an employer may only need to remind an employee of her CNC, and that can often be enough to dissuade the employee from taking a new position. Finally, if a potential new employer finds out about a CNC binding a potential hire, it may be leery of getting entangled in a dispute or litigation.¹⁵ California's experience demonstrates this effect: the state has a strict policy against enforceability of CNCs, yet there are still many employment contracts

¹⁰ Eric A. Posner, *The Antitrust Challenge to Covenants Not to Compete in Employment Contracts* (Working Paper, 2019), available at <https://ssrn.com/abstract=3453433>.

¹¹ Evan Starr, J.J. Prescott & Norman Bishara, *Noncompetes in the US Labor Force 2* (U. Mich. Law & Econ. Research Paper No. 18-013, 2019) (“In terms of the contracting process, we observe significant heterogeneity in the circumstances under which employees enter into noncompetes: roughly 1/3 of noncompetes are first requested after the individual has accepted an employment offer (without any changes in job responsibilities), only 10% of individuals report negotiating over noncompetes, and most individuals who are presented with a noncompete simply agree to it without consulting friends, family, or legal counsel.”).

¹² Posner, *supra* note 10; Opinion Letter, *Metis Group, Inc. v. Allison*, CL 2019-10757 (Va. Cir. 2020), available at <https://www.fairfaxcounty.gov/circuit/sites/circuit/files/assets/documents/pdf/opinions/cl-2019-10757-metis-group-inc-v-stephanie-allison-et-al.pdf> (“The restrictive covenants violate public policy because they are designed to perpetuate a monopoly although the work itself was limited to a particular government project. . . . There was no credible evidence as to why The Metis Group [the employer] needed to create an impermeable barrier preventing others from soliciting their employees or other independent contractors to perform any other work regardless of the nature of the work or location.”);

¹³ FTC Trans., Stutz at 61:18-24 (“But in a given relevant antitrust labor market, if you have most of the employers using non-competes to lock up most of the employees in the market, then collectively the non-competes really do register a very serious anticompetitive effect on the hiring process and the labor market.”).

¹⁴ See e.g., Bourree Lam, *The Special Few Who Are Getting Raises in this Economy*, THE ATLANTIC (Feb. 8, 2016), <https://www.theatlantic.com/business/archive/2016/02/job-switchers-raise/460044/> (“According to ADP’s data, full-time workers who changed jobs saw their paychecks increase an average of 4.5 percent, an improvement over the 3.9 percent average that covers all full-time workers Breaking the data down by age, ADP also found that the wage increase from full-time-job switching was most pronounced for workers aged 25 to 34.”).

¹⁵ See Complaint, *State of Washington v. Mercury Madness, Inc.*, ¶ 4.21 (CNC required an employee to produce to a prospective employer a copy of defendant’s employment agreement for 24 months after leaving employment, even though the CNC itself ended at 18 months); Aruna Viswanatha, *Legal Publisher in Settlement to Drop Noncompete Agreements for Employees*, WALL ST. J. (June 15, 2016), <https://www.wsj.com/articles/legal-publisher-in-settlement-to-drop-noncompete-agreements-for-employees-1465963260> (reporting that “Stephanie Russell-Kraft. . . lost a job with a new employer because of the noncompete agreement she signed with Law360” even though the policy of Law360’s founder was not to enforce CNCs.).

in California that include CNCs.¹⁶ This either means that the employer and/or the employee do not understand that California law bars enforcement of CNCs, or there is a widespread belief that CNCs do not need to be enforced to be useful to the employer.

ii. CNC Harm Crosses State Lines

Harms from CNCs do not stop at state lines. Many significant labor markets overlap state borders. A handful of examples are: Cincinnati with Kentucky, Ohio, and Indiana; the District of Columbia with Maryland and Virginia; Chicago with Illinois and Indiana; Kansas City with Missouri and Kansas; Omaha with Nebraska and Iowa; Portland with Washington and Oregon. Studies show that when there are two neighboring jurisdictions and one has a more restrictive non-compete regime than the other, the jurisdiction that relatively welcomes non-competes can cause harm to workers in the jurisdiction that does not.¹⁷ This reality is one of the reasons that the FTC's participation in this area in terms of setting a floor on CNC enforceability is appropriate and particularly important.

Despite harms crossing state lines, when states have taken steps to limit CNCs, workers have benefitted. When Oregon banned CNCs for low wage workers hourly wages rose up to 6% and job mobility increased by 12-18%.¹⁸ These benefits are observable in higher income industries as well, including the technology industry, which is an industry generally characterized by higher levels of innovation, trade secret concerns. In 2015, Hawaii banned CNCs for high-tech workers, and quarterly earnings for new hires increased by 4% while job mobility increased by 11%.¹⁹ The increase in mobility is noteworthy. It represents employees choosing something better for themselves, whether due to better wages or salary or because of any increased benefits in one of the numerous other ways that employers compete for employees.

iii. Policy Considerations

The harms from CNCs and the ways that they arise present an important policy question. If independent, non-collusive action taken by competing employers makes an antitrust case difficult to litigate but the behavior nonetheless lessens competition,²⁰ is there a need for new analyses and new applications of antitrust law that more closely address realities of labor markets? We appreciate the FTC's desire to evaluate this policy issue and use its rulemaking authority to remedy an anticompetitive harm. Within the antitrust context, the situation may be akin to conscious

¹⁶ FTC Trans., Stutz at 92:13-15; *Id.*, Nunn at 131:6-10; *Id.*, Lobel at 24:3-10; Colvin & Shierholz, *supra* note 8 at 5, 6.

¹⁷ FTC Trans., Lavetti at 145:25-146:15 (“[T]here are in fact spillover impacts on workers across the state border. Those workers are not themselves directly affected by the non-compete laws. These are workers that live and work in a different state, but they share an overlapping labor market with the workers who are affected. We estimated that, on average, about 90 percent of the negative wage effect spills over to the bordering counties. As you go further away from the border, the effects dampen. We can reject that the spillover is smaller than 10 percent. So there is very convincing evidence that there are spillover effects, there are negative externalities on other workers, and I think this is pretty convincing evidence that something ought to be done to protect especially vulnerable workers.”).

¹⁸ FTC Trans., Starr at 163:12-20; Lipsitz & Starr, *supra* note 9, at 3-4.

¹⁹ FTC Trans., Starr at 164:13-18; Natarajan Balasubramanian, et al., *Locked In? The Enforceability of Covenants Not to Compete and the Careers of High-Tech Workers* (US Census Bureau Center for Econ. Studies Paper No. CES-WP-17-09, 2017).

²⁰ FTC Trans., Stutz at 62:9-20.

parallelism, where firms taking similar action effect the same results as an anticompetitive agreement but without the legal liability.²¹ However, unlike in most conscious parallelism pricing cases, the harmful activity (CNCs) is clearly identifiable, and the resulting harms can be remedied simply by curtailing CNCs. Competitive concerns are therefore raised by a single firm's contracts with its employees regardless of whether the firm is actively colluding with its competitors. These reasons have motivated several states to challenge CNCs as well as welcome the FTC's involvement in this issue.²²

Several state attorneys general have made combating abusive CNCs a priority over the past few years.²³ While most of these cases have been brought using state consumer protection laws or laws prohibiting unfair competition, the impetus behind these cases is a concern that employers are using CNCs to *reduce* competition for labor.²⁴

²¹ *E.g.* In re Text Messaging Antitrust Litig., 782 F.3d 867, 879 (7th Cir. 2015).

²² Public Comments of 18 State Attorneys General on Labor in Antitrust, Hearings on Competition and Consumer Protection in the 21st Century (July 15, 2019), https://oag.dc.gov/sites/default/files/2019-07/State_AGs_Comments_to_FTC_on_Labor_Issues_in_Antitrust.pdf; Letter from 19 State Attorneys General to Fed. Trade Comm'n Chairman Simons (Nov. 15, 2019), http://www.ag.state.mn.us/Office/Communications/2019/Documents/20191115_MultistateFTCNonCompeteLetter.pdf.

²³ The publicly announced state settlements regarding CNCs:

1. Washington's Settlement with Mercury Madness Coffee. Oct. 29, 2019. <https://www.atg.wa.gov/news/news-releases/attorney-general-bob-ferguson-stops-king-county-coffee-shops-practice-requiring>
2. Illinois Settlement with Check Into Cash. January 7, 2019. http://www.illinoisattorneygeneral.gov/pressroom/2019_01/20190107b.html (the geographic limitation in the CNC was 15 miles from any location of the corporation, including subsidiaries, resulting in a geographic coverage of over 1,000 stores in 32 states.).
3. New York Settlement with Reliance Star Payment Services. Oct. 26, 2018. <https://ag.ny.gov/press-release/2018/ag-underwood-announces-settlement-payment-processing-firm-end-use-non-compete>
4. New York and Illinois settlement with WeWork. September 18, 2018. <https://ag.ny.gov/press-release/2018/ag-underwood-announces-settlement-wework-end-use-overly-broad-non-competes>.
5. New York settles with Examination Management Services, Inc. (EMSI). Aug. 4, 2016. <https://ag.ny.gov/press-release/2016/ag-schneiderman-agreement-ends-non-compete-agreements-employees-national-medical> (“She was offered a job by a clinical laboratory company that offered more regular hours, higher pay, and no travel requirements, but the offer was rescinded when the company discovered she was subject to a non-compete with EMSI.”)
6. New York and Illinois Settlement with Jimmy John's. December 7, 2016. http://www.illinoisattorneygeneral.gov/pressroom/2016_12/20161207.html
7. New York Settlement with Law360. June 15, 2016. See Aruna Viswanatha, *Legal Publisher in Settlement to Drop Noncompete Agreements for Employees*, Wall St. J. (June 15, 2016), <https://www.wsj.com/articles/legal-publisher-in-settlement-to-drop-noncompete-agreements-for-employees-1465963260> (reporting that “Stephanie Russell-Kraft. . . lost a job with a new employer because of the noncompete agreement she signed with Law360” even though the founder of Law360's policy was to not enforce CNCs.).

²⁴ *E.g.*, Assurance of Discontinuance in the matter of the investigation by Barbara D. Underwood, Attorney General of New York of WeWork Companies Inc., Assurance of Discontinuance 4 (“Whereas, the Attorney General has concluded that WeWork's practice of requiring all employees to sign a Non-Compete – regardless of position or job duties, exposure to confidential information, or compensation – unreasonably restrained competition;”), available at https://ag.ny.gov/sites/default/files/final_aod.09.18_w_exhibits.pdf.

iv. Quality Considerations

Firms compete for workers in a variety of ways beyond wages. Employers compete for employees with benefits ranging from improved health care, flexible schedules, or on-site childcare to less substantive benefits like free snacks and employee lounges. Thus, not only do CNCs tend to reduce wages, but they also very likely impact these benefits. Of course, these qualities are difficult to study and measure, but it is likely that reduced wages are a symptom of reduced competition generally, meaning that workplace environments and other benefits – flexible work hours, increased workplace amenities, health care benefits – are impacted by the prevalence of CNCs.²⁵ The intersection of wage suppression and degradation of the quality of other employment benefits may well track the fact that worker populations tend to experience the effects of CNCs differently. Women and minorities see the largest pay increases when CNCs are relaxed,²⁶ which makes it likely that these groups would also see quality improvements in the workplace when CNCs are limited. Reduced pay should be viewed as only one symptom of a reduced competitive environment.

We encourage the FTC to study this issue seriously. Non-wage benefits can deeply affect workers. Work is where Americans spend most of their time. If CNCs are making the work experience worse than it would be without them, that harm is meaningful even if it's not currently measured.

b. Abusive CNCs Harm Consumers

CNCs not only harm laborers – both laborers bound by them and others in the market who are not bound – but consumers are also harmed through reduced entry and innovation.²⁷ Labor is an essential component of every business. Reducing access to skilled and unskilled labor can prevent an employer from expanding or even entering in the first place. Restricting access for one key category of employee can derail entry that would employ many other types of employees, depriving consumers of the benefits of competition.²⁸

Reduced entry has implications for consumers.²⁹ Namely, higher prices and less innovation. Take the beer market for example, if craft breweries used CNCs more frequently, the competition and

²⁵ FTC Trans., Lobel at 87:9-20 (“The other way is that in cases I’ve been involved with, you see this pattern where an employee is very, very unhappy and that – you know, we’ve seen now we’re at a moment where we know more about hostile work environments, we know about problems in various industries, also with, you know, higher paid employees that just feel that they have no voice . . . and they’re locked in because they have a non-compete. . . .”).

²⁶ Testimony of Professor Evan Starr (Nov. 19, 2019) 5-6 (citing Lipsitz & Starr, *supra* note 9). See also Matthew Johnson, Kurt Lavetti & Michael Lipsitz, *The Labor Market Effects of Legal Restrictions on Worker Mobility*, available at <https://ssrn.com/abstract=3455381> (forthcoming 2020); FTC Trans., Comm. Slaughter at 112; *Id.* Lavetti at 143; *Id.* Starr at 163.

²⁷ See Hyo Kang & Lee Fleming, *Non-Competes, Business Dynamism, and Concentration: Evidence from a Florida Case Study* 33-34 (Searle Center Working Paper, No. 2017-046, 2019) (“[W]e observed an increase in the business concentration in Florida, following strengthened non-compete enforcement.”).

²⁸ Posner, *supra* note 10, at 17-18 (providing the example of a hospital looking to enter a market, but is unable to because it can't hire one specific type of worker, resulting in harm to all the other workers that would have been hired).

²⁹ See Naomi Hausman & Kurt Lavetti, *Physician Practice Organization and Negotiated Prices: Evidence from State Law Changes*, AM. ECON. J.: APPLIED ECON. (forthcoming 2020) (finding 45% of physicians are covered by CNCs and “show[ing] that a judicial decision decreasing NCA enforceability by 10% of the observed policy spectrum (about 0.39 standard deviations) causes physician prices to fall on average by 4.3%. This estimate suggests

collaborative atmosphere that generally exists in that industry today would likely not have yielded the many options the consumers have seen in recent years.³⁰ Because of the proliferation of breweries, competition between has increased, and now breweries may consider CNCs with their employees as a response.³¹

Given the harm to workers and consumers described above, it is worth considering the validity of traditional justification for CNCs.

III. Traditional Justifications for CNCs are Not Supported

Three traditional types of justifications exist for CNCs. The first justification is that the employee may be exposed to trade secrets and permitting the employee to leave with such knowledge could be detrimental to the employer. The second is that employers want to ensure a return on the investment they put into training their employees. The third is that the CNC is a term of the contract between employer and employee and thus reflects the employee's freely bargained-for preferences and benefits. None of these justifications are persuasive in today's labor markets in which a quarter of all workers are covered by CNCs and are particularly inapplicable to low-wage or hourly workers. Moreover, given the availability of alternative arrangements, the use of a CNC is excessive in order to address these concerns.

a. Trade Secret and Training Justifications are Not Persuasive

CNCs are a blunt instrument; trade secret and training justifications are overbroad and inapplicable to the majority of workers subject to CNCs. In reality, low-wage workers and hourly worker are rarely, if ever, exposed to bona fide trade secrets or competitively sensitive information. Even if they are, there are less restrictive ways of addressing this concern than with a CNC. If an employer is worried that a former employee may poach his/her clients, the employer could use a non-solicitation agreement.³² With respect to disclosure of trade secrets, the employer could use a non-disclosure agreement or rely on trade secret laws.³³

that such a policy change at the national level would reduce aggregate medical spending by over \$25 billion annually.”).

³⁰ Kieth Gribbins, *Let's Talk Non-compete Agreements in the Brewing Industry and Reflect on the Toppling Goliath Suit*, CRAFT BREWING BUS. (Aug. 6, 2018), <https://www.craftbrewingbusiness.com/featured/lets-talk-non-compete-agreements-in-the-brewing-industry-and-reflect-on-the-toppling-goliath-suit> (reporting that Toppling Goliath Brewing Co. in Iowa was suing its former head brewer for violating a CNC that bars employment at any brewery within 150 miles for two-years).

³¹ Katherine Carlon, *Trouble Brewing in Iowa's Beer Scene*, CORRIDOR BUS. (Aug. 27, 2018), <https://www.corridorbusiness.com/trouble-brewing-in-iowas-beer-scene/> (“But even as bloggers, fellow brewers and beer enthusiasts criticized Toppling Goliath for attempting to restrict its former employee, others admitted that with 6,600-plus active breweries operating in the U.S. – and more than 80 of them in Iowa, according to the Iowa Brewers Guild (IBG) – competitive pressure is threatening to make relations a little less collegial.

‘Toppling Goliath and Thew are both IBG members, so I’ll not be taking sides,’ said J. Wilson, the IBG’s Minister of Iowa Beer, “but as the industry continues to mature, I’d say more squabbles between breweries is a distinct possibility.”).

³² FTC Trans., Flanagan at 46:12-20.

³³ *Id.*, Lobel at 23:13-24:2.

Likewise, training justifications are overbroad as the CNC is rarely tailored to the training received.³⁴ Low-wage workers most often do not have jobs that provide extensive, specialized training which might justify the restrictions in a CNC. Typically, on-the-job training occurs at the beginning of employment and is particular to the needs of the employer. The employer benefits from its trained employees throughout the entire term of employment, and it should not be able to use CNCs to bind an employee for years only because of some initial training.³⁵

Further, even if you assume CNCs encourage employers to train employees, at the same time, CNCs discourage employees to invest in their own training. If an employee understands that she can't take her skills to another employer for a higher salary or more benefits,³⁶ she is less likely to invest in herself.³⁷

Though some employers may argue there are procompetitive benefits of CNC agreements,³⁸ these alleged benefits are not applicable to many categories of employees and legitimate competitive interests can be protected effectively by less restrictive alternatives which do not impact the labor market as drastically. Currently, employers appear to have little to no incentive to think carefully about tailoring a CNC to a particular employee's specific job functions.³⁹ Instead, boilerplate CNCs have become commonplace even when there are no trade secrets or specialized training to justify such restrictions on worker mobility.

b. Freedom of Contract is Not the Reality

In the past, CNCs may have been included as a specifically bargained-for contract term, but now they are included as a matter of course in many employment agreements.⁴⁰ This contemporary practice of including boilerplate CNCs indicates that CNCs are not bargained for by job-seekers, especially low-wage earners, and suggests that such 'agreements' are actually contracts of adhesion.

³⁴ See e.g., Opinion Letter at N. 12, *Metis Group, Inc. v. Allison*, CL 2019-10757 (Va. Cir. 2020) ("Standard form contracts often suffer from inaccuracies and misrepresentations because they seek to apply general terms to all circumstances instead of addressing the specific parties under the contract.").

³⁵ FTC Trans., Nunn at 128:12-22.

³⁶ At a fundamental level, the argument that a training investment in an employee can mean that the employee's labor is required to work off the debt of the investment, means that, post-training, an employer pays below market wages so the employee can "work off the debt." If it were otherwise, and the employer were paying market wages or above-market wages, the CNC would be of no value because defection to another firm wouldn't occur. It's only through training and then subsequent below market wages that the justification makes any sense. As the data suggests, this seems like an area that would be abused. See FTC Trans., Starr at 165:2-9; Posner *supra* note 10, at 11 (questioning why the law allows employers to have a security interest in a person's human capital when it is clear that a bank would never be given such an interest).

³⁷ Yifat Aran, Note, *Beyond Covenants Not to Compete: Equilibrium in High-Tech Startup Labor Markets*, 70 STAN. L. REV. 1235, 1250 (2018) (citing ORLY LOBEL, TALENT WANTS TO BE FREE: WHY WE SHOULD LEARN TO LOVE LEAKS, RAIDS, AND FREE RIDING 178 (2013); Mark J. Garmaise, *Ties That Truly Bind: Noncompetition Agreements, Executive Compensation, and Firm Investment*, 27 J.L. ECON. & ORG. 376, 383 (2009)).

³⁸ These benefits include protecting trade secrets, poaching clients, and/or protecting investments in training workers. See e.g., *Ellis v. James V. Hurson Assoc.'s, Inc.*, 565 A.2d 615, 619 (D.C. 1989); Matt Marx & Ryan Nunn, *The Chilling Effect of Non-Compete Agreements*, ECONOFACT (May 20, 2018), <https://econofact.org/the-chilling-effect-of-non-compete-agreements>.

³⁹ Posner, *supra* note 10, at 8.

⁴⁰ *Id.*; FTC Trans., Posner at 71:14-72:17; FTC Trans., Comm'n Slaughter at 111:15-20.

Proponents of CNCs may claim that workers could negotiate the CNC out of their agreements. This is not the reality. Recent studies have revealed that there is more market power in employer markets than has been understood. Often, an employer may be taking advantage of a certain moment in time when he/she has market power. Employees are most vulnerable when they need a job – near the beginning of the employment relationship– and that is when the CNC is frequently agreed to.⁴¹ This is especially true with low-wage workers who have very little bargaining power in the first instance. For these reasons, low-wage workers are not in a realistic position to bargain the CNC out of their contracts.

Nonetheless, courts analyzing whether CNCs may be enforced by employers tend to view CNCs as validly agreed to contract provisions. Workers generally lack the resources to seek legal advice before signing a CNC, or to litigate a CNC, either defensively or to seek a court order on the CNC’s enforceability. In terms of addressing this argument for CNCs, State Attorneys General and the FTC, therefore, must work together to protect workers from these harms, especially low-wage workers, while acknowledging the reality of the workplace.

IV. Recommendations

For the reasons articulated above, FTC action is appropriate. The FTC has a number of options available to it, ranging from rulemaking to enforcement actions to its unique ability to study and analyze the issues, to address abusive CNCs.

a. The FTC Should Engage in Rulemaking to Address the Abusive Use of CNCs in the Workplace.

Rulemaking by the FTC represents a balanced approach to addressing the harms caused by abusive CNCs that are so prevalent.⁴² Many of these provisions broadly limit employee mobility, stagnate wages, and generally place more burden on labor and competition than is necessary to protect legitimate business interests.⁴³

Rules promulgated by the FTC will provide clarity to firms and workers regarding when and how CNCs are considered an “unfair method of competition,” or unreasonable restraints of trade. While the exact contours of the FTC’s authority to conduct rulemaking in this area is the subject of some

⁴¹ Evan Starr, J.J. Prescott & Norman Bishara, *Noncompetes in the US Labor Force 2* (U. Mich. Law & Econ. Research Paper No. 18-013 2019) (“That is, noncompetes may serve as intertemporal conduits of monopsony power, translating short-term monopsony power (i.e., the temporary lack of an outside offer and high marginal switching costs) into long-term monopsony power (i.e., a durable right to prevent the employee from joining or starting a competitor).”).

⁴² See *supra* Section II.a.i; Colvin & Shierholz, *supra* note 8, at 4 (“Roughly half, 49.4%, of responding establishments indicated that at least some employees in their establishments were required to enter into a noncompete agreement.”).

⁴³ U.S. DEP’T OF TREASURY OFFICE OF ECON. POLICY, *NON-COMPETE CONTRACTS: ECONOMIC EFFECTS AND POLICY IMPLICATIONS 19* (2016) (“[W]e find stricter non-compete enforcement to be associated with both lower wage growth and lower initial wages”); Evan Starr, *Prepared Testimony for the Hearing on “Antitrust and Economic Opportunity: Competition in Labor Markets 4* (Oct. 29, 2019).

debate, there is reason to feel comfortable that the authority exists.⁴⁴ Setting aside the exact contours of the authority, the general benefits of a rule to the business community and the public at large are easy to identify. Likewise, state regulators benefit from clarity in federal standards that may be adopted by state law or provide persuasive guidance in state actions. State economies benefit from certainty in understanding standards of CNC enforceability within and outside of our state borders. Like rulemaking in many areas of administrative law, rules governing CNCs in the workplace would provide valuable guidance to all stakeholders—employers and employees and regulators alike. Rules would clearly frame how the FTC would view CNCs’ validity and would set forth its role to challenge the use of CNCs in certain employment contracts.

An FTC rule could provide that CNCs are presumptively unreasonable for certain workers in certain situations. This could include, for example, hourly employees who earn below a certain threshold and who typically would not have access to trade secrets or competitively sensitive information, or physicians because of the potential access to care issue. The economic evidence discussed above and at the FTC’s workshop shows that CNCs harm low-wage workers. State legislation and the Illinois Attorney General’s case against the Jimmy John’s sandwich chain provide good examples of how such a rule could apply.⁴⁵ The Illinois Attorney General alleged that workers who did not have access to trade secrets or other confidential proprietary information were unlawfully subjected to a two year CNC prohibiting them from working in any restaurant located within two miles of *any* Jimmy John’s sandwich shop.⁴⁶ Given the number of Jimmy John’s sandwich shops across the country (over 2,800 locations in 43 states across the U.S.),⁴⁷ the CNC prevented a large number of former employees from working in a large swath of the country for a significant period, without any legitimate business justification. Jimmy John’s settled with

⁴⁴ Federal Trade Commission Act (FTC ACT) §6(g)(2006) (“The Commission shall also have power—[f]rom time to time...(except as provided in [unfair or deceptive acts or practices (UDAP) rulemaking proceedings]) to make rules and regulations for the purpose of carrying out the provisions of this subchapter.

As competition enforcers, the State Attorneys General are focusing on the FTC’s ability to address abusive CNCs through its competition authority. We understand that the FTC has rulemaking options relating to a UDAP standard; this Comment does not address those options, or the issues raised at the workshop on such options. See also, Comment of Fed. Trade Comm’n Rohit Chopra 6 Hearing #1 on Competition and Consumer Protection in the 21st Century (September 6, 2018) (“The Commission has in its arsenal a tool that would provide greater notice to the marketplace and that is developed through a...transparent, participatory process: using rulemaking to define “unfair methods of competition” through processes established by the Administrative Procedures Act (APA).”).

⁴⁵ Illinois bans CNCs for workers earning less than \$13 per hour. 820 Ill. Comp. Stat 90/5(c). Maine bans CNCs for persons earning at or below 400% of the federal poverty level. Me. Stat. tit. 26, § 599-A. Maryland bans CNCs for workers earning less than \$15 per hour. MD. CODE ANN., LAB. & EMPL. §3-716(A)(1). New Hampshire bans CNCs for employees earning less than 200% of the federal hourly minimum wage. N.H. REV. STAT. ANN. § 275:70-a(II)(b). Oregon CNCs are voidable and not enforced in the state unless employee’s salary and commission exceed the median family income of a family of four. OR. REV. STAT §653.295(1)(d). Rhode Island CNCs are not enforceable against an employee whose average annual earnings is less than 250% of the federal poverty level. 28 R.I. GEN. LAWS §28-59-3(a)(4). A CNC is void and unenforceable in Washington state for an employee earning less than \$100,000 annually. WASH. REV. CODE ANN § 49.62.020(1)(B). In 2018, Massachusetts enacted a law that bans non-compete agreements against certain low wage and student workers. MASS. GEN. LAWS ch. 149, §24L(c).

⁴⁶ Illinois v. Jimmy John’s Enterprises., No. 2016-CH-07746 (Cook County Cir. Ct. filed June 8, 2016).

⁴⁷ JIMMY JOHN’S, <https://www.jimmyjohns.com/find-a-jjs/> (last visited March 1, 2020).

the Illinois Attorney General and agreed to rescind existing CNCs and to remove all CNCs from its “new hire” packets going forward.⁴⁸

As previously mentioned, the protection of trade secrets is often a proffered justification for the use of CNCs. If protection of trade secrets were really the issue, one would expect CNC to be concentrated among workers with advanced education and in occupations likely to deal with trade secrets.⁴⁹ However, there appears to be little difference in the incidence of CNCs used for more educated workers versus for workers generally.⁵⁰ Indeed, in more than a quarter of workplaces where the typical worker has a high school diploma (and presumably not likely to have access to trade secrets or competitively sensitive information) *all* workers are subject to CNCs.⁵¹

An FTC rule could also require that employers provide separately negotiated consideration for the CNC, either on the front end or on the back end, in order for the CNC to be reasonable. On the front end, consideration could be in the form of a wage or benefits premium provided in order for the worker to take the job subject to the CNC. On the back end, the consideration could be in the form of “garden leave” which, originally espoused under British law, translates to a period when an employer pays an employee’s salary or a part of it when an employee is required to remain out of the labor market due to a CNC.⁵² If, as supporters of CNCs argue, an employee’s agreement not to compete as a condition of employment is a valid contractual term, an argument can be made that the employee’s future agreement not to compete *after employment ends* is worthy of its own consideration. The published research, however, does not document that employees receive *anything* in return for their agreements not to compete.⁵³

Finally, the FTC CNC rule could state that unless the employer provides the employee with the CNC some specified amount of time *before* the employer makes the job offer, the CNC will be presumptively unreasonable.⁵⁴ Empirical data shows that workers presented with CNCs *after accepting a job* experience no wage or training benefits compared to workers presented with CNCs *before accepting a job*.⁵⁵ Significantly, compared to workers with post-employment CNCs, workers presented with CNCs before accepting a job offer have nearly 10% higher wages (in their first few years of work), receive 11% more training and are more than 65% more satisfied in their jobs.⁵⁶

⁴⁸ Press Release, Illinois Attorney General, Madigan Announces Settlement with Jimmy John’s for Imposing Unlawful Non-Compete Agreements (Dec. 7, 2016), http://www.illinoisattorneygeneral.gov/pressroom/2016_12/20161207.html.

⁴⁹ U.S. DEP’T OF TREASURY OFFICE OF ECON. POLICY, NON-COMPETE CONTRACTS: ECONOMIC EFFECTS AND POLICY IMPLICATIONS 11 (2016).

⁵⁰ *Id.*

⁵¹ Colvin & Shierholz, *supra* note 8, at 8.

⁵² Jane Flanagan, *No Exit: Understanding Employee Non-Competes and Identifying Best Practices to Limit their Overuse* (American Constitution Society Issue Brief, Nov. 2019)

⁵³ Evan Starr, J.J. Prescott & Norman Bishara, *Noncompetes in the US Labor Force 2* (U. Mich. Law & Econ. Research Paper No. 18-013 2019)

⁵⁴ This rule could be similar to the rule requiring the pre availability of written warranties under the FTC’s existing rules. 16 CFR Part 702 – Pre-sale Availability of Written Warranty Terms.

⁵⁵ See Evan Starr, *Are Noncompetes Holding Down Wages* (Working Paper, 2019).

⁵⁶ *Id.* at 4.

b. Any FTC Rule Would Not Preempt States From Providing Additional Protections Under State Law for Workers Subject to CNCs.

The FTC has expressed interest in the States' position on the preemptive effect of FTC action. There is an "assumption that the historic police powers of the State [are] not to be superseded by [a] Federal Act unless that was the clear and manifest purpose of Congress."⁵⁷ It is the States' position that any preemption analysis of an FTC rule would implicate conflict preemption and that an FTC rule on CNCs would not in any manner preempt states from more broadly regulating CNCs.

The U.S. Supreme Court has recognized three circumstances in which federal regulations preempt state law: First, where a federal agency indicates express intent to preclude state regulation.⁵⁸ Second, where federal regulation is so comprehensive that it is reasonable to infer that the federal agency intended to occupy the field of regulation.⁵⁹ Third, state law is preempted where a state law actually conflicts with federal law so that compliance with both federal and state regulations is impossible.⁶⁰

The third area of conflict preemption is implicated in preemption analysis of FTC rulemaking because contract law, antitrust, and consumer protection are fields that the FTC traditionally shares with the states.⁶¹ Because of traditional state regulation in these areas, there is an assumption that federal action does not supersede the States unless there is a clear and manifest purpose from Congress.⁶² For example, the Southern District of California ruled that a California law requiring that all components of a product be manufactured in the United States in order for the product to be legally affixed with a "Made in the U.S.A." designation was not preempted because companies that must comply with the California law would also comply with the FTC regulation requiring that a product be made "all or virtually all" within the United States.⁶³

Were the FTC to take action to regulate abusive CNCs, it is the States' position that regulation would set a floor and would not preempt state laws that set restrictions on CNCs that are more protective to workers, or regulate abusive CNCs more rigorously, than the FTC rule.⁶⁴ The FTC

⁵⁷ *Wyeth v. Levine*, 555 U.S. 555, 565 (2009) (quotation omitted); *see also Medtronic, Inc. v. Lohr*, 518 U.S. 470, 485 (1996) ("[B]ecause the States are independent sovereigns in our federal system, we have long presumed that Congress does not cavalierly preempt state-law causes of action.").

⁵⁸ *See, e.g., Sprietsma v. Mercury Marine*, 537 U.S. 51, 58-59 (2002).

⁵⁹ *See, e.g., Rice v. Santa Fe Elevator Corp.*, 331 U.S. 218, 230 (1947).

⁶⁰ *See, e.g., Silkwood v. Kerr-McGee Corp.*, 464 U.S. 238, 248 (1984). Because the Magnuson-Moss Act does not occupy the field, regulations adopted pursuant to the Magnuson-Moss Act only preempt state law to the extent that state law conflicts with the Act. *See Katharine Gibbs Sch. (Inc.) v. FTC*, 612 F.2d 658, 667-68 (2d Cir. 1979).

⁶¹ *See, e.g., Gen. Motors Corp. v. Abrams*, 897 F.2d 34, 40-44 (2d Cir. 1990) (state Lemon Law did not conflict with FTC enforcement); *Automobile Importers of America, Inc. v. Minnesota*, 871 F.2d 717 (8th Cir. 1989) (same); *Nat. Funeral Servs., Inc. v. Rockefeller*, 870 F.2d 136 (4th Cir. 1989) (FTC funeral services regulations did not preempt state law); *Am. Fin. Servs. Ass'n v. FTC*, 767 F.2d 957 (D.C. Cir. 1985) (state consumer protection statutes not preempted by FTC credit practices rule).

⁶² *Abrams*, 897 F.2d at 41-42.

⁶³ *Clark v. Citizens of Humanity, LLC*, 97 F. Supp.3d 1199, 1205 (S.D. Cal. 2015).

⁶⁴ This would be in line with Congressional treatment of employment issues in other circumstances. The FLSA, for example, has an explicit savings clause that makes it clear that States can go above and beyond the protections set by federal law. 29 U.S.C. § 218(a). This type of clause is similar to exemption provisions the FTC has used in the past. *See, e.g., Am. Fin. Servs. Ass'n*, 767 F.2d at 990 (D.C. Cir. 1985).

has the authority to restrict anticompetitive and unfair or deceptive practices.⁶⁵ The FTC regulation may not, however, affirmatively preempt states from their own regulations beyond the FTC rule in this area. We request that any FTC regulation include a savings clause making clear that state legislation offering further protections to workers is not preempted.

This principle aligns with the U.S. Supreme Court’s application of FTC regulation to state action and to subsequent review of FTC action by the D.C. Circuit, including in circumstances where the FTC uses its authority against violations the Sherman Act.⁶⁶ Thus, FTC rulemaking defining CNCs for certain categories of workers as “unfair” (or “unreasonable”) would not preempt a state law forbidding CNCs for additional categories of workers. Instead, any FTC rule would necessarily set a floor above which the States can provide additional protections.

c. The FTC and State Attorneys General Should Work Together to Address Anticompetitive and Abusive CNCs.

CNCs are creatures of state law. State attorneys general have been directly involved in addressing, by litigation or otherwise, abusive and anticompetitive CNCs in employment contracts.⁶⁷ We welcome collaboration with the FTC. As former FTC Chairman Kovacic stated during the January 9, 2020 workshop, “[t]here is a lot of room for state and federal cooperation on this” and “[t]hat cooperation shouldn’t be intermittent; [it] should be a regular element of ongoing work....”⁶⁸ The National Association of Attorneys General (NAAG) Antitrust Committee is comprised of eleven (11) state attorneys general committed to protecting competition for the benefit of consumers and workers.⁶⁹ This would be a natural alliance for the FTC to pursue, and the undersigned States are eager to work with the FTC on this matter.

d. The FTC Should Bring Enforcement Actions to Enjoin the Use of Abusive CNCs.

Litigation is also a key part of the multifaceted approach to addressing the use of the abusive CNCs—along with rulemaking and the studies, reports, and guidance documents discussed below. Use of all will increase the likelihood of successfully combatting these problematic agreements. Some scenarios are so egregious that adjudication on the facts will offer clear lessons on what is unreasonable or unfair and will provide enlightenment to all stakeholders. Clear parameters

⁶⁵ See 15 U.S.C. § 57a(a)(1)(B). We understand that the FTC could also engage in rulemaking under its competition authority. See *supra* n. 45. This Comment does not take a position on whether the FTC should exercise rulemaking under its competition authority or its authority to address unfair or deceptive practices, as we understand that neither option would preempt more protective state regulation in the field.

⁶⁶ See, e.g., *Parker v. Brown*, 317 U.S. 341, 351 (“There is no suggestion of a purpose to restrain state action in the [Sherman] Act’s legislative history.”); *Am. Fin. Servs. Ass’n*, 767 F.2d at 990 (D.C. Cir. 1985) (upholding an FTC regulation does not impermissibly preempt state law because it explicitly allowed for states to “offer protections equal to or greater than” the FTC rule); *Am. Optometric Ass’n v. FTC*, 626 F.2d 896, 910-11 (D.C. Cir. 1980) (remanding a challenge to an FTC rule to district court and questioning whether Congress authorized the Commission to affirmatively preempt state laws).

⁶⁷ See *supra* note 23 (listing state attorneys general public settlements and litigation concerning CNCs).

⁶⁸ FTC Trans., Kovacic at 40:14-22.

⁶⁹ NAT’L ASSOC. OF ATTORNEYS GENERAL, ANTITRUST COMM.,

https://www.naag.org/naag/committees/naag_standing_committees/antitrust-committee.php (last visited March 1, 2020).

regarding the types of CNCs that the FTC will consider unreasonable or unfair will provide useful guidance regarding litigation risks.⁷⁰

e. The FTC Should Study the Anticompetitive Impact of CNCs in Employment Contracts and Issue Guidance.

In carrying out its directive under Section 5 of the Federal Trade Commission Act to prohibit unfair methods of competition, the FTC is empowered to gather and compile industry information.⁷¹ Under Section 6(b) of the FTC Act, the Commission may gather information from market participants regarding certain aspects of their business.⁷² Using this power, the FTC may obtain key information from industry participants about the frequency, use, enforcement and general details relating to CNCs in employment contracts.⁷³ This information could fill in important gaps in current empirical research because the FTC would be able to analyze non-public data to which other researchers do not have access.⁷⁴ This information could also presumably inform the FTC's CNC policy and enforcement decisions going forward.

Additionally, this information could form the basis of a public FTC study report as provided for in Section 6(f) of the Act, further educating industry, policymakers, and the public about the impact of these CNCs on workers and on industries across various sectors. A recent example of this is the FTC's 2016 study and report on patent assertion entity (PAE) activity.⁷⁵ Following a joint workshop with the United States Department of Justice to explore the claimed harms and efficiencies of PAE activity, the Commission initiated a study to investigate the use of the PAE business model and based on findings from that study, made certain recommendations, which likely influenced market-wide conduct. Sometimes, the result of these studies and reports is self-imposed market change without any direct action on the part of the Commission.⁷⁶

The FTC also has the ability to offer impactful guidance, similar to the widely regarded 2010 Horizontal Merger Guidelines or the 2016 Antitrust Guidance for Human Resources Professionals

⁷⁰ FTC Trans., Pierce at 308.

⁷¹ 15 USC §46(b); Jaymar-Ruby, Inc. v. FTC, 496 F.Supp. 838, 846-847 (N.D. Ind. 1980).

⁷² Lesley Fair, *6(b) or not 6(b): That is the Question*, FED. TRADE COMM'N: BUS. BLOG (April 23, 2012), <https://www.ftc.gov/news-events/blogs/business-blog/2012/04/6b-or-not-6b-question>; 15 U.S.C. § 46. Additional powers of Commission; The Commission shall also have power—

(b) To require...persons, partnerships, and corporations, engaged in or whose business affects commerce...to file with the Commission...reports or answers in writing to specific questions...as to the organization, business, conduct, practices, management, and relation to other corporations, partnerships, and individuals of the respective persons, partnerships, and corporations filing such reports or answers in writing....

(f) To make public from time to time such portions of the information obtained by it hereunder as are in the public interest...and to provide for the publication of its reports and decisions in such form and manner as may be best adapted for public information and use....

⁷³ Comment of Fed. Trade Comm'n Rohit Chopra 5, Hearing #1 on Competition and Consumer Protection in the 21st Century (September 6, 2018).

⁷⁴ FTC Trans. at 208, (addressing key information gathering ideas by economists on panel).

⁷⁵ FED. TRADE COMM'N, PATENT ASSERTION ENTITY ACTIVITY (2016).

⁷⁶ See generally FED. TRADE COMM'N, SELF-REGULATION IN THE ALCOHOL INDUSTRY, REPORT OF THE FED. TRADE COMM'N (2008) (making recommendations for the industry to adopt).

issued in conjunction with the DOJ.⁷⁷ The Guidelines and the Guidance provide information about how the federal antitrust agencies will view certain activities, enabling actors to align their conduct to avoid litigation. CNC guidance issued by the FTC could play a similar role—the FTC could communicate its position on CNCs, giving the business community ample opportunity to align its conduct with that position.

f. The FTC Should Engage in a Comprehensive CNC Education Campaign.

The FTC has long been a leader in market and competition education. It has a platform that is easily accessible to the public and to industry—www.FTC.gov. The website provides an opportunity for the FTC to inform employers and workers about abusive CNCs. Through its website, the agency can efficiently educate the public about abusive employment CNCs. Experience has shown that the mere shining of light on the nature and existence of abusive CNCs causes companies to delete them from their employment contracts; this highlights the lack of true business justification for them in the first place.


* * *

We thank the FTC for providing the opportunity to submit this Comment and contribute to the Commission’s review of the use of CNCs in light of current and evolving workplace realities. We look forward to continuing to collaborate with the FTC on antitrust and labor issues, and to address the abusive use of CNCs to stop their harm to our workers, labor markets, consumers and economies.

Respectfully Submitted,



Karl A. Racine
Attorney General for the District of Columbia



Brian Frosh
Maryland Attorney General



Keith Ellison
Minnesota Attorney General



Xavier Becerra
California Attorney General

⁷⁷ FED. TRADE COMM’N & DEP’T OF JUSTICE, HORIZONTAL MERGER GUIDELINES (2010); FED. TRADE COMM’N & DEP’T OF JUSTICE, ANTITRUST GUIDANCE FOR HUMAN RESOURCE PROFESSIONALS (2016).

Kathleen Jennings
Delaware Attorney General

Clare E. Connors
Hawaii Attorney General

Kwame Raoul
Illinois Attorney General

Tom Miller
Iowa Attorney General

Aaron M. Frey
Maine Attorney General

Maura Healey
Massachusetts Attorney General

Aaron D. Ford
Nevada Attorney General

Gurbir S. Grewal
Attorney General of New Jersey

Hector Balderas
New Mexico Attorney General

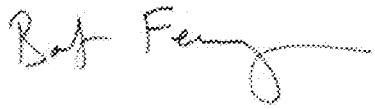
Letitia James
New York Attorney General

Josh Shapiro
Pennsylvania Attorney General

Puerto Rico Secretary of Justice

Peter F. Neronha
Rhode Island Attorney General

Mark R. Herring
Virginia Attorney General

A handwritten signature in black ink, appearing to read "Bob Ferguson". The signature is written in a cursive style with a long horizontal flourish extending to the right.

Bob Ferguson
Washington Attorney General

McKinsey
& Company

The next normal in construction

How disruption is reshaping the world's largest ecosystem

June 2020

Cover image:
© Dong Wenjie/Getty Images

Copyright © 2020 McKinsey &
Company. All rights reserved.

This publication is not intended to
be used as the basis for trading in
the shares of any company or for
undertaking any other complex or
significant financial transaction
without consulting appropriate
professional advisers.

No part of this publication may be
copied or redistributed in any form
without the prior written consent of
McKinsey & Company.

The next normal in construction

How disruption is reshaping the world's
largest ecosystem

June 2020

This article was a collaborative, global effort among
Maria João Ribeirinho, Jan Mischke, Gernot Strube,
Erik Sjödin, Jose Luis Blanco, Rob Palter, Jonas
Biörck, David Rockhill, and Timmy Andersson.

Preface

June 2020

Countervailing factors are reshaping the global economy, and no industry is immune to their impact. Grounded in the built, physical world, construction may seem less vulnerable to the impact of digital technologies and Silicon Valley disrupters. Indeed, the cranes accenting fast-rising urban centers and the workers on commercial and residential projects might lead some executives to believe that as it has been, so it shall be.

In truth, construction is just as susceptible to these disruptions as other industries, but the ways in which the landscape will be affected are different. In 2017, the McKinsey Global Institute (MGI) highlighted that the construction industry needs to evolve and showed ways in which it can change to improve productivity by 50 to 60 percent and deliver \$1.6 trillion a year in incremental global value.¹ The call to action was heard: executives we speak to are thinking through how to prepare for changes ahead—and they increasingly recognize that it's no longer a matter of if or when construction will be affected. Change is already here.

The COVID-19 crisis unfolding at the time of publishing this report will accelerate disruption and the shift to a “next normal” in the construction ecosystem. Many executives are wrestling with the pandemic's economic turmoil, the shifts in demand it entails, and operating restrictions and longer-term safe working procedures. However, it is also critical for executives to lift their view to what the future will hold in terms of changes to business models and industry dynamics. It is in times of crisis that winners segregate from losers, and those who take bold moves fast can reap the rewards.

This research analyzes how the entire ecosystem of construction will change, how much value is at risk for incumbents, and how companies can move fast to adapt to and, in fact, create a new industry structure. We relied on top-down reviews of industry dynamics, bottom-up analysis of company data, and executive surveys to offer an unprecedented look at the entire value chain. In developing the report, we have sought to address the most pressing longer-term strategic

questions for executives in the ecosystem: how their part of the value chain will be affected, by how much, and what they should consider doing to prepare for a future that will differ radically from the present.

Our hope is that these insights will help accelerate a transformation that we believe will and must happen and provide executives around the world with a map to help navigate the rough water ahead.

This research was led by Jan Mischke, partner at the McKinsey Global Institute (MGI) in McKinsey's Zurich office; Jonas Biörck, associate partner based in Stockholm; Gernot Strube, senior partner in Munich and leader of the Capital Projects and Infrastructure Practice; Maria João Ribeirinho, partner in Lisbon; Erik Sjödin, partner in Stockholm; Jose Luis Blanco, partner in Philadelphia; Rob Palter, senior partner in Toronto; and David Rockhill, associate partner in London. We are grateful for the input, guidance, and support of Oskar Lingqvist, senior partner in Stockholm and Steffen Fuchs, senior partner in Dallas and coleader of our Capital Projects and Infrastructure Practice. The project team was led by Timmy Andersson and comprised Nadja Bogdanova, Isak Söderberg, and Richard Karlsson. Many McKinsey partners and colleagues offered helpful expert input, including Alex Abdelnour, Piotr Pikul, Nick Bertram, Subbu Narayanswamy, Marcel Brinkman, Matthew Hill, Gerard Kuperfarb, Priyanka Kamra, Niklas Berglind, Patrick Schulze, Nicklas Garemo, Koen Vermeltfoort, Fredrik Hansson, Ymed Rahmania, Frank Wiesner, Francesco Cuomo, Eric Bartels, and Kathleen Martens. Further, we wish to thank Gunnar Malm and Mats Williamson for their contributions to this report.

This report was edited by Scott Leff and David Peak and designed by Leff, Daphne Luchtenberg, Suzanne Counsell, and Lukasz Kowalik helped disseminate the report.

Maria João Ribeirinho

Jan Mischke

Gernot Strube

Erik Sjödin

Jose Luis Blanco

Rob Palter

Jonas Biörck

David Rockhill

Timmy Andersson

Contents

- 2 **In brief**
- 4 **Executive summary**
- 83 **Related reading**
- 84 **Endnotes**

1

- 16 Historically, the construction industry has underperformed

2

- 24 A changing market environment, technological progress, and disruptive new entrants will trigger industry overhaul

3

- 41 Almost half of incumbent value added is at stake

4

- 55 Transformation will take time, but the COVID-19 crisis will accelerate change

5

- 65 All players must prepare now for a fundamentally different next normal

In brief

The construction industry, and its broader ecosystem, erects buildings, infrastructure, and industrial structures that are the foundation of our economies and are essential to our daily lives. It has successfully delivered ever more challenging projects, from undersea tunnels to skyscrapers. However, the industry also has performed unsatisfactorily in many regards for an extended period of time. The COVID-19 pandemic may be yet another crisis that wreaks havoc on an industry that tends to be particularly vulnerable to economic cycles.

External market factors, combined with fragmented and complex industry dynamics and an overall aversion to risk, have made change both difficult and slow. The COVID-19 crisis looks set to dramatically accelerate the ecosystem's disruption that started well before the crisis. In such times, it is more important than ever for actors to find a guiding star for what the next normal will look like in the aftermath and make the bold, strategic decisions to emerge as a winner.

Many studies have examined individual trends such as modular construction and sustainability. This report provides an assessment of how the full array of disruptive trends will combine to reshape the industry in earnest. Our research builds future scenarios based on more than 100 conversations with experts and executives, firsthand experience serving clients throughout the ecosystem, and reviews of other industries and their transformation journeys. We confirmed the trends and scenarios that surfaced by conducting a survey of 400 global industry leaders. Finally, we quantitatively modeled value and profit pools across the value chain, based on company data today, and formulated future scenarios. We found overwhelming evidence that disruption will touch all parts of the industry and that it has already begun at scale.

Among our findings are the following:

- *Construction is the biggest industry in the world, and yet, even outside of crises, it is not performing well.* The ecosystem represents 13 percent of global GDP, but construction has seen a meager productivity growth of 1 percent annually for the past two decades. Time and cost overruns are the norm, and overall earnings before interest and taxes (EBIT) are only around 5 percent despite the presence of significant risk in the industry.
- *Nine shifts will radically change the way construction projects are delivered—and similar industries have already undergone many of the shifts.* A combination of sustainability requirements, cost pressure, skills scarcity, new materials, industrial approaches, digitalization, and a new breed of player looks set to transform the value chain. The shifts ahead include productization and specialization, increased value-chain control, and greater customer-centricity and branding. Consolidation and internationalization will create the scale needed to allow higher levels of investment in digitalization, R&D and equipment, and sustainability as well as human capital.

- *The COVID-19 crisis will accelerate change that has already started to occur at scale.* Our research suggests that the industry will look radically different five to ten years from now. More than 75 percent of respondents to our executive survey agreed that the nine shifts are likely to occur, and more than 60 percent believe they are likely to occur at scale in the next five years. We already see concrete signs of change: for example, the permanent modular-construction market share of new North American real-estate construction projects has grown by 50 percent from 2015 to 2018, R&D spending among the top 2,500 construction companies globally has risen by approximately 77 percent since 2013, and a new breed of player has emerged to lead the change. Two-thirds of survey respondents believe that COVID-19 will lead to an acceleration of the transformation, and half have already raised investment in that regard.
- *A \$265 billion annual profit pool awaits disrupters.* A value chain delivering approximately \$11 trillion of global value added and \$1.5 trillion of global profit pools looks set for overhaul. In a scenario based on analysis and expert interviews by asset class, strongly affected segments could have a staggering 40 to 45 percent of incumbent value added at risk, even when the economic fallout from COVID-19 abates—value that could shift to new activities such as off-site manufacturing, to customer surplus, or to new sources of profit. If the value at stake is captured by players in the construction ecosystem, total profit pools could nearly double, from the current 5 to 10 percent.² The scale and pace of change and the appropriate response will differ greatly among real-estate, infrastructure, and industrial construction—but all of them will be affected. Players that move fast and manage to radically outperform their competitors could grab the lion’s share of the \$265 billion in new and shifting profits and see valuations more akin to those of Silicon Valley start-ups than traditional construction firms.
- *To survive and thrive, incumbents must respond.* All of the players in the construction value chain will need to develop their strategies for dealing with or leading disruption. This is especially true for engineering and design, materials distribution and logistics, general contracting, and specialized subcontracting, all of which are likely to face commoditization and declining shares of value for parts of their activities. Companies can try to defend their positions and adjust to the changing environment, or reinvent themselves to take advantage of changes in the industry. All will need to invest in enablers like agile organizations.
- *Investors are well advised to use foresight on the respective shifts in their investment activity and will have ample opportunity to generate alpha.* Policy makers should help the industry become more productive and achieve better housing and infrastructure outcomes for citizens. And owners stand to benefit from better structures at lower cost if they play their part in making the shifts happen.



@Getty Images/Mel Melcon

Executive summary

Construction, which encompasses real estate, infrastructure, and industrial structures, is the largest industry in the global economy, accounting for 13 percent of the world's GDP. A closer look at its underlying performance highlights the industry's challenges in good economic times, let alone in times of crisis. We expect a set of nine shifts to radically change the way construction is done. Companies that can adjust their business models stand to benefit handsomely, while others may struggle to survive.

Historically, the construction industry has underperformed

Construction is responsible for a wide range of impressive accomplishments, from stunning cityscapes and foundational infrastructure on a massive scale to sustained innovation. However, in the past couple of decades, it also has been plagued by dismal performance.

Annual productivity growth over the past 20 years was only a third of total economy averages. Risk aversion and fragmentation as well as difficulties in attracting digital talent slow down innovation. Digitalization is lower than in nearly any other industry. Profitability is low, at around 5 percent EBIT margin, despite high risks and many insolvencies. Customer satisfaction is hampered by regular time and budget overruns and lengthy claims procedures.

The industry will feel the economic impact of the COVID-19 strongly, as will the wider construction ecosystem—which includes construction companies' component and basic-materials suppliers, developers

and owners, distributors, and machinery and software providers. At the time of writing, high levels of economic uncertainty prevail worldwide, and the construction industry tends to be significantly more volatile than the overall economy. MGI scenarios suggest that if things go well, construction activity could be back to pre-crisis levels by early 2021. But longer-term lockdowns could mean that it takes until 2024 or even later. In the past, crises have had an accelerative effect on trends, and this crisis is also expected to trigger lasting change impacting use of the built environment, like online channel usage or remote-working practices.

The lagging performance of the construction industry is a direct result of the fundamental rules and characteristics of the construction market and the industry dynamics that occur in response to them. Cyclical demand leads to low capital investment, and bespoke requirements limit standardization. Construction projects are complex, and increasingly so, and logistics need to deal with heavy weight and many different parts. The share of manual labor is high, and the industry has a significant shortage of skilled workers in several markets. Low barriers to entry in segments with lower project complexity and a significant share of informal labor allow small and unproductive companies to compete. The construction industry is extensively regulated, subject to everything from permits and approvals to safety and work-site controls, and lowest-price rules in tenders make competition based on quality, reliability, or alternative design offerings more complicated.

In response to these market characteristics, today's construction industry must grapple with several dynamics that impede productivity and make change more difficult. Bespoke projects with unique features and varying topology have a limited degree of repeatability and standardization. Local market structures and ease of entry have resulted in a fragmented landscape (both vertically and horizontally) of mostly small companies with limited economies of scale. Moreover, every project involves many steps and companies in every project with scattered accountability, which complicates the coordination. Contractual structures and incentives are misaligned. Risks are often passed to other areas of the value chain instead of being addressed, and players make money from claims rather than from good delivery. High unpredictability and cyclicity have led construction firms to rely on temporary staff and subcontractors, which hampers productivity, limits economies of scale, and reduces output quality and customer satisfaction.

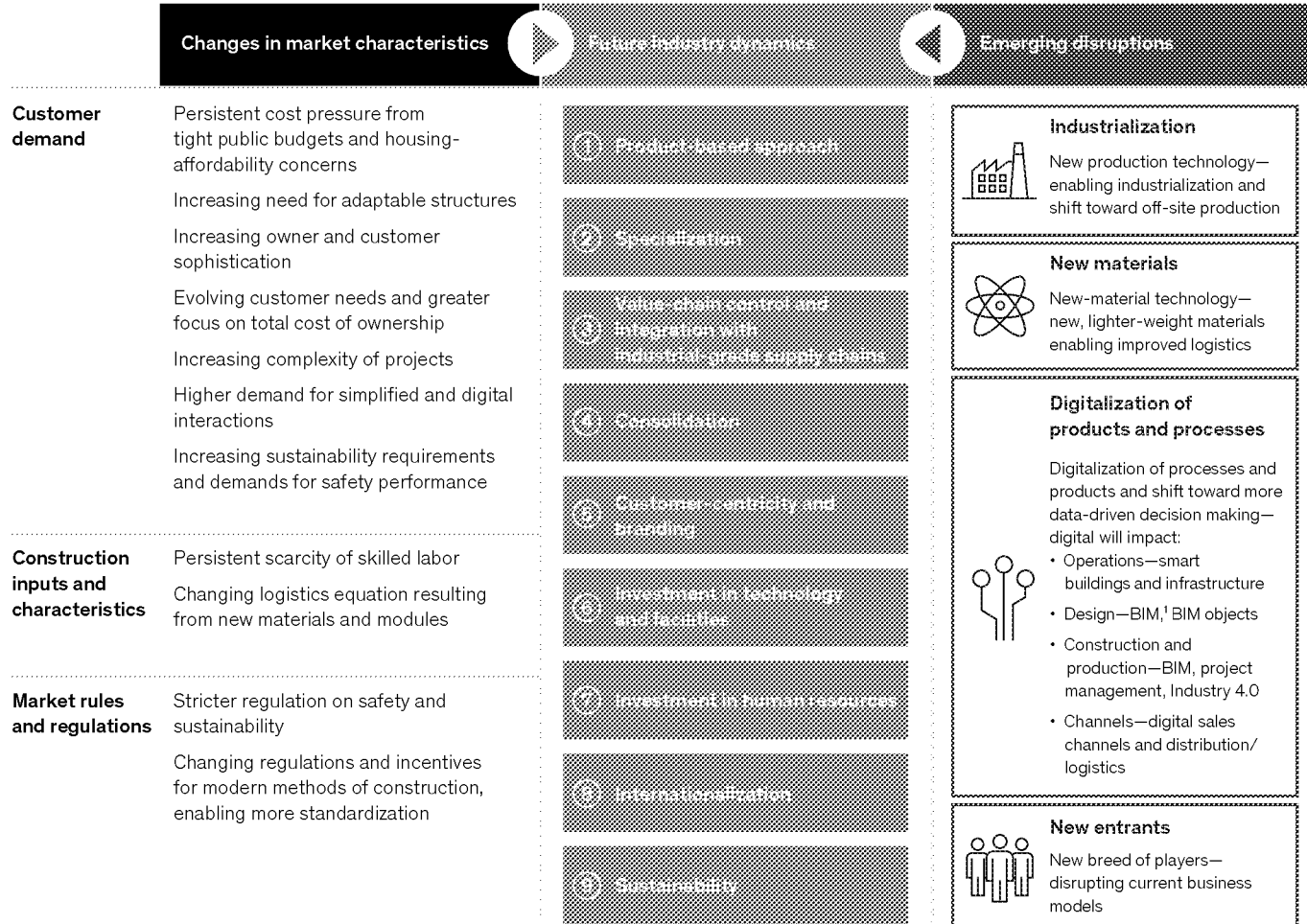
A changing market environment, technological progress, and disruptive new entrants will trigger industry overhaul

The construction industry was already starting to experience an unprecedented rate of disruption before the COVID-19 pandemic. In the coming years, fundamental change is likely to be catalyzed by changes in market characteristics, such as scarcity of skilled labor, persistent cost pressure from infrastructure and affordable housing, stricter regulations on work-site sustainability and safety, and evolving sophistication and needs of customers and owners. Emerging disruptions, including industrialization and new materials, the digitalization of products and processes, and new entrants, will shape future dynamics in the industry (Exhibit A).

Sources of disruption

Rising customer sophistication and total-cost-of-ownership (TCO) pressure. Customers and owners are increasingly sophisticated, and the industry has seen an influx of capital from more savvy customers. From 2014 to 2019, for example, private-equity firms raised more than \$388 billion to fund infrastructure projects, including \$100 billion in 2019 alone, a 24 percent increase from 2018. Client demands are also evolving regarding performance, TCO, and sustainability: smart buildings, energy and operational efficiency, and flexibility and adaptability of structures will become higher priorities. Expectations are also rising among customers, who want simple, digital interactions as well as more adaptable structures.

Changing characteristics and emerging disruptions will drive change in the industry and transform ways of working.



¹Building information modeling.

The industry is facing persistent cost pressure because of tight public budgets and housing-affordability issues. McKinsey analysis found that \$69.4 trillion in global infrastructure investment would be needed through 2035³ to support expected GDP growth and that every third global urban household cannot afford a decent place to live at market prices.⁴ The economic fallout of the COVID-19 crisis magnifies the cost and affordability issues.

Persistent scarcity of skilled labor and changing logistics equations. Skilled-labor shortages have become a major issue in several markets, and retirements will drain talent. For example, about 41 percent of the current US construction workforce is expected to retire by 2031. The impact the COVID-19 crisis will have on this dynamic in the long term is unclear at the time of writing.

Safety and sustainability regulations and possible standardization of building codes. Requirements for sustainability and work-site safety are increasing. In the wake of COVID-19, new health and safety procedures will be required. The global conversation about climate change puts increasing pressure on the industry to reduce carbon emissions.

At the same time, in some markets, governments are recognizing the need to standardize building codes or provide type certificates and approvals for factory-built products rather than reviews of each site. The process, however, is still slow.

Industrialization. Modularization, off-site production automation, and on-site assembly automation will enable industrialization and an off-site, product-based approach. The shift toward a more controlled environment will be even more valuable as the COVID-19 pandemic further unfolds. The next step in the transition to efficient off-site manufacturing involves integrating automated production systems—essentially making construction more like automotive manufacturing.

New materials. Innovations in traditional basic materials like cement enable a reduction of carbon footprints. Emerging lighter-weight materials, such as light-gauge steel frames and cross-laminated timber, can enable simpler factory production of modules. They will also change the logistics equation and allow longer-haul transport of materials and greater centralization.

Digitalization of products and processes. Digital technologies can enable better collaboration, greater control of the value chain, and a shift toward more data-driven decision making. These innovations will change the way companies approach operations, design, and construction as well as engage with partners. Smart buildings and infrastructure that integrate the Internet of Things (IoT) will increase data availability and enable more efficient operations as well as new business models, such as performance-based and collaborative contracting. Companies can improve efficiency and integrate the design phase with the rest of the value chain by using building-information modeling (BIM) to create a full three-dimensional model (a “digital twin”)—and add further layers like schedule and cost—early in the project rather than finishing design while construction is already underway. This will materially change risks and the sequence of decision making in construction projects and put traditional engineering, procurement, and construction (EPC) models into question. Automated parametric design and object libraries will transform engineering. Using digital tools can significantly improve on-site collaboration. And digital channels are spreading to construction, with the potential to transform interactions for buying and selling goods across the value chain. As in other industries, the COVID-19 pandemic is accelerating the integration of digital tools.

New entrants. Start-ups, incumbent players making new bets, and new funding from venture capital and private equity are accelerating disruption of current business models. As the COVID-19-propelled economic crisis unfolds, we also expect an increase in corporate restructuring and M&A activity.

The nine resulting industry shifts

In response, we expect nine shifts to fundamentally change the construction industry. According to our executive survey, more than 75 percent of respondents agree that these shifts are likely to occur, and more than 60 percent believe that they are likely to occur within the next five years. The economic fallout from the COVID-19 pandemic looks set to accelerate them.

Product-based approach. In the future, an increasing share of structures and surrounding services will be delivered and marketed as standardized “products.” This includes developers promoting branded offerings,

with standardized but customizable designs that can improve from one product generation to the next, and delivery using modularized elements and standardized components produced in off-site factories. The modules and elements will be shipped and assembled on site. Production will consist of assembly line–like processes in safe, nonhostile environments with a large degree of repeatability.⁵

Specialization. To improve their margins and levels of differentiation, companies will start to specialize in target niches and segments (such as luxury single-family housing, multistory residential buildings, hospitals, or processing plants) in which they can build competitive advantages. And they will specialize in using different materials, subsegments, or methods of construction. The shift toward specialization will also require companies to develop and retain knowledge and capabilities to maintain their competitive advantages. Obviously, players will need to weigh carefully the effectiveness, efficiency, and brand positioning that greater specialization enables against the potential risk or cyclical benefits of a more diversified portfolio.

Value-chain control and integration with industrial-grade supply chains. Companies will move to own or control important activities along the value chain, such as design and engineering, select-component manufacturing, supply-chain management, and on-site assembly. Companies will be able to achieve this goal through vertical integration or strategic alliances and partnerships by using collaborative contracting and more closely aligned incentives. Digital technology will change the interaction model: BIM models will lead to more decision making early on in the process, distribution will move toward online platforms and advanced logistics management, and end-to-end software platforms will allow companies to better control and integrate value and supply chains. Value-chain control or integration will reduce interface frictions and make innovation more agile.

Consolidation. Growing needs for specialization and investments in innovation—including the use of new materials, digitalization, technology and facilities, and human resources—will require significantly larger scale than is common today. As product-based approaches, with higher standardization and repeatability, further increase the importance of gaining scale, the industry is likely to increasingly see a significant degree of consolidation, both within specific parts of the value chain and across the value chain.

Customer-centricity and branding. With productization—that is, turning development, engineering, or construction services into easy-to-market products or solutions⁶—and specialization in the industry, having a compelling brand that represents an organization's distinctive attributes and values will take on added importance. As in traditional consumer industries, a strong brand can tie customers more closely to the construction company's or supplier's products and help to build and maintain relationships and attract new customers. Similar to brands in other manufacturing industries, such construction brands will encompass, among other aspects, product and service quality, value, timing of delivery, reliability, service offerings, and warranties.

Investment in technology and facilities. Productization implies a need to build off-site factories, which requires investments in plants, manufacturing machinery and equipment (such as robotics to automate manufacturing), and technology. Where modular is not used, the construction site also will likely become more capital intensive, using advanced automation equipment and drones, among other technologies. R&D investment will become more important for specialized or more productized companies, so companies are likely to increase spending to develop new, innovative products and technologies.

Investment in human resources. Innovation, digitalization, value-chain control, technology use, and specialization in end-use segments all increase the importance of developing and retaining in-house

expertise, which will compel players to invest more in human resources. The importance of risk management and other current capabilities will decrease and be replaced by an emphasis on others, such as supply-chain management. To build the necessary capabilities, companies will need to further invest in their workforces. This becomes even more important in light of the transition to the future of work.⁷ Most incumbents struggle to attract the digital talent they need, and will need to raise excitement about their future business models.

Internationalization. Greater standardization will lower the barriers to operating across geographies. As scale becomes increasingly important to gaining competitive advantages, players will increase their global footprints—both for low-volume projects in high-value segments such as infrastructure, as well as for winning repeatable products that will be in demand across the world. The COVID-19 pandemic might slow down this development.

Sustainability. While sustainability is an important decision factor already, we are only at the very beginning of an increasingly rapid development. Beyond the carbon-abatement discussions, physical climate risks are already growing and require a response.⁸ Companies will need to consider the environmental impact when sourcing materials, manufacturing will become more sustainable (for example, using electric machinery), and supply chains will be optimized for sustainability as well as resilience. In addition, the working environments will need to radically change from hostile to nonhostile, making construction safer. Water consumption, dust, noise, and waste are also critical factors.

Today's project-based construction process looks set to shift radically to a product-based approach (Exhibit B). Instead of building uniquely designed structures on the jobsite, companies will conduct their production at off-site construction facilities.⁹ Standardized sub-elements and building blocks will likely be designed in house in R&D-like functions. The elements will be manufactured separately and then combined with customization options to meet bespoke requirements. To produce efficiently and learn through repetition, developers, manufacturers, and contractors will need to specialize in end-user segments. Data-driven business models will emerge. Overall, the process may resemble manufacturing in other industries such as shipbuilding or car manufacturing.

There is reason to believe that a winner-take-most dynamic will emerge, and companies that fail to adjust fast enough risk seeing market shares and margins erode until they eventually go out of business.

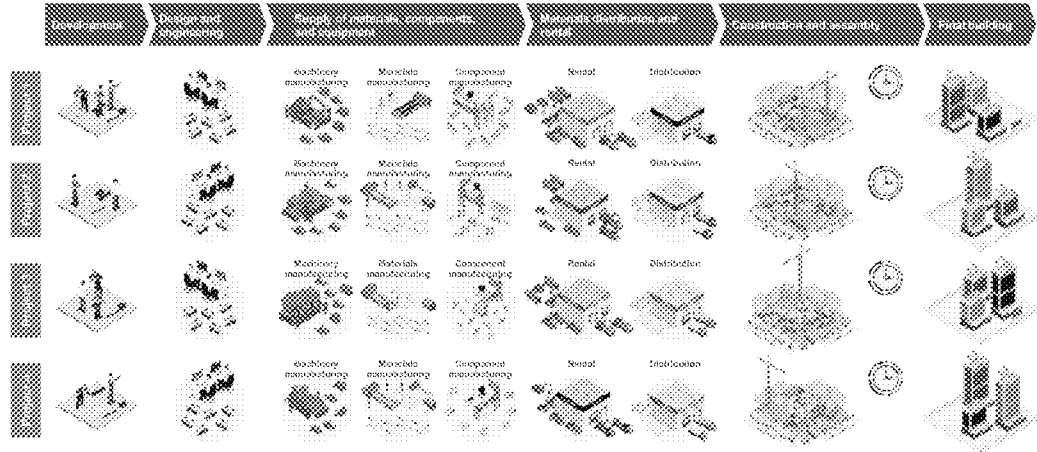
Construction is not the first industry to encounter lagging productivity and disruption across the value chain. Lessons can be learned from others that had similar traits and encountered the same challenges that construction faces now. We have analyzed shifts in four of them: shipbuilding, commercial aircraft manufacturing, agriculture, and car manufacturing. Clear patterns of the shifts are evident in all of them, and value shifted to those handling the change best. Innovation in production technology and new work methods kick-started all four of the industries' journeys. Today, across industries, winners continue to heavily invest in technology, many with focus on digitalization and data-driven products and services.

In commercial aircraft manufacturing, for example, the industry landscape was highly fragmented. Each airplane was built from scratch in a bespoke and project-based-manufacturing setup. Industrialization sparked a shift toward assembly-line manufacturing, which later became highly automated. As a result of the subsequent standardization, the industry entered a phase of consolidation that led to the rise of two major players: Airbus and Boeing. The transformation resulted in a significant shift of value to customers. This transformation journey took roughly 30 years to complete, as commercial aircraft manufacturing faced barriers to change similar to those now confronting construction.

The future construction ecosystem will be radically different.

Today's construction ecosystem

A highly complex, fragmented, and project-based construction process . . .



The construction process is highly **project based**—developed from unique customer specifications, using designs **planned from scratch**, and with limited degree of repetition

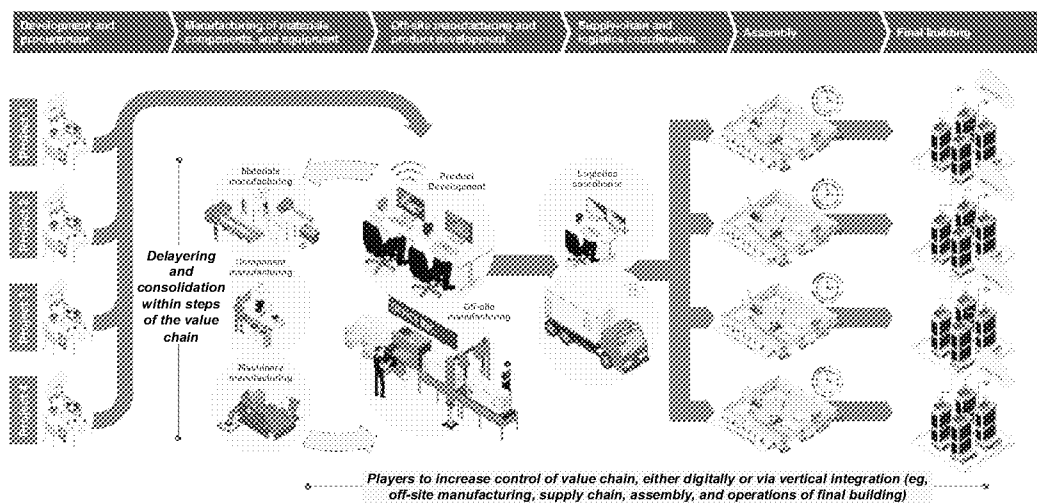
The value chain and player landscape are **local and highly fragmented vertically and horizontally**, resulting in a multitude of players involved at each step and major interface frictions

Construction is performed by generalists **on site in hostile environments**, with a large part of the workforce being **temporary and manual**

Limited use of **end-to-end digital tools and processes** as well as a capital-light delivery approach

The construction ecosystem of the future

. . . A more standardized, consolidated, and integrated construction process



The construction process is increasingly **product based**, meaning structures will be products and manufactured off site by branded product houses **specializing** in certain end-user segments

Developers choose **entire designs or specific components** from a **library** of options developed in house or offered externally on the market

Value chain is more **consolidated, both vertically (delayering) and horizontally**, with increased degree of **internationalization**

Disintermediation takes place through digital marketplaces and direct channels

Contractors focus on **lean, on-site execution and assembly of products**

Data and analytics on customer behavior generated after completion to optimize total cost of ownership and future designs

Almost half of incumbent value added is at stake

The transformation of the industry will create both large opportunities and sizable risks as value and profit pools shift in the next 15 years. Over the past years, approximately \$11 trillion in value added and \$1.5 trillion in profits have been unevenly distributed along the construction value chain and across all asset classes. Looking ahead, up to 45 percent of incumbent value may be at stake in those parts of the market most heavily affected by shifts, such as hotel construction (Exhibit C). Of this total, 20 to 30 percentage points will be kept and redistributed within the ecosystem to enable the shifts to take place. The remaining 15 to 20 percentage points will be value up for grabs as a result of the cost savings and productivity gains generated by the shifts, with the benefits accruing to players or customers (in the form of price reductions or quality increase). If that value is captured fully by players in the ecosystem, total profit pools could nearly double, to 10 percent, from the current 5 percent.¹⁰ Players that move fast and manage to radically outperform their competitors could grab the lion's share of the \$265 billion in new profit pools.

Some players will be more affected than others. For example, software providers are expected to significantly increase their value-added contribution, albeit from a small base of 1 to 2 percent of the value chain. Also, a large share of value is expected to move from construction jobsites to off-site prefabrication facilities. In contrast, general and specialized contractors could face a large decline unless they reposition themselves as companies that go beyond execution alone. Basic design and engineering and materials distribution and logistics may face substantial commoditization and automation risks.

The value at stake could benefit either the players in the ecosystem as profits increase, workers in the form of higher wages, or customers through lower prices and higher quality. Companies that move fast and manage to lower their cost base and increase productivity will have an advantage over the competition. These early movers could translate their productivity gains into profit. In the long term, as other players adjust and competition intensifies, the dynamics in other industries suggest that a large share of the gains will be passed on to customers.

Our baseline scenario estimates that 10 to 12 percent of construction activities will move along shifts outlined in this report by 2035, but change will vary significantly by asset class because of different starting points and abilities to transform. In real estate, for example, we expect that by 2035 an additional 15 percent of new building projects could be completed through a redesigned value chain. This higher-than-average number is partly the result of the potential for standardization in single- and multifamily residential, hotels, offices, and hospitals. For infrastructure, approximately 7 percent of additional new building volume could be delivered in a transformed way—with bridges, airports, and railways, for example, having particular potential. Industrial construction could see an additional penetration of about 5 percent, as several of its subsegments have already made significant progress in the past.

Transformation will take time, but the COVID-19 crisis will accelerate change

The full transformation of the construction industry could take decades, but the process has already begun. Our survey shows that industry leaders largely agree that the shifts outlined in this report are likely to occur at scale within the next five to ten years, and that the COVID-19 crisis will accelerate shifts.

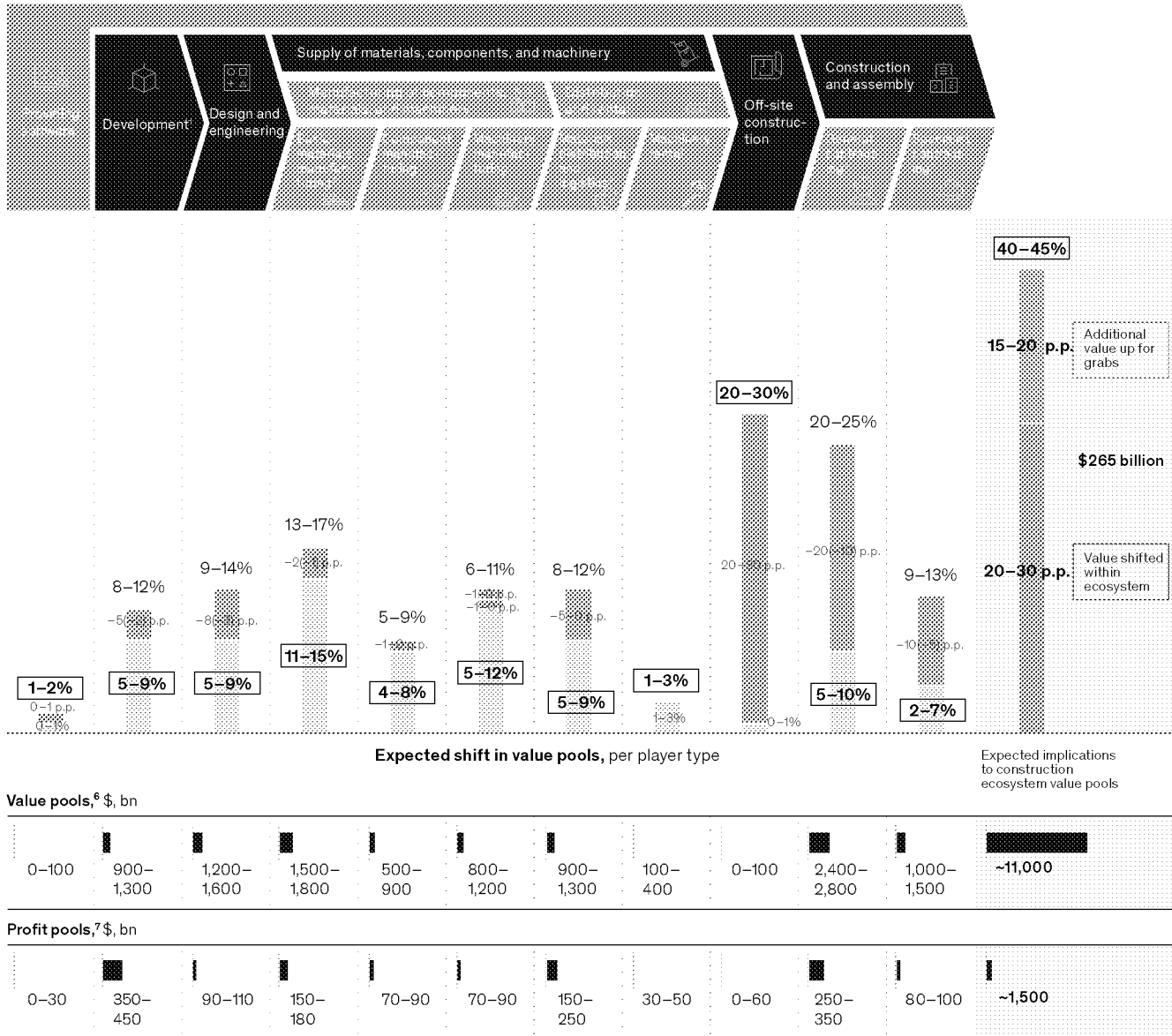
Our executive survey of 400 decision makers in November and December 2019 found that the attitudes of executives have evolved materially since three to five years ago (see sidebar “About the executive survey” in chapter 1 for more details on the survey). In all, 90 percent of the respondents strongly believe that the industry needs to change and that this sentiment has grown in the past ten years. Eighty percent also believe that the construction industry will look radically different 20 years from now.

Exhibit C

Forty to 45 percent of value pools are expected to shift and impact all players along the value chain.

Example of fully productized value chain (eg, real estate new build), current and future value pools, p.p.

Value at risk Remaining value added Value shifted Value captured



¹ Calculated by applying an assessed share of total value of development of output per asset class, allocated on top of total market output, since a limited number of stand-alone, pure-player developers have been identified.
² Looking at players processing raw materials but not the actual manufacturing of raw materials (eg, mining). If all steps of producing and refining raw materials were included, the value pool would be ~2.5x bigger.
³ Adjusted downward to reflect that some things materials distributors sell don't contribute to construction output (eg, clothes, white goods).
⁴ General builders (buildings and other heavy construction).
⁵ Specialized trade construction.
⁶ Defined as value added per player type.
⁷ EBIT pools

Source: CapitaIQ; Euroconstruct; FMI; McKinsey analysis

Beyond our analysis and the overwhelming beliefs of the surveyed executives, we see signs today that the industry had already started to change before the COVID-19 crisis began. For instance, adoption of product-based approaches is increasing. In North America, the permanent modular-construction market share of new real-estate construction projects grew by approximately 51 percent from 2015 to 2018, and revenues for the segment grew (from a small base of \$2 billion) by a factor of 2.4 over the same period. Also, emerging players as well as incumbents are already seeking to control a larger part of the value chain; Katterra, for instance, used new technology to control the value chain, including design and engineering and off-site manufacturing. Indicators suggest the construction industry is increasing its emphasis on R&D, and companies that have invested in construction technology and facilities are gaining traction. Global R&D spending by the top 2,500 construction companies grew by 77 percent from 2013 to 2017.

The COVID-19 crisis looks set to accelerate change (Exhibit D). We conducted an additional survey in early May 2020 to understand the potential implications of the crisis on the disruptions and shifts outlined in the report. Respondents comprised 100 decision makers out of the same sample that responded to our first survey. Nearly two-thirds of respondents believe that the COVID-19 crisis will accelerate industry transformation, and half have already raised investment in line with the shifts. Investments in digitalization and supply-chain control are most pronounced, while respondents believe the crisis will slow down internationalization and the rise of new entrants—giving incumbents a rare opportunity to step in and drive change.

All players must prepare now for a fundamentally different next normal

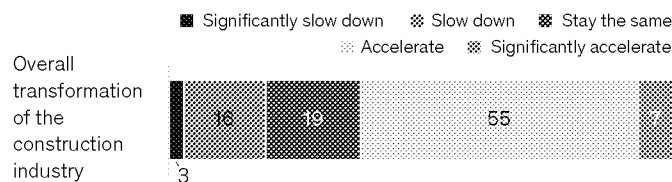
Our research shows that leaders leave laggards behind in times of crisis. Those that go beyond managing their survival to take fast, bold, strategic action tend to emerge as the winners. During past economic cycles, companies that managed to move quickly to improve their productivity (for example, reducing

Exhibit D

Two-thirds of survey respondents believe that the COVID-19 crisis will accelerate industry transformation.

As a result of COVID-19, do you believe that transformation of the construction industry will accelerate, stay the same, or slow down?

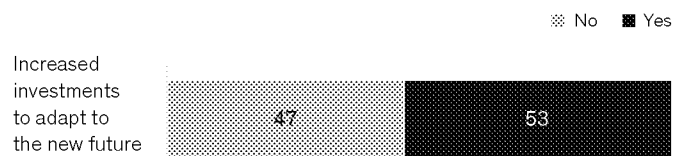
Share of respondents, %



Around two-thirds of respondents believe that the COVID-19 crisis will accelerate the overall transformation of the construction industry

As a result of COVID-19, has your company increased overall investments to adapt to the new future?

Share of respondents, %



More than 50% of respondents' companies have started to invest more to adjust to the new future

Source: Survey of 100 industry CxOs, May 2020

In the face of this transformation, companies all along the value chain need to review where they want to play.

their cost of goods sold through operational efficiency), divest earlier and are more acquisitive during the recovery. They cleaned up their balance sheets ahead of a downturn and outperformed competition in both revenues and earnings before interest, depreciation, taxes, and amortization (EBITDA).

Players in the ecosystem will need to develop strategies to deal with the disruption ahead. Our survey respondents identified four types of players set to face the largest long-term decline: design and engineering firms, materials distributors, general contractors, and specialist contractors. Furthermore, respondents believe that general contractors will be required to move first, as they could experience commoditization and a declining share of value.

In the face of this transformation, companies all along the value chain need to review where they want to play: which asset classes, segments, geographies, and value-chain steps. They will need to assess the impact of each of the disruptions and the nine shifts, decide how they want to act on them, and define new-business models and operating models in line with those decisions. This process is critical whether they aim to defend their core business and adjust to the new environment or fundamentally reinvent themselves and attack. For success, it will be critical for companies to invest in a set of enablers, such as agile organizations. Finally, companies can choose how to implement the new strategy and transformation, whether it's trying to evolve incumbent operations to work within the new setup, starting up new divisions or arm's-length operations, or applying targeted M&A.

In the materials-distribution and logistics segment, for instance, off-site manufacturing facilities will shift demand for shipments to factory hubs, the main logistics nodes, which will increase customer expectations for just-in-time delivery. The segment will be further reshaped by online and direct sales channels (including new competition from online-distribution behemoths), rising customer expectations, and increased use of technologies such as advanced analytics or automated warehouses. A shift in procurement activity, from small specialized trades firms to larger contractors, will affect companies' bargaining power, and internationalization will enable companies to source more from low-cost countries.

In response, companies could try to defend their core by, for instance, focusing on the refurbishment market, becoming leaner, and undertaking category reviews. They could adjust to the changing

environment by, for example, strengthening customer relationships, offering new business solutions to avoid disintermediation, consolidating to gain scale, and developing industrial-grade supply-chain capabilities. Reinvention would entail becoming the logistics hub of the future construction landscape. Strategies could include partnering closely with off-site manufacturers and materials suppliers to optimize logistics and inventory according to their needs, helping with international sourcing, or offering credit financing.¹¹

Companies that familiarize themselves with the next normal and move quickly will be best positioned to both create value and maintain their competitive edge.

Organizations that are adjacent to the construction ecosystem should look to facilitate—and benefit from—the coming changes. Investors are well advised to use foresight to anticipate the respective shifts and generate above-market returns. Insurance companies are already factoring use of modern methods of construction into their terms. Policy makers should help the industry become more productive and thereby attain better housing and infrastructure for citizens. And building owners stand to benefit from better structures at lower costs if they play their part in making the shifts happen.

Construction is already in the perfect storm. Industrialization, globalization, and digitalization have been key drivers of change in all industries. While this change happened in sequential waves—for example, in auto industrialization in the 1970s and 1980s, globalization in the 1990s and 2000s, and digitalization in the 2010s and ongoing—all of these drivers are hitting construction simultaneously. It is a daunting task and will require bold and agile moves to maneuver, but the size of the prize is enormous.



@Getty Images/Jung Getty

1 Historically, the construction industry has underperformed

The construction industry, which encompasses real estate, infrastructure, and industrial structures, is the largest industry in the global economy, accounting for 13 percent of the world's GDP (Exhibit 1). A closer look at its underlying performance highlights the industry's challenges in good economic times, let alone in times of crisis. We expect a set of nine shifts to radically change the way construction is done. Companies that can adjust their business models stand to benefit handsomely, while others may struggle to survive.

Poor long-term performance stems from unfavorable market characteristics and industry dynamics

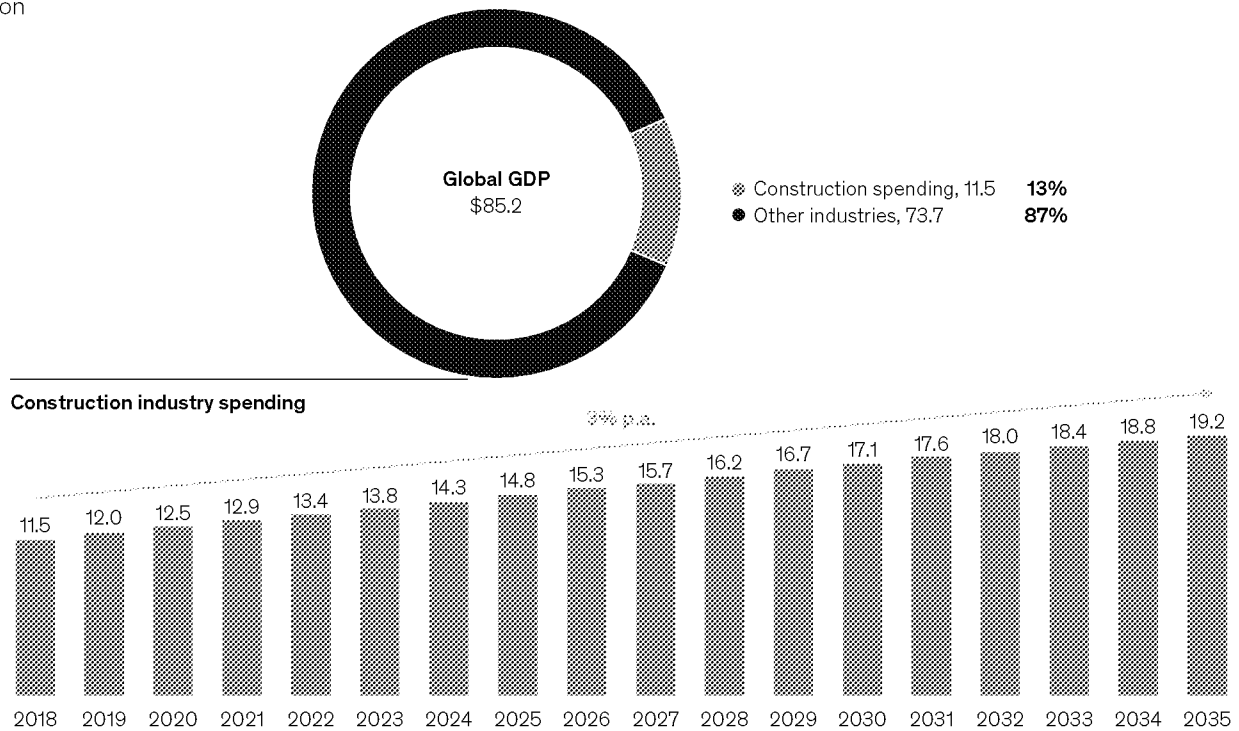
Construction is responsible for a wide range of impressive accomplishments, from stunning cityscapes and foundational infrastructure on a massive scale to sustained innovation. However, in the past couple of decades, it has been plagued by dismal performance, characterized by lagging productivity growth, limited digitalization, frequent insolvencies, and low levels of customer satisfaction. The industry will feel the economic impact of COVID-19 strongly, as will the wider construction ecosystem—which includes construction companies' suppliers, distributors, customers, and competitors, as well as government regulators.

Lagging productivity growth. The industry's global annual labor productivity growth over the past 20 years was less than 1 percent, which is significantly less than the productivity growth of the global economy, approximately 2.8 percent a year.¹²

Exhibit 1

Construction-related spending accounts for 13 percent of global GDP.

\$ trillion



Note: Due to COVID-19, the amount of spending in 2020 and subsequent years is likely to change

Source: IHS Global Insight; ISSA – Infrastructure Stock & Spend Analyzer; World Bank, McKinsey Global Institute analysis

Slow innovation and digitalization. A low degree of standardization combined with a fragmented value chain provides low barriers to entry, and has resulted in limited economies of scale, which impede innovation and digitalization. A 2015 McKinsey Global Institute (MGI) analysis found that the construction industry was among the least digitized industries in the total economy across assets, usage, and labor. Innovation is further hampered, understandably, by risk aversion and limited margins. For example, building-information modeling (BIM) adoption rates have reached just 60 to 70 percent in 35 years. For comparison, many technologies (such as cloud customer relationship management, magnetic resonance imaging, laparoscopic surgery, lithium-ion batteries, and microwaves) reached an adoption rate of 90 percent in eight to 28 years from when they became commercially available.

Low profits and high risks despite strong growth. The industry's overall profitability is about 5 percent, and lower in certain parts of the value chain. The construction industry frequently tops insolvency lists across geographies because of a combination of low profitability and high risk. According to a global Euler Hermes analysis,¹³ construction was the industry with the most major insolvencies in the first three quarters of 2018, approximately 5 percent higher than the second-worst sector, and 70 percent higher than the third. Construction topped a UK insolvency list in the first quarter of 2019, with approximately 3,000 insolvencies over the previous 12 months.¹⁴ Valuation multiples are significantly below the average—EV/EBITDA multiple

Low productivity growth: Less than 1.0 percent per year for the industry over the past 20 years, versus 2.8 percent for the total economy.

for the engineering and construction sector has averaged 5.8 over the past ten years, compared with 12.4 for the S&P 500.

Low customer satisfaction and regular time and budget overruns. A 2016 McKinsey analysis found that construction projects typically take 20 percent longer to finish than scheduled and are up to 80 percent over budget, frequently resulting in litigation.¹⁵ That often leaves customers dissatisfied, resulting in complex and time-consuming claims processes.

For these reasons, awareness was growing even before the COVID-19 crisis struck that the ecosystem must transform to meet the challenges ahead. Our industry survey found that the attitudes of CxOs have evolved materially since three to five years ago (see sidebar “About the executive survey”). In all, 90 percent of executives strongly believe that the industry needs to change and that this sentiment has grown in the past ten years. Eighty percent also believe that the construction industry will look radically different 20 years from now.

20%

Low customer satisfaction:
Typical schedule overrun
of 20 percent for large
construction projects

About the executive survey

Our survey of construction industry leaders was conducted from November to December 2019. Respondents comprised 400 decision makers, the majority of whom were CxOs, owners, and executives, who were distributed over asset classes and geographies as well as across the value chain (exhibit).

Specifically, 63 percent of the respondents were from the real-estate sector; 19 percent, infrastructure; 18 percent, industrial. The geographies break down in the following way: 47 percent are from North America; 39 percent, Europe; 11 percent, Asia–

Pacific (APAC); 2 percent, Middle East and Africa; and 2 percent, Latin America.

We asked questions and let respondents react to our initial hypotheses on the industry, such as on the most important issues in construction today, the change factors that will have the largest impact on the industry, and the most significant disruptions. We also included questions on the future of the construction industry, such as which shifts outlined in this report are most likely to occur, when they will affect the industry at scale, which players will need to move first, and which players will

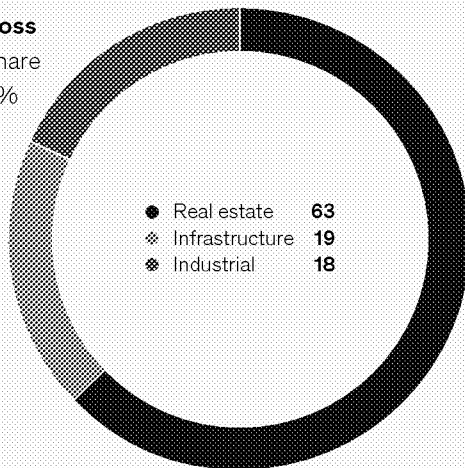
decline the most in the coming ten years.

Given that the COVID-19 crisis was unfolding during the publishing process of this report, an additional survey was conducted in early May 2020 to help us understand the potential implications of the crisis on accelerating the disruptions and shifts outlined in the report. Respondents comprised 100 decision makers from the same sample that responded to the first survey in November and December 2019, with similar distribution over asset classes and geographies as well as across the value chain.

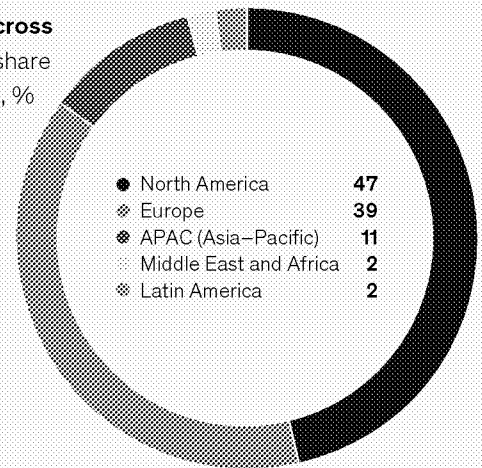
Exhibit

The distribution of survey respondents to our initial survey represents a fair view of the setup of the industry today.

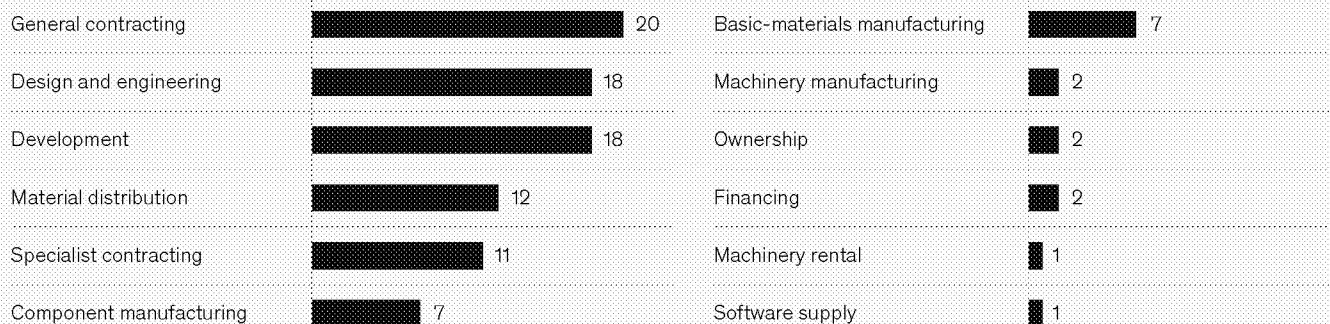
Distribution across asset classes, share of respondents, %



Distribution across geographies, share of respondents, %



Distribution across the value chain, share of respondents, %



Source: McKinsey survey of 400 construction industry CxOs; expert interviews; McKinsey analysis

Rough seas lie ahead for the construction ecosystem. Beyond the economic downturn that will reduce construction demand, the COVID-19 crisis has also shocked supply chains and may lead to lasting shifts in investment patterns (see sidebar “The impact of the COVID-19 crisis on the construction industry’s outlook”).

How companies respond determines industry dynamics and outcomes

The lagging performance of the construction industry is a direct result of the fundamental rules and characteristics of the construction market and the industry dynamics that occur in response to them (Exhibit 2). Our analysis identified the following three broad market characteristics that are now influencing the external environment:

Cyclical demand with bespoke customer requirements

The industry benefits from robust long-term demand, thanks to economic and population growth and rising urbanization around the world. That demand is highly sensitive to economic cycles, however, leading to low capital investment that slows productivity growth. A fragmented market includes customers ranging from individual single-family homeowners making once-in-a-lifetime purchasing decisions to administrators in

The impact of the COVID-19 crisis on the construction industry’s outlook

The economic impact of COVID-19 will be felt strongly throughout the ecosystem. At the time of writing, high levels of economic uncertainty prevail worldwide. MGI scenarios suggest that if things go well, the virus is contained within months, and the right economic policies are implemented, economic activity could be back to pre-crisis levels by early 2021. But longer-term lockdowns, even if intermittent, or other severe restrictions that last until a vaccination is developed, could throw the economy into a severe and sustained downturn—with economic activity returning to 2019 levels only in 2024 or even later.¹

The construction industry is typically significantly more volatile than the overall economy, and it might benefit from public stimulus programs. As economic activity

recedes, there is less need for new commercial or industrial structures, uncertainty dampens investment, and income losses and lower consumer confidence hurt housing construction. As the value of structures in an economy closely tracks GDP, the need for new construction activity is also highly sensitive to GDP growth in longer-term models. A slump as long as five years could substantially reduce construction’s share of GDP beyond the initial contraction, even though this crisis is, in contrast to the global financial recession of 2008–09, not primarily a real-estate crisis.

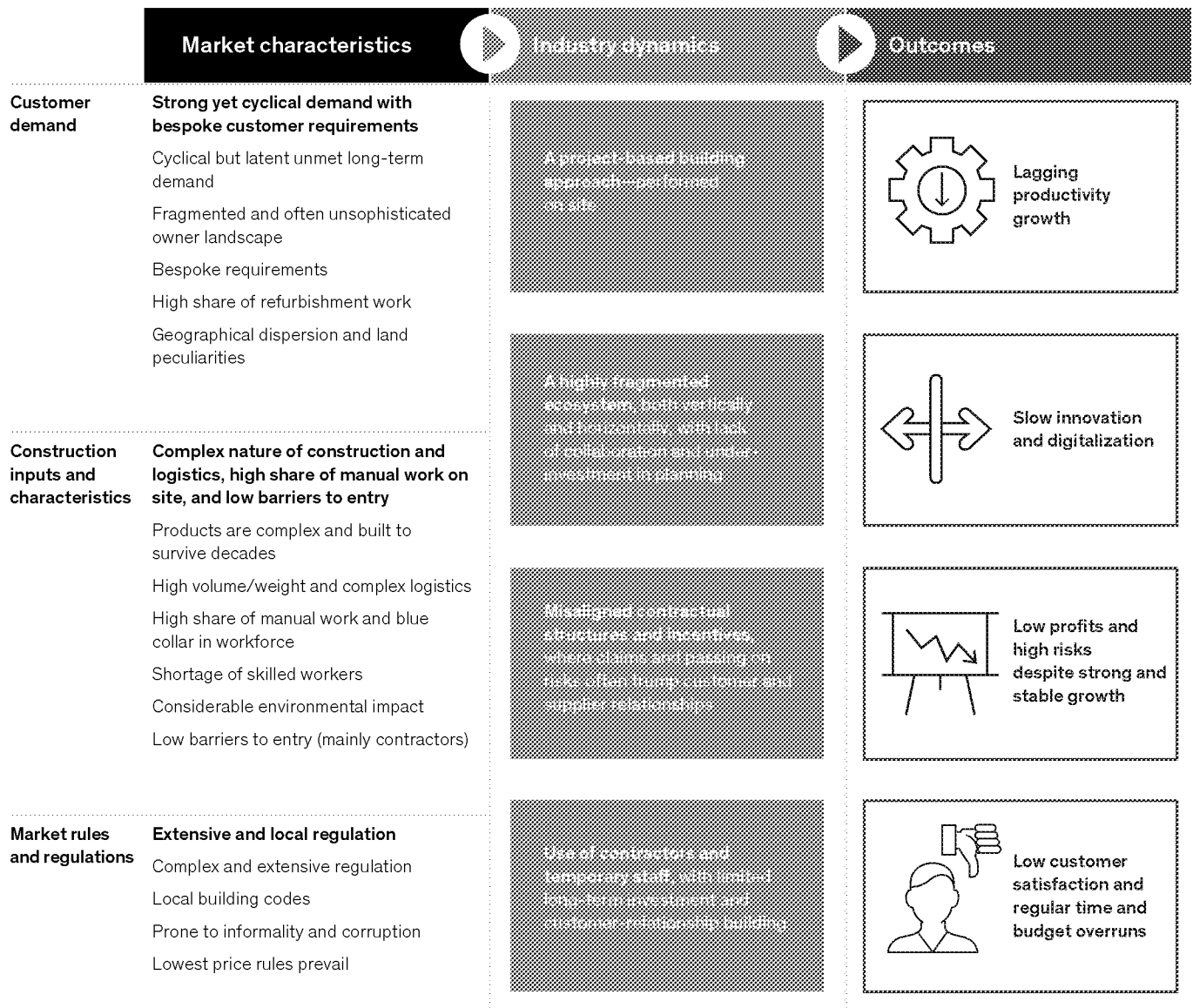
On the upside, unprecedented public stimulus packages passed by Congress could not only help support a V-shaped recovery but also be followed by public investment programs.

COVID-19 also represents a shock to supply. Migrant labor cannot always cross borders, construction workers cannot always get to or work at jobsites, and all employees will need to act in line with clear protocols for some time. Some building-materials supply chains are interrupted.

As economic activity resumes, we may find that the demand for structures has changed permanently. It is too early to judge whether changes in preferences—such as shopping online rather than at malls, employees preferring to work remotely and thus reducing the need for office space, or different airline-terminals or residential layouts—will be lasting. But it is important to monitor such developments closely.

¹ Sven Smit, Martin Hirt, Kevin Buehler, Susan Lund, Ezra Greenberg, and Arvind Govindarajan, “Safeguarding our lives and livelihoods: The imperative of our time,” March 2020, McKinsey.com.

Market characteristics have shaped an industry response with unfavorable outcomes.



city governments overseeing billion-dollar infrastructure projects. Inexperienced or unsophisticated owners often lack the experience and competence to navigate the industry landscape. Further, many projects ask for a significant degree of customization, and small projects and renovations account for a majority of industry revenues. Local-market structures and land peculiarities result in a geographically dispersed industry, with few companies achieving global scale.

Complex nature of construction and logistics, much manual work on site, and low barriers to entry

Construction projects have become increasingly complex, with a growing share of megaprojects that are above \$1 billion. The size and weight of materials and components entail a regionally fragmented setup. Logistics are further complicated by the large number and variety of components and suppliers as well as the delivery by separate parties of different subcomponents of the same products.

Although projects increasingly require more experience and skill to execute, construction work still relies on a large share of manual work being completed by a largely blue-collar workforce. Construction firms typically grapple with a significant shortage of skilled labor, which hampers both existing operations and innovations such as digitalization. Our survey highlighted the labor shortage as one of the main issues in the construction industry, with around 85 percent of respondents in our survey saying pre-crisis that the issue is highly important, although the coming economic turmoil could change that.

Low barriers to entry and a significant share of informal labor allow small and unproductive companies to compete.

Extensive and local regulation

The construction industry is extensively regulated, subject to everything from permits and approvals to safety and work-site controls. As each geography has its own local building codes, companies that operate in multiple geographies must educate themselves on each market, which makes standardizing products, materials, or processes more difficult. In many geographies, the public sector mandates lowest-price rules in tenders, making competition based on quality, reliability, or alternative design offerings more complicated. Such codes and regulations reflect the industry's aversion to risk, as structures need to hold up for decades—or even centuries—and any deficiencies can lead to fatal accidents. However, the regulatory landscape has the potential to change if the industry fundamentals are reformed.

In response to these three broad market characteristics, today's construction industry must grapple with several dynamics that add to the industry's complexity, impeding its productivity and making changes more difficult. Those dynamics include the following:

A project-based building approach

Customers want—and seem willing to pay for—bespoke projects with unique features, and structures need to be designed for the natural environments they are in. The challenge is that such projects have a limited degree of repeatability and standardization and require the coordination of companies across the value chain. This dynamic is a main cause of the industry's complexity and low productivity growth. There are obviously differences across asset classes.

A highly fragmented ecosystem

Local market structures and ease of entry have resulted in a fragmented landscape of mostly small companies with limited economies of scale. Moreover, the project-based construction process involves many steps with scattered accountability and a multitude of active entities in every project—from several specialist engineering and planning companies to multiple subcontractors and sub-subcontractors and myriad material suppliers. Since the level of collaboration across the value chain is low, the result is a siloed ecosystem where companies tend to manage their own risk and frictions at the interfaces are high.

Misaligned contractual structures and incentives

The multitude of stakeholders in a project rarely collaborate well because of misaligned incentives. Owners

often tender projects at the lowest cost and pass on risks such as soil properties or rising prices for materials that they might better handle or absorb themselves. Engineers are often paid as a percentage of total construction cost, limiting their desire to apply design-to-cost and design-to-constructability practices. General contractors are often only able to make profits via claims, so rather than highlighting design issues early in a project they often prefer charging for change orders later. Incentives and discounts from distributors and material suppliers to subcontractors obscure material prices.

Extensive use of contractors and temporary staff

Cyclicality and the project-based approach to building create high volatility in activity from day to day. Materials and components can arrive at sites late, and companies must manage restrictions on activities that can be performed only during certain hours of the day. Rather than solving these underlying issues, construction firms rely on temporary staff and subcontractors—which hampers productivity, limits economies of scale, and reduces output quality and customer satisfaction. Furthermore, the need for temporary staff is often solved by contracting foreign workers, which can result in additional language challenges.

All of these market characteristics and industry dynamics are interconnected, so the construction industry must change the underlying fundamentals to meet the challenges ahead. To date, the wider industry response has been to maintain the status quo, which is the root cause of many of the challenges mentioned above that construction companies are facing. However, executives recognize the necessity to adapt. Our survey found 90 percent of respondents strongly believe that the industry needs to change—and that this need has increased over the past ten years (Exhibit 3).

Exhibit 3

Industry leaders expect change.

Survey of 400 industry CxOs across asset classes, geographies, and value chain, share of respondents, %

■ Yes ▨ No

Do you think that there is a need for change in the construction industry setup and the current ways of working?



Do you think that the need for change in the construction industry setup and the current ways of working is higher compared to five to ten years ago?



Source: McKinsey survey of 400 construction industry CxOs, expert interviews, McKinsey analysis



@Getty Images/SerhiBobyk

2 A changing market environment, technological progress, and disruptive new entrants will trigger industry overhaul

The construction industry was starting to experience an unprecedented rate of disruption before the COVID-19 pandemic. In the coming years, fundamental change is likely to be catalyzed by changes in market characteristics, and emerging disruptions will shape the future dynamics of the industry (Exhibit 4). The COVID-19 crisis amplifies these dynamics.

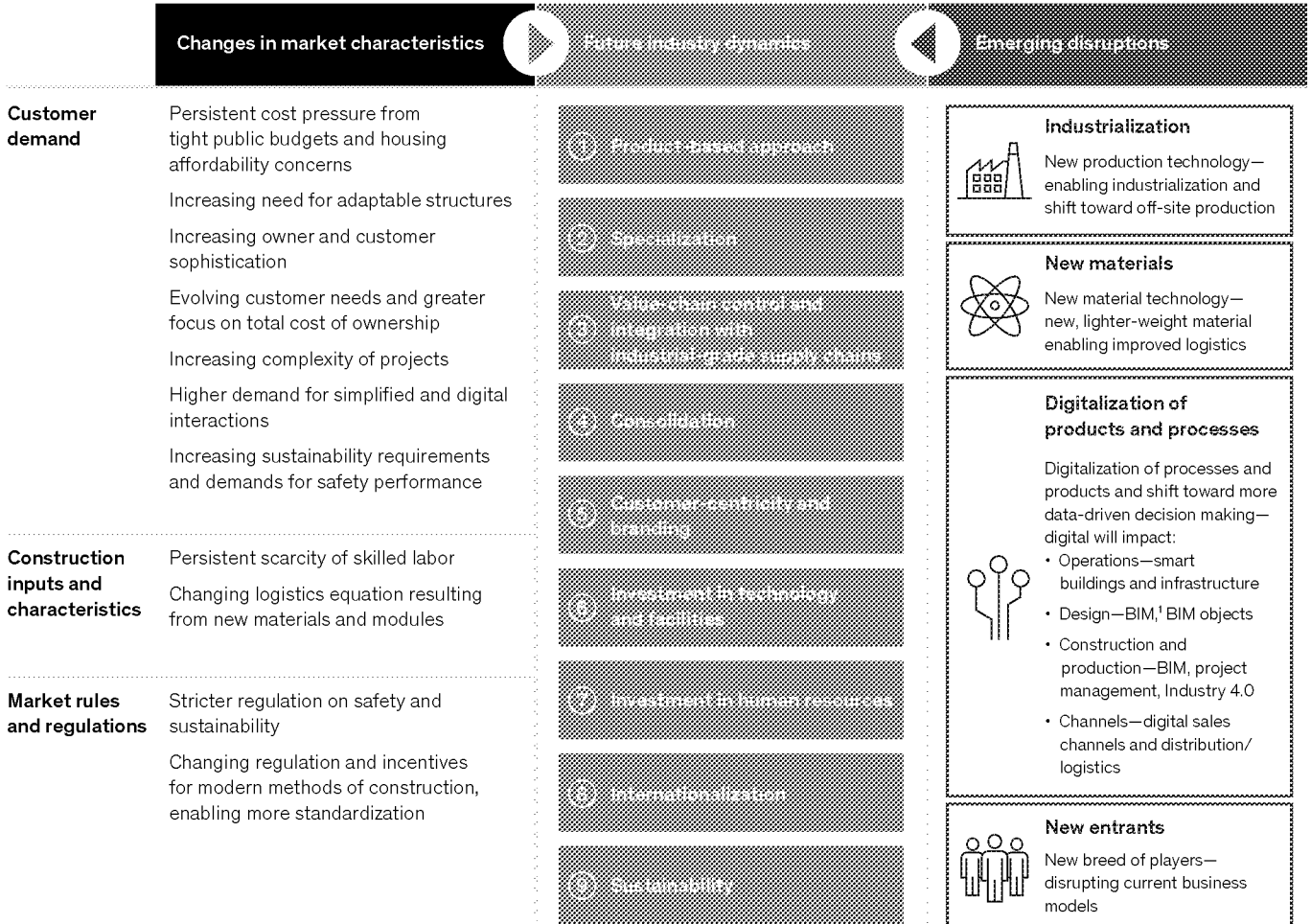
Evolving market characteristics will be a catalyst for change

Throughout the industry, several evolving market characteristics—including the following—will challenge construction companies to highlight both the heightened need for change and the pace at which it must occur.

1. Rising customer sophistication and total-cost-of-ownership (TCO) pressure

The industry has seen an influx of capital from more savvy customers. From 2014 to 2019, for example, private-equity firms raised more than \$388 billion for infrastructure projects, including \$100 billion in 2019

Changing characteristics and emerging disruptions will drive change in the industry and transform ways of working.



¹Building information modeling.

alone, a 24 percent increase from 2018. Institutional investment in multifamily homes as an asset class has also soared. Because these investors will have different expectations of everything from delivery times to budgets, construction companies will have to be prepared for that level of engagement.

Client expectations regarding performance are evolving: smart buildings, energy and operational efficiency, and flexibility of structures will become higher priorities. Our survey found that 69 percent of respondents believe that the transition of customers toward TCO would have a major impact on the construction industry, and more than 90 percent expect this shift to occur in the next ten years.

Our analysis suggests the industry will see increased cost pressure from gaps in infrastructure and affordable housing. In our survey, 74 percent of respondents believe that a greater focus on costs (mainly in

infrastructure and affordable housing) would have a large impact on the industry. This trend will be felt in the near term: more than 90 percent expect to feel its influence within the next ten years.

McKinsey analysis found that \$69.4 trillion in infrastructure investment (at 2018 prices) will be required by 2035 to support expected GDP growth.¹⁶ Investment will need to increase from current levels by 0.3 percentage points of GDP, or an average of approximately \$300 billion a year. Government budgets are tight, however, as debt levels have soared during the pandemic. In real estate, our research found that about a third of the global urban population cannot afford a decent place to live.¹⁷ The supply of housing will need to increase—at lower price points. The COVID-19 crisis magnifies cost and affordability issues.

2. Persistent scarcity of skilled labor and changing logistics equations

The shortage of qualified workers has become a major issue in several markets. Our survey, which was conducted before the COVID-19 outbreak, shows that 87 percent of respondents believe that skilled-labor scarcity will have a high impact on the industry, and almost 50 percent of respondents expect that it will become even more of an issue at scale over the next two decades (Exhibit 5). For example, about 41 percent of the current US construction workforce is expected to retire by 2031. The impact the COVID-19 crisis will have on this dynamic in the long term is unclear at the time of writing.

New, lighter-weight materials and structurally stronger modules will change the logistics equation and allow longer-haul transport of materials and modules, and hence greater centralization.

3. Sustainability and safety regulation and possible standardization of building codes

Our survey revealed that 69 percent of the respondents believe that stricter regulation on work-site sustainability and safety would affect the industry. While 90 percent expect sustainability and safety to make a significant impact in the next ten years, 19 percent believe the shift will occur far more rapidly—potentially within the next year. Because of COVID-19, new health and safety procedures will be required for some time.

The global conversation about climate change, exemplified by the implementation of UN sustainability targets, will compel construction companies and materials suppliers to factor sustainability into their products, construction processes, and designs. The current pace of urbanization will require significant investment in infrastructure and housing to accommodate regional population shifts, highlighting the need for urban sustainability. A 2015 McKinsey analysis found that green districts (densely populated areas located in a city that use technologies and design elements to reduce resource use and pollution) can reduce energy consumption by 20 to 40 percent and freshwater consumption and wastewater production by up to 65 percent.¹⁸

The physical impact of climate change will shape demand (through mass migration and physical mitigation such as flood defense), while construction will come under pressure to mitigate carbon emissions. For example, cement production currently accounts for 8 percent of global carbon emissions, according to the Royal Institute of International Affairs.¹⁹ The rate of new regulations could also accelerate in line with the perceived threat.

Indications are emerging that regulations and incentives are changing to accommodate modern methods of construction, enabling more standardization across the building process. One such method is modular construction. Type certificates can replace individual on-site approvals to alleviate the shift, or governments can actively mandate modern methods of construction. For example, all government housing projects in Singapore must use prefabricated volumetric modules.

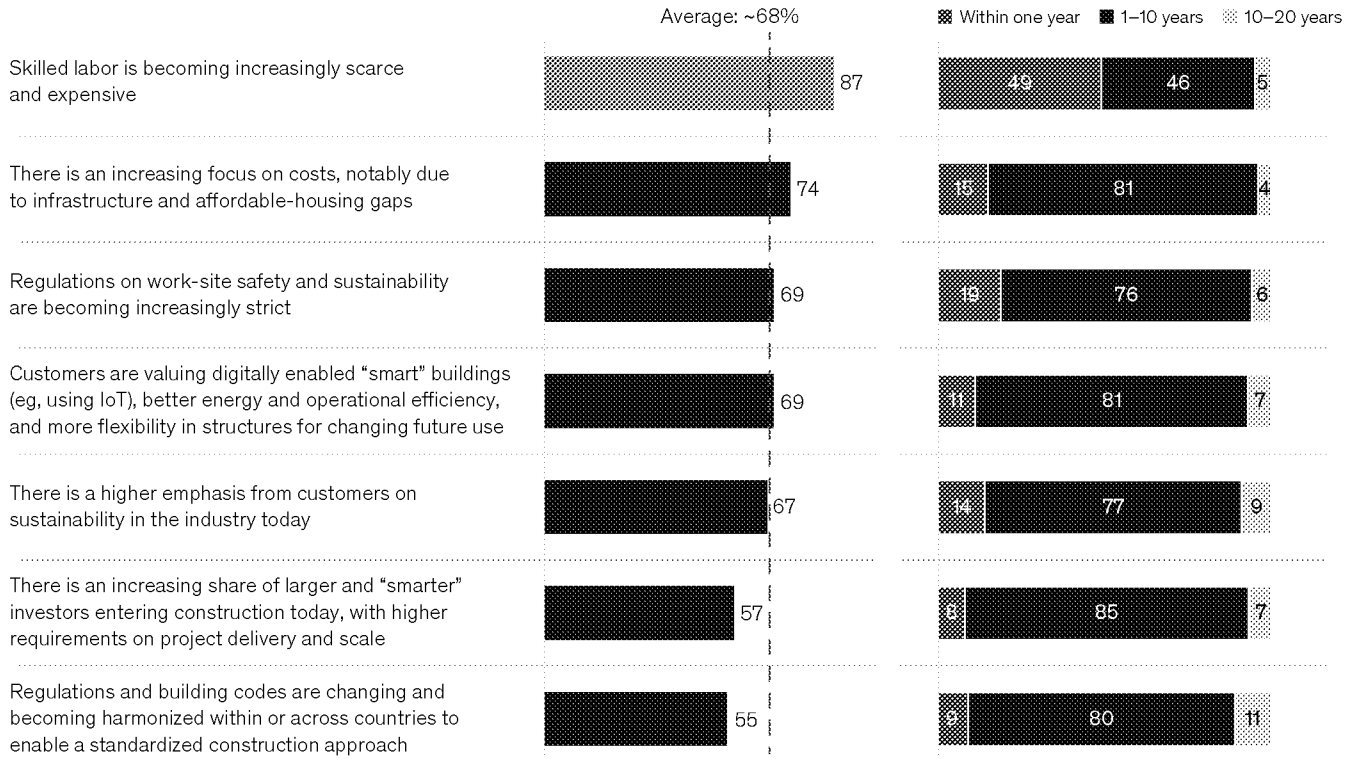
The industry believes that market characteristics will change at scale.

Which of [these changes in market characteristics] do you think will have the highest impact on the construction industry?

Share of respondents rating changes in characteristics “high impact,”¹ %

When do you think [the changes in market characteristics] will impact at scale?

Share of respondents, %



Almost 90% of the industry believe that a further shortage of skilled labor would have a high impact on construction...

... and almost half of respondents think that this will happen within the coming year

¹High impact equals a 7 or higher (where 10 is highest impact). Figures may not sum to 100%, because of rounding.

Source: McKinsey survey of 400 construction industry CxOs; expert interviews; McKinsey analysis

Emerging disruptions will fuel the transformation of the value chain

The demands put on the construction ecosystem have continued to change, but the fundamental business models and approaches of most players have not. The result has been instability in the ecosystem and difficulty meeting growing demand efficiently. In addition, the compounding effects of new production technologies and materials, digitalization of processes and products, and new breeds of players across the value chain have the potential to radically transform the construction industry.

Industrialization

Advancements in technology have enabled industrialization and a shift toward an off-site, product-based approach. A more controlled environment will be even more valuable for the duration of the COVID-19

pandemic. Elements include a production system that uses increasingly modular components, automated manufacturing, and robotics-supported on-site execution.

Modularization has the potential to enable standardization and drastically improve productivity in construction. Prefabrication and off-site manufacturing have been around for a long time in construction, but the combination of lighter-weight materials and digital planning and production technologies could enable the industry to attain new levels of quality, variability, and efficiency. A 2018 survey of UK home builders found that 40 percent of the respondents were already investing in manufacturing facilities or intended to do so in the near future. In Scandinavia, 45 percent of housing is currently built using off-site manufacturing.²⁰

The next step in the transition to efficient off-site manufacturing involves integrating automated production systems—essentially making construction more like automotive manufacturing. Parts of the industry are already moving in this direction. Modular construction firm Lindbäcks uses Randek’s industrial construction machinery to perform manufacturing tasks such as nailing, milling of openings, sheet cutting, and gluing.

On-site execution that uses automation technology could become significantly more efficient. Construction has already started to explore automation in three areas: additive construction (3-D printing), autonomous navigation technology for construction machinery enabled by light detection and ranging (LiDAR, a remote-sensing method for examining the Earth’s surface), and robotics and drone technology.

New materials

Increased use of lighter-weight materials, such as light-gauge steel frames and cross-laminated timber, is reducing costs and allowing longer-haul logistics and more centralization in production and distribution. New types of versatile, lighter-weight, and flexible concrete are self-compacting and have a higher viscosity that eliminates vibration and finishing while enabling single-point pouring and more intricate formwork. Other varieties materially reduce carbon footprint. Alternative materials (both structural and nonstructural) are also being developed: for example, ethylene tetrafluoroethylene (ETFE) is 99 percent lighter, stronger, more eco-friendly, better at transmitting light, and more flexible than glass.

Digitalization of products, design, processes, and channels

Digital technologies are enabling better collaboration, greater control of the value chain, and a shift toward more data-driven decision making. In turn, companies are adopting 5-D building-information modeling (BIM), advanced analytics, and digital-procurement and supply-chain management throughout their organizations.

Smart products and operations. Smart buildings and infrastructure that integrate the Internet of Things (IoT) will increase data availability and enable more efficient operations as well as new business models, such as performance-based contracting. IoT sensors and communication technology give companies the ability to track and monitor utilization, energy efficiency, and maintenance needs. By using BIM, owners and operators can create a virtual 3-D model with precise transparency on all components used in a completed building, which can increase efficiency as well as reduce maintenance costs.

Design. BIM has been used in construction for many years now. Over time, additional features and components (such as scheduling and budgeting) have been added to create a full-scale project-management software solution. However, the industry has struggled to adopt and successfully integrate BIM in operations. By using BIM to create a full 3-D model (a “digital twin”) early in the project rather than finishing design once the project is already initiated, companies can improve efficiency and integrate the

design phase with the rest of the value chain. This capability improves coordination and communication with materials and component suppliers and allows early clash detection and design and planning improvements. In a 2017 McKinsey survey, contractors indicated that their expected use of BIM could increase by 50 percent and that they were planning significant investment to expand their BIM programs until 2020.²¹ Such capabilities will materially change risks in construction projects and put traditional engineering, procurement, and construction (EPC) models into question.

Construction and production processes. Using digital tools can significantly improve on-site collaboration. These solutions include mobile project-management apps and cloud-based project control towers that integrate communication among teams on site and sync with sensors, wearable devices, and desktop machines to constantly track progress and utilization. Advanced analytics can help to further enhance construction efficiency. Industry 4.0²² approaches allow greater flexibility in factories and thus more customization of modules. As in other industries, the COVID-19 pandemic is accelerating the integration of digital tools.

Channels. Digital channels are spreading to construction, with the potential to transform interactions for buying and selling goods across the value chain. Online marketplaces, which have optimized supply chains in other industries, could significantly improve the efficiency of buying and selling goods along the construction value chain and improve interactions—both between customers and suppliers of whole projects as well as among industry participants along the value chain during projects. Digital channels can also radically disrupt distribution and reshape construction logistics. Across the value chain, start-ups have emerged to establish online marketplaces for buying and selling goods such as heavy equipment, construction materials, and professional services. Over the past two years, these marketplaces have received 27 venture-capital funding rounds. This activity represents about 40 percent of the total venture-capital funding rounds to online marketplaces. As in other industries, the COVID-19 pandemic is accelerating the integration of digital tools.

New entrants

Start-ups, incumbent players making new bets, and new funding from venture capital and private equity are disrupting current business models. In addition, a new breed of player, backed by capital from investment funds, is entering the construction industry. Indeed, the number of annual venture-capital funding rounds grew about 30 percent a year from 2012 to 2018. Katerra, for example, raised \$1.2 billion to develop a business model based on new technology to control the value chain, including design and engineering and off-site manufacturing. As the economic crisis unfolds, we also expect an increase in corporate restructuring and M&A activity.

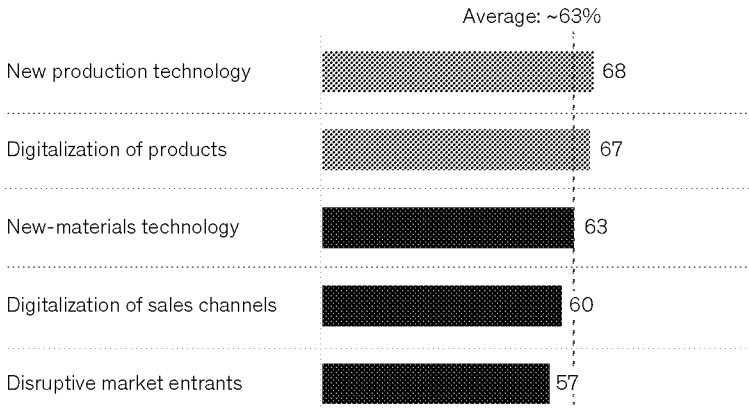
In our survey of industry leaders, a majority of the respondents believe that the disruptions outlined in this report would have a large impact on the industry (Exhibit 6). A significant majority expect the shifts to occur in the next five years and to fuel the industry's transformation journey in the near term. As the COVID-19-propelled economic crisis unfolds, we also expect an increase in corporate restructuring and M&A activity.

Nine shifts will disrupt the construction industry ecosystem

We expect nine shifts to fundamentally change the entire construction ecosystem. According to our executive survey, more than 75 percent of respondents agree that these shifts are likely to occur, and more than 60 percent believe that they are likely to occur within the next five years. The economic fallout from the COVID-19 pandemic looks set to accelerate them.

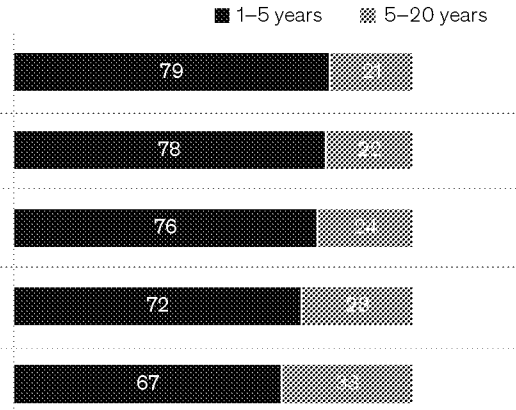
Industry leaders expect disruption to occur.

Which [of these emerging disruptions] do you think will have highest impact on the construction industry? Share of respondents rating that emerging disruptions will have “high impact,”¹ %



More than two-thirds of respondents think that industrialization and digitalization will have the highest impact of the emerging disruptions

When do you think the emerging disruptions will impact construction at scale? Share of respondents, %



More than two-thirds of respondents expect disruptions to impact construction in the near term

¹ High impact equals a 7 or higher, where 10 is highest impact.

Source: McKinsey survey of 400 construction-industry CEOs; expert interviews; McKinsey analysis

A refusal to adapt to this upheaval will only worsen performance, while developing new business and operating models could allow companies to generate more value and profit. The new characteristics will combine to erect higher barriers to entry, which will be positive for companies that manage to strengthen their market position and to invest.

Product-based approach. In the future, a large share of construction projects will be built using customizable, modularized elements and components produced using standardized processes in off-site factories. The modules and elements will be shipped and assembled on site. Production will consist of assembly line–like processes in safe, nonhostile environments with a large degree of repeatability.²³ In addition, common, industry-wide standards for elements and components may emerge. There will likely be a balance between simple elements and components (manufactured according to common, industry-wide standards) and tailored, customizable ones (such as exteriors) to fit bespoke customer needs.

Each player will develop its own design library of elements and components that can be assembled according to customer requirements. A portion of the market might be composed of prefabricated volumetric modules, and customizable, LEGO-like, modularized elements and components could become the industry standard. With this shift, the creativity in designing bespoke products will remain—developers and product manufacturers will collaborate, possibly through strategic alliances, to design products according to unique circumstances.²⁴

The product-based approach will not be confined to construction: developers and owners are also likely to increasingly develop and market branded “products” comprising a standardized structure and contractual terms with integrated service offerings.

Digital technologies will be a critical factor in the shift to a product-based approach. Therefore, companies that own the digital model will be able to control the process without actually owning any factories and to price products based on TCO rather than using today’s simple cost-plus approach.

Specialization. To improve their margins and levels of differentiation, companies will likely increasingly specialize in target niches and segments (such as luxury single-family housing, multistory residential buildings, hospitals, or processing plants) in which they can build a competitive advantage. And they will specialize in using different materials, subsegments, or methods of construction. The shift toward specialization will also require companies to develop and retain knowledge and capabilities to maintain their competitive advantages. Obviously, players will need to carefully weigh the effectiveness, efficiency, and brand positioning that greater specialization enables against the potential risk and cycle-hedging benefits of a more diversified portfolio.

Value-chain control and integration with industrial-grade supply chains. Companies will move to own or control important activities along the value chain, such as design and engineering, select-component manufacturing, supply-chain management, and on-site assembly. Companies will be able to achieve this goal through vertical integration or strategic alliances and partnerships by using collaborative contracting and more closely aligned incentives. Similar to other manufacturing industries, controlling the supply of key components will be critical to securing just-in-time delivery of right-sized inputs to manufacturing as well as the supply of goods to on-site assembly. Digital technology will change the interaction model: BIM models will lead to more decision making early on in the process, distribution will move toward online platforms and logistics management, and end-to-end software platforms will allow companies to better control and integrate value and supply chains. By successfully integrating a five-dimensional BIM model with the value chain, for example, companies will be able to link activities from formulating the initial concept to producing the finished product. Value-chain control or integration will reduce interface frictions and make innovation more agile.

Consolidation. Growing needs for specialization and investments in innovation—including the use of new materials, digitalization, technology and facilities, and human resources—will require significantly larger scale than is common today. In addition, larger and more professional investors will seek more sizable, more sophisticated companies to be their counterparties. As product-based approaches, with a greater amount of standardization and repeatability, further increase the importance of gaining scale, the industry is likely to increasingly see a significant degree of consolidation, both within specific parts of the value chain and across the value chain. Globalization will further increase scale effects as future winning products will be fashionable and in demand across the world.

Customer-centricity and branding. With productization—that is, turning development, engineering, or construction services into easy-to-market products or solutions²⁶—and specialization in the industry, having a compelling brand that represents an organization’s distinctive attributes and values will take on added importance. As in traditional consumer or B2B industries, a strong brand can tie customers more closely to the construction company’s or supplier’s products and help to build and maintain relationships and attract new customers. Similar to brands in other sectors, such a brand will encompass, among other aspects, product and service quality, value, timing of delivery, reliability, service offerings, and warranties.

Investment in technology and facilities. Productization implies a need to build off-site factories, which requires investments in plants, manufacturing machinery and equipment (such as robotics to automate manufacturing), and technology. Where modular is not used, the construction site also will likely become more capital intensive, using advanced automation equipment and drones, among other technologies. R&D investment will become more important for specialized or more productized organizations, so companies are likely to increase spending to develop new, innovative products and technologies. All across the value chain, investment in digitalization will continue to rise.

Investment in human resources. Innovation, digitalization, value-chain control, technology use, and specialization in end-use segments all increase the importance of developing and retaining in-house expertise, which will compel players to invest more in human resources. In addition, the shifts outlined in this report will likely require companies to reskill their workforce. The importance of risk management and some other current capabilities will decrease and be replaced by an emphasis on others, such as supply-chain management. To build the necessary capabilities, companies will need to invest further in their workforces. This becomes even more important in light of the transition to the future of work.²⁶ Most incumbents struggle to attract the digital talent they need and will need to raise excitement about their future business models.

Internationalization. Greater standardization will lower the barriers to operating across geographies. As scale becomes increasingly important to gaining competitive advantages, players will increase their global footprints—especially for low-volume projects in high-value segments such as infrastructure—although the COVID-19 pandemic might slow down this development.

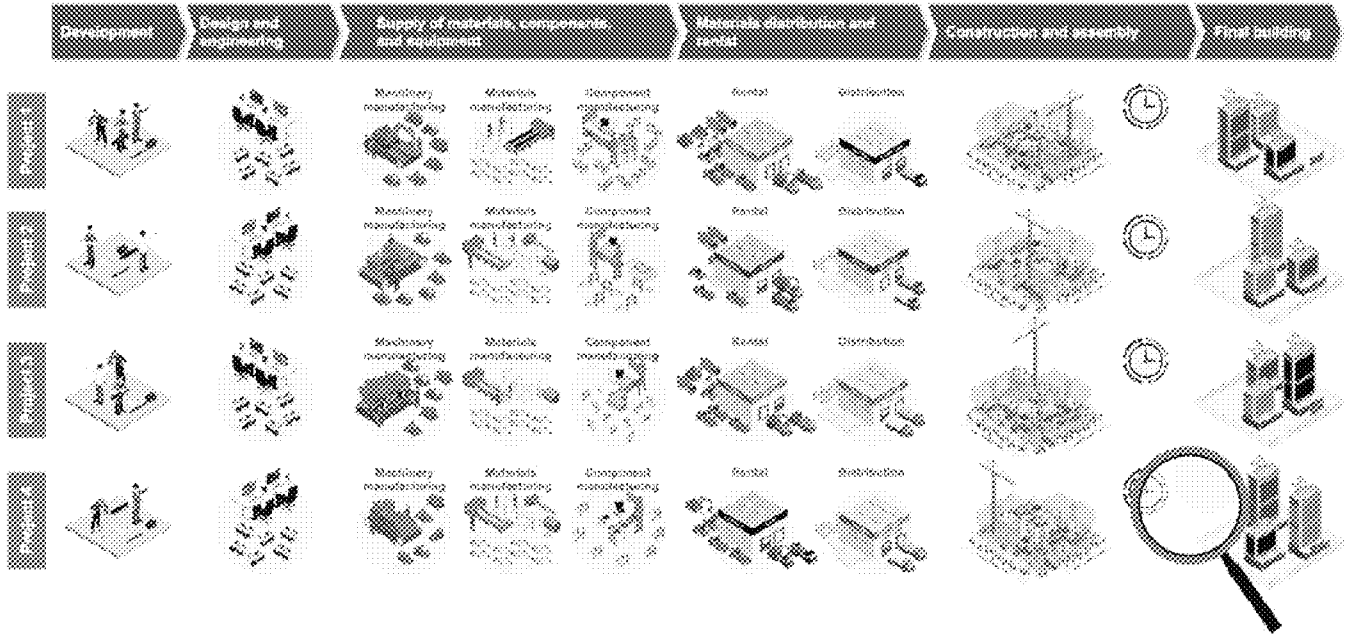
Sustainability. While sustainability is an important decision factor already, we are only at the very beginning of an increasingly rapid development. Beyond the carbon-abatement discussions, physical climate risks grow as the scale increases.²⁷ Companies will need to consider the environmental impact when sourcing materials, manufacturing will become more sustainable (for example, using electric machinery), and supply chains will be optimized for sustainability as well as resilience. In addition, working environments will need to radically change from hostile to nonhostile, making construction safer. Water consumption, dust, noise, and waste are also critical factors.

The construction process is expected to undertake a radical shift toward an industrialized setup by moving from a project- to a product-based approach (Exhibit 7). The current complex and fragmented construction ecosystem will transition to a more standardized, consolidated, and integrated construction process. Not all parts of the construction industry will be equally affected by the shifts. A large proportion of projects will still be unique, low-volume builds carried out in a conventional manner.

According to our survey, more than 75 percent of respondents believe that each of the nine shifts outlined in this section is likely to occur (Exhibit 8), and a majority of those respondents believe that each of the shifts is likely to make an impact on the industry at scale in the next five years. While 75 percent of respondents indicated that the industrialization shifts (product-based approach, technology and facility, human resources, control of the value chain, and customer-centricity) will occur within the next five years, around 40 percent believe shifts around scale (consolidation, internationalization, and specialization) will occur over the next five to 20 years.

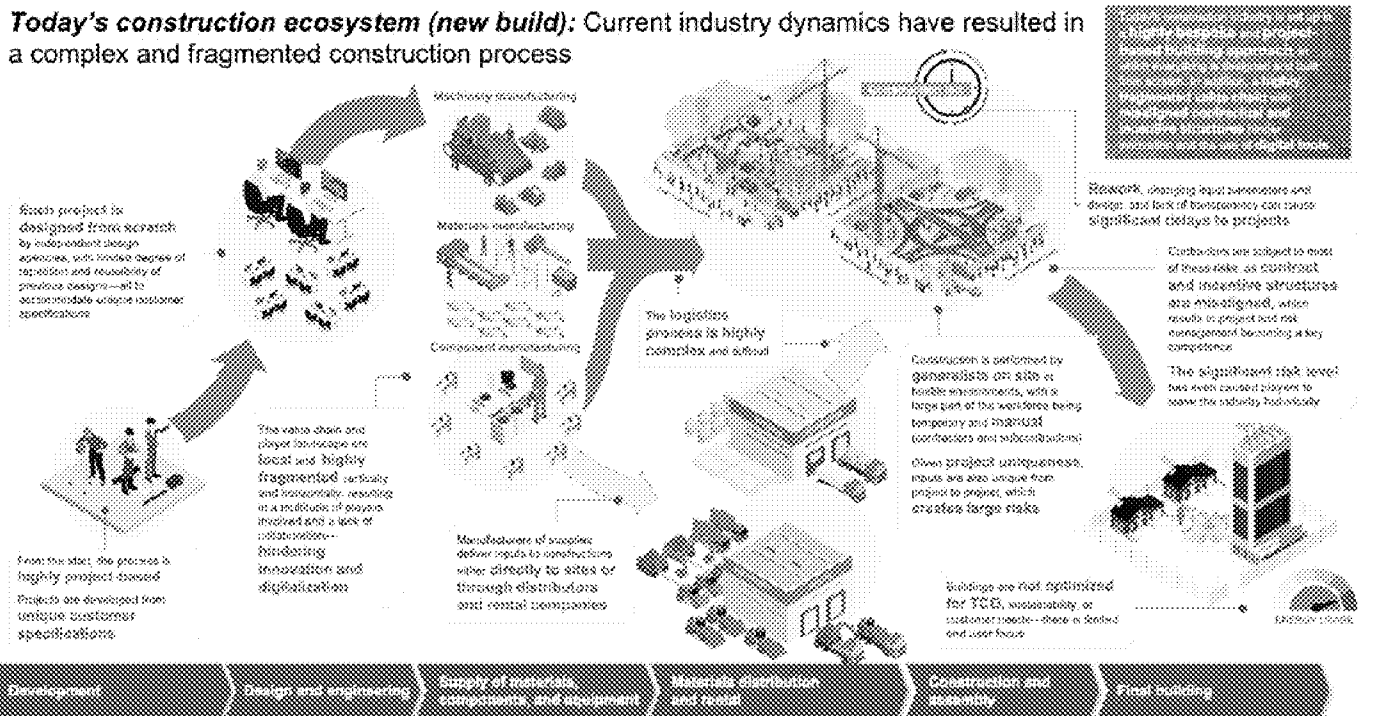
FROM TO

Today's construction ecosystem (new build): A highly complex, fragmented, and project-based construction process

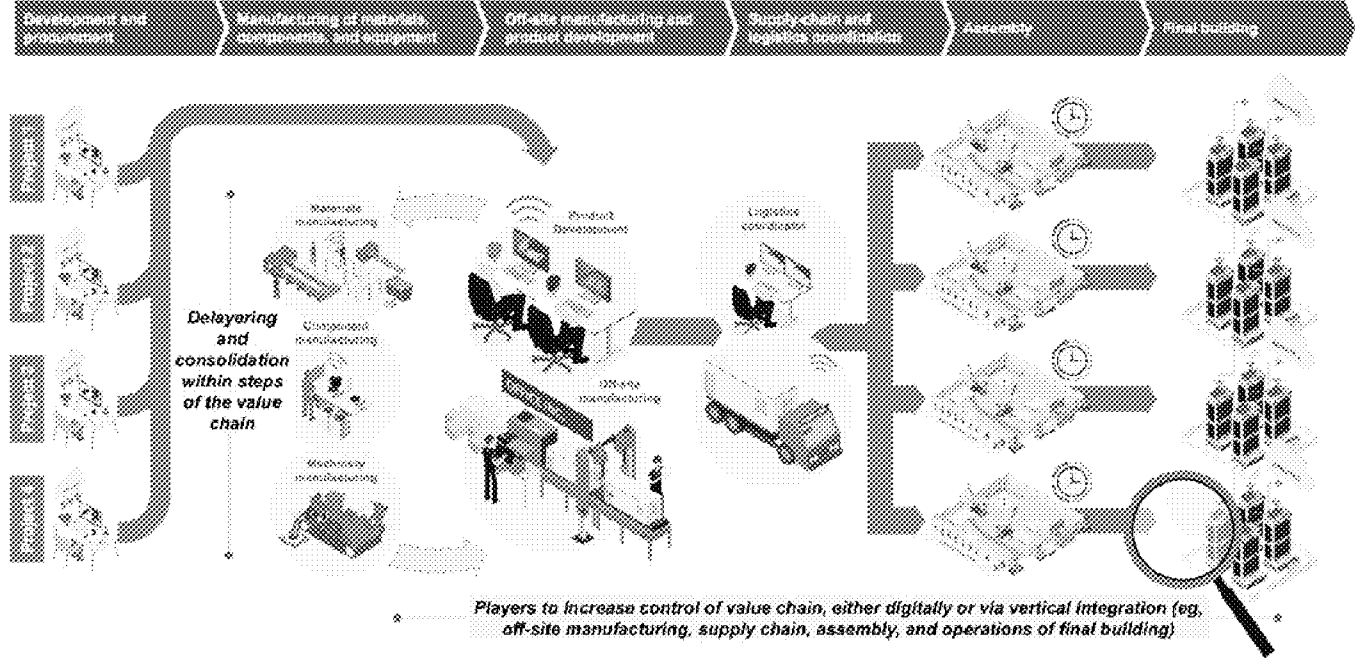


FROM TO

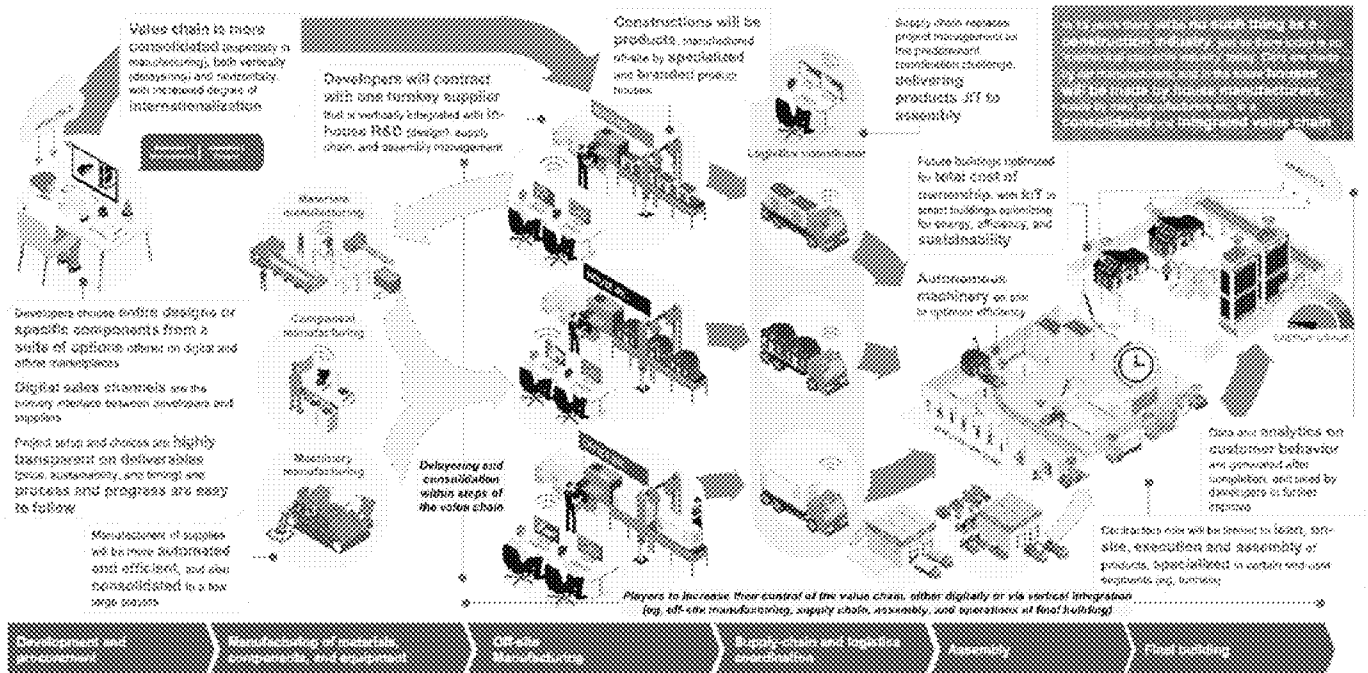
Today's construction ecosystem (new build): Current industry dynamics have resulted in a complex and fragmented construction process



The construction ecosystem of the future (new build): A more standardized, consolidated, and integrated construction process



The construction ecosystem of the future (new build): However, industry shifts will drive a transformation of the construction process



Construction can draw lessons from other industries that have faced disruption

Construction is not the first industry to encounter low productivity and disruption across the value chain. Lessons can be learned from others that had similar traits and encountered the same challenges. We have analyzed shifts in four industries with similar attributes: shipbuilding, commercial aircraft manufacturing, agriculture, and car manufacturing. By studying these industries, clear patterns emerge regarding shifts and changes to the industry value pools. In each case, an eventual transformation was preceded by a set of underlying industry issues (such as low productivity and dissatisfied customers). Innovation in production technology and new best-practice work methods kick-started the journey.

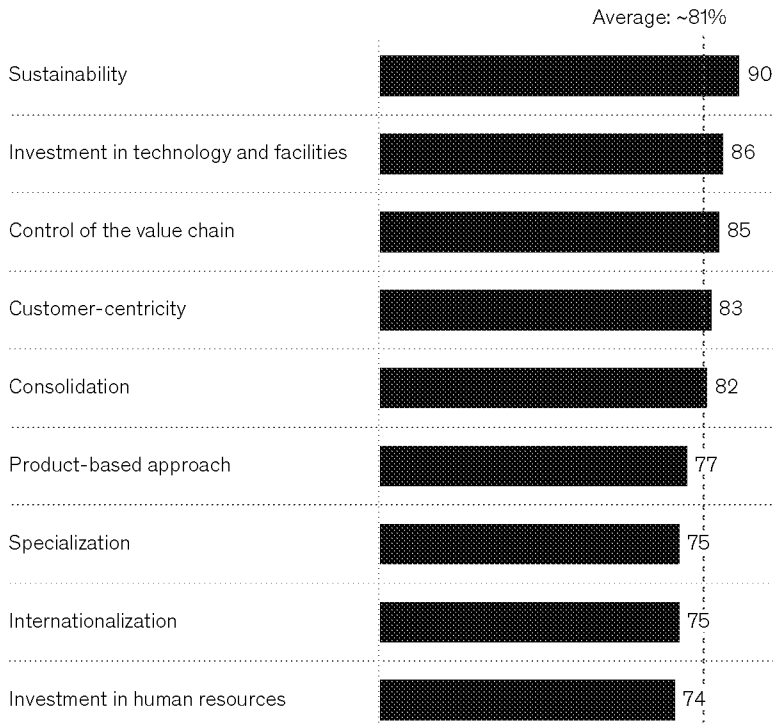
Shipbuilding. A geographically fragmented market structure meant that local shipyards covered the full process. Ships were manually produced in a bespoke and project-based setup over long periods of time and with a limited degree of repetition and standardization.

Exhibit 8

Industry leaders expect shifts to occur in the short term.

How probable do you think [the listed shifts] are to occur?

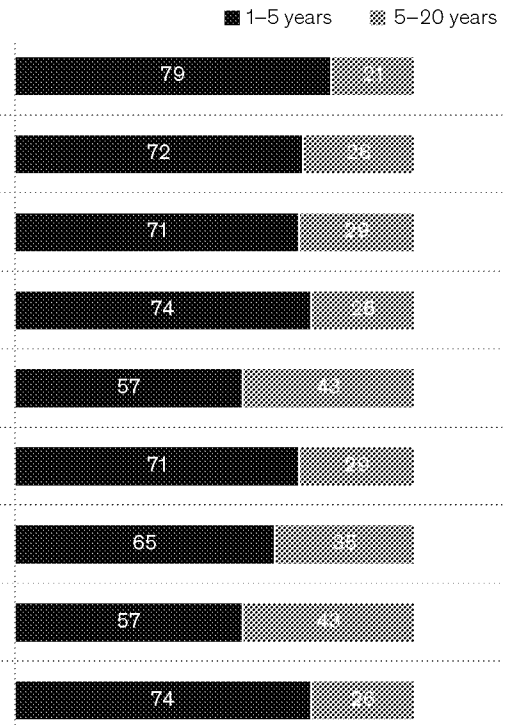
Share of respondents rating shifts as “probable,”¹ %, n = 400



More than 75% of respondents believe that the shifts will probably occur—sustainability shift seen as most likely

When do you think the shifts will impact at scale?

Share of respondents who rated shifts as “probable,”¹ %, n = 370



More than 70% of the respondents who believe that shifts will occur also believe that industrialization will occur in the short term

Note: N = 400 - whereof 88% real estate, 19% infrastructure, 18% industrial; 47% North America, 33% Europe, 1% APAC, 2% Middle East and Africa, 2% Latin America

Probable equals a 5 or higher, where 10 equals the highest certainty that the shift will occur

Source: McKinsey survey of 400 construction-industry CEOs, expert interviews, McKinsey analysis

Commercial aircraft manufacturing. In addition to sharing many of the same characteristics as shipbuilding, commercial aircraft manufacturing required a high degree of specialized trade skills (for example, engineering and physics), and the cost of failure was (and still is) very high.

Agriculture. Each agricultural organization was confined to a specific plot of land, and the amount of manual work in production (for example, seeding and farming) was high.

Car manufacturing. Historically, automakers produced cars one by one with limited use of best practices and standardization. Design and production required a high level of specialized trades such as engineering.

While none of these industries is fully comparable to construction, they shared a number of characteristics: Most also were historically highly fragmented and adopted a largely bespoke and project-based approach, with limited standardization and repetition of processes. Productivity was low, and as demand picked up companies had trouble increasing production. All of them but agriculture also share the high complexity with construction and a high cost of failure. We included agriculture because its geographic dispersion and reliance on land as a key input are similar to construction.

Further, customer satisfaction was often low, caused by long delivery timelines, costly products, and limited assurance on quality. All these industries went through multidecades-long transformation journeys along similar dimensions as the nine shifts we lay out for construction (Exhibit 9).

In commercial aircraft manufacturing, for example, the industry landscape was highly fragmented. Each airplane was built from scratch in a bespoke and project-based-manufacturing setup. Industrialization sparked a shift toward assembly-line manufacturing, which later become highly automated. As a result of the subsequent standardization, the industry entered a phase of consolidation that led to the rise of two major players: Airbus and Boeing. The transformation resulted in a significant shift of value to customers. According to an analysis based on data compiled by Airline Monitor, the realized prices of airlines have been decreasing, on average, at a compound annual growth rate (CAGR) of nearly 2 percent, and today's models have significantly improved safety, TCO, and technology.

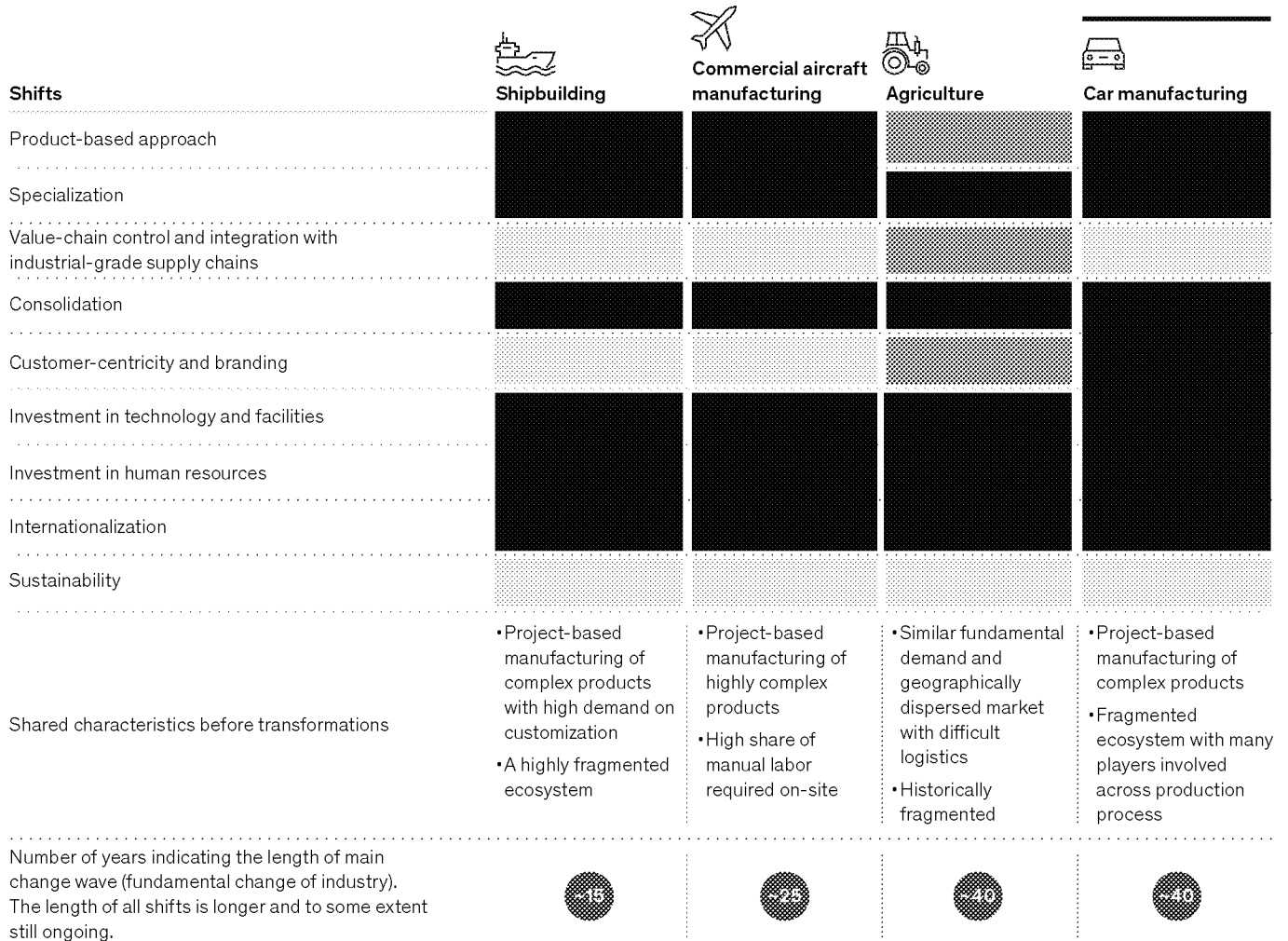
This transformation journey took roughly 30 years to complete, as commercial aircraft manufacturing faced barriers to change similar to those now confronting construction: risks associated with product innovation, relatively strict regulation, often-limited scale of projects on which to apply innovation, and a value chain that requires many stakeholders to be involved and closely aligned. Commercial aircraft was able to navigate these variables and meet the challenges, which should give construction reason to be optimistic today.

Product-based approach. In shipbuilding, commercial aircraft manufacturing, and car manufacturing, players shifted to a product-based approach for which production facilities became assembly sites. The most famous example is Ford's innovation of the assembly-line manufacturing process for its Model T. Most of the auto-manufacturing industry adopted the process within ten years. In this model, prefabricated and modularized subcomponents are inputs, and ships, airplanes, and cars are outputs. While the manufacturing process was significantly standardized, products remained customizable because subcomponents could take various forms and sizes within an industry-wide, standardized framework. When early movers boosted their productivity and profit margins, competitors adopted the innovation over time. Toyota's lean manufacturing and use of robotics, further innovations in the assembly-line manufacturing process, boosted the company from a small player to one of the largest in the industry.

Exhibit 9

The expected shifts in construction have already occurred in other industries that show some (albeit imperfect) similarities.

■ Shift observed in industry ▨ Shift not observed in industry ▩ Shift somewhat observed/ongoing in industry ⊗ Indicative length of main change wave, years



Specialization. As industrialization started to reform these industries and processes became standardized, companies targeted specific niches and segments (for example, tankers, freight ships, and cruise ships in shipbuilding and budget, luxury, and utility autos in car manufacturing). As a result of this specialization, players created a competitive advantage by developing knowledge and scale in their market segment.

Value-chain control and integration with industrial-grade supply chains. As ship, aircraft, and car manufacturing shifted to assembly lines, the supply of critical components was increasingly important. In many cases, those components were the basis of differentiation: in car manufacturing, for example, the quality of the engines could be a distinctive factor. Therefore, it was important to control the supply.

Vertical integration or partnerships along the value chain were common shifts in the industries. In commercial aircraft manufacturing, engines were, and are, produced by external suppliers, but, in order to develop better-quality and more efficient engines than their competitors', manufacturers hold integrated partnerships in R&D and testing. Also, Boeing recently decided to build the 777X wing internally (which was formerly outsourced) and has also set up an internal avionics division to reduce reliance on suppliers of navigation, flight controls, and information systems.

Consolidation. As industrialization emerged in the four industries, companies started to consolidate in order to gain scale. In agriculture, land reforms in combination with industrialization (such as standardized seeding and harvesting) resulted in the industry's transformation from a large set of small and local farms to one dominated by regional and global players. In manufacturing industries, standardization spurred a large wave of consolidation. In commercial aircraft manufacturing, several companies consolidated into Airbus and Boeing. The defense sector also consolidated over the past 50 years, with several large deals made to align companies' services and product portfolios.

Customer-centricity and branding. Following specialization in end-use segments, companies invested heavily to build strong brands within their market niches and segments. In car manufacturing, brands tell stories that are centered on the customers—and customers let the products shape their lifestyles. Given changes in how consumers acquire and use cars, automakers have emphasized their use of technology and innovation to enhance the customer experience.

Investment in technologies and facilities. Industrialization created the need to invest in technology and facilities: manufacturing plants needed to be built, machinery to be acquired. Product and manufacturing innovation became important sources of competitive advantage, which led players to boost R&D spending significantly. In the four comparable industries, greater R&D spending led to short-term gains and advantages for the companies, while customers have benefited over the long term. Consider that the current cost of a car or airplane has changed little in the past ten to 20 years, but both cars and airplanes have significantly more value-adding technologies and other features. The trend has continued with investments by original-equipment manufacturers in the electric-vehicle-battery market—from R&D and packaging to cell production. Volkswagen recently invested in a battery-cell factory that it is developing in partnership with SK Innovation in Germany. It has also struck major supply deals with LG Chem, Samsung, and Chinese battery maker CATL. Overall, the company's ratio of R&D spending to total revenues is now close to 6 percent compared with an average across the construction sector of less than 2 percent. Indeed, Volkswagen alone invested more than \$13 billion in R&D in 2019, the same amount the 25 largest construction and building materials players together spent on R&D, according to the 2019 EU Industrial R&D Investment Scoreboard. And although that level of R&D spending may converge to the current automotive-sector average of almost 5 percent, it would still represent a significantly higher commitment to R&D than is typical in construction.²⁸ In sum, across industries winners continue to heavily invest in technology, many with a focus on digitalization and data-driven products and services.

Investment in human resources. Employee attraction and retention became a priority when industrialization affected the four comparable industries at scale. First, players built up their technical knowledge in order to create a competitive advantage. Second, improved production processes have, over time, resulted in a need for constant retraining of the workforce.

Internationalization. Industrialization ushered in the standardization of processes, which was adopted across geographies. Internationalization enabled companies to expand beyond their borders in pursuit

of scale, gave them access to new markets, and resulted in operations cost savings. For example, in an attempt to increase commercial aircraft sales in China and the Mideast, Airbus and Boeing set up local final-assembly lines in China.

Sustainability. The growing global emphasis on sustainability is being felt across industries. Most notably, automotive has already embarked on a material transformation toward zero-emission vehicles. In Norway, airport operator Avinor and Widerøe Airlines vowed to fully electrify all domestic flights by 2040.

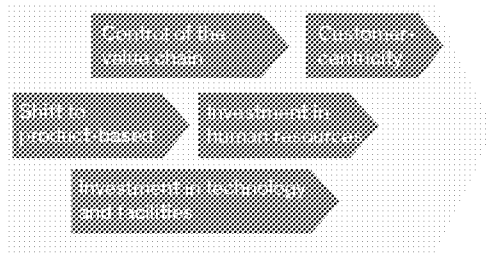
Exhibit 10

The construction industry expects sequencing of shifts similar to comparable industries.

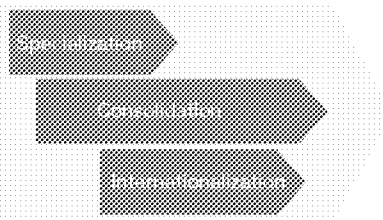
Transformation journeys in comparable industries have typically followed the same pattern

Length of phases highly indicative

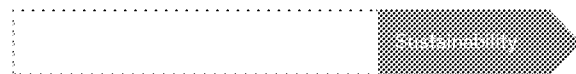
First main change wave—industrialization, 20–25 years



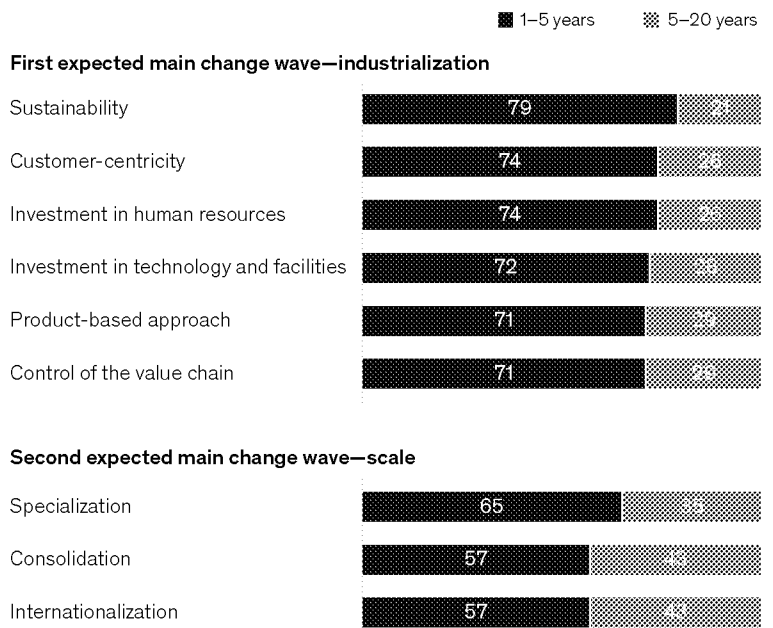
Second main change wave—scale, 10–15 years



Sustainability shift to occur sooner in the transformation of the construction industry



In construction, industry practitioners expect a similar transformation journey, share of respondents, %



Source: McKinsey survey of 400 construction-industry CEOs; expert interviews; McKinsey analysis

Across industries, winners continue to heavily invest in technology, many with a focus on digitalization and data-driven products and services.

The transformation journeys of the comparable industries took decades to complete (Exhibit 10). Survey respondents indicate that shifts in construction will occur in two main waves, similar to the transformation journey observed in those industries. In the first wave, industrialization will standardize processes and increase sector productivity. The second wave will focus on scale, where players will specialize in end-use segments, consolidate vertically in the value chain, and also expand internationally. By drawing on these lessons, construction companies can begin to position themselves for the coming upheaval.



@Getty Images/SJasmin Merdan

3 Almost half of incumbent value added is at stake

The transformation of the industry will create both large opportunities and sizable risks as value and profit pools shift in the next 15 years. Over the past years, approximately \$11 trillion in value added and \$1.5 trillion in profits have been unevenly distributed along the construction value chain and across all asset classes. Looking ahead, up to 45 percent of incumbent value may be at stake in those parts of the market most heavily affected by shifts, such as hotel construction. Of this total, 20 to 30 percentage points will be kept and redistributed within the ecosystem to enable the shifts to take place. The remaining 15 to 20 percentage points will be value up for grabs as a result of the cost savings and productivity gains generated by the shifts, with the benefits accruing to players or customers (in the form of price reductions or quality increase). If that value is captured fully by players in the ecosystem, profitability could nearly double, to 10 percent of revenues, from the current 5 percent.²⁹ Players that move fast and manage to radically outperform their competitors could grab the lion's share of the \$265 billion in new profit pools. (For more on our methodology, see sidebar "How we measure value and profit pools.")

Value is distributed unevenly along the construction value chain

Value and profit pools have remained stable, with only minor changes, for a long period, as the overall industry has maintained its status quo. While a multitude of players are active within and across the construction value chain, few manage to grab significant shares of value—not to mention profits (Exhibit 11).

For example, software providers and off-site manufacturers often command high earnings before interest and taxes (EBIT) but remain relegated to niche pockets in the value chain. In contrast, developers manage to capture a large share of the value pool. General contractors and subcontractors typically have low margins (especially considering the high risks of the industry). Given their sheer number, however, they still collectively attract a significant share of overall value. Top generalist contractors have revenues of about \$60 billion, while average revenues for contractors in our database are much lower, at \$27 million. The value declines to just \$10 million for specialist contractors and declines further when the long tail of individual owner-operators and small firms (which is not included in our database) is included (Exhibit 12).

Materials distribution and logistics still represents a relatively high share of both value added and profits, as it plays a central role in connecting a large number of suppliers with project sites on which subcontractors are active on each site. The best-performing basic-materials providers manage to attain the typical scale of manufacturers and achieve EBIT margins of 15 to 25 percent, but a long tail of less-profitable players brings down average margins.

Minor differences exist across regions

The differences among regions are somewhat limited (Exhibit 13). However, a few things should be taken into consideration, such as whether contractors specialize, to what degree materials distributors are able to add

How we measure value and profit pools

We analyzed financial data on more than 240,000 global companies in the Capital IQ database, in an effort to size value and profit pools across the ecosystem. We then built a database that categorizes industry players according to their principal activities, using the Standard Industrial Classification (SIC) system. Revenue pools were determined by adding up revenue, and profit pools were determined by earnings before interest and taxes (EBIT). Averages for the 2015–17 period were then used to create a consistent view over time.

Value added is defined as revenue excluding externally procured cost. This approach calculates the value added by each step (and player type) in the value chain and helps avoid “double counting” what has been procured from previous steps in the chain or players in the same

step. Drawing on a sample of 10,000 companies where data was available as well as expert interviews, we determined the typical ratio of value added per step in the value chain and applied these calculations to companies for which exact data were not available. Individual player type and size were considered when approximating value added.

The resulting data set covered three regions and six countries: Europe (mainly Spain, Sweden, and the United Kingdom); Asia (China and Japan); and North America (the United States). Individual perspectives were developed for each country and were then aggregated to create regional and global perspectives. Together, these three regions account for approximately 90 percent of global spending on construction, while the group of countries

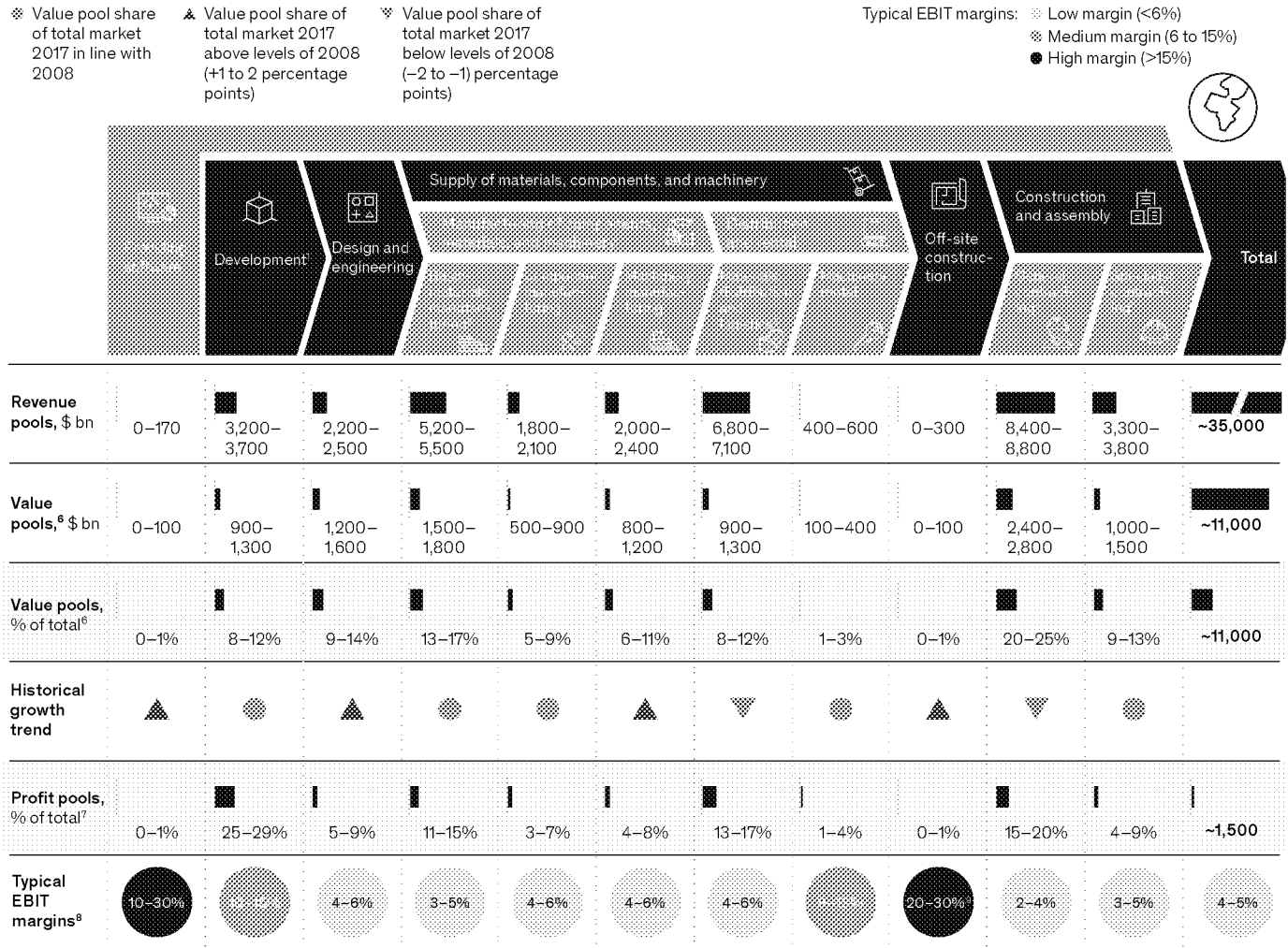
adds up to approximately 65 percent of global spending.

Data quality and coverage varied by country. Adjustments were made to account for gaps in data. On average, the data coverage included in the data set was estimated at 60 to 65 percent. In the United States, for example, EBIT coverage was particularly poor. Therefore, EBIT margins were based on what was available in the sample and then augmented using other sources, including research on annual reports and other publicly available information.

Finally, adjustments were made to remove revenue that did not contribute to industry output. For example, figures for materials distribution and logistics companies don't include revenues from the distribution of white goods.

Value pools are fragmented across the value chain and profitability levels are low.

Value and profit pools per player type in ecosystem (new build and renovation); average, 2015–17



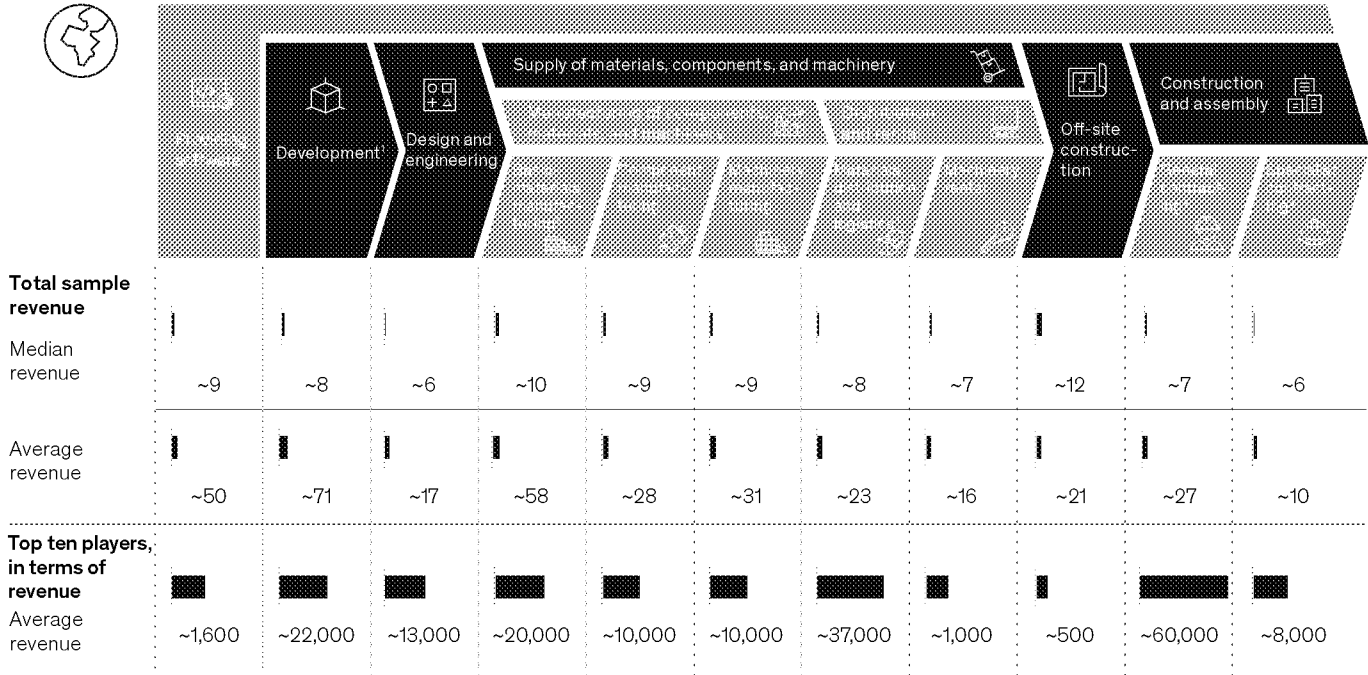
value, and to what extent players rent machinery and tools. Moreover, a country's tendency to import goods or add value domestically also affects the distribution of value added.

The following takeaways highlight some of the differences by region:

- In China, general contracting has higher value added since contractors tend to perform specialist work in an integrated manner.
- UK value pools are skewed toward materials distribution because the United Kingdom is a more indirect market than some other countries. In the United States, however, the massive size of the domestic market allows for distributors to reach larger scale and therefore create more value.

The construction value chain is fragmented with a significant number of small-scale players.

Refers to average for 2015–17, \$ mn



¹ Calculated by applying an assessed share of total value of development of output per asset class, allocated on top of total market output, since a limited number of stand-alone, pure-player developers have been identified.
² Looking at players processing raw materials but not the actual manufacturing of raw materials (eg, mining). If all steps of producing and refining raw materials were included, the value pool would be ~2.5x bigger.
³ Adjusted downward to reflect that some things materials distributors sell don't contribute to construction output (eg, clothes, white goods).
⁴ General builders (buildings and other heavy construction).
⁵ Specialized trade construction.

Source: CapitalIQ; Euroconstruct; FMI; McKinsey analysis

— The US rental market is fairly sophisticated, and companies tend to achieve significant scale as a result of the sheer size of the domestic market.

Almost half of submarket value-added pools could be reshuffled

As the industry transforms, segments in which the nine shifts have the highest potential to materialize (such as hotels or single- or multifamily housing in new real-estate projects) could see a reshuffling of 40 to 45 percent of value added in the next 15 years (Exhibit 14). Of this total, 20 to 30 percentage points will be kept and distributed within the ecosystem to enable the shifts to take place. The remaining 15 to 20 percentage points will be value up for grabs as a result of the cost savings and productivity gains generated by the shifts, with the benefits accruing to companies as profits, workers as wage rises, or customers in the form of better quality or price reductions. As competition catches up with early movers, companies will lower their prices to win individual projects and pass more of the value on to customers—a pattern that has been observed in other industries.

There are minor regional differences in value added per chain step among regions and countries.

Compared to global results



▲ Somewhat higher ▼ Somewhat lower ● Similar



¹ Calculated by applying an increased share of total value of development of output per asset class, allocated on top of total market output, since a limited number of stand-alone, pure-player developers have been identified.
² Looking at players processing raw materials but not the actual manufacturing of raw materials (eg. mining). If all steps of producing and refining raw materials were included, the value pool would be +2.5% bigger.
³ Adjusted downward to reflect that some things materials distributors sell don't contribute to construction output (eg. clothes, white goods).
⁴ General builders (buildings and other heavy construction)
⁵ Specialized trade construction.

The rate of change will differ across the construction value chain

Some players will be more affected than others. For example, software providers are expected to significantly increase their value-added contribution, albeit from a small base of 1 to 2 percent of the value chain. Also, a large share of value is expected to move from construction jobsites to off-site prefabrication facilities. In contrast, general and specialized contractors could face a large decline unless they reposition themselves as companies that go beyond execution alone. Basic design and engineering and materials distribution and logistics may face substantial commoditization risks.

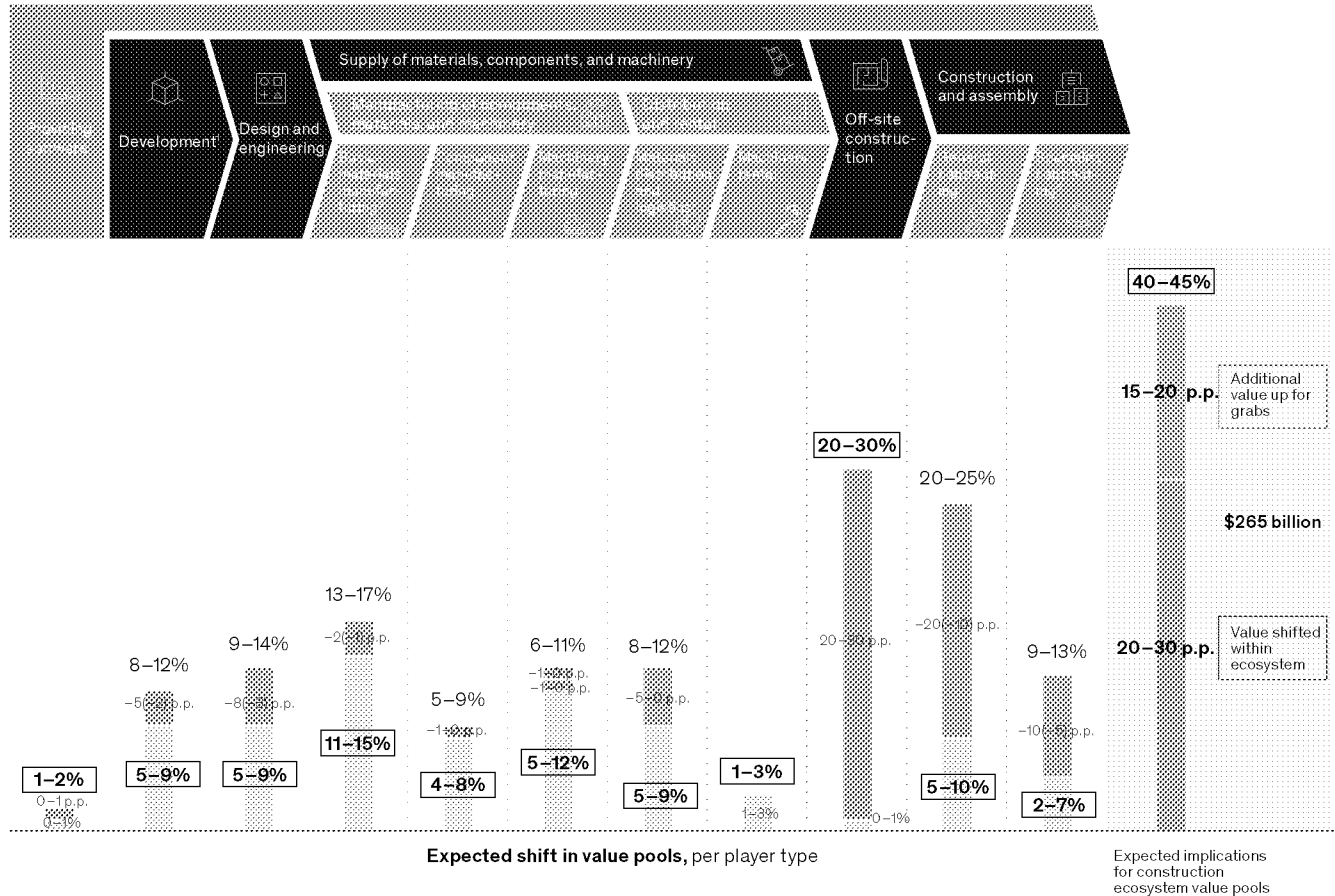
Our estimates are based on expert interviews and analysis. However, the rate of change in the industry could play out slower or faster, depending on overall dynamics and adoption rates.

Exhibit 14

Forty to 45 percent of value pools are expected to shift and impact all players along the value chain.

Example of fully productized value chain (eg, real estate new build), current and future value pools, p.p.

Value at risk Remaining value added Value shifted Value captured



Value pools, ⁶ \$, bn											
0-100	900-1,300	1,200-1,600	1,500-1,800	500-900	800-1,200	900-1,300	100-400	0-100	2,400-2,800	1,000-1,500	~11,000
Profit pools, ⁷ \$, bn											
0-30	350-450	90-110	150-180	70-90	70-90	150-250	30-50	0-60	250-350	80-100	~1,500

¹ Calculated by applying an assessed share of total value of development of output per asset class, allocated on top of total market output, since a limited number of stand-alone, pure-player developers have been identified.
² Looking at players processing raw materials but not the actual manufacturing of raw materials (eg, mining); if all steps of producing and refining raw materials were included, the value pool would be ~2.5x bigger.
³ Adjusted downward to reflect that some things materials distributors sell don't contribute to construction output (eg, clothes, white goods).
⁴ General builders (buildings and other heavy construction).
⁵ Specialized trade construction.
⁶ Defined as value added per player type.
⁷ EBIT pools

Source: CapitaIQ; Euroconstruct; Fitch; McKinsey analysis

Providing software. Software and platforms increasingly will be built to integrate and serve companies throughout the ecosystem. The industry will see the launch of analytics services and software, while online marketplaces will offer entire designs or specific components from a suite of options. Emerging digital sales channels will become the primary interface between builders and suppliers.

Software providers are expected to be strongly affected by changes in the ecosystem, and their value added will increase accordingly. Software will enable cost savings across the value chain—and grab significant share of the gains as a result. Overall software usage is expected to increase as technology advances—for example, through more sophisticated data analysis and increased IoT connectivity. According to IDC, IT-related software and infrastructure will grow 5.3 percent and 1.7 percent, respectively, despite the COVID-19 impact. Some estimates indicate it could even double as the construction industry starts to catch up with the manufacturing industry in terms of IT spending as a share of revenue, which is currently at two to three percentage points.³⁰ Therefore, the associated value pools should rise considerably as well.

When taking into account the doubling of spending on software, the value-pool increase of software providers could rise by one percentage point.

Development. Public- and private-sector developers of real-estate, infrastructure, and industrial projects orchestrate the development process from beginning to end: securing financing, sourcing land, and scoping and overseeing value-adding projects. As customer expectations continue to evolve rapidly, developers look set to increasingly specialize and invest in productizing and branding their offerings, which will increasingly require multinational scale, leaving behind those working in traditional ways. A deep understanding of customer needs, by segment and subsegment, will be increasingly important.

The most successful players are already bringing together these customer insights with product and supply-chain innovation to deliver high-performing projects. For example, leading commercial real-estate developers are already building direct relationships with end users and creating spaces that can be rapidly repurposed as demand changes. Similarly, highway developers are connecting directly with drivers to better understand usage patterns to inform future highway design and operation. In many ways, developers set the tone for the whole industry.

Since developers sit at the top of the value chain, they can strongly influence how and how fast disruption in other parts of the value chain takes place, including actively embracing industrial production of their offerings as well as library-based designs and subsystems.

Such approaches can reduce the cost, time, and riskiness of projects—all improvements in the financial viability of projects—which may translate into higher profits, greater volume, or value shifting to customers. Specifically, significantly shortened project duration will be the key driver for cost reduction in development as all indirect and financing costs will be reduced. We expect two to five percentage points of value added to be at stake.

Design and engineering. Disruption could fundamentally change what it means to be an engineer or an architect in construction. Historically, these professionals have applied their considerable expertise to create designs and specifications for individual projects: each design optimized to meet the project's unique requirements. The coming years will see these stand-alone professional-services firms closely collaborating with productized and branded developers, off-site construction firms, and highly specialized contractors as an integrated R&D-like function. These firms will increasingly add value through the standardization

of structure and subsystem designs and develop standardized design libraries of products in their target segments. This modular design will be reused across a large set of construction projects. In this way, design and engineering firms could influence industry standards. As the industry shifts to a more product-based approach, the challenge for engineering and architecture firms will be to retrain their existing workforces and hire the right talent.

Of course, modularization and automation will not apply to all projects: highly architectural or complex projects will have limited amounts of standardization. Similarly, renovation-maintenance-improvement (RMI) projects are likely to continue to follow more traditional design approaches for some time. In affected segments, however, design and engineering firms are likely to improve their efficiency by using standardized products and libraries as well as software-based design automation—though the extent will depend on how the transformation plays out. The best performers will still stand to gain. There will often be a premium for modular or customizable product design, as only a few firms have the required capabilities and experience. But as the industry adjusts to designs that can be replicated and adapted multiple times, the volume of work is likely to decrease in the affected segments. The need for redesign is also expected to decrease drastically with more specialized and productized approaches, generating significant cost savings for design and engineering and putting the respective value of activities at stake.

Currently, design and redesign account for 14 percent of total value added. The developments discussed here could lower cost by three to eight percentage points—and shift it to the best performers or other parts of the value chain.

Basic-materials manufacturing. A large share of the inputs used in construction projects involve processing raw materials such as cement, steel, wood, or glass. Many players in this sector are already large, global firms with slower-moving shifts in value-chain dynamics. The most pronounced impact might arise from a transition to new, lighter-weight materials, as well as satisfying a growing number of sustainability requirements including less waste and more recycling. While the industry should benefit from long-term (post-cycle) growth of the construction market, the volume of traditional materials per structure, such as the amount of cement per building, looks set to decline. Digitization and consolidation of the distribution and contracting landscape may alter logistics and customer interfaces.

Overall, we expect that about one to two percentage points in value generated in this sector might be at stake.

Component manufacturing. Today, components such as elevators, HVAC equipment, and pipes are often produced using a silo-like approach, which limits the ability to integrate the components once they have been installed in buildings. Installers tend to have strong affinities for specific suppliers—due to either personal experience or suppliers' incentive schemes—while the brand affiliation of end customers is not as strong.

Several of the nine shifts will most strongly affect component manufacturers: productization and standardization (including through BIM object libraries), as well as online channels, will increase price transparency and lead to commoditization. Consolidation will improve the bargaining power of large contractors or modular construction firms, and internationalization will lead to low-cost-country sourcing. In turn, the best companies will achieve further economies of scale and offer solutions with a TCO advantage and value-adding digital services. The companies might also shift from manufacturing components to entire subsystems and taking direct-sales approaches.

Since developers sit at the top of the value chain, they can strongly influence how and how fast disruption in other parts of the value chain takes place.

Overall, we expect that about one percentage point of value added from component manufacturing will be grabbed by other parts for the value chain or the best-performing players.

Machinery manufacturing. Over the next two decades, manufacturers will transition from producing traditional heavy machinery and tools used in the construction process to highly automated, connected products used in the ecosystem. The new equipment will be integrated with robotics that could be used in the ecosystem's manufacturing processes—for example, in plants for building materials, components, and buildings. Rather than simply selling products, manufacturers will offer services that are completed with their products. This will increase the value added of machinery, and increasing automation will support greater volume. In addition, demand for machinery used in off-site fabrication is expected to rise.

In turn, improved efficiency and reduced time requirements for on-site work are expected to have a negative impact on the volume of machines sold.

Overall, we estimate a range of an increase or decrease by one percentage point in value added for machinery manufacturing.

Materials distribution and logistics. Distributors procure, store, and transport basic materials, components, and equipment and resell them to consumers and businesses. Some distributors also provide credit. Part of this model is the organization of logistics and inventory, primarily for construction sites and installers.

Several of the nine shifts may hit distributors in a negative way. Productization, standardization, and consolidation will move decisions and procurement upstream from small specialized subcontractors to large contractors and product-based developers, increasing bargaining power and reducing the breadth of materials needed. Better and earlier planning using BIM and digital twins will reinforce those shifts and reduce the need for local stock. Off-site manufacturing facilities will shift demand for shipments to factory hubs, with more predictable levels of demand, which will be the main logistics nodes and decrease the need for a dense storage network close to construction sites, while also raising the expectation of just-in-time delivery. Internationalization will enable more sourcing from low-cost countries. Online and direct-sales

channels, including new competition from online distribution behemoths—which serve customers with high expectations and use increasing amounts of technology, such as advanced analytics or automated warehouses—will further reshape this segment.

While these shifts pose a major threat to distributors that lack scale and logistics capabilities, they provide opportunities to companies that have them. There is an opportunity to consolidate the sector, supported by lean efficiencies, category reviews, and new business solutions. Direct-to-customer digital interaction channels and interfaces connecting to BIM and building-management systems allow better integration into the value chain. Distributors can fill the roles of the logistics hubs of the future construction landscape by using advanced analytics in logistics, demand forecasting, and inventory management to allow just-in-time delivery from suppliers to modular-construction factories to construction sites. Distributors can create new value for customers by helping with international sourcing, by offering credit finance, packing in assembly order, offering in-room delivery, making deliveries before the working day, providing on-site logistics planning and operations, or even handling simple pre-assembly.³¹

Overall, we expect zero to five percentage points of value added could be at stake and could move to other parts of the value chain or to the best-performing companies that are fastest to adjust to the new world.

Machinery rental. Currently, rental businesses provide companies in the ecosystem with an efficient capital-expenditure option for yellow machinery and tools, which help to optimize equipment utilization. Digital technologies will enable greater efficiency through on-site logistics, either providing the right machinery at the right time or operating and using the machinery for customers. By offering digital services, rental companies gather usage data from products to enable best-in-class employment of machinery for customers.

Some machinery—particularly tools—is expected to be moved off site, trending toward higher utilization and limited rentals rather than direct purchases. This shift will affect machinery-rental players negatively, although only to a small extent. In turn, service-based business models and IoT-connected machinery and tools could compensate for these shifts. Machinery for groundwork will likely be less affected by the upcoming shifts.

Overall, we do not expect a significant net change in value into or out of the machinery-rental sector.

Off-site construction. Today, off-site construction companies primarily manufacture building elements, structures, or modules for real estate—but also for industrial structures and infrastructure like bridge segments. Overall, off-site construction is still a relatively young and immature part of the larger construction ecosystem, with high fragmentation and smaller-scale players that use mostly manual labor.

All nine of the shifts described in this report are expected to positively shape future demand for off-site construction in one way or another, and the collective effect is expected to be the most significant throughout the ecosystem. Led by a product-based approach, standardization, and sustainability, the coming years will see a shift to manufacturing a broad range of products off site, typically on a manual or automated production line.

Off-site construction will see much deeper integration even in flat-pack design such as doors, windows, and fully preinstalled mechanical, electrical, and plumbing systems. Data will be flexibly transferred from BIM models to the automated factory controls for customization within standard designs. The level of integration

and connector technology aims to require no skilled labor on the final site and enable extremely fast building times. There will be a mix of flat-pack (2-D) and volumetric (3-D) preconstruction.

Through close collaboration with designers or integrated R&D departments, companies will build standardized libraries of subsystems that allow mass customization. We expect off-site construction companies to specialize by end-user segments, such as hospitals or certain types of bridges. Branding will also become important to differentiate offerings.

While off-site construction is expected to enjoy strong demand growth, owner expectations and requirements will increase. As a consequence, the future landscape is likely to look very different from today's. Players that can differentiate at scale will stand head and shoulders above others. Some of those players might already exist in today's ecosystem, while others might arise from new entrants that see opportunities in areas such as real estate or infrastructure being the next platform for the deployment of smart technology. Either way, future winners will look very different from the fragmented players that exist today.

Our analysis suggests that, as a direct result of the shift to an industrialized approach, the off-site construction sector could gain a share in value added of 20 to 30 percentage points in the most affected segments. This amount represents the cost of off-site labor (typically 10 to 15 percent of a modular project) and the cost of investing in and operating the factory. Recently constructed facilities suggest an initial capital outlay of \$50 million to \$100 million.

General and specialist contracting. With a focus on managing overall construction projects, general contractors perform construction work, coordinate subcontractors and suppliers, and handle risk. Many of them currently have a limited degree of specialization. Specialist contractors perform specialized tasks in a construction project, such as mechanical, electrical, and plumbing work—but often do so for a broad spectrum of projects.

The generalist and specialist contracting sector faces some of the most imminent threats to their business model as work shifts from jobsites to off-site facilities. Up to 80 percent of the traditional labor activity in a modular building project can be moved off site to the manufacturing facility. Some of the most skill-intensive and expensive types of work, such as mechanical, electrical, and plumbing, can be handled by lower-cost manufacturing workers.

Better and earlier specification with BIM and other digital tools, as well as specialization and productization, will further reduce risks and project-management needs and simplify supply-chain management.

Overall, in the most affected segments, we expect ten to 20 percentage points of value added in general contracting and an additional five to ten percentage points of value added in specialized contracting to be at stake. Competitors include module manufacturers and the developers that employ them or the contractors that are best able to adjust their business models in line with the shifts ahead.

Companies that move fast may reap disproportionate rewards

Whether the share of value at stake benefits players in the larger construction ecosystem or their customers depends on industry dynamics and conduct. Drawing on expert interviews and industry analysis, we created a scenario for overall industry profitability. In the short term, profit pools in some segments of the industry could double if the value of the shifts benefits the ecosystem as whole. Profits could even increase for

companies that make fast and radical changes to outperform their competition and that capture a majority of value shifts in the value chain (Exhibit 15). In the long term, as competition intensifies, we expect gains from cost savings to benefit customers through price reductions and quality improvements.

The impact of these shifts will differ significantly by asset class and project type

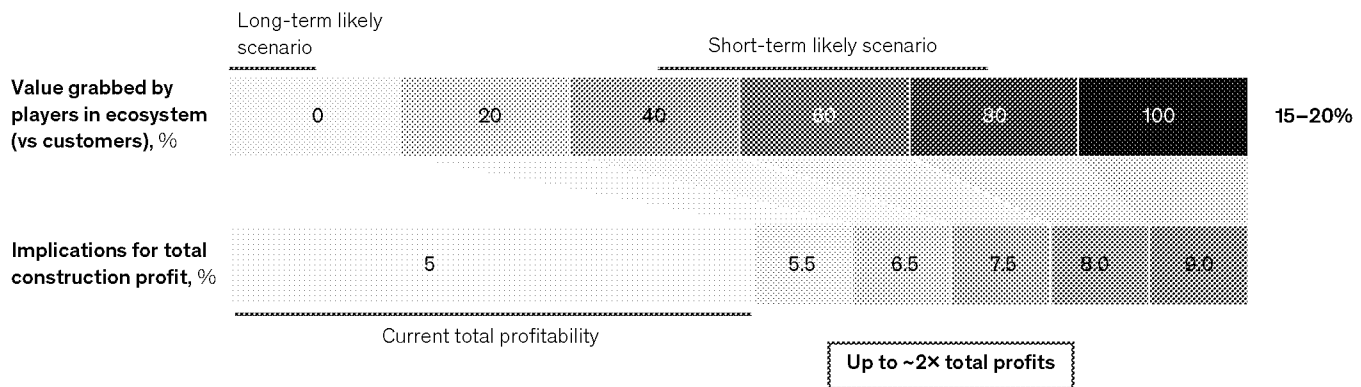
Across segments, shifts in the industry are expected to affect new building projects more than renovation projects. Therefore, our estimates have focused on new projects, though renovation could certainly still have some potential (see sidebar “The potential of renovation”).

That said, emerging evidence in the market indicates that the renovation segment is subject to shifts that are similar to those of new projects and has significant potential for change. The next 15 years could bring considerable improvements.

Exhibit 15

Profit margins could increase for players that move fast and capture value that is up for grabs.

Fully productized value chain in the subsegments of real-estate new build where shifts are the most applicable



The potential of renovation

In this report, we focus on new building projects because we believe this is the category for which change will occur most quickly. In fact, many shifts in new projects have already begun and are moving particularly fast in this segment.

The renovation segment will likely transform at a later time. First, renovation projects are more frequently carried out by fragmented small and medium-size enterprises, which tend to have limited funding available for pursuing innovation. Subsystems manufactured off site could

also be more difficult to integrate into renovation projects than new building projects, as a greater degree of customization would be expected. The reason is that customized building projects often do not follow cogent standards of measurement or setup.

Shifts have different levels of applicability within asset classes and their respective subsegments, and value redistribution is expected to affect them differently in the foreseeable future. Our baseline scenario estimates the adoption rate for the shifts outlined in this report at about 11 percent across asset classes by 2035. Exhibit 16 shows an illustrative timeline of how the shifts are expected to affect new building projects in the various asset classes.

In real estate, for example, we expect an additional applicable volume of 15 percent of new building projects by 2035. This higher number is partly the result of the potential for standardization in single- and multifamily residential, hotels, offices, and hospitals.

Across segments, the shifts could generate profits of up to approximately \$265 billion for innovative players and customers (Exhibit 17).

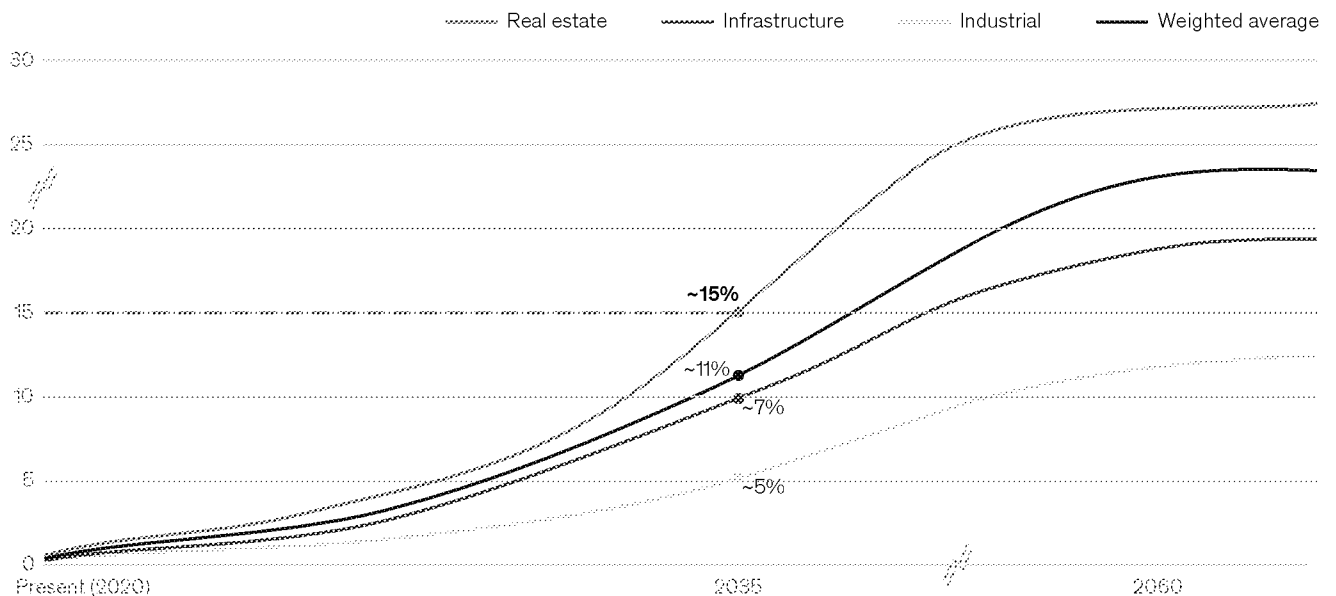
The shifts are expected to affect infrastructure to a lesser degree than real estate. Structures are more difficult to transport as subsystems than in real estate, and repeatability is often more difficult. Some segments—such as airports and railways, but also bridges or tubing rings in tunnels—show more potential, but overall this volume is small compared with roads, which are expected to be less affected. In our scenario, additional applicable volume for new building projects within the infrastructure segment could be approximately 7 percent by 2035.

Exhibit 16

Asset classes will be impacted at different rates of speed.

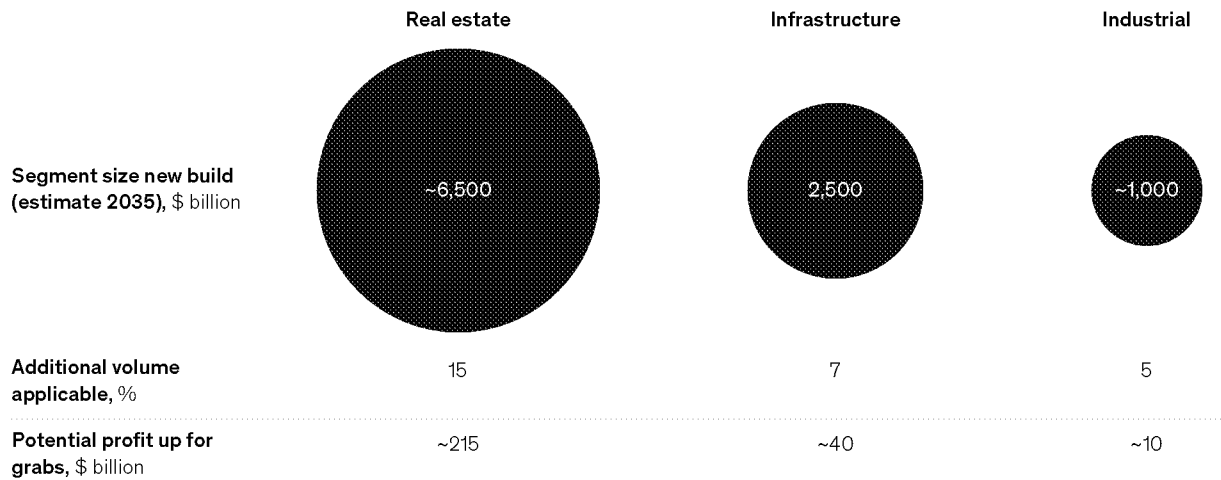
Fully productized value chain in the subsegments of new build

Additional applicable volume for the potential in new build, % of total output per asset class



About \$265 billion in new profits is at stake for fast movers.

Fully productized value chain in the subsegments of real-estate new build where shifts are the most applicable



The industrial segment is expected to have a lower incremental penetration than real estate and infrastructure. The oil and gas sector already has high productivity and is advanced in terms of off-site production, and so are some other industrial structure projects. We estimate that an additional applicable volume of 5 percent for new industrial building projects could be affected by the shifts.



@Getty Images/Morsa Images

4 Transformation will take time, but the COVID-19 crisis will accelerate change

The full transformation of the construction industry will take time. However, the process has already begun. The industry leaders who responded to our survey largely agree that the shifts outlined in this report are likely to occur at scale within the next five to ten years. The COVID-19 pandemic is likely to accelerate the impending changes.

Industry leaders emphasize that the need for drastic change is greater today than it was five to ten years ago. In fact, approximately 80 percent of the survey respondents believe that the industry will look radically different in 20 years (Exhibit 18). In addition, more than 75 percent of respondents believe that the shifts outlined in this report are likely to occur, while a majority expect them to materialize in the next five years.

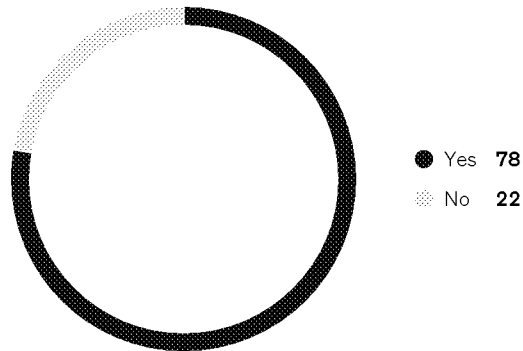
Disruption has started to occur at scale

Both incumbent players and emerging start-ups have been pushing for changes in the industry. While similar transformation journeys have taken decades in other industries (as described in Section 3), construction could harness new digital technologies to speed up its process. In addition, the construction industry has attracted a significant amount of capital in the past few years from private equity and venture capital.

Many players believe that the construction industry is ripe for disruption.

Do you think that the construction industry will change radically in the coming 20 years?

Share of respondents, %



Selected quotes from respondents who answered 'yes'

*"We've been building the same way for 200 years ... but now **assumptions are changing that will drive** demand for automation, ownership, scalability, speed, and the need for zero carbon emissions—**all of these drivers are unprecedented in the construction industry.**"*

— Real-estate architect in US

*"The industry is ripe for disruption. . . . Think Kodak resisting the move from film to digital. **The extrinsic risk for traditional contractors doing things tomorrow the way they do them today is extremely high.**"*

— COO for a real-estate general contractor in US

*"All resources (raw material as well as qualified labour) are limited. At the same time, there is increasing demand for [housing with sufficient living standards]—which drives a gap that needs to be filled. **There must be new methods of construction [to be able to fill the gap]—there is no other way.**"*

— CEO for an infrastructure general contractor in Switzerland

Sources: McKinsey survey of 400 construction-industry CxOs; expert interviews; McKinsey analysis

Tomorrow's construction industry will be radically different from today's. Beyond our analysis of the next normal, and the overwhelming belief of the surveyed executives, we see signs today that the industry had already started to change before the COVID-19 crisis began. There has been activity by both new and existing players against each of our nine shifts.

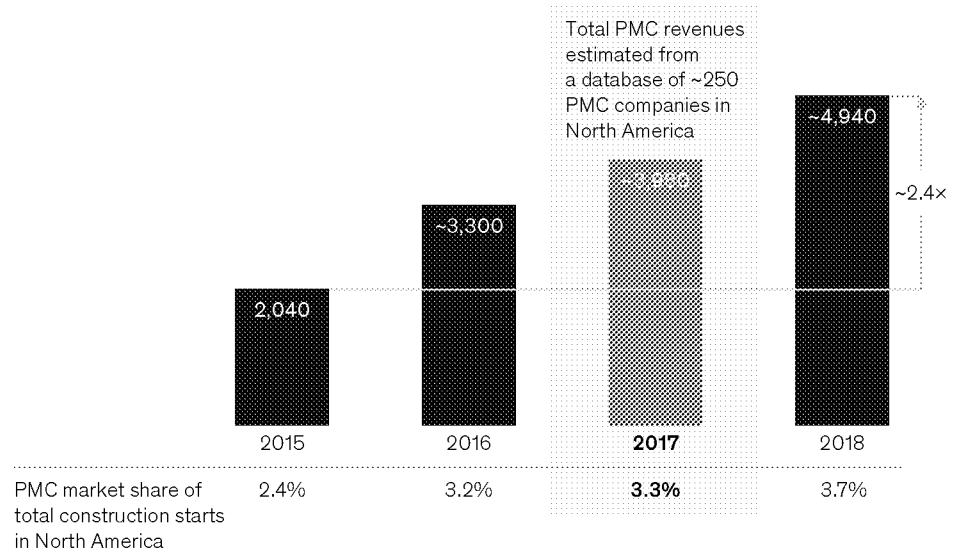
Product-based approach. While prefabrication and modularization have been present in construction for a long time, they have only recently started to regain traction with new materials, better quality, and more design flexibility. Modular construction offers several advantages over traditional methods: it reduces the need for labor (a particularly appealing benefit in markets facing labor shortages), reduces costs, and decreases construction timelines by 30 to 50 percent.

The product-based approach is gaining traction. In North America, for example, growth in the total market share for permanent modular-construction real-estate projects was 51 percent during the period of 2015–18, and total revenue growth for the segment increased by a factor of 2.4 (Exhibit 19). In the United Kingdom and the United States, modular construction accounts for approximately 20 percent of total hotel construction projects.

Seventy-seven percent of our survey respondents believe that a shift to a product-based construction approach is likely to occur at scale, and around 71 percent of those believe such a shift will take place in the next five years.

There are emerging indications that permanent modular construction is taking off in North America.

Revenue of permanent modular construction (PMC) companies in North America 2015–18,¹ millions, \$



¹ Estimated from Modular Building Institute database of 250 modular construction companies engaged in PMC. Source: Construction Connect Insights; Modular Building Institute

Specialization. Players have already begun to specialize in end-user segments, and this shift is expected to gain further traction as construction processes become standardized and the industry adopts an increasingly product-based approach.

For example, in the marine-construction sector, top players are highly specialized contractors such as Royal Boskalis Westminster, which over the years has become one of the world leaders in dredging, land reclamation, and offshore energy construction (Case Study 1).

Approximately 75 percent of survey respondents believe that players will specialize in the future construction-industry ecosystem, and approximately 65 percent of those believe that this change will happen during the next five years.

Case Study 1

Boskalis has focused on becoming the leading global dredging contractor and marine services provider through strategic acquisitions and R&D investments

With more than 100 years of history, Boskalis has become a global leader in the field of dredging and offshore energy. Over the years, the company made strategic acquisitions of both dredging companies and

marine-services providers to maintain leading positions in specific markets as well as to expand its fleet for optimal deployment across projects. Investment in R&D have been essential to Boskalis as a way of increasing both efficiency and effectiveness of operations. The company has dedicated in-house research team and test facilities, and it also collaborates with peer companies and academia. With its multipurpose vessels, Boskalis won the Innovation Award in the category “Dredging Support Vessel” in 2016.

Value-chain control and integration with industrial-grade supply chains. Emerging players as well as incumbents are already seeking to control a larger part of the value chain, particularly those currently moving to adopt a product-based construction approach.

For example, Katerra used new technology to productize and to control the value chain, including design and engineering and off-site manufacturing (Case Study 2).

Nearly 85 percent of the survey respondents believe that it is likely that players will move to control the value chain in the future construction industry, and around 71 percent of those believe that this will take place within the next five years.

Case Study 2

Katerra aims to control the construction value chain by leveraging technology

US-based off-site construction company Katerra has raised \$1.2 billion to integrate and control the construction value chain. To harness technology, the company integrates activities throughout the construction process, including design and engineering, components supply, and off-site manufacturing. A Katerra-designed and -developed, end-to-end integrated digital platform serves as the backbone from project initiation to delivery. As a result, Katerra’s modular component factories can work in tandem with the company’s automated precast-concrete factories.

Consolidation. The construction industry has already started to consolidate, particularly in specific segments and certain parts of the value chain. Specifically, M&A activity in the engineering and construction industry increased approximately 9 percent a year from 2011 to 2017, reaching approximately \$180 billion, while its growth in global M&A volume has been nearly 7 percent a year, reaching \$3.2 trillion (see Case Study 3 for an example of consolidation in the equipment-rental industry).

From 2014 to 2017, the industry experienced an average of nearly 100 more M&A transactions a year than it did from 2009 to 2014.

Approximately 82 percent of survey respondents believe that the industry is likely to consolidate, and approximately 57 percent believe that consolidation will occur in the next five years.

Case Study 3

Equipment rental has been significantly consolidated in Europe and North America

As equipment-rental companies seek scale, this industry segment has experienced substantial consolidation. In Europe, for example, Loxam and Boels began in 2015 to apply buy-and-build strategies—

and recently bid to acquire the two foremost equipment-rental companies in northern Europe. Similar movements have been observed in North America, especially in the United States. These moves affect several asset classes, including real estate, infrastructure, and industrial.

Customer-centricity and branding. While greater focus on customers and branding is likely to gain scale after the shift to a product-based approach, companies are already intensifying their customer focus. Companies that have incorporated the voice of the customer into their product design have seen an uptick in occupancy rates and satisfaction. For example, the European office developer HB Reavis developed a customer-centric and branded approach for its building projects. The company has earned more than 60 awards since 2015 (Case Study 4).

Eighty-three percent of survey respondents believe that players in the future construction industry will shift their focus toward customers, and approximately 74 percent believe that this shift is likely to take place in the next five years.

Case Study 4

HB Reavis is pursuing a customer-centric and branded business model

The office market has evolved dramatically over the past decade. Technological advancements accelerated shifts in employees' work patterns, leading organizations to seek greater flexibility in their office-space arrangements. European real-estate-office developer HB Reavis sought to take advantage of this opportunity by adopting a people-centric approach to the design and construction of flexible work-space solutions. The company used design thinking, ethnographic research, and video prototyping to identify the features and amenities most important to workers. This process produced detailed design briefs that informed project development and construction.

The new brand was named Qubes, and its offices were 95 percent booked before opening. Standardized contract terms for an office package helped to streamline the process for prospective tenants. As important, the extensive user testing resulted in no additional investment in remodeling. Further, 100 percent of HB Reavis buildings under development are expected to become certified by the International WELL Building Institute, a global certification of health and wellness for buildings and interior spaces.

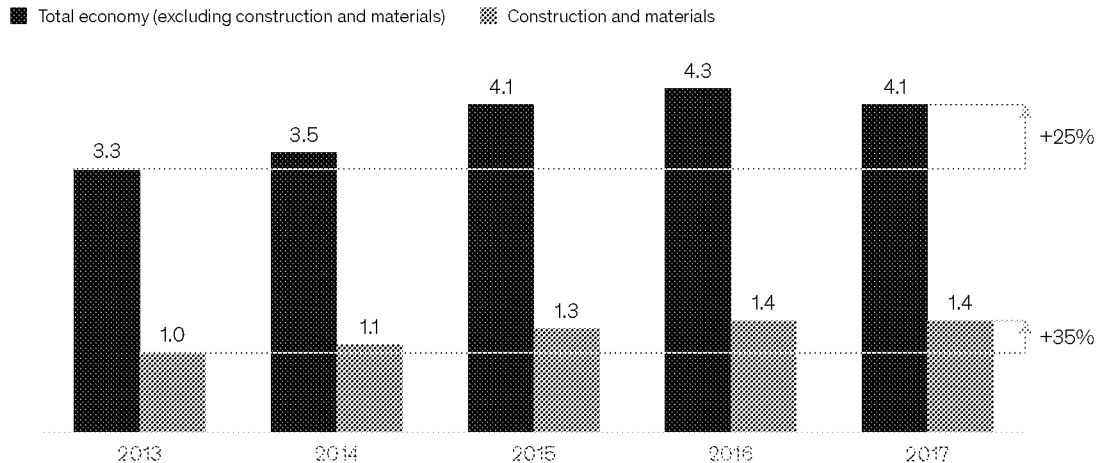
Investment in new technology and facilities. R&D spending in construction, at around 1.4 percent of net sales (based on a sample of the 2,500 companies), still lags behind other industries, which spend approximately 4.1 percent. However, indicators suggest the construction industry is increasing its emphasis on R&D, and companies that have invested in construction technology and facilities are gaining traction. In fact, R&D spending in construction has increased by 35 percent since 2013 compared with 25 percent for the total industry (Exhibit 20).

From 2012 to 2018, funding for construction-technology companies increased at a compound annual growth rate (CAGR) of nearly 40 percent versus nearly 27 percent for other companies throughout the total economy.³²

Nearly 86 percent of respondents believe that a shift toward more investment in new technologies and facilities is likely to occur, and nearly 72 percent believe that this is likely to happen at scale within the next five years.

R&D spending in the construction industry has increased faster than the rest of the economy.

Benchmark of R&D spend of top 2,500 companies globally, R&D spend as % of net sales



Investment in human resources. Players have started to implement more robust HR activities aimed at attracting and retaining employees. Seventy-five percent of engineering and construction players in the United States have made changes to training and development programs in the past two years, to better enable employee retraining and ensure continuous learning of their workforces (Case Study 5).

Approximately 74 percent of our survey respondents believe that the industry will invest more in HR going forward, and 74 percent believe that it is likely to take place at scale in the next five years.

Case Study 5

Bechtel has launched a learning and development academy to support continuous development

A leading engineering, construction, and project-management company, US-based Bechtel focuses on serving both industries and the government. As part of the company's in-house virtual learning and development program, Bechtel University provides more than 1,000 instructor-led and online courses on leadership, professional development, technical skills, sustainability, languages, safety, and ethics. For its efforts, the academy was awarded two Brandon Hall Awards for technology and innovation in 2016.

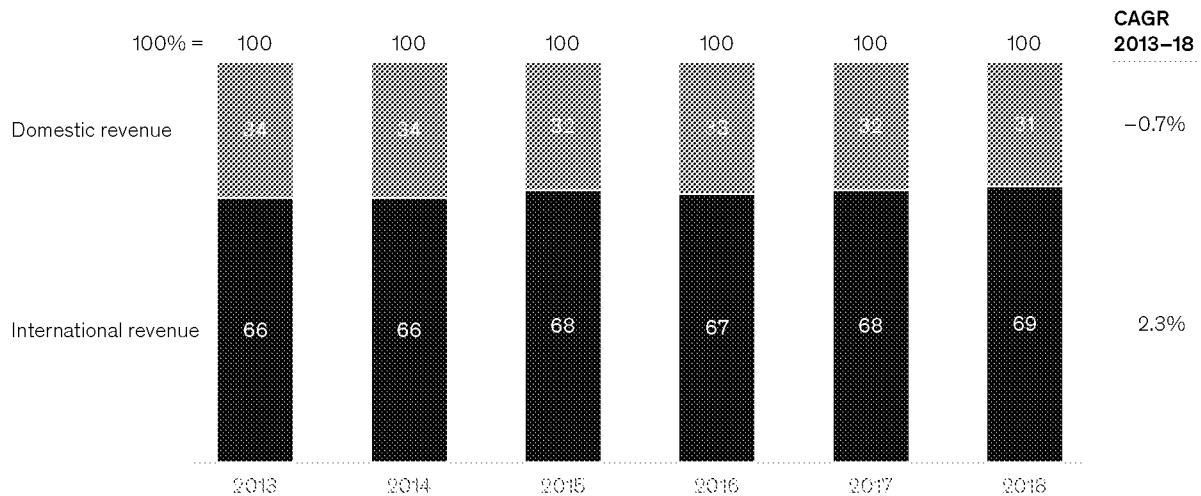
Internationalization. This shift has already started to show signs of traction in the construction industry, with larger players moving internationally to gain scale. Our analysis of the top ten contractors by revenue found that their international revenues increased by a CAGR of 2.3 percent from 2013 to 2018, while domestic revenues decreased by a CAGR of 0.7 percent during the same period (Exhibit 21). International revenues accounted for almost 70 percent of their total in 2018.

Nearly 75 percent of our survey respondents believe that players will expand internationally, and 57 percent believe that this is likely to take place at scale in the next five years.

Exhibit 21

Among the top ten contractors globally, international revenues have increased faster than domestic revenues.

Domestic vs international average revenue breakdown for top 10 contractors,¹ %



¹Balfour Beatty, Bouygues Construction, Grupo ACS, Hochtief, Royal BAM, Skanska, Strabag, Vinci, Bechtel, Leing O'Rourke
 Source: CapitalIQ

Sustainability. There has been a major change in attitudes toward reducing carbon emissions across industries—and construction is one of them. Green building activity continues to rise, driven by both client demands and environmental regulations, as well as a push to create “healthier” buildings and improve occupants’ health. According to the World Green Building Council, building and construction are together responsible for 39 percent of all carbon emissions in the world. Operational emissions (from energy used to heat, cool, and light buildings) account for 28 percent, while the remaining 11 percent comes from embodied carbon emissions associated with materials and construction processes throughout the whole building life cycle.

The United Nations has announced several ambitious targets for sustainability in construction. For example, the target reduction rate for energy intensity per square meter in buildings will be approximately 30 percent by 2030, as defined by the Paris Agreement. Furthermore, several of the largest players have already established ambitious carbon-reduction targets for the future (Case Study 6). However, the fragmented and project-based nature of the construction sector creates additional challenges for the adoption of the sustainable practices, often coupled with the lack of educated green-construction professionals.

Nearly 90 percent of respondents believe that sustainability in construction will be important at scale going forward, and 79 percent believe that the shift will take place in the next five years.

Case Study 6

Several of the largest construction and materials players have set ambitious sustainability targets

Many large construction companies have set ambitious carbon-reduction targets for the coming years. For example, both BAM and Skanska have announced plans to reduce CO₂ emissions by 50 percent by 2030. Other companies have set a goal of using 100 percent renewable fuel as early as 2022. Sustainability is high on the agenda for cement players as well, thanks in part to pressure coming from investors. Germany's HeidelbergCement plans to produce carbon-neutral concrete by 2050, with overall CO₂ reduction targets being assessed against the criteria of the Science Based Targets initiative.¹

¹See sciencebased-targets.org.

The COVID-19 crisis will accelerate the transformation

According to our additional survey conducted in light of the Covid-19 outbreak, around two-thirds of respondents expect the pandemic to accelerate the impending changes and transformation of the industry. Fifty-three percent of the respondents have also started to invest more to adjust to the new future (Exhibit 22).

Specifically, a majority of the respondents expect the crisis to accelerate the disruptions ahead (Exhibit 23). An exception is the expectation regarding disruptive market entrants, as many start-ups and tech companies struggle with financing in the unfolding situation. Around one-third of the respondents also state that their companies have started to invest more in the respective dimensions since the outbreak (especially in digitalization of sales channels and products).

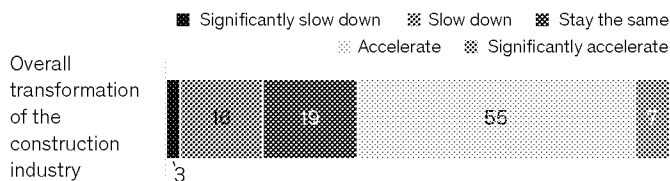
Also, around two-thirds of respondents believe that the COVID-19 crisis will accelerate the nine shifts. They also stated that their companies have started to invest more in these dimensions (especially in technology and facilities, as well as value-chain control). However, one-third of the respondents expect the shifts toward more industrialization and investments in human resources to slow down (Exhibit 24).

Exhibit 22

Two-thirds of survey respondents believe that the COVID-19 crisis will accelerate industry transformation.

As a result of COVID-19, do you believe that transformation of the construction industry will accelerate, stay the same, or slow down?

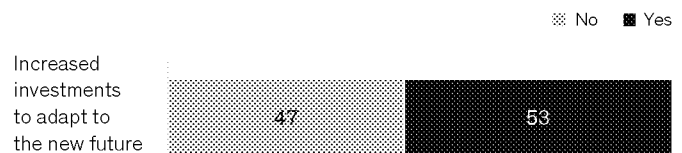
Share of respondents, %



Around two-thirds of respondents believe that the COVID-19 crisis will accelerate the overall transformation of the construction industry

As a result of COVID-19, has your company increased overall investments to adapt to the new future?

Share of respondents, %



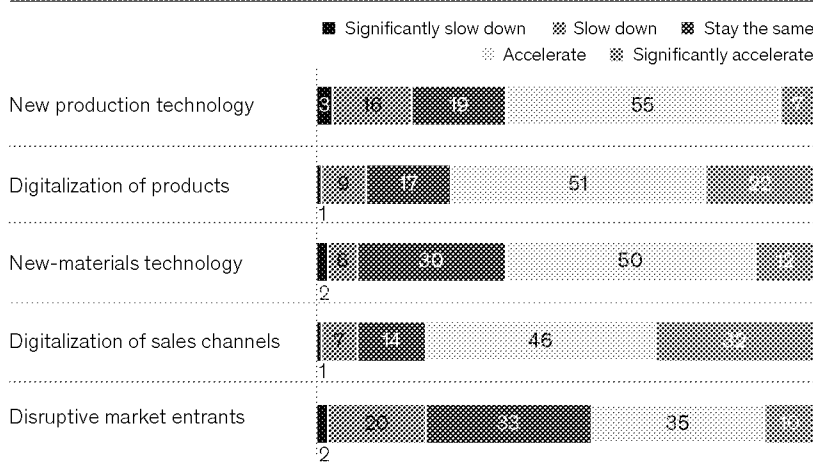
More than 50% of respondents' companies have started to invest more to adjust to the new future

Source: Survey of 100 Industry CEOs, May 2020

A majority of survey respondents believe that the COVID-19 crisis will accelerate disruptions—and have increased investments accordingly.

As a result of COVID-19, which [of these emerging disruptions] do you believe will accelerate, stay the same, or slow down?

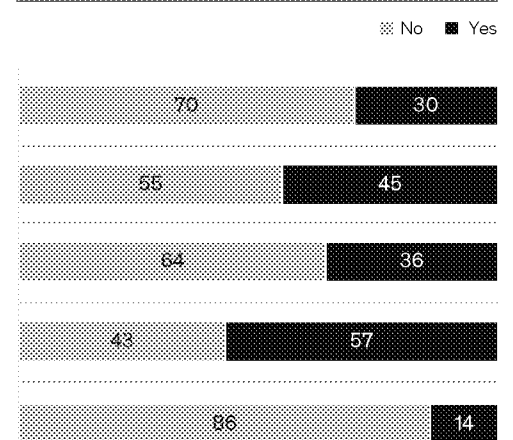
Share of respondents, %



Around two-thirds of respondents believe that the COVID-19 crisis will accelerate virtually all emerging disruptions (disruptive market entrants being the exception)

Source: Survey of 100 Industry CxOs, May 2020

As a result of COVID-19, has your company increased investments in the respective disruptions? Share of respondents, %



Around one-third of respondents' companies have invested more in disruptions (except in market entrants), especially in the digitalization of sales channels and products

Owners and financial institutions need to support change

The industry cannot change in isolation. Rather, change must be a joint effort by both players in the ecosystem and those they interact with. In particular, both customers that commission construction projects and the institutions that finance them must adapt in order to enable change. Demand must exist for players to invest in the capabilities that are required to enable the shifts, and that may require stakeholders to alter how they define value as well as to whom it accrues.

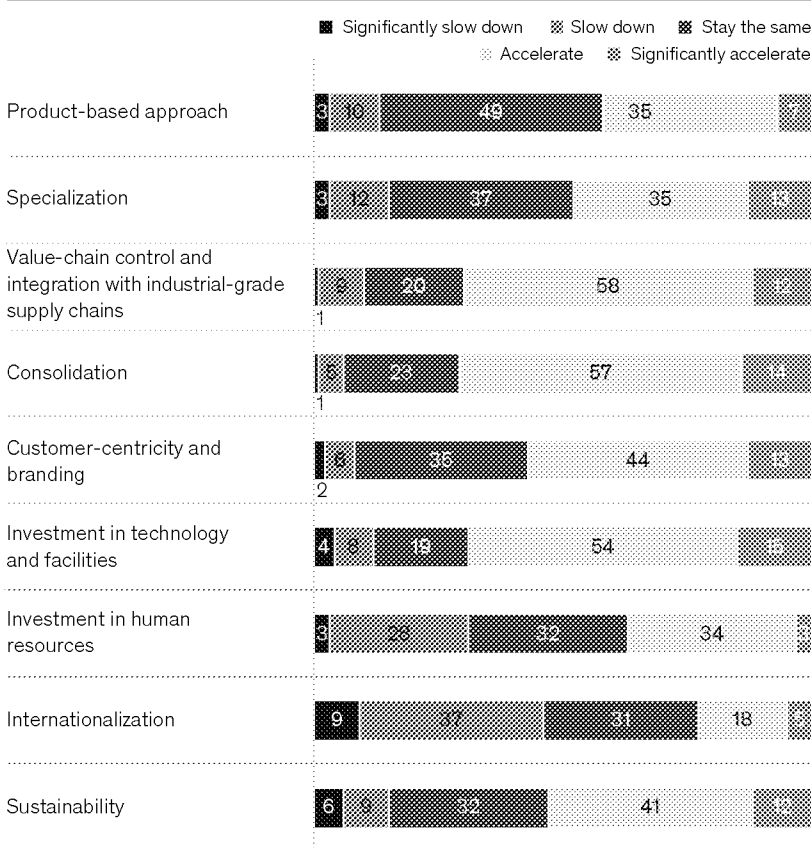
Historically, owners have opted for customized solutions, but some level of standardization is necessary for the efficiency gains outlined in the future ecosystem to be viable. Owners must embrace new, more modular approaches to design to create demand for it in the marketplace. Such a shift would not be altruistic. Indeed, owners should expect to reap multiple benefits in the future ecosystem. Long-term-efficiency gains are expected to materialize as customer surplus, which would result in the generation of additional output by customers' spending. Moreover, the future ecosystem is expected to focus on and improve TCO, thus further benefiting owners. Last, shifts could result in greater overall transparency, which owners could use to improve performance per dollar spent as well as reap other benefits.

The future of construction also requires new financing solutions as well as a willingness to change the risk profile of investments. A shift to a product-based approach would affect project schedules and the total

Around two-thirds of respondents believe that most industry shifts will accelerate as a result of the COVID-19 crisis, although internationalization and investment in people are expected to slow down.

As a result of COVID-19, which [of these shifts] do you believe will accelerate, stay the same, or slow down?

Share of respondents rating shifts, %

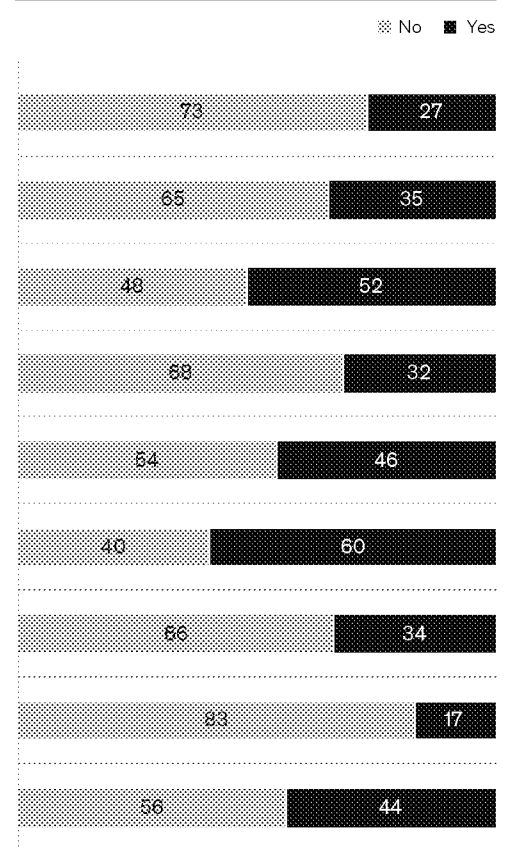


Around two-thirds of respondents believe that the COVID-19 crisis will accelerate virtually all emerging disruptions (disruptive market entrants being the exception)

Source: Survey of 100 Industry CxOs, May 2020

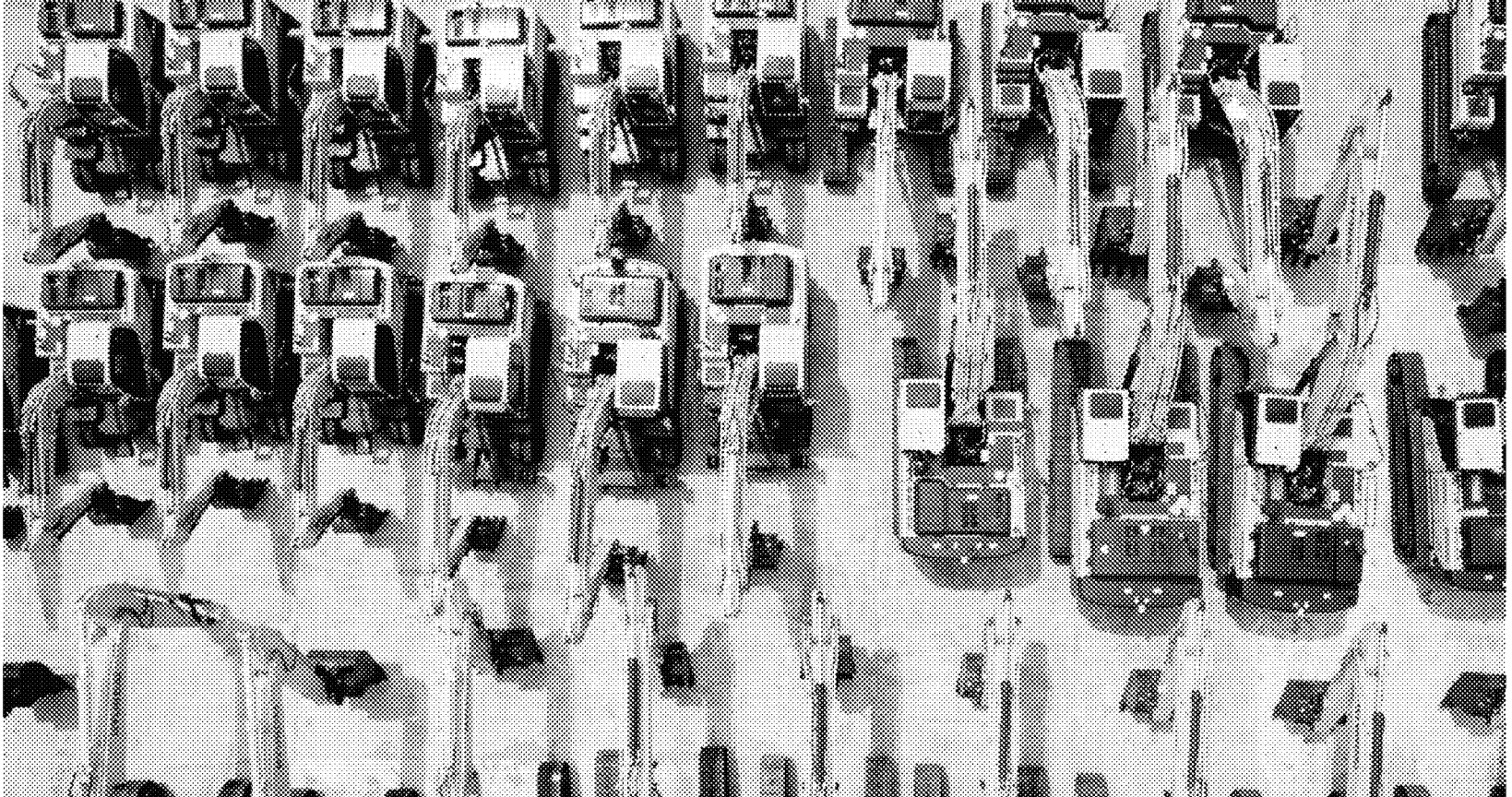
As a result of COVID-19, has your company increased investments in the respective shifts?

Share of respondents, %



Around one-third of respondents' companies have invested more in most shifts, especially in technology and value-chain control

time required for financing. The required overall construction period would be shorter, thus decreasing the overall amount of time for which financing is required. However, a larger share of financing could also be required up front. To create incentives for players to embrace change, financial institutions must therefore develop products tailored to the new demands. However, since market risk and project complexity are also expected to decline, financial institutions will need to reexamine how they price risk into their products to ensure an attractive return. This also holds true for other closely linked institutions such as insurance companies, where some are already factoring use of modern methods of construction into their terms.



@Getty Images/Bloomberg

5 All players must prepare now for a fundamentally different next normal

To effectively manage industry disruption, companies throughout the construction ecosystem must change their strategies, business models, and operating models. They will need to put the enablers in place to survive in the new world and choose their own transformation approach. Some segments will be more affected than others, but each will have its own winning moves.

COVID-19 makes bold strategic action yet more important. During crises, companies that take fast, bold strategic action beyond managing survival tend to emerge as the winners. Looking at the past economic cycles, companies that managed to move fast on productivity (such as reducing cost of goods sold through operational efficiency), divested earlier and pursued more acquisitions in the recovery, and cleaned up their balance sheets ahead of the downturn outperformed competition in both revenues and EBITDA.³³

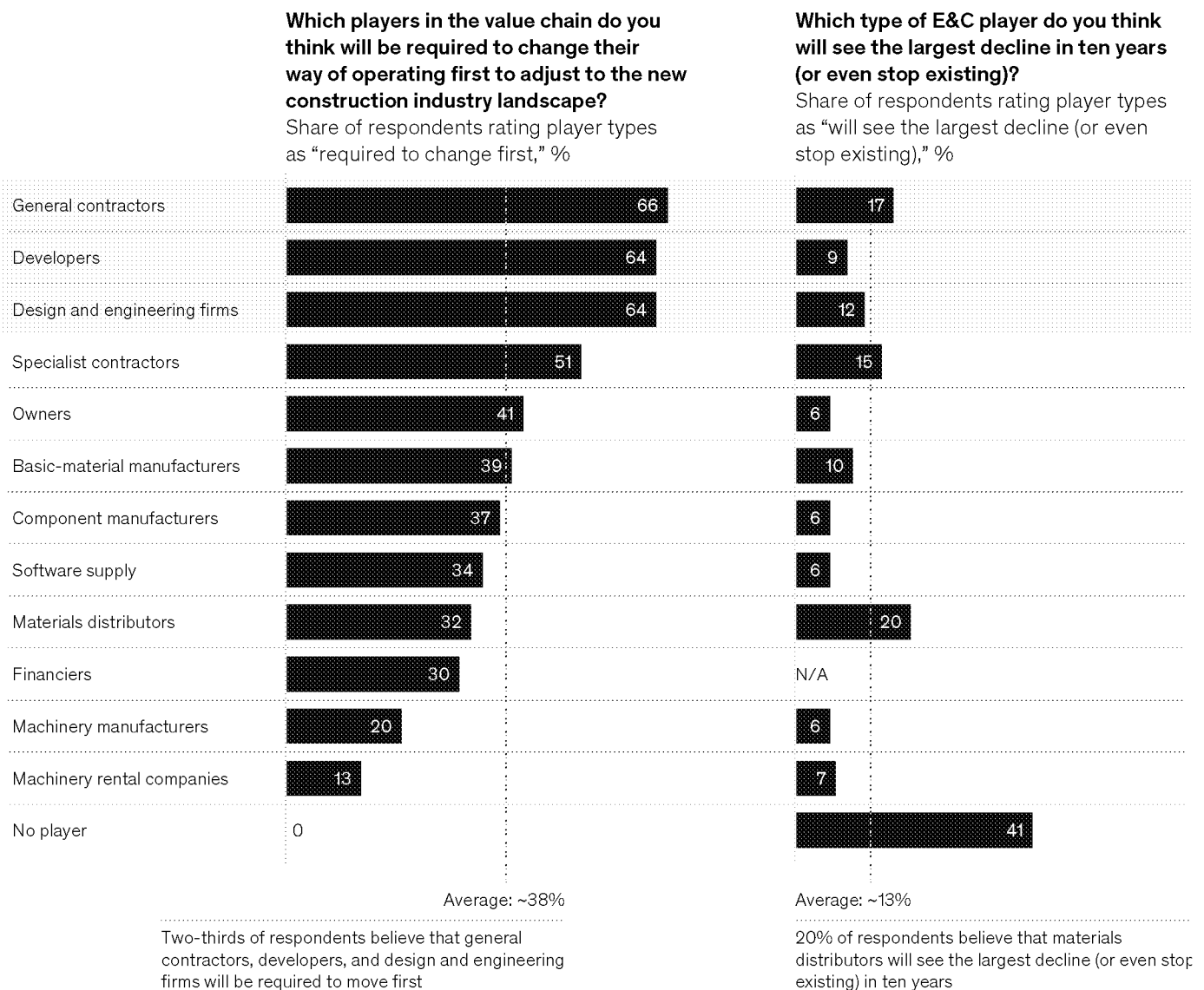
Players adjacent to the construction ecosystem should react to the changes in various ways, to both facilitate and benefit from change. First, investors are advised to use foresight on the respective changes when evaluating opportunities. Second, policy makers should have a high level of interest in making the entire

industry more productive and in achieving better housing and infrastructure outcomes. Third, owners are expected to benefit as the industry changes but will need to play an active part in making the shifts happen.

Four segments of companies are set to face the largest long-term decline independent of the COVID-19 impact: materials distributors, generalist contractors, specialist contractors, and design and engineering firms. These players could face commoditization as well as a declining share of value. In addition, respondents believe that general contractors will be required to move first (Exhibit 25).

Exhibit 25

While all players will be affected, industry leaders believe general contractors will be required to move first and materials distributors will see the largest decline.



Source: McKinsey survey of 400 construction-industry CEOs; expert interviews; McKinsey analysis

All players will need to act on the nine shifts that are ahead

In response to the industry transformation detailed throughout this report, companies will need to reinvent themselves and redefine their strategies and business and operating models (Exhibit 26). A typical journey might start with an initial strategy-resetting effort including a review of where to play and how to win, and include an assessment of how—and how strongly—the nine shifts will affect focus markets, and what requirements are put on future winning business and operating models. Companies will also need to put in place a set of enablers to make them successful for the future and choose how they want to implement the new strategy. Given the highly dynamic of the changing construction industry, successful strategy formulation and direction setting need to be approached as a constant reiterative process.

The strategy (resetting) effort needs to determine “where to play.” Which asset classes, segments, geographies, and value-chain steps are large, growing, profitable, and at acceptable levels of risk? In which ones can a company build a sustained competitive advantage and barriers to entry? The spectrum may range from global development of high-end office towers through the regional provision of precision cross-laminated timber for residential prefab to continent-wide steel bridge construction or owner-operation of logistics facilities.

Following the question of “where to play,” companies must assess the impact of the nine shifts on “how to win”; that is, the implications for future winning business and operating models. The answer will greatly differ between segments—for example, small residential refurb projects compared with multibillion-dollar oil and gas projects or between basic-materials supply versus machinery rental. Executives in all sectors, however, will need to answer questions related to the expected impact of each of the nine shifts. Following are a few thought starters:

Product-based approach. How far can I productize and standardize my product and service offering? Which target customer groups do I want to address? What kind of customization will be necessary and economical to offer to them? Should I try to follow current customer demands or attempt to act as a market maker? Do I want to explicitly avoid productization and focus on high-end, individual projects?

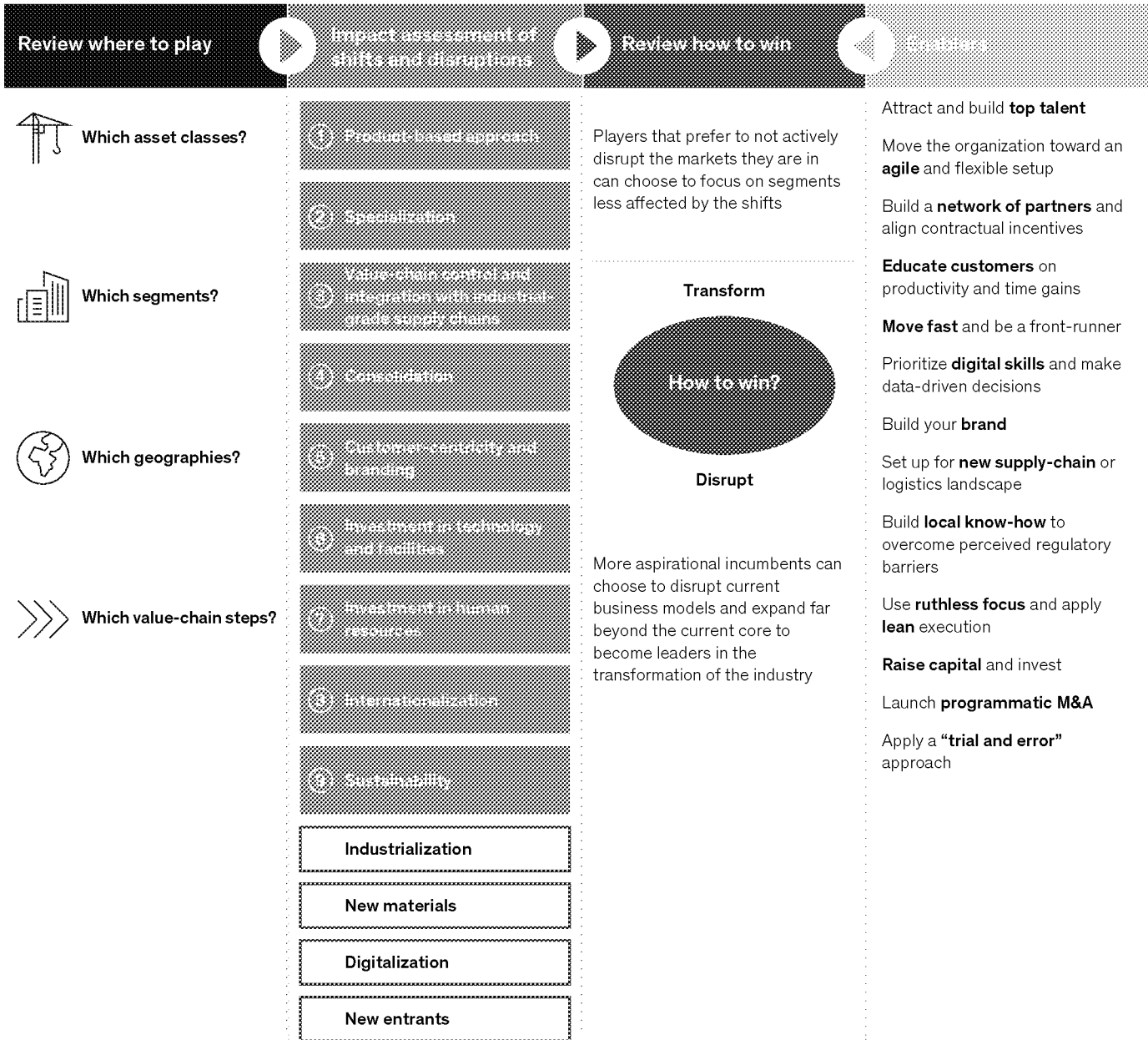
Specialization. How much do I want to specialize in niches in the markets I've chosen to play in? What is the right balance between creating competitive advantage and experience versus balancing the order pipeline and market cyclicalities? Can I be a regional leader across many segments?

Value-chain control and integration with industrial-grade supply chains. What interface complexities do I have today, and to what extent do they constrain innovation and efficiency improvements? Can and should I enter—or disintermediate—adjacent value-chain steps? Do I need M&A to do that? Can I achieve end-to-end control of the value chain through strategic alliances and partnerships? Do my chosen partners share the same vision of how to innovate value delivery and business models?

Customer-centricity and branding. If I aim to be a recognized leader in my chosen markets, how can I build brands to support that goal—and what values should those brands convey? How can I balance branding at the corporate level with products or services and by region? How can I shift my business model to create a better outcome and experience for my business partners as well as eventual customers, even if the current ecosystem's incentive structure might be set up to benefit from overengineering to increase project volumes, change orders and claims, fudging market transparency, or unjustifiably passing on risks?

Consolidation. What scale do I need to enable the required massive increase in investment in technology, facilities, and human resources and to build critical scale in all target markets? Can I achieve that scale organically, or do I need M&A? Would sell-side M&A be an option for me?

All players will need to reinvent themselves and redefine their strategies and business and operating models.



Strategy implementation and transformation approach: Broad transformation, new divisions in new setup, arm's-length subsidiary attacker, and M&A

Investment in technology and facilities. What activities can I automate or digitize? What are the right technology suppliers? How can I build out or migrate my IT platforms to be future-proof? Which activities can be taken from project sites to factories? What is the right production-footprint balance versus logistics cost?

Investment in human resources. How can I revamp the culture of my organization to embrace change? What skill-building activities for my staff will be needed? What sort of new talent will I need? And, particularly in technology, how can I successfully compete on the recruiting market? Might an attractive new strategy and brand be part of the answer?

Internationalization. In which target geographies can I be successful? How do I balance governance at the regional, business-segment, and functional levels? Can I use low-cost country-sourcing or production-footprint strategies? How can I allow for regional variation in building codes and industry dynamics?

Sustainability. How can I raise the environmental and social sustainability of my operations, my products, and the inputs I source? How far can I go—and how fast—to meet regulations as well as my own ethical standards? How can I create a business case for investments in sustainability?

Incumbents also need to decide to what extent they should focus on defending their core business and adjusting to the changing environment compared with actively reinventing themselves to attack and disrupt the markets they operate in. Disruptive plays require risk taking and investment, but successful moves can be rewarded with step changes in profitability and valuation multiples.

Defend the core and transform. Players that prefer not to actively disrupt the markets they are in may choose to focus on segments that will be less affected by the impending shifts, such as renovation or unique, iconic buildings. These segments may face declining market share and increasing levels of price competition, however. Consequently, players will need to continually improve their capabilities and competitiveness as well as their ability to manage risks. And they will need to adjust to the changing environment. Contractors, for example, will need to invest in on-site automation, digitization of processes, end-to-end adoption of BIM, or sourcing of high-value pods that simplify construction on site.

Reinvent to attack and disrupt. More aspirational incumbents may choose to reinvent current business models and expand far beyond the current core to be leaders in the industry's transformation. They would explicitly target the segments where change is imminent, consolidate value chains, embrace online direct channels with advanced logistics management, apply programmatic M&A to achieve scale, or invest in manufacturing capabilities and capacities well beyond the scale of their current project pipeline.

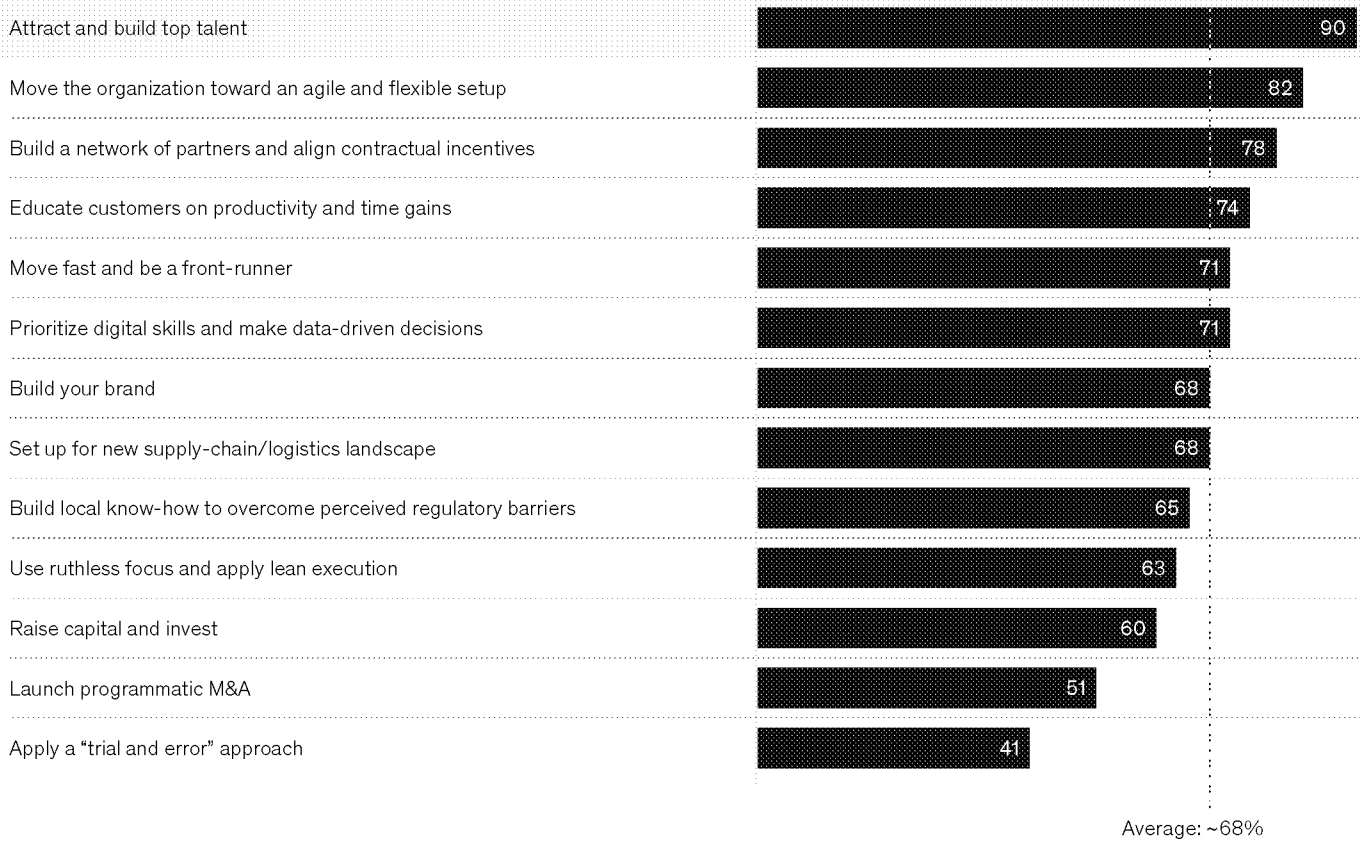
Regardless of chosen strategies, a set of enablers can benefit all players in the ecosystem. As indicated by our survey, attracting and building top talent is of primary importance (Exhibit 27). In fact, 90 percent of the surveyed industry experts believe this action will be the most important one in the future ecosystem. In addition, 82 percent of the respondents consider moving toward an agile organizational setup critical, and 78 percent cite building a network of partners and aligning contractual incentives.

Finally, companies need to define how to implement and achieve a corporate transformation. At the most basic level, several options exist. Some companies will build new business units in line with the target business, operating model, and new-market environment—and, over time, transfer more and more volume and resources to that unit. Other companies will attempt to transform their incumbent or core operations.

The industry believes that winning moves are centered on talent attraction and retention.

What will be the winning moves in the future construction-industry ecosystem?

Share of respondents rating move as “very beneficial,”¹ %



90% of respondents believe that attracting and building talent will be the winning move in the future

¹Very beneficial equals a 7 or more, where 10 equals the most beneficial.

Source: McKinsey survey of 400 construction-industry CEOs; expert interviews; McKinsey analysis

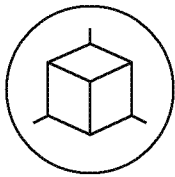
Yet others might opt to start new attacker businesses at arm’s length from the incumbent operations. And some will use buy- or sell-side M&A to jump-start a move to new business and operating models. Depending on the circumstances, all of these approaches can succeed.

Each industry type has its own winning moves

In the years to come, each player in the value chain must react to or drive the shifts and reinvent business and operating models. We provide short examples and vignettes of the type of changes ahead, as inspiration for the development of individualized winning strategies.



Thought starters for Developers



Public- and private-sector developers of real-estate, infrastructure, and industrial projects orchestrate the development process from beginning to end: securing financing, sourcing land, and scoping and overseeing value-adding projects. As customer expectations continue to evolve rapidly, developers look set to increasingly specialize and invest in productizing and branding their offerings, which will increasingly require multinational scale and leave behind those working in traditional ways. Specialization will be key to attract customers, and developers need to decide whether they want to focus on slivers, whether they want to develop individual opportunities or also establish quarter or precinct capabilities, and which of the required capabilities they need to own versus source. Further, a deep understanding of customer needs, by segment and subsegment, will be at the heart of the most successful strategies, whatever the sector or asset type; without it, the assets that have been built can quickly become obsolete.

The most successful players are already bringing together these customer insights with product and supply-chain innovation to deliver high-performing projects. For example, leading commercial real-estate developers are building direct relationships with end users and creating spaces that can rapidly be repurposed as demand changes. Similarly, highway developers are connecting directly with drivers to better understand usage patterns to inform future highway design and operation. In many ways, developers set the tone for the whole industry.

Since developers sit at the top of the value chain, they can strongly influence how and how fast disruption in other parts of the value chain takes place, including actively steering companies toward industrial production of their offerings as well as library-based designs and subsystems. Such approaches can reduce the cost, time, and risk of projects—all improvements in the financial viability of projects—which may translate into higher profits, greater volume, or value shifting to customers. Only with the right signals from developers will change spread through the rest of the ecosystem. Here the public sector has an important role to play given that in aggregate, the sector is a very sizable developer (and owner).

Transform

Double down on segments less affected by shifts where you have a distinctive expertise?

- Continue with the current business model, emphasizing segments facing less disruption in the medium term (e.g., renovation, maintenance, improvement)?
- Invest in deeper understanding of customer segments?
- Identify and invest in core capabilities (e.g., land access and acquisition)?

Augment current offerings with technology and module sourcing?

- Identify opportunities to use prefabricated and module sourcing (e.g., through standardized structural-frame modules for all buildings a standard catalog of kitchens) in a targeted way, to substantially reduce project risk and improve time and cost parameters?

- Invest in digital technology to accelerate and reduce the risk of target segments (e.g., identification of land parcels using geographic information system, real-time progress tracking using lidar scanning, digital twin to streamline operations and maintenance)?

Disrupt

Create a product house (with or without manufacturing) and build a vertically integrated platform business with close links to customers?

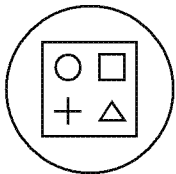
- Develop a carefully composed, specialized product and service portfolio that is tailored to end users and purposefully designed for industrialized manufacturing at scale (ideally, influencing the establishment of industry standards in the long term)?

- Aim to be a global leader in offerings by scaling operations and entering international markets?
- Build a digital backbone to support end-to-end customer journeys, from acquisition through specification, design, execution, and handover?
- Look for opportunities for vertical integration and control of the value chain by integrating design and engineering and manufacturing?
- Establish new funding and risk-management structures to accommodate new delivery models?
- Capture operational data to continuously refine and improve designs for future projects?
- Use develop-and-hold strategies to capture TCO advantages of product or service offerings?



Thought starters for

Designers and engineers



Disruption could fundamentally change what it means to be an engineer or an architect in the construction industry. Historically, these professionals have applied their considerable expertise to create designs and specifications for individual projects: each design optimized to meet the project's unique requirements. The coming years will see these stand-alone professional-services firms closely collaborating with productized and branded developers, off-site construction firms, and highly specialized contractors as an integrated R&D-like function. The firms will increasingly add value through the standardization of structure and subsystem designs, by developing standardized design libraries of products in their target segment that are highly integrable to allow for a customizable whole. This modular design will be reused for a large set of construction projects. In this way, design and engineering firms could influence industry standards. As the industry shifts to a more product-based approach, the challenge for engineering and architecture firms will be to reskill their workforces and hire the right talent to design in this new world.

Of course, modularization and automation will not apply to all projects: highly architectural or complex projects will have limited degrees of standardization. Similarly, refurbishment-maintenance-improvement (RMI) projects are likely to continue to follow more traditional design approaches for some time. In affected segments, however, design and engineering firms are likely to improve their efficiency by using standardized products and libraries, as well as more software-based automation and parametric design—though the extent will depend on how the transformation plays out. The best performers will still stand to gain. Large owners may pay a premium for modular or customizable product design, as only a few firms have the required capabilities and experience. But as the industry adjusts to designs that can be replicated and adapted multiple times, the volume of work is likely to decrease in the affected segments. The need for redesign is also expected to decrease drastically with more specialized and productized approaches, generating significant cost savings for design and engineering and putting their respective value at stake.

Transform

Continue to focus on traditional approaches in core segments?

- Concentrate on projects such as renovation, maintenance, and improvement that are not expected to face disruption in the short term?
- Double down in segments with strong demand or where the company has the competitive advantages of strong customer relationships or a deep understanding of customer requirements?

Invest in technology to streamline parts of the design process?

- Partner with a software company to automate simple parts of the design processes by using parametric and generative design software and BIM workflows; retrain existing employees and seek out new talent?
- Invest in building (or buying) object libraries and standardized design elements to be used both off and on site?

Change the commercial model and charge for advisory services and outcomes—not hours or percentages of cost?

- Invest in getting closer to customers to better understand what determines value, and adjust offerings to suit (e.g., design a facade panel optimized for speed of insulation, create a valve optimized for minimal service intervals)?
- Calculate and articulate the value add from services and switch the business model from charges per hour to an outcome-based model?

Become the go-to partner for design-to-manufacture?

- Build capabilities in design-to-manufacture?
- Invest heavily in manufacturing talent, R&D, and design software?
- Build relationships with off-site manufacturers?

Consolidate to gain scale and internationalize the company footprint?

- Consolidate with other players and internationalize the company's presence to

gain benefits from increased scale and to gain access to the largest projects?

- Aggressively target international markets?

Disrupt

Become the design arm of an integrated manufacturer or product house including via mergers, acquisitions and alliances (MA&A)?

- Integrate with a modular manufacturer and/or developer to become an in-house engineering arm that optimizes design for manufacture?
- Integrate with a software house to build engineering capability?

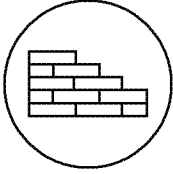
Build business around data?

- Support clients in quickly defining the most adequate concept and then delivering it, and support operations and maintenance (O&M) through advanced analytics (AA)?



Thought starters for

Basic-materials manufacturers



A large share of inputs used in construction projects involve processing raw materials such as cement, steel, wood, or glass. Many players in this sector are already large, global firms with slower-moving shifts in value-chain dynamics. The most pronounced impact might arise from a transition to new, lighter-weight materials, as well as higher requirements on sustainability, recycling, and waste reduction. While the industry should benefit from long-term (post-cycle) growth of the construction market, the volume of traditional materials, such as the amount of cement, looks set to decline on a per-building basis (including due to a rise of lighter-weight off-site module production and, consequently, less strong foundations). Digitalization and consolidation of the distribution and contracting landscape may alter logistics and customer interfaces. It will be critical for the industry to stay ahead of those developments, raising the pace of innovation and the level of customer intimacy.

Transform

Double down on operational excellence by, among other actions, digitalizing processes and plants, adopting agile and lean principles, and using advanced analytics in the core business?

- Raise the level of customer intimacy, understanding better how to address customers' primary challenges and improving (digital) interactions?

Invest in sustainability and higher pace of innovation?

- Develop, brand, and advocate for products that are sustainable or made from zero-carbon materials?
- Invest in materials recycling, and find a strong position vis-à-vis contractors?
- Increase use of alternative fuels?
- Invest in new types of talent with backgrounds in product development, construction, and digital—and avoid the

pure-cost-cutting culture that prevails in the industry?

- Manufacture materials that enable lower TCO through energy efficiency?
- Carefully monitor materials choices in module production, and build a presence or alternative?

Disrupt

Invest in a portfolio of products and technologies that will eventually cannibalize your core business?

- In cement, replace Portland cement with zero-carbon binders like ultra-sulfated cement, as slower processes are less problematic in prefabrication?
- Invest in carbon capture and storage/ carbon capture, and use technologies such as lightweight aggregates?
- Move from "reactive" precast facilities to the development of entire prefab construction

and building systems that include engineering or BIM modules, construction solutions, precast modules, etc?

Proactively disrupt company's footprint?

- Disrupt your own production footprint with advanced digital and carbon-neutral plants?
- Plan downscaling and exit road maps for outdated facilities?
- Create further regional consolidation, including through distressed assets and corporate failures?

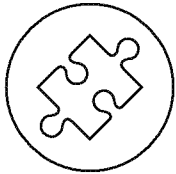
Integrate in value chain by developing own solutions?

- Create modular-construction business units; for instance in infrastructure like bridges?
- Become the one-stop shop for building materials?



Thought starters for

Component manufacturers



Today, components such as elevators, HVAC equipment, and pipes are often produced using a silo-like approach, which limits the ability to integrate the components once they have been installed in buildings. Installers tend to have strong affinities for specific suppliers—due to either personal experience or suppliers' incentive schemes—while the brand affiliation of end customers is usually not as strong.

Several of the nine shifts will most strongly affect component manufacturers: productization and standardization (including through BIM object libraries) and online channels will increase price transparency and lead to commoditization. Consolidation will improve the bargaining power of large contractors or modular-construction firms, and internationalization can lead to low-cost-country sourcing. In turn, the best companies can push consolidation and market share gains themselves to achieve further economies of scale, and offer solutions with a TCO advantage and value-adding digital services. They might also shift from manufacturing components to entire modules or subsystems, direct-sales approaches, and orchestrating just-in-time logistics to prefabrication or eventual construction sites.

Transform

Gain scale and standardize across borders?

- Consolidate with other players and internationalize presence to gain benefits from increased scale?
- Standardize production across borders?

Invest in sustainable solutions and optimize energy efficiency for customers?

- Develop products and solutions that help customers optimize their operations in terms of energy usage throughout the entire system?
- Offer services to help customers evaluate energy usage and improve efficiency where possible?

Invest in digital channels while avoiding commoditization?

- Invest in digital and BIM object libraries and shift go-to-market strategies to influence

component choices earlier in the value chain?

- Build direct channels to emerging product houses, module suppliers, and consolidated contractors?
- Invest in R&D, brand, and differentiation to decrease the potential effects of standardization and increased transparency, including value-adding features and services and value-based pricing models?
- Build or integrate into industrial-grade digital supply chains from factory to site?

Modularize products and service offerings to enable standardization and reduce variances?

- Seek to modularize product offerings to be able to provide standardized submodules that can be combined into entire systems and solutions?

Disrupt

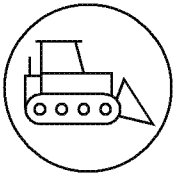
Adjust the business model to be able to charge for output instead of actual products?

- Change business and pricing models to charge for the actual savings or benefits that are generated (e.g., optimized energy systems for heating, ventilation, lighting, etc) instead of charging for the actual products?
- Take shared responsibility for optimizing customer costs and benefits (e.g., energy usage) in the entire system?
- Move forward along the value chain to producing modules?



Thought starters for

Machinery manufacturers



Over the next two decades, manufacturers will transition from producing traditional heavy machinery and tools used in the construction process to highly automated, connected products used in the ecosystem. The new equipment will be integrated with robotics that could be used in the ecosystem’s manufacturing processes—for example, in plants for building materials, components, and buildings. Rather than simply selling products, manufacturers will offer services that are completed with their products. This will increase the value added of machinery, and increasing automation will support greater volume. In addition, the demand for machinery used in off-site fabrication is expected to rise. In turn, improved efficiency and reduced time requirements for on-site work are expected to have a negative impact on the volume of machines sold.

Transform

Consolidate and internationalize the company footprint as products become more standardized?

- Focus on internationalizing as preferences and the way machines are used become more harmonized across regions?
- Engage in M&A activity and identify attractive targets for acquisition?

Future-proof the product portfolio in line with trends and general development of the market—for example, IoT or vehicle electrification?

- Develop a product portfolio that seamlessly integrates into an end-to-end, controlled, industrial-grade supply chain (requiring IoT connectivity)?
- Adopt an electric fleet and product portfolio to prepare for lower carbon-emission requirements?
- Build supporting infrastructure, such as charging stations for electric vehicles and 5G connectivity on site as well as training and certification programs for operators?
- Help customers to evaluate their energy usage and optimize energy efficiency in operations?

Right-size areas expected to be negatively affected by shifts to a product-based approach?

- Identify the areas in the business that could be negatively affected by the various shifts (for example, tower cranes used in residential real estate, as project timelines are expected to get shorter)?
- Divest or “milk” different areas of the business, and invest the generated cash in other parts of the business?
- Shift the product portfolio toward the tools and machinery used in off-site construction facilities?

Invest in digital and develop agnostic software to optimize the entire fleet on site?

- Develop machines and tools that are compatible with each other and able to run on different types of software, since customers want to optimize their entire fleets of mixed brands?
- Develop products that are compatible and can be connected to different types of site protocols (e.g., BIM)?
- Invest in bulletproof systems to handle and access data (e.g., to avoid unauthorized people taking control of machines)?

Disrupt

Build automated machinery for use in future off-site manufacturing facilities?

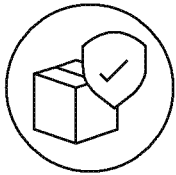
Develop robotics and autonomous vehicles used on site for select, repetitive tasks (e.g., bricklaying)?

Adjust the business model to be able to charge for output (e.g., machine uptime, tons of earth moved) instead of actual products?



Thought starters for

Materials distributors and logistics players



Distributors procure, store, and transport basic materials, components, and equipment and resell them to consumers and businesses. Some distributors also provide credit. Part of this model is the organization of logistics and inventory, primarily for construction sites and installers.

Several of the nine shifts may hit distributors in a negative way. Productization, standardization, and consolidation will move decisions and procurement upstream from small specialized subcontractors to large contractors and product-based developers, increasing bargaining power and reducing the breadth of materials needed. Better and earlier planning using BIM and digital twins will reinforce those shifts and reduce the need for local stock. Off-site manufacturing facilities will shift demand for shipments to factory hubs, with more predictable levels of demand, which will be the main logistics nodes and decrease the need for a dense storage network close to construction sites, while also raising the expectation of just-in-time delivery. Internationalization will enable more sourcing from low-cost countries. Online and direct-sales channels, including new competition from online distribution behemoths—which serve customers with high expectations and use increasing amounts of technology, such as advanced analytics or automated warehouses—will further reshape this segment.

While these shifts pose a major threat to distributors that lack scale and logistics capabilities, they provide opportunities to companies that have them. There is an opportunity to consolidate the sector, supported by lean efficiencies, category reviews, and new business solutions. Direct-to-customer digital interaction channels and interfaces connecting to BIM and building-management systems allow better integration into the value chain. Distributors can fill the roles of the logistics hubs of the future construction landscape by using advanced analytics in logistics, demand forecasting, and inventory management to allow just-in-time delivery from suppliers to modular-construction factories to construction sites. Distributors can create new value for customers by helping with international sourcing, offering credit finance, packing in assembly order, offering in-room delivery, making deliveries before the working day, providing on-site logistics planning and operations, or even handling simple pre-assembly.³⁴

Transform

Focus on segments that the shifts will affect less than others?

- Stay in a traditional materials-distributor role but focus on projects expected to be produced on site in the foreseeable future, such as refurbishment work or in other segments that are expected to have low levels of standardization?

Concentrate on operational excellence, better customer-value propositions, and integrated systems, to avoid disintermediation?

- Focus on operational improvements, lean execution, category reviews, commercial and pricing excellence, and procurement optimization?
- Reduce the carbon footprint of operations and materials?

- Create a digital user experience to match those of leading online players?
- Allow integration with BIM and building-management solutions, allowing direct ordering from models as well as improving forecasting?
- Provide flexible and just-in-time logistics and value-adding logistics services (e.g., packing in assembly order, in-room delivery, delivery before the next working day)?
- Focus on the core value proposition, superior availability, inventory transparency, best-in-class logistics, and demand forecasting using advanced analytics?
- Improve customer relationships with advice and financing services?

Consolidate to gain scale in each region of presence?

- Consolidate with other players to gain benefits from increased scale?

Disrupt

Become the logistics hub of the future construction landscape?

- Integrate or partner with other players in the value chain to create and control an industrial-grade supply chain?
- Partner closely with off-site manufacturers and materials suppliers to optimize logistics and inventory according to their needs and prioritize just-in-time delivery?



Thought starters for

Machinery- rental players



Currently, rental businesses provide companies in the ecosystem with an efficient capital-expenditure option for yellow machinery and tools, which help to optimize equipment utilization. Digital technologies will enable greater efficiency through on-site logistics, either providing the right machinery at the right time or operating and using the machinery for customers. By offering digital services, rental companies gather usage data from products to enable best-in-class employment of machinery for customers.

Some machinery—particularly tools—is expected to be moved off site, trending toward higher utilization and limited rentals rather than direct purchases. This shift will affect machinery-rental players negatively, although only to a small extent. In turn, service-based business models and IoT-connected machinery and tools could compensate for these shifts. Machinery for groundwork will likely be less affected by the upcoming shifts.

Transform

Generate data-driven productivity insights?

- Collect data critical for equipment to generate “customer-ready” insights, through collaboration that improves project planning, productivity, and safety, including equipment usage (encouraging early returns and pay-as-you-go-type services), equipment and labor productivity (based on hours of operation), or unsafe behavior?

Capitalize on major trends and build supporting infrastructure?

- Lead the way on deployment and support of the new wave of equipment in the field, such as electric vehicles, autonomous vehicles, and robotics?
- Build supporting infrastructure, such as charging stations for electric vehicles and 5G connectivity on site, as well as training and certification programs for operators?

- Help customers evaluate their energy usage, and optimize energy efficiency in operations?
- Increase service offerings to help customers maintain a focus on their core expertise?
- Take on responsibility for tasks that are unrelated to contractors’ and owners’ core businesses (e.g., logistics, waste handling) to help them focus on their main sources of value add?

Invest in digital and agnostic software to optimize entire fleet at site?

- Build fleet of machines and tools that are compatible and able to run on different types of software, since customers want to optimize their entire fleet of mixed brands?
- Offer products that are compatible with each other and can be connected to different types of site protocols (e.g., BIM)?
- Invest in bulletproof systems to handle and access data (e.g., to avoid unauthorized people taking control of machines)?

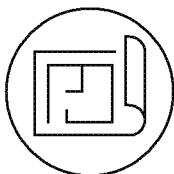
Disrupt

Adjust the business model to be able to charge for output (e.g., machine uptime, tons of earth moved) instead of actual products?



Thought starters for

Off-site construction players



Today, off-site construction companies primarily manufacture building elements, structures, or modules for real estate—but also for industrial structures and infrastructure like bridge segments. Overall, off-site construction is still a relatively young and immature part of the larger construction ecosystem, with high fragmentation and smaller-scale players that use mostly manual labor.

All nine of the shifts described in this report are expected to positively shape future demand for off-site construction in one way or another, and the collective effect is expected to be the most significant throughout the ecosystem. Led by a product-based approach, standardization, and sustainability, the coming years will see a shift to manufacturing a broad range of products off site, typically on a manual or automated production line.

Off-site construction will see much deeper integration even in flat-pack design such as doors, windows, and fully preinstalled mechanical, electrical, and plumbing systems. Data will be flexibly transferred from BIM models to the automated factory controls for decent customization within standard designs. The level of integration and connector technology aims to require no skilled labor on final site and enable extremely fast building times. There will be a mix of flat-pack (2-D) and volumetric (3-D) preconstruction. These products will then be assembled on site.

Through close collaboration with designers or integrated R&D departments, companies will build standardized libraries of subsystems that allow mass customization. We expect off-site construction companies to specialize by end-user segments, such as hospitals or certain types of bridges. Branding will become important to differentiate offerings.

While off-site construction is expected to enjoy massive demand growth in the near term, owner expectations and requirements will increase. For instance, the capability to integrate sustainability and new materials, and flexible automation (minimal work on final site) will become increasingly important criteria. As a consequence, the future landscape is likely to look very different from today's. Players that can differentiate at scale will stand head and shoulders above others. Some of those players might already exist in today's ecosystem, while others might arise from new entrants that see opportunities in areas such as real estate or infrastructure being the next platform for the deployment of smart technology. Either way, future winners will look very different from the players that exist today.

We expect off-site construction companies to specialize by end-user segments, such as hospitals or certain types of bridges. Branding will become important to differentiate offerings.



Thought starters for

Off-site construction players



Transform

Achieve economies of scale?

- Build a sufficiently large factory and secure sufficient throughput for ensuring repeatability, learning, and volume savings on procurement?
- Secure a robust pipeline of projects to ensure capacity utilization, possibly via strategic partnerships?
- Specialize to target individual segments and align operating model accordingly?
- Select a segment to pursue and develop products accordingly?
- Set up production system in line with the targeted segment, such as maximizing cost efficiency or allowing flexibility?

Optimize design for production and logistics?

- Apply design thinking to enable production efficiencies?
- Develop standardization that allows for mass customization?
- Design and choose materials to make transport and assembly easier?

Build digital end-to-end processes?

- Set up digital front ends for customers to allow mass customization?
- Deploy digital platform that supports processes—including customer interaction, design, machinery and tools, logistics, assembly, quality control, and facility management?

Optimize production footprint in line with markets and characteristics of products or logistics?

- Depending on the volume and weight of modules and the geographic scale, consider (1) setting up one large plant to cover an entire region; (2) building the plant close to project sites (if serving megaprojects); (3) focusing on metropolitan areas (large capitals); or (4) serving all sites by setting up offshore manufacturing in low-cost locations?

Build ecosystem of partners?

- Build close relationships with banks to ensure customer access to financing?
- Build partnerships with contractors for assembly and project risk management and with developers for early involvement in projects?

Automate production in factory, including Industry 4.0 setup and mass customization?

- Introduce manufacturing robotics and other automation technology?
- Ensure a steady demand pipeline to utilize factory, and plan facilities to use automation effectively?
- Procure materials and components that allow for automated production (with higher accuracy)?
- Build automated factory with inherent flexibility to allow for customization while maintaining cost efficiency?

Disrupt

Integrate and own the entire value chain?

- Acquire design capabilities to manage the development of standardized products tailored for the manufacturing process?
- Integrate materials supply at the back end of the value chain to help capture gains from standardization, and internalize distributor and original-equipment manufacturers' markups?
- Leverage increased control to become best in class at logistics, and apply just-in-time delivery to avoid unnecessary on-site storage?
- Influence the establishment of industry standards for the design of subsystems?
- Build own development capabilities, and raise funds for ownership and operations of assets?

Build a library of subsystems and try to influence the emergence of standards (after reaching scale)?

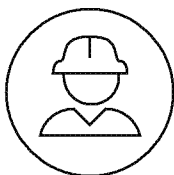
Make a data and TCO play?

- Capitalize on TCO advantages by including and managing sensors and smart building components; also combine scale and as-built BIM models for predictive maintenance?
- Capture and leverage building-use data?



Thought starters for

General contractors



General contractors face major commoditization risk. Part of the construction work looks set to shift off site. And digitalization, specialization, and productization will reduce risks and project-management needs and thus hollow out the role of the general contractor. However, the demand for the traditional contractor role is expected to continue for projects less affected by shifts, such as renovation or other areas with low levels of standardization. To differentiate themselves from the competition, general contractors will focus on becoming lean in execution to manage costs, closely integrate with the rest of the value chain, and specialize on end-user segments and subsegments within an asset class.

Transform

Specialize and win in segments in which shifts are less prevalent and expertise creates an edge?

- Identify an end-user segment in which expertise generates a competitive edge, such as hospitals, and become best in class?
- Focus on segments expected to be less affected than others by shifts, such as refurbishment, or segments expected to have low levels of standardization?
- Leverage off-site manufacturing to the extent applicable in the segment, but maintain the contractor role and learn to manage risk very well?

Focus on operational excellence and lean execution of on-site assembly?

- Focus on the assembly of construction products that are manufactured off site using in-house labor?

- Create partnerships with off-site manufacturers to leverage relationships and increase efficiency in execution?
- Build best-in-class on-site logistics capabilities?
- Clean up project portfolio in terms of end markets, services offered, and geographic footprint to enable a focus on operational excellence?
- Improve risk-management and value-assurance processes and lean capabilities?
- Reconsider the balance of different types of projects in portfolio (e.g., size of contracts, types of contracts) to adjust risk in portfolio?

Develop value-added offerings and move forward along the value chain?

- Implement capabilities for performance-based contracting and offer consulting services to customers?

Disrupt

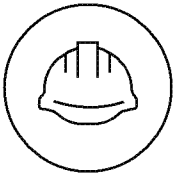
Drive integration along the value chain and lead in module development and production or new building approaches?

- Benefit from role as main point of contact for owners or developers to integrate design and engineering, module production, and related logistics activities?
- Design new building systems and use innovative building approaches using a mixture of prefabricated standardized elements and on-site solutions to sustainably raise competitiveness?
- Strive to become responsible from the beginning to the end, taking care of everything from customers' capex portfolios to managing assets?



Thought starters for:

Specialist contractors



Today, **specialist contractors perform** discrete tasks in construction projects, such as mechanical, electrical, and plumbing work. In the future, a significant share of activity for specialist contractors will move to factories and lower-cost workers. Digitalization and direct-procurement approaches may remove markups and incentives on materials purchases common in some markets in the industry. Much like general contractors, they will have to expand their skills, excel in execution, integrate with companies throughout the value chain, and consider further specialization (for example, particular types of buildings)—or prepare for downscaling.

Transform

Master digitalization and operational excellence?

- Prepare for seamless integration into value chain via digital interfaces and process support?
- Focus on operational excellence, including lean and logistics or procurement, to stay ahead of competition?

Specialize and consolidate or rightsize?

- Achieve scale in the market, matching or exceeding the scale of contractors, to sustain bargaining power vis-à-vis general contractors, distributors, and component manufacturers?
- Consider pursuing sell-side M&A?
- Specialize in high-value-adding segments and activities or those less affected by the shifts, such as refurbishment work or complex projects?

Disrupt

Enter subsystem integration or production?

- Partner with manufacturers or build own capabilities in subsystem supply—e.g., 3-D printing of facade elements or module connectors in water and electricity?
- Increase level of subassembly off site to enable standardization of installation process?

Become an adviser to help customers optimize their energy usage?

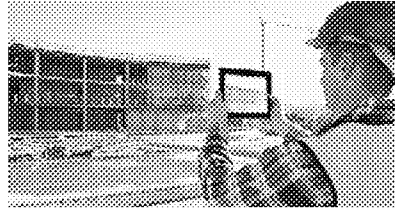
- Become an adviser to those selecting components, materials, and solutions that optimize energy usage for customers?

Disruptive change is coming to the construction industry, and it will affect every segment of the value chain. Revenues and value added will be redistributed, in some cases dramatically. Disruption has already begun and will only accelerate in the coming years, and the COVID-19 crisis has created an overriding urgency to act—and act decisively. Executives, investors, and policy makers alike will need to step up the scale and pace of their responses and make sure they stand on the right side of a transformation that will create new winners but also leave many behind.

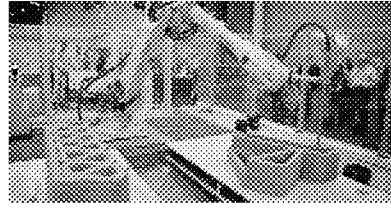
Related reading

Further insights

McKinsey publishes on issues of interest to industry executives. Our recent articles include:



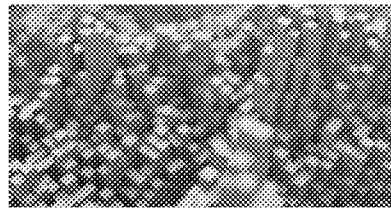
How construction can emerge stronger after coronavirus



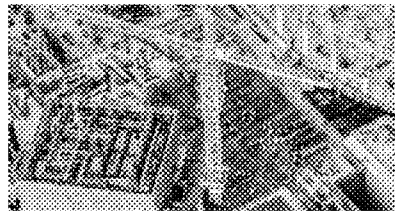
The impact and opportunities of automation in construction



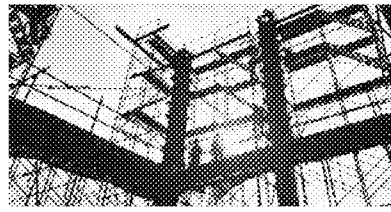
Collaborative contracting: Moving from pilot to scale-up



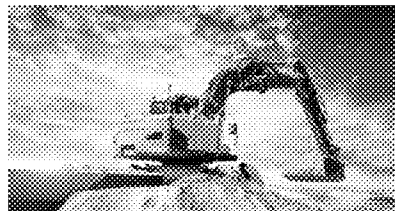
Decoding digital transformation in construction



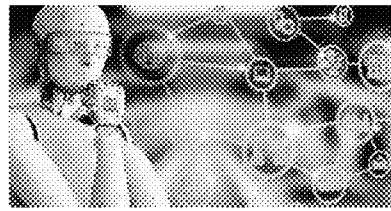
How analytics can drive smarter engineering and construction decisions



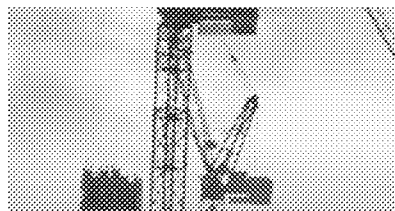
Fine-tuning the growth engine: M&A in engineering and construction



Seizing opportunity in today's construction technology ecosystem



Artificial intelligence: Construction technology's next frontier



Bridging infrastructure gaps: Has the world made progress?

Endnotes

- ¹ “Reinventing construction through a productivity revolution,” McKinsey Global Institute, February 2017.
- ² EBIT margins are calculated on revenues, not value pools. We see large variances in EBIT margins among players; the profitability of some segments and regions is considerably higher than that of others..
- ³ “Bridging infrastructure gaps: Has the world made progress?” McKinsey Global Institute, October 2017.
- ⁴ “Tackling the world’s affordable housing challenge,” McKinsey Global Institute, October 2014.
- ⁵ Hostile environments include outdoor construction sites, places where weather conditions vary, and dangerous settings such as high-rise buildings. Nonhostile environments are typically indoors, on the ground, and measurably safer.
- ⁶ An example would be a student residency “product,” with a standardized design and service package, built from modules produced in factories.
- ⁷ “Skill shift: Automation and the future of the workforce,” McKinsey Global Institute, May 2018.
- ⁸ “Climate risk and response: Physical hazards and socioeconomic impacts,” McKinsey Global Institute, January 2020.
- ⁹ Nick Bertram, Steffen Fuchs, Jan Mischke, Robert Palter, Gernot Strube, and Jonathan Woetzel, “Modular construction: From projects to products,” June 18, 2019, McKinsey.com.
- ¹⁰ EBIT margins are calculated on revenues, not value pools. We see large variances in EBIT margins among players; the profitability of some segments and regions is considerably higher than that of others.
- ¹¹ Alex Abdelnour, Amna Hussein, Nick Malik, Steve Reis, Nick Santhanam, and Sri Swaminathan, “The coming shakeout in industrial distribution: And five keys to outperformance in the years ahead,” McKinsey, May 2019, on McKinsey.com.
- ¹² “Reinventing construction through a productivity revolution,” February 2017.
- ¹³ “The view: Insolvency outlook,” Euler Hermes, January 2019.
- ¹⁴ “Insolvencies soar as business confidence dwindles,” *Financial Times*, July 30, 2019.
- ¹⁵ Rajat Agarwal, Shankar Chandrasekaran, and Mukund Sridhar, “Imagining construction’s digital future,” June 2016, McKinsey.com.
- ¹⁶ “Bridging infrastructure gaps,” October 2017.
- ¹⁷ “Tackling the world’s affordable housing challenge,” October 2014.
- ¹⁸ Shannon Bouton, David Newsome, and Jonathan Woetzel, “Building the cities of the future with green districts,” May 2015, McKinsey.com.
- ¹⁹ Joanna Lehne and Felix Preston, “Making concrete change: Innovation in low-carbon cement and concrete,” Chatham House, June 13, 2018, chathamhouse.org.
- ²⁰ “Housing in Sweden: An overview,” Turner Center for Housing Innovation, UC Berkeley, November 2017.
- ²¹ Filipe Barbosa, Jan Mischke, and Matthew Parsons, “Improving construction productivity,” July 2017, McKinsey.com.
- ²² Andreas Behrendt, Nicolai Müller, Peter Odenwälder, and Christoph Schmitz, “Industry 4.0 demystified—lean’s next level,” March 30, 2017, McKinsey.com.
- ²³ Hostile environments include outdoor construction sites, places where weather conditions vary, and dangerous settings such as high-rise buildings. Nonhostile environments are typically indoors, on the ground, and measurably safer.
- ²⁴ “Modular construction,” June 2019.
- ²⁵ An example would be a student residency “product,” with a standardized design and service package, built from modules produced in factories.
- ²⁶ “Skill shift,” May 2018.
- ²⁷ “Climate risk and response: Physical hazards and socioeconomic impacts,” McKinsey Global Institute, January 2020.
- ²⁸ *EU Industrial R&D Investment Scoreboard*, December 2019.
- ²⁹ EBIT margins are calculated on revenues, not value pools. We see large variances in EBIT margins among players; the profitability of some segments and regions is considerably higher than that of others.
- ³⁰ Gartner IT Key Metrics Data Summary report, December 2019.
- ³¹ “The coming shakeout in industrial distribution,” May 2019.
- ³² Includes funding from venture-capital firms, private-equity firms, and investment banks but does not include Chinese companies.
- ³³ Chris Bradley, Martin Hirt, and Sven Smit, “Strategy to beat the odds,” *McKinsey Quarterly*, February 13, 2018, McKinsey.com.
- ³⁴ “The coming shakeout in industrial distribution,” May 2019.

Copyright © McKinsey & Company. All rights reserved.

McKinsey.com

FTC_AR_0000970

The Labor Market Effects of Legal Restrictions on Worker Mobility*

Matthew S. Johnson
Duke University

Kurt Lavetti
Ohio State University

Michael Lipsitz
Miami University

June 9, 2020

Abstract

We analyze how the legal enforceability of Noncompete Agreements (NCAs) affects labor markets. Using newly-constructed panel data, we find that higher NCA enforceability diminishes workers' earnings and job mobility, with larger effects among workers most likely to sign NCAs. These effects are far-reaching: examining local labor markets that cross state borders reveals that enforceability affects workers' earnings in different legal jurisdictions. Revisiting a classic model of wage-setting, we find that—in contrast to prior evidence—workers facing high enforceability are unable to leverage tight labor markets to increase their wage. Finally, higher NCA enforceability exacerbates gender and racial wage gaps.

JEL Codes: J31, J42, K31, M55.

*Emails: matthew.johnson@duke.edu; lavetti.1@osu.edu; malipsitz@gmail.com. We thank Kevin Lang, Eric Posner, Johannes Schmieder, and Evan Starr for their helpful comments, as well as audience members at Miami University, Carolina Region Empirical Economics Day, the Annual Meeting of the Society of Labor Economists, UC Davis, University of Hawaii, Duke University, the Russell Sage Foundation Non-Standard Work Meeting, the Federal Trade Commission, the Department of Justice, the Department of the Treasury, and the Bureau of Labor Statistics. We appreciate Tristan Baker for invaluable research assistance. This work has been supported by Grant # 1811-10425 from the Russell Sage Foundation and the W.K. Kellogg Foundation. Any opinions expressed are those of the principal investigators alone and should not be construed as representing the opinions of either Foundation.

1 Introduction

There is growing consensus that the US labor market has failed to produce economic gains for the majority of workers in recent years. Average real hourly wages have changed little over four decades,¹ and the share of income accruing to labor declined from 65 percent in the late 1940s to 63 percent in 2000, before accelerating downward to 58 percent in 2016.² Various forces have been posited to underlie these trends, including the decline of labor unions, the rise of superstar firms (Autor et al., 2017), and the rise of domestic outsourcing (Weil, 2014; Goldschmidt and Schmieder, 2017).

Another potential explanation that has received increasing scrutiny is firms' use of postemployment restrictions, the most salient of which are noncompete agreements (NCAs). NCAs contractually limit a worker's ability to enter into a professional position in competition with his or her employer in the event of a job separation. NCAs are common: Starr et al. (2018) find that 18 percent of workers in 2016 were bound by NCAs, and nearly 40 percent of workers had signed an NCA at some point in their career. NCAs may hinder wage growth by limiting workers' ability to seek higher-paying jobs, by suppressing their ability to negotiate higher wages at current jobs, or by decreasing labor market churn. At the same time, NCAs might increase incentives for firms to invest in training, knowledge creation, and other portable assets (Rubin and Shedd, 1981) that could increase their workers' productivity and earnings.

The extent to which NCAs are legally enforceable is determined by state law. Despite growing momentum from policy makers at state and national levels³ to amend the enforceability of NCAs, there remains an incomplete understanding of the labor market effects of NCAs. At least two factors have limited existing research. One factor is a lack of comprehensive panel data on NCA enforceability. Researchers have, to date, relied largely on cross-sectional measures of states' enforceability or a small handful of changes. This approach has drawbacks: cross-sectional variation in enforceability might be correlated with other unobserved differences across states, cross-sectional measures based on a single year can introduce measurement error if NCA laws change over time, and a small handful of changes may not generalize

¹Desilver, Drew, "For Most U.S. Workers, Real Wages Have Barely Budged in Decades," *Pew Research Center*, August 7, 2018.

²President's Council of Economic Advisors Issue Brief "Labor Market Monopsony: Trends, Consequences, and Policy Responses" October 2016.

³The Workforce Mobility Act of 2018 (US Senate Bill 2782, introduced by Chris Murphy) states "No employer shall enter into, enforce, or threaten to enforce a covenant not to compete with any employee of such employer" (<https://www.congress.gov/bill/115th-congress/senate-bill/2782/text?r=6>). The Freedom to Compete Act of 2019 (US Senate Bill 124, introduced by Marco Rubio) has similar language (<https://www.congress.gov/bill/116th-congress/senate-bill/124/all-info>).

to the population. A second limitation to existing research is that identifying the mechanisms through which NCAs affect labor markets has proven elusive. Without a clear understanding of *why* NCA enforceability affects workers, it is difficult to translate empirical evidence into specific policy recommendations.

We present new evidence on the effect of NCA enforceability on workers' wages and job mobility. First, we construct a new panel dataset to use *within-state* changes in NCA laws to identify the overall labor market effects of NCA enforceability. Second, we illustrate one mechanism through which NCA enforceability—by increasing the costs of mobility—affects earnings. Finally, informed by prior evidence of differences in bargaining power and wage-setting across worker demographics, we show that the earnings effect of NCA enforceability exhibits economically meaningful heterogeneity across demographic groups, contributing a new insight to the determinants of wage inequality in the United States.

To identify the effects of NCA enforceability, we create a new dataset with annual measures of NCA enforceability for each of the 50 US states and the District of Columbia from 1991 to 2014. These data include both judicial and legislative decisions that change state-level NCA enforceability, coded to match the criteria developed by legal scholars to quantify enforceability. The vast majority of these law changes (91.4%) occur due to judicial decisions via court rulings. An important component of the judicial process is *stare decisis*, or the doctrine of precedent. A consequence is that judges are more constrained than legislators in allowing economic or political trends to affect decisions, a fact that is useful for our research design. We combine our enforceability dataset with earnings and mobility outcomes from the Current Population Survey, the Job to Job Flows dataset, and the Quarterly Workforce Indicators dataset, all from the US Census Bureau.

We find that increases in NCA enforceability decrease workers' earnings and mobility. Moving from the 10th to 90th percentile in enforceability is associated with a 3-4% decrease in the average worker's earnings. The earnings effects are almost entirely driven by declines in implied hourly wages. The effect is even stronger among occupations, industries, and demographic groups in which NCAs are used more frequently (according to Starr et al. (2018)). We also find that NCA enforceability reduces worker mobility, particularly when NCAs are used more frequently. Under some assumptions, our estimates imply that rendering NCAs unenforceable nationwide would increase average earnings among *all* workers by nearly 8.5%. This wage increase is approximately equal to the estimated wage premium that accrues to workers who enter occupations with government-mandated licensing, and roughly half the size of the wage premium associated with membership in a labor union.

Though only a fraction of workers actually sign NCAs, the use of NCAs might create externalities on other workers by reducing labor market churn or increasing recruitment costs (Starr et al., 2018). To test whether such externalities exist, we show that NCA enforceability laws generate spillover effects within local labor markets on workers in different legal jurisdictions. Focusing on local labor markets that are divided by a state border, we test whether a change in NCA enforceability in one state indirectly affects the earnings and mobility of workers located in an adjoining state. We find an indirect effect of enforceability that is nearly 90% of the average direct effect on workers in the state that experienced the law change. This large spillover effect suggests that the “treatment” of NCA enforceability affects a larger population than the relatively small share of workers actually bound by NCAs.

To investigate the mechanisms through which NCA enforceability reduces earnings, we posit that stricter enforceability hinders workers’ ability to leverage improvements in labor market conditions to negotiate wage increases. We revisit a longstanding theory that wages are determined under a model of implicit contracts between workers and employers. Beginning with the seminal paper of Beaudry and DiNardo (1991), this literature has consistently found that workers’ wages rise when their outside option improves: a worker’s current wage is more strongly affected by the minimum unemployment rate over the course of her job spell than by the initial unemployment rate at the beginning of the spell. This finding implies that the cost of job mobility for workers is low. Because (enforceable) NCAs by construction raise the costs of job mobility, however, it is plausible that this relationship is dependent on states’ NCA policies.

Using more recent CPS data, we show that the result from Beaudry and DiNardo (1991) still holds *on average*, but it flips in states with strongly enforceable NCA policies. In these states, the minimum unemployment rate has essentially no effect on a worker’s current wage, and the initial unemployment rate has a much stronger effect (consistent with contracts negotiated under costly mobility). In contrast, in states with weakly enforceable NCA policies, the effect of the minimum unemployment rate over a job spell on the current wage is even more pronounced, even conditioning on the initial unemployment rate (consistent with very costless mobility). These findings imply that strict NCA enforceability erodes workers’ ability to leverage tight labor markets to achieve higher earnings, and is consistent with the hypothesis that NCAs “undermine workers’ prospects for moving up the income ladder” (Krueger, 2017).

Finally, we document economically meaningful heterogeneity in the earnings effect of NCA enforceability across demographic groups. NCAs might restrict outside options more for women than for men, due for example to gender differences in will-

ingness to commute (Le Barbanchon et al., 2019), and differences in men and women’s willingness to negotiate (Bertrand, 2011) could imply that women are less willing to violate the terms of their NCA than are men. Similar evidence has been found for non-white workers relative to white workers. More broadly, to the extent that enforceable NCAs decrease the competitiveness of labor markets, they may endow firms with monopsony power to price discriminate among their workers (Robinson, 1933). Consistent with this evidence, we find that stricter NCA enforceability reduces earnings for female and for non-white workers by twice as much as for white male workers. Neither of these effects are explained by differences in occupations, industries or education across groups. Using a standard wage decomposition, our estimates imply that the 90-10 differential in NCA enforceability accounts for 3.6-9.1% of the earnings gaps between white men and other demographic groups.

Our findings contribute to a growing literature on the effects of NCA enforceability. Recent studies using cross-sectional variation have estimated that greater NCA enforceability reduces workers’ earnings (Starr, 2018; Starr et al., 2018), though others have found opposing evidence in some high-skilled labor markets such as doctors (Lavetti et al., 2018) and CEOs (Kini et al., 2019). A more consistent finding in this literature is that NCA enforceability reduces mobility (Marx et al., 2009; Garmaise, 2011; Starr et al., 2018). Other papers have used cross-sectional variation to test how enforceability moderates the employment effects the minimum wage (Johnson and Lipsitz, 2019) or studied specific law changes to analyze effects of enforceability on subgroups of workers, like knowledge workers (Marx et al., 2015, 2009), managers (Garmaise, 2011), and hourly workers (Lipsitz and Starr, 2019). Our paper is the first to examine labor market effects using a comprehensive set of all NCA law changes between 1991 and 2014, and, to the best of our knowledge, the first study to empirically demonstrate one way that NCA policies fundamentally change how workers and employers bargain over wages.

Our findings also contribute to several other literatures. First, a growing literature investigates the effects of employer power in labor markets. Many studies have found evidence consistent with local employer concentration affecting wages (for example, Azar et al. (2017), Benmelech et al. (2018), Prager and Schmitt (2019)), and other studies find other ways that employers have monopsony power (Dube et al., 2018). Our results imply that NCAs skew power dynamics in employment relationships in favor of the employer by diminishing the worker’s outside options, giving employers effective market power over the workers that they have hired. Second, our findings provide new insight to a longstanding debate in law and economics regarding freedom of contracting (see, e.g., Bernstein (2008) for an overview). Advocates of the freedom

of contract argue that the ability to freely enter into contracts increases economic efficiency, as contracts that decrease economic efficiency would leave one or both actors worse off, and would therefore not be signed in the first place. While we are unable to examine the efficiency effects of NCA enforceability on those that sign NCAs, our findings of substantial externalities of NCA enforceability reveals a shortcoming in this argument to justify the allowance of NCAs from an efficiency standpoint.⁴ Finally, our work complements a literature analyzing the effects of NCA enforceability on outcomes outside of the labor market, such as corporate investment (Jeffers, 2018), entrepreneurship (Marx, 2018), and knowledge spillovers from patents (Belenzon and Schankerman, 2013).

2 Conceptual Framework

To develop a conceptual framework to guide our empirical analysis of how NCA enforceability affects wages, we must first be explicit about two things.

First, it is useful to define what NCAs *do*. NCAs, by construction, limit workers' mobility in a specific way. While the exact terms of an NCA are contract-specific, a central reason that an employer has a worker sign an NCA is to prevent her from moving jobs to a competing firm. The definition of "competing" depends on the nature of production. In an industry in which client lists are essential to production, an NCA might dictate that the worker cannot depart for another employer in the same industry and within a specified geographic radius (e.g. within 25 miles or the same state); in an industry in which trade secrets are essential for firms to retain a competitive edge, the NCA might dictate that the worker cannot depart for another employer in the same industry anywhere in the country. No matter the specifics, the NCA is intended to limit the worker's outside options.

Second, we must be explicit that our focus is on effects of NCA *enforceability*, not NCA *use*. If use of enforceable NCAs is the relevant treatment, our analysis presumes that changes in enforceability will alter the *intensity* of that treatment on the labor market. This change could happen two ways.

First, changes in enforceability could impact treatment intensity by changing the share of workers bound by NCAs. We do not observe panel data on individual-level NCA use, so we cannot directly test whether this is the case. However, there is existing evidence that enforceability shifts the incidence of NCA use. In the cross-section, states with higher NCA enforceability have a larger share of physicians (Lavetti et al., 2018), CEOs (Kini et al., 2019), and hair stylists (Johnson and Lipsitz, 2019) that sign

⁴In this regard, we complement the findings from Starr et al. (2018), who illustrate externalities of NCA *use* in a different approach.

NCAAs. This evidence is not unanimous, however: Starr et al. (2018) find essentially no difference in NCA use by states’ enforceability in a representative sample of US workers.

NCA enforceability could also change treatment intensity on the intensive margin—by altering the effect of treatment (signing an NCA) on the treated (those who have signed one). Indeed, though NCAs are used in states that render them unenforceable (e.g. Starr et al. (2018)), it is plausible that the extent to which an employer can leverage a worker’s NCA depends on enforceability.⁵

Having articulated what NCAs are intended to do, and what we are measuring by focusing on enforceability, we turn to the focus of this section: how we expect NCA enforceability to affect wages. For this purpose, we embed NCAs into the job search model of the labor market developed in Bagger et al. (2014). This model decomposes workers’ wage growth over the course of their careers into contributions of human capital accumulation and job search, the latter including both “between-job” (moving from lower-paying to higher-paying firms) and “within-job” (using competing offers to negotiate for pay increases from their current employer) wage growth. This model provides a natural framework to consider how NCAs affect wages, as its focus on the role of human capital accumulation versus job search highlights two competing channels through which NCAs could affect wages, which we elaborate on below.

In the remainder of this section, we provide economic intuition for multiple channels through which NCAs and earnings might interact. In Appendix A, we formalize this intuition by extending the model in Bagger et al. (2014) to embed NCAs into a search model of the labor market.

We organize our framework by considering:

1. *Direct* effects of NCA enforceability on earnings of workers that sign NCAs
2. *Indirect* (spillover) effects of NCA enforceability on earnings of other workers
3. *Differential* effects of NCA enforceability across different groups of workers

2.1 Direct Effects on Workers that Sign NCAs

To the extent that enforceability affects the incidence of NCA use, the direct effect of signing an (enforceable) NCA on wages reflects multiple competing channels.

⁵Note that this argument holds even if a worker is not fully informed about the enforceability of the NCA she has signed. As long as employers *are* informed, and there is some probability that workers can learn, then employers will know the NCA has less bite in expectation when it is not legally enforceable.

Human Capital Accumulation

A common justification for NCAs is that they increase employers' incentives to make productivity-enhancing investments in their employees. Employers might be reluctant to invest in "transferable" assets, such as general human capital or client lists, that an employee could take with her in the event of a departure, since the employer is unlikely to recoup the full value of the investment (Grossman and Hart, 1986; Williamson, 1975). By preventing an employee from departing to a competitor or founding her own competing firm, an NCA can alleviate this investment "hold-up." NCAs would then clearly benefit firms by increasing quasi-rents associated with higher investment and productivity. As long as a portion of these quasi-rents are passed through to employees, then NCAs—by inducing firm-sponsored investment—would translate to higher wages.⁶

In Appendix A, we model the human capital effects of NCAs by extending Bagger et al. (2014) to incorporate faster human capital accumulation for workers who sign NCAs. All else equal, faster human capital accumulation drives faster wage growth for workers with NCAs, contributing to Proposition A.2, which presents conditions under which wage growth is faster with or without an NCA.

Reducing Workers' Outside Options

Workers' outside options have a large effect on their wage (Caldwell and Harmon, 2019). An NCA, by construction, limits a worker's options for future job mobility. Workers bound by enforceable NCAs are unable to move to competing firms (or, at least, face higher costs of doing so). This reduction in expected *realized* mobility reduces workers' expected future wage trajectory, in so much as it prevents them from accepting jobs at higher-paying firms (Haltiwanger et al., 2018) and prolongs exposure to potential negative match-specific wage shocks (Liu, 2019).

Similarly, because an NCA reduces a worker's *threat* of departure, it reduces her ability to leverage that threat to get a pay increase from her current employer. This reduced threat point likely meaningfully affects a worker's wage trajectory: Bagger et al. (2014) find that the wage effects of such *within-job* search dominate the wage effects of *between-job* search.

Moreover, the wage penalty associated with an inability to leverage competing offers is likely to be particularly pronounced under certain labor market conditions. Workers that begin job spells during a period in which the labor market is weak can leverage subsequent improvements in labor market conditions as long as their costs of mobility are low: this is because a worker cannot commit to a wage contract

⁶The extent of such pass-through of quasi-rents to wages is likely to depend on many factors, such as the ease with which a worker can be replaced (Kline et al., 2019).

that reflects a weak outside option at the time of hire if her outside option suddenly improves (Beaudry and DiNardo, 1991). Because an NCA raises a worker's cost of mobility, it would leave her unable to bargain for a pay increase commensurate with the strength of the labor market.

In Appendix A, we extend Bagger et al. (2014) by limiting the offer arrival rate for employed workers with an NCA, which decreases wage growth for those workers.⁷ This modification contributes to Proposition A.2, which describes when NCAs increase or decrease earnings: by limiting workers' ability to bargain for earnings increases by leveraging outside offers, and their ability to accept new, higher-paying jobs, (enforceable) NCAs reduce the wage gains that accrue through job search.

Compensating Differentials

A forward-looking and well-informed worker, recognizing the future diminished value of job search she would bear by signing an NCA, would require an initial compensating wage differential to accept one. In a frictionless model, competition drives the size of the compensating differential close to the expected discounted opportunity cost. In a model with search frictions, a worker receiving a job offer with an NCA will accept it if the total value of the job offer is greater than the value of remaining unemployed; in this case the size of the compensating differential would depend on parameters such as the arrival rate and value of unemployment. Any of these models would predict that due to a limitation of outside offers, workers with NCAs would have lower wage *growth*, but potentially higher initial *levels* of compensation.

A prerequisite for the existence of a compensating differential is that workers must be aware that they are signing an NCA at the time of initial negotiation. There is evidence that this condition is not always met: firms frequently ask workers to sign NCAs *after* accepting their job (Marx, 2011; Starr et al., 2018), which renders workers unable to bargain for a compensating differential for accepting an NCA. Furthermore, workers must fully anticipate the costs of future decreased mobility to bargain for a compensating differential, which might not be the case if workers are myopic or have very high discount rates (Greenwald, 1986).

Under the assumption that workers understand that they sign an NCA, and understand its costs, a compensating differential can arise in the model of Bagger et al. (2014). In that model, workers' pay when entering a new match is set to guarantee them a proportion, β , of future match-specific rents. Since an NCA reduces their future value stream (by limiting their gains from job search), the initial wage

⁷Another way to represent how NCAs reduce the value of job search would be to impose a cost (e.g., a buy-out payment) on workers for moving to a new firm. As discussed in Appendix A, both approaches generate the same qualitative prediction.

will increase to guarantee the same proportion of match-specific rents accrues to the worker.⁸

2.2 Indirect (Spillover) Effects on Other Workers

The use of (enforceable) NCAs by some firms could affect the wage not just of the workers that sign them, but also have spillover effects on other workers in the same labor market.⁹ Such spillover effects may arise for several reasons.

First, the prevalence of NCAs in the population might affect offer arrival rates to workers searching for jobs. One reason that this might happen is that NCAs thin labor markets: when a firm hires a worker with an (enforceable) NCA, the worker becomes effectively removed from the the pool of potential hires for other firms. The firm also, to an extent, becomes removed from the pool of searching firms; this is because the firm has a worker who a) it might have already made a costly investment in, and b) is unable to leverage outside options to negotiate for wage increases—both of which make that worker more valuable moving forward than hiring a new worker. If the worker has been paid a compensating differential, that cost is likely sunk, as well. In thinner labor markets, workers and firms match less often, which drives down equilibrium wages (Bleakley and Lin, 2012; Gan and Li, 2016).

Another reason that NCA enforceability could reduce arrival rates for job seekers is that NCA use by some firms can increase recruitment costs for all firms (Starr et al., 2018). Firms are unlikely to be able to directly observe whether a job applicant is currently bound by an NCA, which would (in expectation) slow down the recruiting process and decrease the value of posting vacancies (Starr et al., 2018). More generally, NCA use might reduce overall labor market dynamism, which can decrease wages by slowing workers' ability to find higher-paying employers (Haltiwanger et al., 2018) and be decreasing the offer arrival rate (Bagger et al., 2014).

Relatedly, NCA use could decrease the number of searching firms, increasing local firms' market power (in, e.g., a Cournot sense). Enforceable NCAs decrease entrepreneurship (Starr et al., 2018) and new firm entry (Jeffers, 2018). These effects would increase local labor market concentration, which a wide literature has shown depresses wages (e.g. Azar et al. (2017); Arnold (2019); Jarosch et al. (2019)).¹⁰

⁸We note that, while the Bagger et al. (2014) easily accommodates an NCA initial compensating differential for the reduced value of job search, it does *not* accommodate one for workers' increased growth rate of human capital development. In other words, in this model, if worker A has a faster projected rate of human capital development than worker B, but the two workers are otherwise identical, both workers will have identical starting wages. See the paragraph immediately following Equation 6 in Bagger et al. (2014) for discussion.

⁹Our discussion in this section draws heavily from Starr et al. (2018).

¹⁰It is also possible that enforceable NCAs could have *positive* spillover effects on wages: if NCAs

As we describe in Appendix A.4, we formalize this spillover effect by extending Bagger et al. (2014) to assume that the job offer arrival rate depends positively on market thickness. NCAs, by causing labor markets to thin, decrease arrival rates to other workers, decreasing the steady-state distribution of wages for workers. Proposition A.4 describes how this spillover effect affects initial wages, and the ensuing discussion describes how it affects wage growth.

2.3 Differential Effects of NCA Enforceability Across Groups of Workers

NCA enforceability could have a stronger impact on earnings for some groups of workers than others—even holding prevalence of NCA use constant. For example, even though NCA use is not systematically different for male and female workers (Starr et al., 2018), NCA enforceability may have a differential effect on the wages of men and women for several reasons.

First, evidence suggests that NCAs are implicitly more binding for women than for men. Marx (2018) finds that strict NCA enforceability decreases entrepreneurship among women at a greater rate than among men, which the author attributes to male employees facing lower relative costs of violating an (enforceable) NCA than female employees, e.g. via different expected costs or benefits of litigation.

Second, the reduction to workers’ outside options that is caused by NCAs could result in a larger wage penalty for women than for men. Most directly, NCAs might reduce female workers’ outside options more than male workers’: NCAs often have a geographic component, and men tend to be more willing to commute far distances for their job than women (Le Barbanchon et al., 2019). Furthermore, to the extent that NCAs create positive quasi-rents for employers (e.g. by increasing productivity-enhancing investment), there is evidence that firms share these rents to a greater extent with male workers than with female workers (Black and Strahan, 2001; Card et al., 2015; Kline et al., 2019). On the other hand, improvements in outside options lead to higher wages for both male and female workers (Caldwell and Harmon, 2019; Caldwell and Danieli, 2018). Thus, any positive effects of NCAs on wages is likely to be more muted for females than for males, whereas the negative effects should be felt by both. Strict NCA enforceability may thus exacerbate bargaining disparities between male and female workers.¹¹

reduce competition in the goods markets by limiting the supply of entrepreneurs who are potential entrants in those markets, workers’ marginal revenue product of labor would be higher because goods market prices are higher, potentially increasing wages.

¹¹Put another way, prior literature has shown that men are more likely than women to negotiate for salary increases when the “rules of wage negotiation” are ambiguous, but are equally likely to do so when employers signal clearly that negotiation is allowed (Leibbrandt and List, 2014). To the extent that receiving an outside job offer provides an unambiguous rationale to negotiate with one’s

Third, if strict NCA enforceability reduces the competitiveness of labor markets, it may give firms monopsony power to discriminate between male and female workers (Robinson, 1933; Black and Brainerd, 2004; Barth and Dale-Olsen, 2009). For each of these reasons, NCA enforceability might reduce wages more for female workers than for similar male workers.

For these same three reasons, NCA enforceability could depress earnings of racial minorities more than white workers. There is anecdotal evidence that NCAs can be more implicitly binding for black workers than their white co-workers.¹² As with gender, there is evidence that black workers face lower negotiating power in the workplace than white workers (Hernandez et al., 2018). Furthermore, the reductions in labor market competitiveness that enable firms to discriminate between workers would have a similarly negative effect on wages of racial minorities as for women.

2.4 Takeaways

The framework presented in this section organizes the empirical analyses we report in Sections 4, 5, 6, and 7. The qualitative takeaways of the model are:

1. The overall effect of NCA enforceability on earnings is ambiguous.
2. Enforceability could have spillover effects on wages of workers not bound by NCAs, and the effect is likely negative.
3. By reducing workers' ability to threaten to change jobs, NCAs reduce workers' ability to secure wage gains, particularly in tight labor markets.
4. The negative earnings effects of NCA enforceability are likely more pronounced for women and racial minorities.

Before reporting these results, we describe the institutional background of NCA enforceability, how we quantify states' history of enforceability, and the data sources we use.

employer, removing this option will also exacerbate inequalities in bargaining power between men and women in the workplace.

¹²An illustrative example is a suit brought forward by plaintiff Tracy Miller, an African American worker employed by Illinois Central Railroad. After receiving an employment offer from a competitor, Miller was told by his employer that he could not take it because he had signed an NCA. The plaintiff alleged that the same NCA went unenforced multiple times when several of his white co-workers accepted employment with other industry competitors, in clear violation of NCAs they had also signed. More information available at <https://www.bsjfirm.com> (accessed July 2019)

3 Data

3.1 State-Level NCA Enforceability Measures

The cornerstone of our project is a state-level panel dataset with annual measures of states' NCA enforceability. As documented by Bishara (2010), NCA laws vary along seven quantifiable dimensions across states and over time (see Table B.1 for a list of the dimensions). For example, one dimension (Q3a) indicates the extent to which employers are legally required to compensate workers that sign NCAs at the beginning of a job spell. Another dimension (Q8) reflects whether the NCA is enforceable when the employer terminates the employee who signed the NCA (as opposed to a voluntary separation).

Our dataset contains values representing the stringency of the law on each of these seven legal dimensions for every state between 1991 and 2014. This dataset builds from Bishara (2010), who quantified how each state's law treated each of these seven dimensions on a scale from 0 (completely unenforceable) to 10 (easily enforceable) in the years 1991 and 2009. We created the panel version by first replicating the cross-sectional scores from Bishara (2010) in 1991 and 2009 using the same primary sources: a series of legal texts titled "Covenants Not to Compete: A State by State Survey," updated annually by Malsberger. We used detailed notes and decision rules provided by Bishara (2010) to ensure our approach to quantify enforceability followed that of Bishara (2010). After replicating the cross-sectional scores, we filled in the timing of all intervening changes using the same quantification methodology, and extended the data through 2014. Our approach mirrors that of Hausman and Lavetti (2017), who created an analogous dataset for NCA enforceability specific to physicians from 1991–2009. Using the seven dimensions of enforceability, we construct a composite *NCA Enforceability Score* for each state-year from 1991–2014.¹³ These data have never previously been used to study the general labor-market effects of NCA laws.

Differences in how states interpret these dimensions lead to substantial differences in the *NCA Enforceability Score* across states. At the extreme ends of this spectrum, Florida Statute 542.335 explicitly allows the use of NCAs as long as a legitimate business interest is being protected, the agreement is in writing, and the agreement is reasonable in time, area, and line of business.¹⁴ The law allows for a large variety of

¹³Following Bishara (2010), for questions in states where no legal precedent exists, we mark the value as missing. The composite index is a weighted average of scores on each of the seven legal dimensions. When the score for a question is missing, we omit it from the calculation of that weighted average, as in Bishara (2010). Out of 8,568 year-state-question observations (24 years, 51 states, 7 questions), a total of 900 (10.5%) are missing.

¹⁴Florida Statute 542.335. Full text available at http://www.leg.state.fl.us/statutes/index.cfm?App_mode=Display_Statute&URL=0500-0599/0542/Sections/0542.335.html

protectable interests (such as trade secrets, training, and client relationships), permits the beginning of employment or continued employment to act as “consideration” (i.e., compensation) for an NCA, allows the courts to modify NCAs to make them enforceable, and renders NCAs enforceable even when an employer terminates an employee. At the other end of the spectrum, North Dakota Century Code 9-08-06 explicitly bans all NCAs in employment contracts.¹⁵ Quantifying these statutes, Florida has the highest NCA Enforceability Score during our time period (which we normalize to 1), and North Dakota has the lowest score (which we normalize to 0).

Furthermore, law changes have led to sizable changes in the NCA Enforceability Score *within* states over time. Consider, for example, a state Superior Court case in Pennsylvania: *Insulation Corporation of America v. Brobston* (1995). The case concerned an employee of an insulation sales company who had signed an NCA. After being terminated for poor performance, he was hired by a competitor of his original employer, in alleged violation of the NCA. While the NCA in question was ultimately not enforced, the court’s decision set new precedent that NCAs may generally be enforced following employer termination: “...the circumstances under which the employment relationship is terminated are an important factor to consider in assessing... the reasonableness of enforcing the restrictive covenant.”¹⁶ This case resulted in the component of the NCA Enforceability Score specific to treatment following employer termination (Q8) to change from 4 (out of 10) to 7 in Pennsylvania; the resulting change in Pennsylvania’s overall NCA Enforceability Score was equal to roughly a third of a standard deviation in the distribution across our sample period.

Table 1 summarizes differences in levels of NCA enforceability across the country and within states over time, between 1991 and 2014. There are 82 within-state NCA law changes over our sample period, and these are dispersed roughly evenly across the Northeast, Midwest, South, and West regions. The average law change results in a change in the magnitude of the NCA Enforceability Score that is about 7% of the average score over this period, and the within-state standard deviation in enforceability is equal to roughly 17% of the overall standard deviation. Our analyses rely on these within-state changes in enforceability.

Figure 1 shows the timing of NCA law change events. Changes were relatively evenly dispersed throughout the study time period. There are a few more enforceability increases than decreases, though both are well-represented. Figure 2 shows the sample-weighted mean NCA Enforceability Score across states over the sample

¹⁵North Dakota Century Code 9-08-06. Full text available at <https://www.legis.nd.gov/cencode/t09c08.pdf>

¹⁶*Insulation Corp. of America v. Brobston*, 667 A.2d 729, 446 Pa. Superior Ct. 520, 446 Pa. Super. 520 (Super. Ct. 1995).

period. NCA enforceability has been generally flat or increasing over time, with an especially steep increase during the mid to late 1990s.

3.1.1 Testing the Exogeneity of NCA Law Changes

Our ability to use within-state changes in NCA enforceability to identify its causal effect on earnings and mobility would be compromised if legal changes to NCA enforceability were correlated with states' underlying political, labor, or business characteristics that may also impact earnings growth. For instance, changes to enforceability could potentially be spurred by strong labor unions on the one hand, mobilized business interests on the other, or a general change in the business climate.

Ex ante, we expect this concern to be minimal. The majority of law changes in our sample are due to judicial decisions. In most cases, these decisions are initiated due to a legal case that is idiosyncratic to a particular occupation, industry, or employment relationship; however, the consequences of these decisions affect the state's labor law much more broadly. Relative to legislators, judges are less influenced by stakeholder pressure that could sway their decision-making because of the doctrine of *stare decisis*.¹⁷

To formally test whether NCA law changes are exogenous to such underlying forces, we estimate whether states' political, social and economic characteristics predict NCA law changes. We use a variety of data sources. These include the University of Kentucky Center for Poverty Research's National Welfare Data (University of Kentucky Center for Poverty Research, 2018) on population, workers compensation beneficiaries, an indicator for whether the state governor is a member of Democratic party, the share of state house and senate representatives (respectively) in the Democratic party, minimum wage, and the number of Medicaid beneficiaries. We also use the database constructed in Caughey and Warshaw (2018) to obtain measures of policy liberalism (liberalism in the state as reflected by government policy) and mass liberalism (liberalism in the state as reflected by responses of individuals to policy questions), both of which are measured separately on social and economic dimensions. From this dataset we also obtain the percentage of voters who identify as Democrats. For more details on the construction of these measures, see Caughey and Warshaw (2018). Next, we gather data on the ideologies of state legislatures from McCarty and Shor (2015), including the State House and State Senate ideology scores, in aggregate as well as separately by Democrats and Republicans. Finally, we include data on union membership from Hirsch and Macpherson (2019).

Table 2 presents the results from a regression in which the dependent variable is

¹⁷For a discussion of *stare decisis*, see Knight and Epstein (1996).

a state’s annual NCA enforceability, and the independent variables are each of the 20 characteristics noted above (lagged by one year), as well as state and Census division by year fixed effects (we use these same fixed effects in our subsequent analysis). Out of 20 variables, the vast majority have coefficients that are both economically and statistically insignificant. Only two of these 20 variables are statistically significant at the 10% level (the minimum wage and the State Senate Democrats ideology score), none are significant at the 5% level, and a joint F test on the statistical significance of these predictors is insignificant at the 10% level ($p = 0.184$).¹⁸ Furthermore, the partial R^2 of the model, after residualizing on division by year and state fixed effects, is 0.113, meaning that these predictors collectively explain only 11% of the variance in within-state changes to NCA policy. Thus, these results provide supportive evidence that NCA law changes are indeed exogenous to underlying economic, political, or social trends.

To complement the evidence in Table 2, we further assess the concern that NCA law changes might not be exogenous in our subsequent analysis. We use an event study analysis in Section 4.2.2 to check for pre-trends in the outcome variable, and we show that our results are qualitatively robust to controlling for all the economic and political controls used in this section.

3.2 Data on Earnings and Mobility

We gather data on earnings, employment, mobility, and other labor market outcomes from four sources: the Current Population Survey (CPS) Annual Social and Economic Supplement, the Job-to-Job Mobility dataset, the Quarterly Workforce Indicators (QWI) dataset, and the CPS Occupational Mobility and Job Tenure Supplement (JTS). We describe each of these datasets, and how they fit into our analysis, in turn.

First, we gather individual-level data on earnings and employment from the CPS ASEC (otherwise known as the March Supplement).¹⁹ The ASEC is a CPS supplement collected each March that contains information about the wage and salary income of respondents. The CPS also includes respondents’ demographic and geographic information.²⁰ We restrict the ASEC sample to include individuals who

¹⁸It is not surprising that two out of twenty predictors are statistically significant. The probability of finding two or more significant predictors (at the 10% level) out of twenty, conditional on each of the predictors having zero true effect and each being independent (which is surely not true in practice, but provides an adequate benchmark) is approximately 0.88 ($1 - 0.90^{20}$).

¹⁹Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren. Integrated Public Use Microdata Series, Current Population Survey: Version 6.0 [dataset]. Minneapolis, MN: IPUMS, 2018. <https://doi.org/10.18128/D030.V6.0>

²⁰While the ASEC is relatively small compared with, for example, the American Communities Survey (ACS), its existence precedes our earliest data on NCA enforceability (whereas the ACS

reported having worked for a private-sector employer (not self-employed) in the year prior to being surveyed. We include the years 1991 to 2014, restrict to individuals who were between the ages of 18 and 64 at the time they were surveyed, and remove observations for which earnings or hours variables have been topcoded. The resulting ASEC dataset contains approximately 1.5 million observations, 1.2 million of which represent full-time workers. We deflate earnings and wages in the ASEC using the Consumer Price Index. We match NCA enforceability measures by state and year.

Our second dataset is the Job-to-Job Flows (J2J) dataset from the U.S. Census Bureau, which we use to examine the effect of enforceability on job mobility. Derived from the Longitudinal-Employer Household Dynamics dataset,²¹ these data contain aggregate job flows between cells defined by combinations of age, sex, quarter, origin job state, destination job state, origin employer industry, and destination employer industry. We aggregate these data to the level of the state-industry-year, and we create three measures of job mobility that could potentially be affected by NCA enforceability: (1) the *total rate* of job-to-job separations per worker, (2) the share of job-to-job separations in which the separating worker's destination job is in a different state than his or her origin job, and (3) the share of job-to-job separations in which the separating worker's destination job is in a different industry than his or her origin job.

Third, we use the Quarterly Workforce Indicators (QWI) dataset from the Census Bureau. Like the J2J, the QWI is a public use file that aggregates data from the LEHD, and it contains data on earnings, as well as numbers of hires and separations, at the county-quarter level for the near-universe of private workers, stratified by sex and age group. We use the QWI both to complement the CPS in our estimation of the earnings effects of NCA enforceability, and also to investigate spillovers from enforceability. One drawback with the QWI for our purposes is that the QWI is not a balanced panel over our sample period, as some states did not begin reporting the necessary data until the late 1990s or later. For this reason, we are left with only 64 legal changes (instead of the universe of 82 legal changes) when using the QWI.

Fourth, in our investigation of the mechanism underlying the relationship between enforceability and earnings, we use data from the CPS Occupational Mobility and Job Tenure Supplement (JTS) over the years 1996 to 2014. The JTS is conducted biannually in either January or February. Among other things, it includes questions

does not). We are therefore able to leverage all changes in NCA enforceability from 1991-2014. Our results are quite similar if we instead use the ACS.

²¹U.S. Census Bureau. (2019). Job-to-Job Flows Data (2000-2019). Washington, DC: U.S. Census Bureau, Longitudinal-Employer Household Dynamics Program, accessed on April 7, 2020 at <https://lehd.ces.census.gov/data/#j2j>. Version R2019Q1.

about the respondent’s history of employment, such as “How long have you been working [for your present employer]?”²² We use responses to this question to calculate the year that the worker began his or her job spell, which allows us to match individuals to the enforceability score at the time of hire. Our outcome variable of interest is weekly earnings, and we use additional variables as controls. We merge in annual national unemployment rates between 1947 and 2014 from the Bureau of Labor Statistics website for the analysis, which we describe in Section 6.

4 The Effect of NCA Enforceability on Workers’ Earnings and Mobility

In this section, we examine the effect of NCA enforceability on earnings and mobility. We then consider whether these effects are more pronounced among workers who are most likely to have signed an NCA, and we then show that our estimates are stable to numerous robustness checks and sensitivity analyses.

4.1 Main Results on Earnings

We use a difference-in-difference design to estimate the effects of NCA enforceability on earnings, leveraging intra-state variation in NCA enforceability over time. Our basic regression model is

$$Y_{ist} = \alpha + \beta * Enforceability_{st} + X_{it}\gamma + \rho_s + \delta_{d(s)t} + \varepsilon_{ist}, \quad (1)$$

where Y_{ist} is the outcome of interest, $Enforceability_{st}$ is a state’s annual composite NCA enforceability score across the 7 dimensions described in Section 3, X_{it} is a vector of individual-level controls, ρ_s is a fixed effect for each state, and $\delta_{d(s)t}$ is a fixed effect for each Census division by year.²³ The coefficient of interest, β , is identified from changes in earnings in states that change their NCA enforceability, relative to other states in the same Census division over the same period. Standard errors are clustered by state. A key identifying assumption is $E(Enforceability_{st}\varepsilon_{ist}|\rho_s, \delta_{d(s)t}) = 0$: conditional on state and division-year effects, changes in enforceability are uncorrelated with the error term. The evidence in Section 3.1.1 supports this assumption.

Results are reported in Table 3. Columns 1-4 use data from the ASEC, restricted to full-time workers between the ages of 18 and 64 who reported working for wage

²²Note that “for your present employer” may alternatively be “for company name from basic CPS/as a self-employed person/at your main job.” See <http://www.nber.org/cps/cpsjan2016.pdf>.

²³There are 9 Census divisions that partition the United States. We include division-year fixed effects to account for potential time-varying shocks to different areas of the country. The estimated effect of *Enforceability* on earnings, corresponding to Column 1 of Table 3, is roughly 50 percent larger (-0.177, $p < .01$) if we use year fixed effects in lieu of Division by year fixed effects.

and salary income at a private employer the prior year.²⁴ The coefficient in Column 1 suggests that going from NCA enforceability of 0 (completely unenforceable) to 1 (the strictest enforceability observed in our sample) leads to an 11.0 percent decline in earnings ($\exp(-.117) - 1, p = .002$). Adding fixed effects for broad occupation codes in Column 2 diminishes the point estimate slightly but improves its precision ($p < .001$). To get a sense of the magnitude of this estimate, the 10th and 90th percentiles of *Enforceability* observed in our sample are 0.55 and 0.9, respectively. The estimates thus imply that moving from the 10th to the 90th percentile in *Enforceability* leads to a 3.5 percent average decline in annual earnings ($\exp(-.101 * .35 - 1 = 0.035)$).

A negative effect of *Enforceability* on annual earnings could reflect either a decline in hours worked or a decline in workers' implied hourly wage. In Column 3, the dependent variable is instead the log of a worker's reported weekly hours:²⁵ while the point estimate is negative, it is relatively small and statistically insignificant ($p = 0.16$). In Column 4 the dependent variable is the individual's implied log hourly wage (calculated as annual earnings divided by fifty-two times usual weekly hours). The estimated coefficient is identical to the coefficient on annual earnings.

Finally, in Column 5, we corroborate the estimates in Columns 1–4 that used the CPS ASEC sample by using data from the QWI. We run essentially the same regression specification as Column 1, except that we are able to include fixed effects for each county (rather than state)²⁶ and each Division-Year-Quarter (rather than Division-Year). We weight the regression by county-level employment. The estimate is very similar in magnitude to that in Column 1 and highly statistically significant.

It is instructive to benchmark our results against the estimated wage effects of other labor market institutions. For example, the household income premium associated with membership in a labor union is an estimated 15-20 log points (Farber et al., 2018); the income premium for workers in an occupation that requires a government-issued occupational license is estimated to be 7.5% Gittleman et al. (2018).²⁷ To compare the effects of NCA enforceability against these institutions, we can extrapolate our estimates to consider what would happen to earnings under a national policy that rendered all NCAs unenforceable. We generate predicted earnings for each individual in the 2014 ASEC sample using coefficients from Column 1 of Table 3, for two different levels of NCA score: first, the NCA score observed in 2014 in that individual's state, and second, at the lowest observed NCA enforceability level (0). These

²⁴All results are very similar if we include part-time workers.

²⁵We include part time workers in this regression to avoid selecting the sample based on the dependent variable.

²⁶The estimate is essentially unchanged if we instead use state fixed effects.

²⁷Estimates of the wage premium associated with occupational licensing vary widely: for example, Redbird (2017) finds no wage premium using a 30-year comprehensive panel of licensing laws

predictions imply that average earnings among *all* workers would increase by 8.5% nationally if NCAs were made unenforceable.²⁸ This magnitude is roughly one half the household premium from labor union membership and is slightly larger than the premium attained by workers in occupations with government-mandated licenses.

Figure 3 visually illustrates the relationship between annual earnings and NCA enforceability using binned scatterplots. Each graph plots earnings and NCA enforceability, net of state and census division by year effects. Panel (a) includes no additional controls, and panel (b) includes the additional controls used in Column 2 of Table 3 (1-digit occupation codes and individual-level demographic controls). Both figures clearly depict a strongly negative, roughly linear relationship between enforceability and earnings, corroborating the regression estimates.

Our NCA Enforceability Score pools seven dimension of NCA enforceability, but these dimensions might differ in their earnings effects. In Table B.2, we reestimate the effect of changes in NCA law on earnings in a specification analogous to Column 1 of Table 3, but focusing on each individual component of the composite NCA score separately. The first seven rows represent separate regressions identical to Equation 1, except that $Enforceability_{st}$ is replaced with each respective element of the NCA score described in Table B.1.²⁹ With two exceptions (which are both insignificant at the 10% level), the effect of each score is negative, and is significant at the 5% level for two out of seven components, and at the 10% level for one additional component. The dimensions yielding the greatest negative earnings effect are those requiring consideration, both at the outset of employment (Q3a) and after employment has already begun (Q3bc), consistent with evidence in Starr (2018). The existence of a state statute (Q1) has a negative but insignificant earnings effect. This lack of an effect for Q1 is perhaps not surprising: some states that do not have explicit statutes regarding NCA enforceability (e.g., Kansas and Connecticut) nonetheless enforce NCAs more readily than many other states. Given this ambiguity of the Q1 dimension, in the

²⁸Specifically, let X_i be the vector of the values of all variables (including fixed effects), except for NCA enforceability score, that are present in the regression in Column 1 of Table 3 for each individual, i , in 2014. Let γ be the vector of respective coefficients estimated in the same regression, and let β be the coefficient on $Enforceability_i$, the NCA Enforceability Score for individual i 's state of residence in 2014. Then, if $\hat{Y}_{i,1} = \gamma X_i + \beta Enforceability_i$ represents predicted earnings for individual i , and $\hat{Y}_{i,2} = \gamma X_i$ represents predicted earnings for individual i when $Enforceability_i = 0$, the predicted earnings increase is calculated as the average of $\hat{Y}_{i,2}$ minus the average of $\hat{Y}_{i,1}$, divided by the average of $\hat{Y}_{i,1}$.

²⁹Estimating a model with each component of the score separately likely introduces some omitted variable bias, as elements of the score are correlated with each other. However, including all individual components of the score in the same regression causes the sample size to shrink significantly due to missingness in some of the components (where missingness indicates that the question has not been legally settled). That model, however, generates coefficients qualitatively similar to those shown in Table B.2.

final row of Table B.2 we replace $Enforceability_{st}$ with a modified version of the NCA Enforceability Score that omits the component related to existence of a state statute (Q1). The resulting coefficient is, if anything, stronger than that estimated in Table 3. Thus, no single dimension drives our results, and the dimensions with the largest effects are consistent with what one might expect based on theory and on prior results.

4.2 Assessing the Causal Interpretation and Robustness of the Estimated Earnings Effect

We conduct three distinct tests to assess the causal interpretation of our results, which we describe in turn below.

4.2.1 Heterogeneous Earnings Effects Based on Prevalence of NCA Use

The results in Table 3 imply that stricter NCA enforceability leads to lower earnings for the average worker. This relationship should be stronger in settings in which NCAs are used more often; in the limiting case, if NCAs are never used for a certain group of workers, we should expect no effect of NCAs on earnings for those workers (unless spillover effects are sufficiently large).

In this section, we examine heterogeneity in the effect of enforceability by prevalence of NCA use. This exercise serves two useful purposes. First, it serves as a test of the robustness of the results reported in Section 4.1. If we find that enforceability has larger earnings effects among groups less likely to be bound by NCAs, it might raise questions about the research design. Second, this exercise allows policymakers to assess the impact that changes in NCA enforceability will have on the earnings of groups more likely to be exposed to NCAs.

While we do not observe whether individual workers have or have not signed an NCA, Starr et al. (2018) report several sources of heterogeneity in NCA use by worker characteristics. We focus on three sources: workers' education, occupation, and industry. First, Starr et al. (2018) find that workers with a Bachelor's degree or higher are significantly more likely to sign NCAs than workers without a college degree. Second, Starr et al. (2018) find heterogeneity in use across 22 occupation categories and 19 industry categories. We use the occupation and industry in which an individual reports working to the CPS to classify workers as working in *High or Low NCA Use Occupations* and *High or Low NCA Use Industries*.³⁰ We replicate

³⁰We define Low NCA Use Occupations as Farm, Fish and Forestry; Legal Occupations; Grounds Maintenance; Food Preparation and Serving; Construction; Extraction; Transport and Materials Moving; Office Support; and Community and Social Services, and High NCA Use Occupations as all

our main difference-in-difference specification, Equation 1, except that we now add an interaction term of *Enforceability* with an indicator for *College Educated Worker*, *High NCA Use Occupation*, or *High NCA Use Industry* (as well as an indicator for the respective main effects).

Table 4 reports these heterogeneity estimates. Column 1 reports the baseline average effect on earnings, corresponding to Column 1 in Table 3. Column 2 includes an interaction of NCA Enforceability Score with an indicator for whether a worker has a college degree (*College Educated Worker*). The main effect on *NCA Enforceability Score* is close to zero and statistically insignificant, implying that enforceability has little to no effect on earnings for non-college educated workers. On the other hand, the interaction term ($-0.143, p < .01$) implies that enforceability has a much stronger effect on earnings of college-educated workers. The sum of the main effect on *NCA Enforceability Score* and the interaction effect implies that going from the 10th to 90th percentile of enforceability leads to a 6.0% decrease in earnings for college-educated workers ($\exp((-0.035 + -0.143) * 0.35) - 1 = -0.06, p < .01$), an earnings effect that is 70 percent larger than the earnings effect for the whole population implied by Column 2 of Table 3.

Column 3 reports heterogeneity by occupational use of NCAs. The estimates imply that going from the 10th to 90th percentile of enforceability leads to a 4.9% decrease in earnings in high-use occupations ($\exp((-0.083 + -0.061) * 0.35) - 1 = -0.049, p < .01$); the effect for low-use occupations is roughly half as large ($p = .02$), and the difference is statistically significant ($p < .01$). Finally, Column 4 reports heterogeneity by industries' use of NCAs. Going from the 10th to 90th percentile of enforceability leads to a 5.5% decrease in earnings in high-use industries ($p < .01$); the effect for low-use industries is roughly 60% as large ($p < .01$), and the difference is statistically significant ($p < .01$).

In Column 5, we simultaneously estimate the heterogeneous impacts of NCA enforceability along these three categories. The coefficients on the interactions of NCA Score with *High Use Occupation* and *High Use Industry* attenuate, but remain negative and significant. The interaction of NCA Score with *College Educated* changes little and remains statistically significant.³¹

others. Low NCA Use Industries are Agriculture and Hunting; Accommodation and Food Services; Arts, Entertainment, and Recreation; Construction; Real Estate; Transportation and Warehousing; Retail Trade; Other Services; and Management of Companies. These occupations and industries represent those with NCA use below or above the national average, according to Figures 5 and 6 in Starr et al. (2018).

³¹Since college-educated workers tend to get paid more than those without a college degree, this stability of the *College Educated* estimate is consistent with the evidence in Starr et al. (2018) that NCA use is increasing in workers' annual earnings.

4.2.2 Event-Study Estimates on Earnings

Two concerns are common with difference-in-difference designs. The first is evaluating the plausibility of the assumption that treatment and control groups would counterfactually follow common trends. In our context, this assumption might be violated if, for example, business or labor advocacy organizations change lobbying effects that influence both earnings as well as judges’ decision-making. Such effects would constitute a form of reverse causality. A second concern is imbalance in treatment timing. Our regression design leverages changes in NCA laws that occurred in different states in different years. This variation in treatment timing can give differential weight to states depending on the distribution of event times within the sample; this weighting could cause the interpretation of our estimand to differ from that of an average treatment effect (Goodman-Bacon, 2018).

To address these concerns, we complement our difference-in-difference estimates with an event-study analysis. Following Cengiz et al. (2019), for each state that experiences an NCA enforceability change, we isolate a four year window before and after the law change, and we identify a set of “control” states in the same Census division that did not experience a law change at any point during that window. We then stack the data for each treatment window and the corresponding set of control states, and estimate the difference in outcomes between treated and control states in each year relative to the law change. The regression equation is:

$$\ln w_{is\bar{s}t} = \sum_{k=-4}^{k=4} \beta_k \Delta Enforceability_s^{tk} + X_{it}\gamma + \rho_{s\bar{s}} + \delta_{d(s)t} + \varepsilon_{is\bar{s}t} \quad (2)$$

where $\ln w_{is\bar{s}t}$ is log hourly earnings of worker i in state s , matched to treatment state \bar{s} in year t . $\Delta Enforceability_s^{tk}$ is equal to the magnitude of the change in state s ’s composite NCA enforceability score from a law change that occurred at year $t + k$ and is zero otherwise (and therefore zero for all k when $s \neq \bar{s}$, *i.e.*, when the state is a control state). X_{it} is a vector of individual-level controls, $\rho_{s\bar{s}}$ is a fixed effect for each state by treatment-state combination, and $\delta_{d(s)t}$ is a fixed effect for each Census division by year. We cluster standard errors by state.

Figure 4 shows the event study estimates for each $\hat{\beta}_k$ (we normalize the coefficient $\hat{\beta}_{-1}$ to be zero). The first pattern that emerges from the figure is that there is little evidence of a pre-trend in earnings, supporting assumption (and the evidence in Section 3.1.1) that NCA law changes were largely exogenous to underlying economic trends. The second pattern in the figure is a decline in hourly earnings the year of the law change that generally grows in magnitude over next four years. These estimates imply

that moving from the 10th to 90th percentile of *Enforceability* leads to a reduction in hourly wages that is as high as 5.5 percentage points four years after the law change. This magnitude is larger than our difference-in-differences estimates, but given the wide confidence intervals in these event study results the annual estimates are not statistically significantly different from each other nor from the overall difference-in-difference estimate.

4.2.3 Addressing Other Threats to Identification

In Section 3.1.1, we provided evidence that economic, social, and political factors do not collectively predict changes in NCA enforceability, and this argument was corroborated by the lack of pre-trends in the event study graph just shown. Still, we can ensure that coincidental changes in these factors are not driving our estimated effect of enforceability. We assess the robustness of our estimates to this concern in Table B.3. We replicate the structure of Table 3, but we include additional controls for each of the predictors included in Table 2.³² While there are minor changes in the magnitudes of estimates, the qualitative conclusions are all unchanged, supporting the causal interpretation of the *Enforceability* coefficient in our regressions.

Focusing on the institutions underlying our identifying variation, we argued above that judicial decisions, which make up the vast majority of NCA law changes, are less prone to endogeneity than are statutory changes from legislative action. However, there is some evidence that judges' decision-making can be swayed by external forces like business interests, particularly for judges that are elected rather than appointed (Katz, 2018). To ensure that our results are not driven by confounding influences on elected judges we obtained data on how judges are selected across states from Bannon (2018). We recreate our main analyses a) excluding the 6 states that have partisan judicial elections (i.e., judges are selected via election and the judge's political party is listed on the ballot) and b) excluding the 21 states in which judges are elected (whether or not the elections are partisan), in Tables B.4 and B.5, respectively. If anything, our point estimates are *larger* in magnitude with these restricted samples (they become more imprecise in the latter table, which is to be expected since we are eliminating over 40% of the states in our sample). Since judicial elections are a key mechanism through which political or economic preferences of voters might affect judicial decisions, this evidence provides further reassurance against this potential form of endogeneity.

³²We omit the ideology variables gathered by McCarty and Shor (2015), which were only calculated since 1993. Inclusion of those variables (which limits the sample period) does not substantively change the estimates.

4.3 Effects of Enforceability on Job Mobility

While the main focus of our analysis is the earnings effect of NCA enforceability, we also estimate its effect on worker mobility. This analysis is useful because it serves as validation that our variation in enforceability is capturing what NCAs are designed to do—restrict workers’ mobility.

Table 5 presents estimates based on job mobility data from the J2J dataset. We explore three measures of job mobility, each measured within state-year-quarter-sex-age group-industry cells: the overall job-to-job separation rate,³³ the share of job-to-job flows across state lines, and the share of job-to-job flows across two-digit NAICS industries. For each of these measures, we estimate the overall effect of NCA enforceability, as well as the differential effect for *High NCA Use Industries*, which we defined in Section 4.2.1.

In Column 1 we estimate the effect of the origin state NCA enforceability score on the overall job-to-job separation rate and find a small and statistically insignificant effect. However, in Column 2 we interact NCA enforceability with an indicator for whether the origin job was in a high NCA use industry, and find that NCA enforceability substantially reduces job-to-job separations in high use industries. The coefficient on *High NCA Use Ind* \times *NCA Score* is negative (-0.199) and highly significant ($p < .01$). The estimate implies that moving from the 10th to the 90th percentile of NCA enforceability decreases the rate of job-to-job separations by 6.0% in high use industries.

In Columns 3 and 4 we test for effects on the share of job-to-job transitions that occur across state borders. In high NCA use industries, stricter enforceability increases the geographic distance associated with job changes. Moving from the 10th to the 90th percentile of NCA enforceability increases the share of job changes that cross state lines by 0.6%. This estimate suggests that greater NCA enforceability forces workers bound by NCAs to move further to escape restrictions imposed by NCAs, which typically include a geographic component. In the case of labor markets that are defined by industries rather than geography, escaping an NCA may require changing industries. In Columns 5 and 6 our dependent variable is the share of job-to-job transitions in which a worker switches industries. The coefficient on *High NCA Use Ind* \times *NCA Score* is negative, but it is small and not statistically significant.

This evidence reveals that NCA enforceability has meaningful effects on both the

³³We define the overall job-to-job separation rate as the number of new hires in a cell with no nonemployment spell or a short nonemployment spell, divided by the total employment in that cell. At first glance it might seem more appropriate to name this measure job-to-job *hiring* rate, but it nonetheless represents all of the separation events that resulted in a job-to-job transition.

level and direction of workers' job mobility, and it illustrates that our measures of NCA enforceability capture actual changes to the *effective* use of NCAs. The results also motivate our investigation into one mechanism through which NCA enforceability affects earnings, which we describe in Section 6.

5 Spillover Effects of NCA Enforceability

It is important to remember that we do not observe which workers actually *sign* NCAs. If NCA enforceability only affects the earnings of workers actually bound by an NCA, one could interpret our results thus far as “intent-to-treat” effects, in which the “treatment” is signing an (enforceable) NCA. However, in Section 2, we described theoretical reasons why the earnings effects of NCA enforceability might not be limited only to those workers signing NCAs. In this section we examine whether such “spillover effects” are present by testing whether enforceability in one state affects the earnings of workers in a different legal jurisdiction.

We examine whether legal changes to NCA enforceability in a “donor” state impact workers who share a local labor market with that state but work in a different state. Consider the St. Louis metro area, which includes counties in Missouri but also several counties across the state border in Illinois. If Illinois experiences an NCA law change, does it affect the earnings of workers employed on the Missouri side of the St. Louis metro area? And vice versa if Missouri experiences a law change?

We measure local labor markets as commuting zones, which are clusters of counties that have strong commuting ties and have been used in many prior studies as measures of local labor markets (David et al., 2013). We identify commuting zones that straddle state borders: these commuting zones are local labor markets that include business establishments in two states and are therefore subject to two different NCA enforcement regimes (as well as changes therein). We remove 8 commuting zones that contain counties in more than 2 states (to ensure clarity of measurement for *Donor State NCA Score*, the variable measuring enforceability in the cross-border state). These restrictions leave us with a set of 137 commuting zones. In our main analysis, we focus on the 545 counties in these commuting zones that themselves lie directly on state borders; with this restriction, we avoid counties such as Los Angeles County, which shares a commuting zone with counties in Arizona but is nearly 200 miles driving distance from anywhere in Arizona.

We employ data from the QWI dataset, which as described in Section 3 includes quarterly earnings and employment flows at the county level, separated by various firm characteristics and worker demographics. Each observation in the dataset represents a unique year, quarter, county, sex, and age group cell, as defined in the QWI.

To test for spillovers, we use an analog of the difference-in-difference model corresponding to Equation 1 to estimate the impact of a change in NCA enforceability across a state border, among workers employed in a commuting zone that straddles the state border. The outcome variable is the log of average quarterly earnings within each cell for all private sector employees. We estimate the model:

$$Y_{ctga} = \phi_0 + \phi_1 * Enforce_{ct} + \phi_2 * BorderEnforce_{ct} + \phi_3 * Female_g + \psi_a + \zeta_c + \Omega_{d(c)t} + \varepsilon_{ctga}, \quad (3)$$

where c indexes county, t indexes year-quarter, g indexes sex, a indexes age group, and $d(c)$ indexes the Census division in which county c is located. $\Omega_{d(c)t}$ is a Census division by year-quarter fixed effect. The primary coefficient of interest is ϕ_2 , which is an estimate of the spillover effect on workers in county c of enforceability in the state that borders the commuting zone in which county c is located. ϕ_1 estimates the direct effect of enforceability in a worker's own state, analogous to our estimates thus far. We cluster standard errors two ways by state and commuting zone.

We report results in Table 6. Column 1 verifies that the direct relationship between (own) state NCA scores and earnings holds in this restricted sample. The coefficient on *Own State NCA Score* is negative (-0.178), highly significant ($p < .01$), and slightly larger than the overall estimate reported in Table 3. The model in Column 2 additionally includes *Donor State NCA Score*. The direct effect of *Own State NCA Score* increases very slightly to -0.207, $p = .003$. The coefficient on *Donor State NCA Score* reveals evidence of meaningful spillover effects: the coefficient is negative (-0.181, $p = .021$) and is nearly as large as the direct effect of an increase in a worker's own state.

The spillover effect of a border-state's enforceability is only likely to be relevant if the border state's counties make up a large share of the local labor market's total employment. Intuitively, a law change in a donor state should have little to no effect on earnings in a focal county if the focal county makes up the vast majority of the commuting zone's employment; in this case, any spillover effects of NCA enforceability (whether via thinning the labor market, reducing dynamism, increasing employer market power) will be minimal. On the other hand, a law change in a donor state would have a larger impact on a focal county that comprises a very small part of its commuting zone's employment.

We test for such heterogeneity in Column 3. Along with their main effects, we include interactions of *Own State NCA Score* and *Donor State NCA Score* with the

ratio of sex- and age-group-specific employment in the focal county to sex- and age-group-specific employment in the commuting zone as a whole (*Own Cty Emp / CZ Emp*).³⁴ The results demonstrate that spillover effects of donor-state enforceability are heterogeneous in a manner consistent with the logic above. The main effect of *Donor State NCA Score*, representing the spillover effect in a county that comprises zero percent of its commuting zone’s employment, is negative (-.210, $p < .01$) and has a slightly *larger* magnitude than the coefficient on main effect of *Own State NCA Score*. In contrast, the spillover effect is substantially smaller for counties that contribute a large share of employment in their commuting zone: adding the main effect of *Donor State NCA Score* with its interaction with *Own Cty Emp / CZ Emp* (0.163, $p < .01$) implies a small spillover effect ($-0.210 + 0.163 = -0.047$), that is statistically insignificant ($p = .511$) for a county comprising 100% of its CZ employment.³⁵

The evidence of spillover effects on earnings is further demonstrated in Figure 5, which uses the methodology in Section 4.2.2 applied to changes in *Donor State NCA Score*. The event study graph exhibits no evidence of any pre-trend in earnings, supporting the causal interpretation of our estimates. The effect of border state NCA Enforceability appears only after a few years: this lag (relative to Figure 4) could be due to adjustment times by firms that result in a relatively slow spread of labor market conditions across the commuting zone.

Our analysis thus far has considered spillover effects of NCA enforceability in counties across state lines that lie on the state border. It would be less plausible for NCA enforceability to affect earnings in counties across state lines that are much further from the state border. To check whether our estimated spillover effects do, in fact, attenuate with distance to the state border, in Table B.7 we consider a falsification test for counties that should be progressively more immune to NCA law changes in border states. We calculate coefficients on *Donor State NCA Score* for regressions run on (1) the main border county sample; (2) counties not physically on state borders and not in commuting zones that straddle state borders; (3) the sample from Column (2), further restricting to only counties that lie over 50 miles from any state border; and (4) the sample from Column (2), further restricting to

³⁴We also include the main effect of this ratio but do not report its coefficient in the table.

³⁵Unlike the analysis with the J2J dataset, we leave the regressions in Table 6 unweighted. We do this for two reasons. First, we weight the J2J analysis by employment to estimate an average treatment effect for the US population; because the QWI sample in Table 6 is limited to border counties, weighting serves no such purpose. Second, spillover effects (as discussed) are likely to be more pronounced in counties with a small share of employment. Therefore, an estimate that weights observations by employment may reveal little to no impact of Donor State NCA Score. We report a weighted version of Table 6 in Table B.6, which indeed shows an attenuated average effect. However, Column 3 reveals that the heterogeneity based on employment shares in the CZ in Column 3 persists in the weighted specification, as expected.

only counties that lie over 100 miles from a state border. We assign to each county a donor state NCA score that corresponds to the state geographically closest to that county.³⁶ Reassuringly, the point estimate on *Donor State NCA Score* decays across the table, and is effectively zero at unreasonably long commuting distances.³⁷

In Section 2.2, we discussed several reasons why strict NCA enforceability could generate the negative externalities on earnings documented in this section, including by thinning labor markets or giving firms wage-setting power. Other explanations are possible. For example, workers may decide to find a job across state lines if their own state increases NCA enforceability. Such behavior would cause an outward shift in labor supply in border states, causing the market-clearing wage to decline. We find no evidence, however, that such worker behavior can explain the spillover effects on earnings. In Table B.8, we present estimates of the spillover effects of enforceability on workers' *mobility*. The structure mimics Table 6, except that our dependent variables are the log quarterly number of hires and separations from QWI in Columns 1 to 3 and 4 to 6, respectively. Across all six columns, enforceability in a worker's *own* state has a negative effect—of roughly the same magnitude—on hires and separations, corroborating the mobility results we found in Section 4.3 using the J2J dataset. The spillover effects (reported in Columns 2 and 5) are imprecisely estimated, though they are negative and of a magnitude that is 40-50 percent as large as the direct effect.³⁸ Thus, there is no evidence that workers move across state lines in response to an NCA law change in their own state; if anything, these estimates suggest that strict NCA enforceability reduces cross-border mobility.

Collectively, these results on earnings and mobility provide evidence that NCA enforceability reduces earnings and labor market churn, even across state borders. Though we do not measure which workers do and do not sign NCAs, these results suggest that NCA use has external effects on workers and firms that do not use them, consistent with the theoretical considerations discussed in Section 2.2.

³⁶Specifically, we calculate the distance between county centroids. If the centroid of a county in a different state is less than m miles from the centroid of the focal county, we exclude that focal county from the relevant regression. Donor state NCA scores are similarly assigned by finding the county in a different state whose centroid is closest to the focal state's centroid, and using that donor state's NCA score.

³⁷Reassuringly, however, the point estimate on *Own State NCA Score* reveals that the direct effect of own-state NCA score remains stable across these various geographic restrictions.

³⁸Additionally, Columns 3 and 6 document an identical pattern of heterogeneity to that observed on earnings: an NCA law change in a donor state has a larger effect on mobility in a focal county among counties comprising a small portion of the commuting zone's total employment, compared to counties comprising a large share.

6 Why Does NCA Enforceability Reduce Earnings? The Effects of Costly Mobility

Our results so far have shown that stricter NCA enforceability reduces earnings and mobility, and that these effects are present even for workers whose contracts are not directly affected by the enforceability regime. In Section 2, we discussed channels through which enforceable NCAs, by restricting workers' mobility, could reduce workers' earnings. In this section, we conduct a test to examine one such channel, and show that NCAs diminish workers' abilities to take advantage of favorable labor markets over the course of their job tenure.

A longstanding theory in labor economics is that wages are determined by “implicit contracts” in which firms insure workers against declines in their wage. This theory implies that wages are not determined in a spot market, but rather set by implicit contracts with terms that depend in part on the worker's outside option. The seminal paper by Beaudry and DiNardo (1991) (hereafter, BDN) theorized that wages will behave differently depending on whether or not workers' mobility across jobs is costly. If mobility is costly—that is, it is difficult for workers to find another job once they have begun a job spell—then labor market conditions at the time a worker begins her spell will determine her wage for the duration of her spell. If labor market conditions improve, the worker's costly mobility means she cannot take advantage of new job opportunities and her employer has no incentive to increase her wage. Alternatively, if mobility is costless, a worker cannot commit to a contract if her outside option subsequently improves; because the worker can threaten to quit, improvements in labor market conditions induce employers to raise wages. Therefore, costless mobility implies that the best labor market conditions over the course of a worker's job spell will be correlated with her current wage.

BDN develop a simple empirical method to test between these models, and they find strong evidence consistent with a model of implicit contracts with costless mobility: the effect of the most favorable labor market conditions over a worker's job spell exceeds and washes out any effect of the contemporaneous condition (predicted by a spot market) or the condition at the time of hire (predicted by an implicit contracts model with costly mobility). This result has been replicated numerous times with different datasets and time periods (e.g., Schmieder and Von Wachter, 2010).

An immediate prediction is that NCAs—by making mobility more costly—change the nature of implicit contracts in the labor market. When NCAs are more enforceable, workers will no longer be able to leverage improvements in their outside option during a job spell, and their wage will be determined in much larger part by the initial

labor market condition than in states where NCAs are less enforceable.

We begin by replicating BDN. We use the CPS JTS, limiting our analysis to full-time, private sector workers, for the years 1996-2014 (compared to BDN, who used the years 1976 to 1984).³⁹ We estimate the model:

$$\ln w(i, t + j, t) = \Omega_1 X_{i,t+j} + \Omega_2 C(t, j) + \varepsilon_{i,t+j}, \quad (4)$$

where $w(i, t + j, t)$ is the wage of individual i at time $t + j$ who began her job spell at time t . $C(t, j)$ is a vector of unemployment rates which, depending on the model, include *Initial UR* (the unemployment rate at the beginning of the individual's job spell) and/or *Minimum UR* (the lowest unemployment rate between the beginning of the job spell and the time of measurement of the wage). Following BDN, we use annual national unemployment rates from the Bureau of Labor Statistics. $X_{i,t+j}$ is a vector of individual level characteristics. Again following BDN, in $X_{i,t+j}$ we include race, Hispanic status, sex, marital status, age, age squared, tenure, tenure squared, education, and industry dummies. We depart from the BDN specification in three minor ways to accommodate our analysis. First, we do not include Metropolitan Statistical Area (MSA) fixed effects: doing so decreases our sample size by approximately 25% (due to individuals whose MSA has been omitted from public use extracts of CPS supplements). In their stead, we use dummy variables for metropolitan area status (as used in Equation 1). Second, we include Census division by year fixed effects to harmonize with the main estimates of the effects of NCA enforceability. Third, we do not consider the contemporaneous unemployment rate, which would be collinear with Division by year fixed effects. Each of these adjustments ultimately has little bearing on our estimates.⁴⁰

We report these results in Table 7. Columns 1–3 replicate the BDN main results for our sample period. In Column 1 we include only the unemployment rate at time of hire (*Initial UR*): our estimated coefficient has a smaller magnitude than that estimated in BDN (ours: -0.008; BDN: -0.030), but it is negative and highly statistically significant ($p < .01$). Column 2 uses, instead, the minimum unemployment rate over the course of the worker's job spell (*Minimum UR*). Similar to BDN, we find a negative and statistically significant effect. Column 3 mimics the main finding of BDN:

³⁹We omit years prior to 1996 due to a lack of data availability: though BDN use CPS data collected prior to 1996, the dataset we employ (the CPS JTS) has only been collected since 1996.

⁴⁰Inclusion of MSA fixed effects (unreported) has little effect on our estimates. Our estimates are also robust to excluding Census division by year fixed effects, and to using state-level unemployment rates in lieu of national unemployment rates, which allows us to include contemporaneous unemployment rates in our regressions (since they are not collinear with division-year fixed effects). We choose to use national rates to follow BDN, and also because state-level unemployment rates could in theory be an outcome of NCA enforceability policies.

including both *Initial UR* and *Minimum UR* attenuates the coefficient on *Initial UR* close to zero but leaves the coefficient on *Minimum UR* negative and highly significant ($p < .01$). In other words, on average, wages are consistent with a model of implicit contracts with costless mobility—just as Beaudry and DiNardo (1991) and the subsequent literature have found.

To test the hypothesis that NCA enforceability shifts the labor market from an implicit contracts model with costless mobility to one with costly mobility, we estimate the model:

$$\ln w(i, t + j, t, s) = \Omega_1 X_{i,t+j} + \Omega_2 C(t, j) + \Omega_3 Enf_{t,s} + \Omega_4 C(t, j) * Enf_{t,s} + \varepsilon_{i,t+j}, \quad (5)$$

where $Enf_{t,s}$ is the NCA enforceability score in state s at time t , the beginning of the worker’s job spell. This model allows the effect of labor market conditions to vary with the strength of NCA enforceability at the time the worker was hired. If NCA enforceability affects the cost of mobility in an implicit contracts environment, we expect two effects. First, we expect the coefficient on $Enf_{t,s} \times \text{Minimum UR}$ to be positive, indicating that employees have *less* ability to leverage favorable labor markets over the course of their job spell when NCA enforceability is high. Second, we expect the coefficient on $Enf_{t,s} \times \text{Initial UR}$ to be *negative*, indicating that wages are *more* responsive to labor market conditions at the time of hire when NCA enforceability is high.

We report the results in Columns 4 and 5. Column 4 mirrors Column 3, but includes an additional control: NCA enforceability at the employee’s time of hire ($Enf_{t,s}$). Encouragingly, the coefficients on *Initial UR* and *Minimum UR* do not change, indicating that NCA enforceability is not acting as a de facto proxy for one of the unemployment rates.⁴¹

In Column 5, we include the interactions demonstrating the change in the cost of mobility. First, consider the main effects of *Initial UR* and *Minimum UR*, which indicate the effect of initial and most favorable labor market conditions, respectively, for a state with the lowest NCA enforceability. These coefficients mirror, and amplify, the findings from Beaudry and DiNardo (1991): a higher initial unemployment rate for

⁴¹Caution should be taken when interpreting the coefficient on *Initial NCA Score* in Column 4, which is smaller in magnitude than in our prior results and statistically insignificant. This specification includes controls for tenure and tenure squared: these are important controls in the BDN framework but may bias the magnitude of the coefficient on *Initial NCA Score* towards zero. This is because tenure may be affected by NCA laws, especially given our prior results that NCA enforceability impedes worker mobility. Omitting tenure and tenure squared as controls in the regression in Column 4 slightly increases the coefficient on *Initial NCA Score* to -0.054, though it is not statistically significant. Excluding the tenure controls does not meaningfully affect the magnitude or significance of the coefficients of interest in subsequent regressions.

a worker in a low-enforcing state does not reduce her wage today—if anything it leads to a *higher* wage—whereas the main effect of *Minimum UR* indicates that a worker’s wage today is strongly responsive to her most favorable labor market condition over her tenure. In other words, wages in a state with low NCA enforceability are *even more* aligned with an implicit contracts model of costless mobility than the overall population.

Next, consider the two interaction terms, indicating the differential effects of these conditions for a worker in the highest enforcing state. The coefficient on $Enf_{t,s} \times Initial\ UR$ (-0.017 ; $p < .05$) shows that a higher unemployment rate at time of hire affects earnings much more negatively when NCAs are more enforceable. In contrast to workers in a state with the lowest enforceability, the interaction term $Enf_{t,s} \times Minimum\ UR$ (0.020 ; $p < .05$) shows that the most favorable labor market condition over job tenure has a much more muted effect on the current wage in states with higher enforceability: for workers in a state with the highest observed enforceability, the most favorable labor market condition over the course of their tenure has essentially no effect on their wage ($-0.028 + 0.020 = -0.008$, $p = .20$).

These estimates imply that NCA enforceability fundamentally changes the way that workers and employers negotiate wages. To visualize the real implications of these findings, we can use our estimates from Table 7 to predict how the wage path of a worker beginning a job spell in a particular year differs depending on the NCA enforceability in her state. We consider the predicted earnings path for two hypothetical workers—one in a low- and the other in a high-enforceability state—who each began job spells in 2009 at identical wages and held their job through 2019. This period is of particular interest given that 2009 was the onset of the Great Recession and witnessed a large increase—then decrease—in the unemployment rate. We predict earnings each year based on a regression identical to that reported in Column 5 of Table 7, and we plot the path of normalized predicted earnings⁴² for both individuals in Figure B.1, alongside the monthly national unemployment rate. The two wage paths move in perfect tandem until 2013—the year that the unemployment rate begins to drop below the initial unemployment rate at the start of the workers’ tenure (2009). Beginning in 2013, the paths diverge. The worker in the low-enforcing state is able to take advantage of the improvement in her outside option and increase her earnings above and beyond the initial earnings path. The path of earnings for the worker in the high enforceability state, on the other hand, is significantly less responsive to the labor market tightening: this worker’s earnings continue to rise at a relatively

⁴²The normalization simply subtracts the difference in low enforceability versus high enforceability state earnings from each predicted value for high enforceability in order to consider two hypothetical workers with identical initial wages. The initial gap (in January, 2009) is \$3.57.

constant rate according to her tenure at the firm. Ultimately, this divergence leads normalized earnings in the high enforceability state to be 2.2% lower than in low enforceability states (\$934.34 vs. \$913.53) in 2019.

These results reveal one mechanism through which NCA enforceability—by increasing the costs of worker mobility—affects earnings. When NCAs are strictly enforced, individuals are less able to increase their earnings as their outside option improves over the course of their job tenure. An important implication of this result is that NCA enforceability can affect earnings even if enforceability does not directly affect a worker’s *realized* job mobility: by shutting down a worker’s *threat* of mobility, NCAs reduce workers’ bargaining power. Prior evidence has highlighted how important this threat of mobility is for wage growth: Bagger et al. (2014) show that the wage gains from job search *within* job spells dominates the gains from search *across* job spells. In other words, enforceable NCAs change the terms that govern how workers and employers bargain over wages.

More broadly, these estimates illustrate a means through which NCA enforceability has plausibly contributed to the declining labor share of income and wage stagnation in recent decades. General improvements in the labor market do not translate into wage gains for workers in states that enforce NCAs more strictly. A realistic implication is that productivity gains translate less into wage gains when NCAs are more enforceable, which would lead strict NCA enforceability to reduce the labor share of income.

7 NCA Enforceability Reduces Earnings More for Women and Racial Minorities

In Section 2.3, we discussed reasons why the earnings effect of NCA enforceability would be unevenly distributed across demographic groups, and in particular be more pronounced for women and racial minorities. Motivated by this discussion, we investigate whether the earnings effect of NCA enforceability is heterogeneous on the basis of sex and race.

Figure 6 displays results from two regressions that add demographic group indicators, alone and interacted with NCA Score, to the regression reported in Column 1 of Table 3. We make two additional modifications: first, we remove the restriction that workers must be working full-time to avoid selecting the sample on an outcome that is known to differ across men and women.⁴³ Second, whereas before we simply controlled for whether a respondent is white or not, and male or female, we include

⁴³The results do not meaningfully change if we reimpose the full-time restriction.

the more detailed demographic categories presented in the figure. The displayed coefficients, which are on the interaction of the relevant group indicator with the *NCA Enforceability Score*, represents the impact of NCA enforceability on the earnings of individuals in that group. We report coefficients from two models: our “main estimate” that makes no further modifications, and one that additionally includes interactions of the *NCA Enforceability Score* with dummies for college-educated, high use occupation, and high use industry.

First, consider the coefficients from the “main estimate” model. The effect of NCA enforceability on earnings is negative and significant for each demographic group, or close to significant in the case of white men. However, the effect is much more negative all female groups (White Female, Black Female, Other Female) and for Black Men than it is for White Men. The differences between the coefficient for each group and for white men is each statistically significant at least at the 10% level, as displayed as the upper p-value in square brackets above each coefficient.

These estimated differences might be misleading if sex or race correlate with differences in education or occupational choice, which Section 4.2.1 showed moderates the effect of NCA enforceability on earnings. To address this concern, the second set of estimates depicted in the figure additionally controls for the interactions of the *NCA Enforceability Score* with dummies for college-educated, high use occupation, and high use industry.⁴⁴ While the estimates do attenuate somewhat, they remain negative and mostly statistically significant. Furthermore, the earnings effect of NCA enforceability remains statistically significantly different for nonwhite women and black men when compared with white men, though the difference for white women loses statistical significance ($p = 0.137$).

These results suggest that NCA enforceability not only reduces earnings *on average*, but it also exacerbates existing disparities across demographic groups. This point is illustrated two ways. First, the coefficients in Column 2 of Table B.9 imply that moving from the 10th to 90th percentile of the NCA Score distribution (NCA score = 0.55 and 0.9, respectively) would decrease average earnings of white men by approximately 3.2%, vs. decreases ranging from 3.7% to 7.7% for the other demographic groups. Together with the estimates in Column 1, these results imply that if a state that enforces NCAs at the 90th percentile of the distribution were to switch to enforcing NCAs at the 10th percentile of the distribution, the earnings gap between white men and each other demographic group would close by 3.6% for nonblack, non-white men, 4.6% for black women, 5.6% for white women, 8.7% for black men, and

⁴⁴For a full accounting of the two regressions depicted in the figure, as well as regressions which control separately for each additional control, see Table B.9.

9.1% for nonblack, nonwhite women.

The evidence provided in this section shows that, in addition to affecting average earnings across workers in the US workforce, strict NCA enforceability specifically harms workers who have historically faced disadvantages in the labor market. Thus, limiting the enforceability of NCAs would not only likely raise earnings on average, but also help close racial and gender wage gaps.

8 Conclusion

We estimate the impact of NCA enforceability on workers' earnings, and we investigate the mechanism underlying this relationship. Using newly-assembled panel data on state-level NCA enforceability, we show that stricter NCA enforceability leads to a decline in workers' earnings and mobility. The earnings effect is greater for workers more likely to work under NCAs, and greater for females and racial minorities. We also find that the earnings effect of NCA enforceability spills over across legal jurisdictions, illustrating that NCA enforceability has far-reaching consequences on labor market outcomes, with effects that likely extend far beyond the subset of workers that actually sign NCAs.

Furthermore, we identify and find evidence of one mechanism underlying the relationship between earnings and NCA enforceability: stricter NCA enforceability undermines workers' ability to negotiate for pay increases when labor market conditions improve. This finding suggests that making NCAs enforceable fundamentally changes the way that workers and employers negotiate wages. Rather than setting wages consistent with a model of implicit contracts and *costless* mobility of workers (which a long literature has found to be the case), wages under strict NCA enforceability are instead consistent with a model of implicit contracts with *costly* mobility. This finding is not just important for academic reasons: given that on-the-job wage growth accounts for a meaningful share of workers' earnings growth over their career, our findings imply that NCA enforceability shuts down a primary way that workers can otherwise negotiate for higher pay over their job tenure.

Our results inform a longstanding debate regarding freedom of contract. An argument frequently cited in this debate is that workers would not sign NCAs if they were made worse off by doing so. However, at the *market* level, our findings imply that freedom to contract harms workers. This relationship arises due to the negative externalities from NCA use, which we find are economically meaningful. This relationship could also suggest that there are frictions in the labor market that mean that NCAs do not always enhance efficiency among firms that use them, though since we do not observe actual NCA use we cannot say so definitively.

Our results inform policy centered around NCAs on two major dimensions. First, enforceability of NCAs seriously inhibits growth of labor income. In theory, decreases in employee income may be more than compensated by increases in employer income. Therefore, short of limiting NCA enforceability, policymakers could consider other levers to override this transfer of income when NCAs are enforceable. Second, enforceability of NCAs causes negative externalities that affect workers in nearby states, and (plausibly) workers who have not signed NCAs. Internalization of externalities is one of the most widely accepted roles of government, suggesting that the existence of such external effects is an especially compelling rationale for government intervention in the use and enforceability of NCAs.

A limitation of our study is that we do not observe whether an individual worker has signed an NCA. However, our results inform what is the actual lever at policymakers' disposal: the enforceability of NCAs. Thus, our paper provides insight to state and federal lawmakers that are considering laws that amend, or even outright ban, employers' ability to use NCAs. As more data becomes available that measures NCA use at the worker- or firm-level over time, this will allow further studies into the earnings effects of NCA *use*, as opposed to *enforceability*. We look forward to future work in this domain.

References

- Arnold, D. (2019). Mergers and acquisitions, local labor market concentration, and worker outcomes. *Local Labor Market Concentration, and Worker Outcomes (October 27, 2019)*.
- Autor, D. H., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2017). The fall of the labor share and the rise of superstar firms.
- Azar, J., I. Marinescu, and M. I. Steinbaum (2017). Labor market concentration.
- Bagger, J., F. Fontaine, F. Postel-Vinay, and J.-M. Robin (2014). Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics. *American Economic Review* 104(6), 1551–96.
- Balasubramanian, N., J. W. Chang, M. Sakakibara, J. Sivadasan, and E. Starr (2018). Locked in? the enforceability of covenants not to compete and the careers of high-tech workers.
- Bannon, A. (2018). Choosing state judges: A plan for reform. *Brennan Center For Justice at NYU School of Law*.
- Barth, E. and H. Dale-Olsen (2009). Monopsonistic discrimination, worker turnover, and the gender wage gap. *Labour Economics* 16(5), 589–597.
- Beaudry, P. and J. DiNardo (1991). The effect of implicit contracts on the movement of wages over the business cycle: Evidence from micro data. *Journal of Political Economy*, 665–688.
- Belenzon, S. and M. Schankerman (2013). Spreading the word: Geography, policy, and knowledge spillovers. *Review of Economics and Statistics* 95(3), 884–903.
- Benmelech, E., N. Bergman, and H. Kim (2018). Strong employers and weak employees: How does employer concentration affect wages?
- Bernstein, D. E. (2008). Freedom of contract. *Liberty of Contract, in Encyclopedia of the Supreme Court of the United States (David S. Tanenhaus, 08–51)*.

- Bertrand, M. (2011). New perspectives on gender. In *Handbook of labor economics*, Volume 4, pp. 1543–1590. Elsevier.
- Bishara, N. D. (2010). Fifty ways to leave your employer: Relative enforcement of covenants not to compete, trends, and implications for employee mobility policy. *U. Pa. J. Bus. L.* 13, 751.
- Black, S. E. and E. Brainerd (2004). Importing equality? the impact of globalization on gender discrimination. *ILR Review* 57(4), 540–559.
- Black, S. E. and P. E. Strahan (2001). The division of spoils: rent-sharing and discrimination in a regulated industry. *American Economic Review* 91(4), 814–831.
- Bleakley, H. and J. Lin (2012). Thick-market effects and churning in the labor market: Evidence from us cities. *Journal of urban economics* 72(2-3), 87–103.
- Caldwell, S. and O. Danieli (2018). Outside options in the labor market. *Unpublished manuscript*.
- Caldwell, S. and N. Harmon (2019). Outside options, bargaining, and wages: Evidence from coworker networks. *Unpublished manuscript, Univ. Copenhagen*.
- Card, D., A. R. Cardoso, and P. Kline (2015). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics* 131(2), 633–686.
- Caughey, D. and C. Warshaw (2018). Policy preferences and policy change: Dynamic responsiveness in the american states, 1936–2014. *American Political Science Review* 112(2), 249–266.
- Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics* 134(3), 1405–1454.
- David, H., D. Dorn, and G. H. Hanson (2013). The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review* 103(6), 2121–68.
- Diamond, P. A. (1982). Wage determination and efficiency in search equilibrium. *The Review of Economic Studies* 49(2), 217–227.
- Dube, A., J. Jacobs, S. Naidu, and S. Suri (2018). Monopsony in online labor markets.
- Exley, C. L. and J. B. Kessler (2019). The gender gap in self-promotion.
- Farber, H. S., D. Herbst, I. Kuziemko, and S. Naidu (2018). Unions and inequality over the twentieth century: New evidence from survey data. Technical report, National Bureau of Economic Research.
- Gan, L. and Q. Li (2016). Efficiency of thin and thick markets. *Journal of econometrics* 192(1), 40–54.
- Garmaise, M. J. (2011). Ties that truly bind: Noncompetition agreements, executive compensation, and firm investment. *The Journal of Law, Economics, and Organization* 27(2), 376–425.
- Gittleman, M., M. A. Klee, and M. M. Kleiner (2018). Analyzing the labor market outcomes of occupational licensing. *Industrial Relations: A Journal of Economy and Society* 57(1), 57–100.
- Goldschmidt, D. and J. F. Schmieder (2017). The rise of domestic outsourcing and the evolution of the german wage structure. *The Quarterly Journal of Economics* 132(3), 1165–1217.
- Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing. *Unpublished*.
- Greenwald, B. C. (1986). Adverse selection in the labour market. *The Review of Economic Studies* 53(3), 325–347.
- Grossman, S. J. and O. D. Hart (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *The Journal of Political Economy*, 691–719.
- Haltiwanger, J., H. Hyatt, and E. McEntarfer (2018). Who moves up the job ladder? *Journal of Labor Economics* 36(S1), S301–S336.
- Haltiwanger, J. C., H. R. Hyatt, L. B. Kahn, and E. McEntarfer (2018). Cyclical job ladders

- by firm size and firm wage. *American Economic Journal: Macroeconomics* 10(2), 52–85.
- Hausman, N. and K. Lavetti (2017). Physician concentration and negotiated prices: Evidence from state law changes.
- Hernandez, M., D. R. Avery, S. D. Volpone, and C. R. Kaiser (2018). Bargaining while black: The role of race in salary negotiations. *Journal of Applied Psychology*.
- Hirsch, B. and D. Macpherson (2019). Union membership and coverage database from the cps.
- Jarosch, G., J. S. Nimczik, and I. Sorkin (2019). Granular search, market structure, and wages. Technical report, National Bureau of Economic Research.
- Jeffers, J. S. (2018). The impact of restricting labor mobility on corporate investment and entrepreneurship. *Unpublished*.
- Johnson, M. S. and M. Lipsitz (2019). Why are low-wage workers signing noncompete agreements? *Unpublished*.
- Katz, A. (2018). The chamber in the chambers: The making of a big-business judicial money machine. *DePaul Law Review* 67.
- Kini, O., R. Williams, and S. Yin (2019). Ceo non-compete agreements, job risk, and compensation. *Available at SSRN 3170804*.
- Kline, P., N. Petkova, H. Williams, and O. Zidar (2019). Who profits from patents? rent-sharing at innovative firms. *The Quarterly Journal of Economics* 134(3), 1343–1404.
- Knight, J. and L. Epstein (1996). The norm of stare decisis. *American Journal of Political Science* 40(4).
- Krueger, A. B. (2017). The rigged labor market. *Milken Institute Review*.
- Lamadon, Thibaut, M. M. and B. Setzler (2019). Imperfect competition, compensating differentials and rent sharing in the u.s. labor market.
- Lavetti, K., C. Simon, and W. D. White (2018). The impacts of restricting mobility of skilled service workers: Evidence from physicians. *Unpublished*.
- Le Barbanchon, T., R. Rathelot, and A. Roulet (2019). Gender differences in job search: Trading off commute against wage. *Available at SSRN 3467750*.
- Leibbrandt, A. and J. A. List (2014). Do women avoid salary negotiations? evidence from a large-scale natural field experiment. *Management Science* 61(9), 2016–2024.
- Lipsitz, M. and E. Starr (2019). Low-wage workers and the enforceability of non-compete agreements. *Unpublished*.
- Liu, K. (2019). Wage risk and the value of job mobility in early employment careers. *Journal of Labor Economics* 37(1), 139–185.
- Marx, M. (2011). The firm strikes back: non-compete agreements and the mobility of technical professionals. *American Sociological Review* 76(5), 695–712.
- Marx, M. (2018). Punctuated entrepreneurship (among women). *Unpublished*.
- Marx, M., J. Singh, and L. Fleming (2015). Regional disadvantage? employee non-compete agreements and brain drain. *Research Policy* 44(2), 394–404.
- Marx, M., D. Strumsky, and L. Fleming (2009). Mobility, skills, and the michigan non-compete experiment. *Management Science* 55(6), 875–889.
- McCarty, N. and B. Shor (2015). Measuring american legislatures aggregate data, v4.0.
- Molloy, R., R. Trezzi, C. L. Smith, and A. Wozniak (2016). Understanding declining fluidity in the us labor market. *Brookings Papers on Economic Activity* 2016(1), 183–259.
- Prager, E. and M. Schmitt (2019). Employer consolidation and wages: Evidence from hospitals. *Washington Center for Equitable Growth Working Paper*.
- Redbird, B. (2017). The new closed shop? the economic and structural effects of occupational licensure. *American Sociological Review* 82(3), 600–624.
- Robinson, J. (1933). *The economics of imperfect competition*. London: MacMillian.
- Rubin, P. H. and P. Shedd (1981). Human capital and covenants not to compete. *The Journal of Legal Studies* 10(1), 93–110.

- Schmieder, J. F. and T. Von Wachter (2010). Does wage persistence matter for employment fluctuations? evidence from displaced workers. *American Economic Journal: Applied Economics* 2(3), 1–21.
- Starr, E. (2018). Consider this: Training, wages, and the enforceability of covenants not to compete. *Unpublished*.
- Starr, E., N. Balasubramanian, and M. Sakakibara (2018). Screening spinouts? how non-compete enforceability affects the creation, growth, and survival of new firms. *Management Science* 64(2), 552–572.
- Starr, E., J. Prescott, and N. Bishara (2018). Noncompetes in the us labor force. *Unpublished*.
- Starr, E. P., J. Frake, and R. Agarwal (2018). Mobility constraint externalities.
- Topel, R. H. and M. P. Ward (1992). Job mobility and the careers of young men. *The Quarterly Journal of Economics* 107(2), 439–479.
- University of Kentucky Center for Poverty Research (2018). Ukcpr national welfare data, 1980-2017.
- Weil, D. (2014). *The fissured workplace*. Harvard University Press.
- Williamson, O. E. (1975). Markets and hierarchies. *New York* 2630.

9 Tables and Figures

Table 1: Descriptive Statistics on NCA Law Changes, 1991-2014

Region	Northeast	Midwest	South	West	Total
Average NCA Score	0.75	0.79	0.77	0.41	0.69
Standard Deviation of NCA Score	0.09	0.12	0.12	0.34	0.24
Maximum NCA Score	0.91	0.97	1.00	0.94	1.00
Minimum NCA Score	0.63	0.00	0.50	0.08	0.00
Number of Law Changes	21	20	25	16	82
Number of States in Region	9	12	17	13	51
Number of NCA Score Increases	13	15	15	9	52
Number of NCA Score Decreases	8	5	10	7	30
Average Magnitude Positive NCA Score Change	0.04	0.05	0.08	0.05	0.05
Maximum Positive NCA Score Change	0.15	0.11	0.24	0.17	0.24
Average Magnitude Negative NCA Score Change	-0.05	-0.04	-0.04	-0.03	-0.04
Maximum Negative NCA Score Change	-0.06	-0.06	-0.17	-0.09	-0.17
Between-State Standard Deviation	0.08	0.25	0.11	0.21	0.18
Within-State Standard Deviation	0.03	0.03	0.04	0.03	0.03

Notes: Statistics in the table represent data from 1991–2014, and the unit of observation is a state-year. The minimum and maximum of the NCA Score are normalized to 0 and 1, respectively.

Figure 1: Timing of NCA law changes from 1991 through 2014

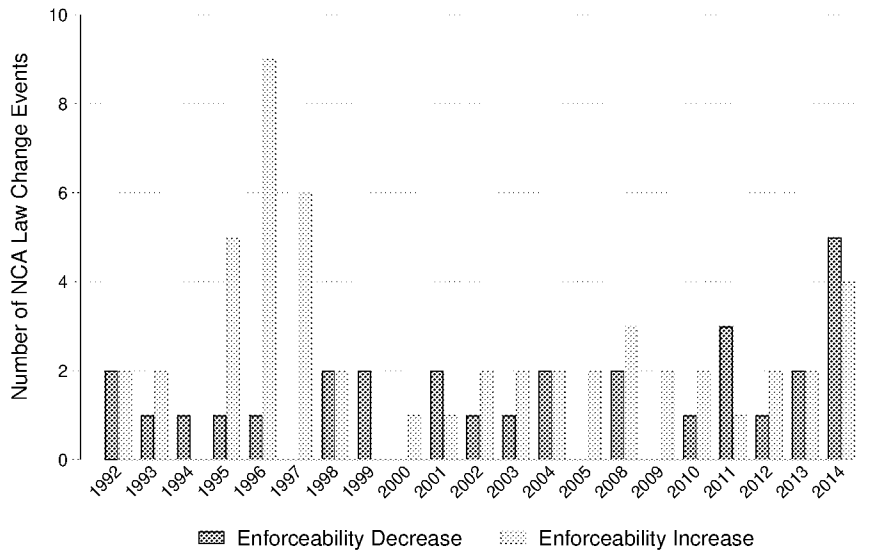
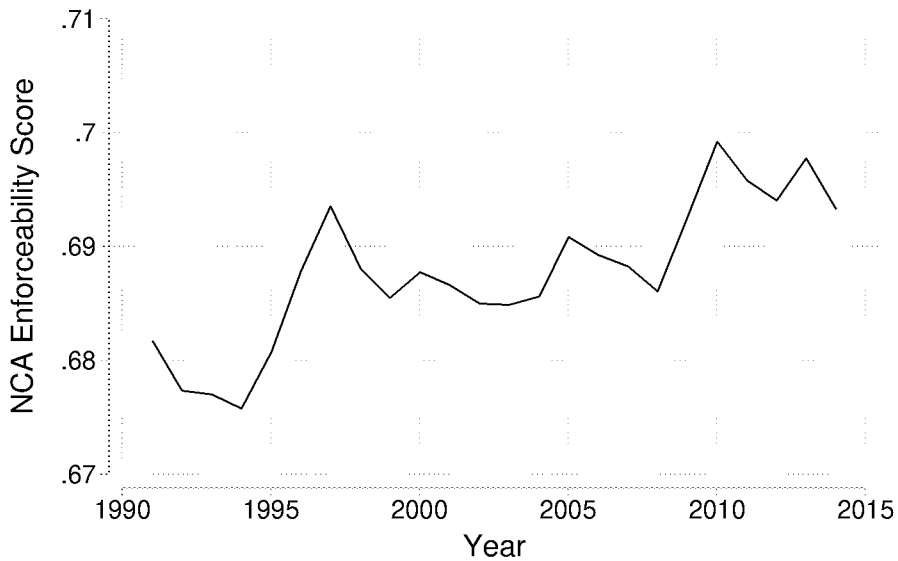


Figure 2: Average NCA Enforceability Score from 1991 to 2014



Notes: The series in this figure represents the population-weighted average NCA Score in the US in each year.

Table 2: Can Economic and Political Factors Explain Changes in NCA Enforceability?

Dependent Variable:	NCA Enforceability	
Population (100,000s)	-0.00	(0.00)
Unemployment Rate	0.00	(0.00)
Number of Workers Compensation Beneficiaries	-0.00	(0.00)
Democratic Party Governor	-0.01	(0.00)
% of State House from Democratic Party	0.01	(0.07)
% of State Senate from Democratic Party	0.04	(0.03)
State Minimum Wage	-0.01*	(0.01)
Number of Medicaid Beneficiaries (100,000s)	-0.00	(0.00)
Social Policy Liberalism Score	-0.00	(0.02)
Economic Policy Liberalism Score	-0.02	(0.01)
Social Mass Liberalism Score	-0.00	(0.02)
Economic Mass Liberalism Score	0.03	(0.04)
Democratic Party ID Count	-0.09	(0.31)
State House Ideology Score	-0.00	(0.01)
State Senate Ideology Score	0.00	(0.01)
House Democrats Ideology Score	-0.04	(0.04)
House Republicans Ideology Score	0.04	(0.05)
Senate Democrats Ideology Score	-0.03*	(0.02)
Senate Republicans Ideology Score	-0.00	(0.02)
Union Membership	-0.00	(0.00)
N	829	
R^2	0.113	
F-Test p-Value	0.184	

Notes: Models also include state and year fixed effects. Reported R^2 calculated after residualizing on state and year fixed effects. Standard errors reported in parentheses are clustered by state.

*** p<0.01, ** p<0.05, * p<0.1

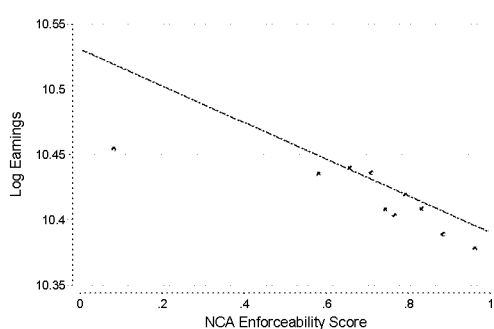
Table 3: The Effect of NCA Enforceability on Earnings

	Log Earnings		Log Hours	Log Wage	Log Average Earnings
	(1)	(2)	(3)	(4)	(5)
NCA Enforceability Score	-0.117*** (0.036)	-0.101*** (0.027)	-0.025 (0.017)	-0.101*** (0.027)	-0.135*** (0.031)
Observations	1216726	1216726	1545874	1216726	3548388
R^2	0.275	0.357	0.132	0.346	0.942
Geographic FE	State	State	State	State	County
Time FE	Div x Year	Div x Year	Div x Year	Div x Year	Div x Quarter
Occupation FE	N	Y	Y	Y	N
Sample	ASEC	ASEC	ASEC	ASEC	QWI

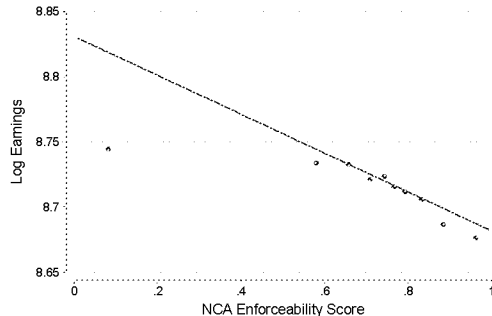
ASEC samples use years from 1991-2014 and include individuals between ages 18-64 who reported working for wage and salary income at a private employer. All ASEC regressions include controls for male, white, Hispanic, age, age squared, whether the individual did not complete college, and indicators for the metropolitan city center status of where the individual lives. Column (5) includes controls for male, age group, and county fixed effects. The dependent variable in Column (4), log hourly wage, is calculated as the log of total annual wage and salary income last year divided by (usual weekly hours last year times 52). Columns (1), (2), and (4) include full-time workers only, while Column (3) includes part-time workers to avoid selection on the dependent variable.

SEs clustered by state in parentheses. *** $P < .01$, ** $P < .05$, * $P < .1$

Figure 3: The Relationship between NCA Enforceability and Earnings: Binned Scatterplots



(a) All state-years with no controls.



(b) All state-years with 1-digit occupation code and demographic controls.

Each figure represents a binned scatterplot that relates an individual's log annual earnings to the NCA Enforceability score in his or her state that year. In each graph, both variables are residualized on state and Census division by year fixed effects. In panel (b), the variables are further residualized on broad occupation class fixed effects, age and age-squared, and indicators for white, Hispanic, male, not having completed college, and metro area status.

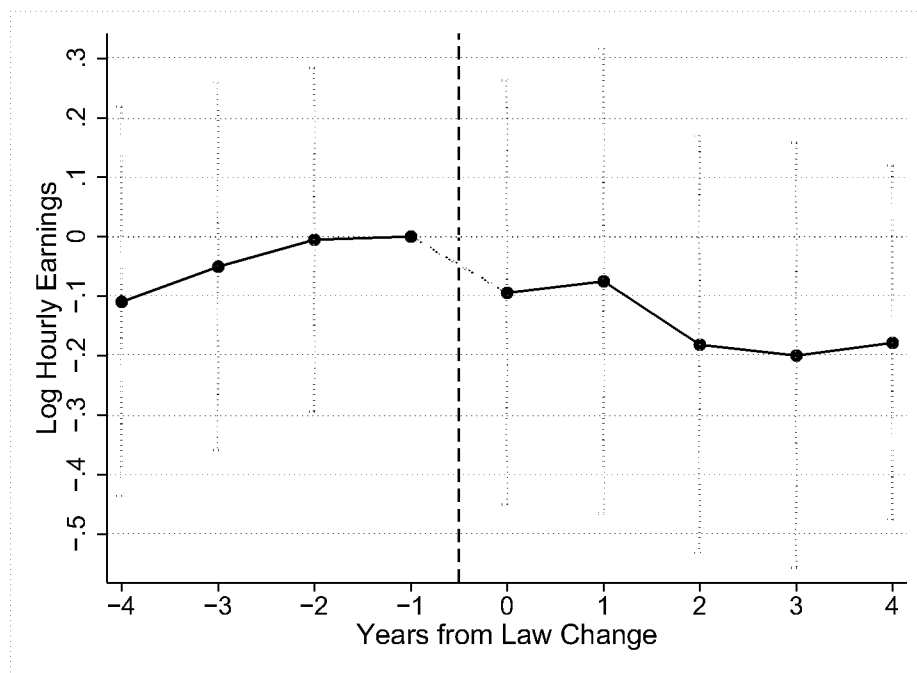
Table 4: Heterogeneous Effects of NCA Enforceability on Earnings by Education, Occupation, and Industry

	(1)	(2)	(3)	(4)	(5)
NCA Enforceability Score	-0.117*** (0.036)	-0.035 (0.039)	-0.083** (0.034)	-0.092*** (0.034)	-0.025 (0.036)
College Educated Worker	0.415*** (0.013)	0.514*** (0.021)	0.376*** (0.012)	0.391*** (0.010)	0.445*** (0.015)
College Educated Worker \times NCA Score		-0.143*** (0.032)			-0.122*** (0.023)
High NCA Use Occ			0.256*** (0.008)		0.194*** (0.005)
High NCA Use Occ \times NCA Score			-0.061*** (0.014)		-0.015* (0.008)
High NCA Use Ind				0.270*** (0.008)	0.220*** (0.007)
High NCA Use Ind \times NCA Score				-0.068*** (0.013)	-0.037*** (0.010)
Observations	1216726	1216726	1216726	1216726	1216726
R^2	0.275	0.275	0.290	0.292	0.304

The sample in all columns is the CPS ASEC from 1991-2014 and includes individuals between ages 18-64 who reported working for wage and salary income at a private employer the prior year. All regressions include fixed effects for state, fixed effects for Census region by year, fixed effects for broad occupational class, and individual controls for male, white, Hispanic, age, age squared, whether the individual did not complete college, and indicators for the metropolitan city center status of where the individual lives. In Columns (3) and (4), High NCA Use Occupations are occupations with NCA use greater than the national average, as tabulated by Starr et al. (2018).

SEs clustered by state in parentheses. *** $P < .01$, ** $P < .05$, * $P < .1$

Figure 4: Event Study Estimates of the Effect of NCA Enforceability Changes on Log Hourly Earnings



The sample includes four-year windows around NCA law change events, as well as control states in the same Census division with no corresponding event in the four-year window. The estimating equation includes controls for sex, age, age squared, level of education, race, Hispanic status, and whether or not the respondent lives in a metropolitan area, as well as state and Census division-by-year fixed effects. Coefficient estimates and 90% confidence intervals pictured (normalized to coefficient estimate one year prior to law change).

Table 5: The Effects of NCA Enforceability on Job Mobility

	J2J Separation Rate		Share J2J Across State		Share J2J Across Industry	
	(1)	(2)	(3)	(4)	(5)	(6)
NCA Enforceability Score	-0.006 (0.064)	0.062 (0.065)	0.009 (0.025)	0.003 (0.025)	-0.004 (0.012)	-0.005 (0.012)
High NCA Use Ind × NCA Score		-0.199*** (0.065)		0.017** (0.008)		0.004 (0.013)
Observations	677272	677272	659380	659380	668807	668807
R^2	0.866	0.867	0.602	0.602	0.667	0.667
Mean Dep Var	1.12	1.12	0.16	0.16	0.60	0.60

The sample is the J2J from 1991-2014. An observation is a state-sex-age group-quarter-industry cell. All regressions include controls for sex, age group, and industry, as well as division by year by quarter and state fixed effects.

Regressions are weighted by employment, and standard errors are clustered by state. ***P<.01, **P<.05, *P<.1

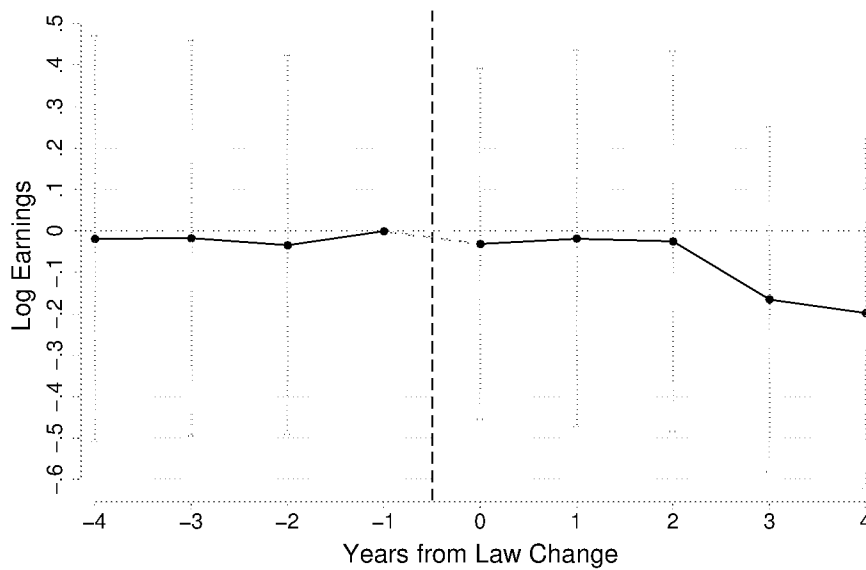
Table 6: The External Effects of NCA Enforceability on Earnings

	(1)	(2)	(3)
Own State NCA Score	-0.178*** (0.057)	-0.207*** (0.066)	-0.184*** (0.068)
Donor State NCA Score		-0.181** (0.076)	-0.210** (0.079)
Own Cty Emp/CZ Emp × Own State NCA Score			-0.124 (0.151)
Own Cty Emp/CZ Emp × Donor State NCA Score			0.163*** (0.054)
Observations	615097	615097	613679
R^2	0.898	0.898	0.901

The dependent variable is log earnings. The sample is the QWI from 1991-2014 restricted to counties directly on state borders in commuting zones that straddle a state border. An observation is a county-sex-age group-quarter. All regressions include controls for sex, age group, as well as division by year by quarter and county fixed effects. Own Cty Emp/CZ Emp is the ratio of sex- and age-group-specific employment in own county divided by sex- and age-group-specific employment in the entire commuting zone.

Standard errors are clustered by own state in Column (1), and two-way clustered by own state and commuting zone in columns (2) and (3). ***P<.01, **P<.05, *P<.1

Figure 5: Event Study Estimates of the Effect of Bordering State NCA Enforceability Changes on Log Earnings



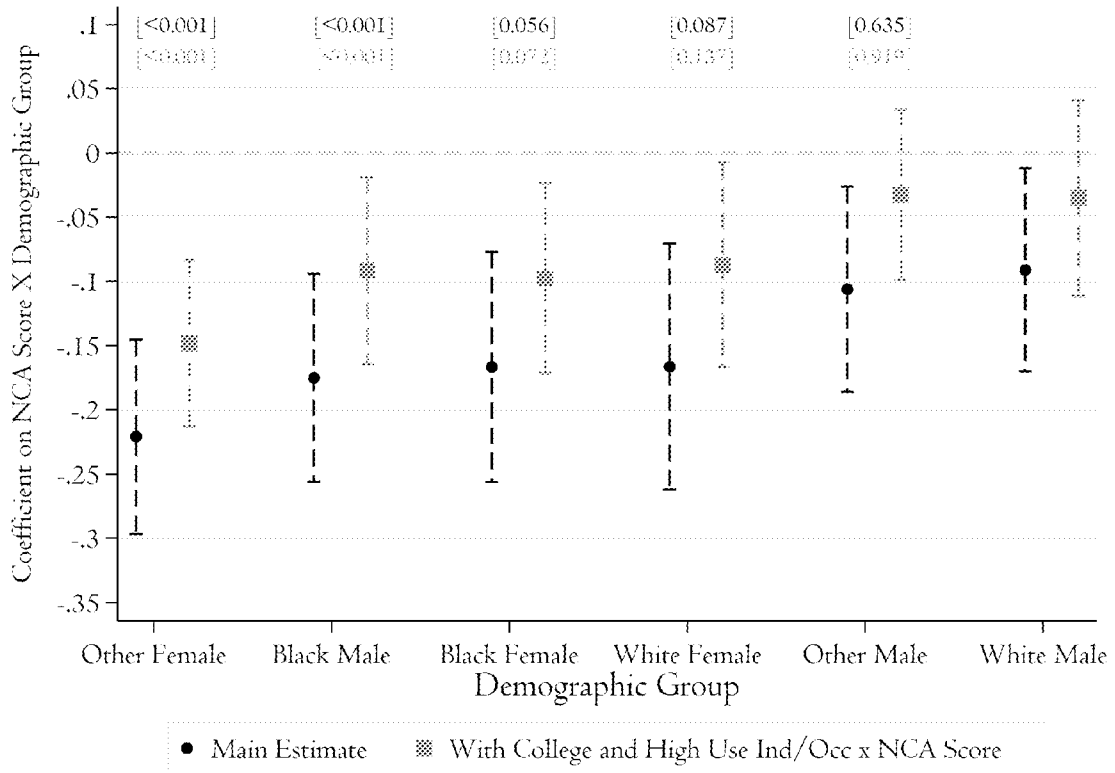
The sample includes four year windows around NCA law change events in the border county sample, as well as control counties in the same Census division with no corresponding event in a bordering county in the four year window. The estimating equation includes controls for sex and age group, as well as state and Census division-by-year fixed effects. Coefficient estimates and 90% confidence intervals pictured (normalized to coefficient estimate one year prior to law change).

Table 7: NCA Enforceability Changes How Workers and Employers Negotiate Implicit Contracts

	Log Earnings				
	(1)	(2)	(3)	(4)	(5)
Initial UR	-0.008*** (0.002)		-0.002 (0.003)	-0.002 (0.003)	0.010** (0.005)
Minimum UR		-0.017*** (0.003)	-0.014*** (0.005)	-0.014*** (0.005)	-0.028*** (0.006)
Initial NCA Score				0.007 (0.068)	-0.019 (0.082)
Init. NCA Score × Init. UR					-0.017** (0.006)
Init. NCA Score × Min. UR					0.020** (0.009)
No. Obs.	76350	76350	76350	76350	76350
R ²	0.364	0.364	0.364	0.364	0.364

The dependent variable is log weekly earnings. All regressions include state, Census division by year, and industry fixed effects, as well as controls for quadratics in age and tenure, and indicators for high school or less, black, Hispanic, married, union member, metro center status, and female. SEs clustered by state in parentheses. ***P<.01, **P<.05, *P<.1

Figure 6: Heterogeneous Effects of NCA Enforceability on Earnings by Race and Sex



The figure depicts coefficients from two regressions of earnings on NCA Score, interacted with demographic groups. The first regression builds on Column 1 of Table 3, adding indicators for each demographic group, as well as interactions of those indicators with NCA Score (the coefficients on which are depicted in the figure, along with 90% confidence intervals). The second regression adds controls for college education, high use occupation, and high use industry, as well as each interacted with NCA Score. The values in brackets report p-values for the *difference* between each coefficient and the coefficient for white males, with the main estimate above and the estimate including the extra controls below.

A Formalization of Theory

This appendix considers an augmentation of the model of Bagger et al. (2014). Bagger et al. (2014)’s baseline model of workers’ wage growth over their career uses a search and matching framework with human capital accumulation and on-the-job search. We consider a modification in which some workers sign NCAs with a firm, preventing their job mobility while employed by that firm. We consider channels linking wages and NCAs posited in Section 2, and derive conditions under which those channels would lead to the expected relationships in the model.

A.1 Summary of Bagger et al. (2014)

First, we introduce and summarize the model of Bagger et al. (2014). In that model, unemployed and employed workers match with prospective employers at rates λ_0 and λ_1 , respectively. Workers produce according to their human capital: a worker with human capital level h_t produces, in log terms, $y_t = p + h_t$, where p is the productivity of the firm, drawn from exogenous distribution $F(p)$. Workers are paid according to a piece rate: their wage is (again, in log terms) $w_t = r + p + h_t$, where $R = e^r \leq 1$ is the piece rate. The logged piece rate, r , is actually negative, meaning that it represents the amount of productivity that is “returned” to the employer. When exponentiated, the piece rate, R , therefore represents the *share* of productivity that is “returned” to the employer.

When unemployed workers match with a new employer, their wages are determined by setting the piece rate such that the worker receives a share, β , of the value of their match above and beyond the value of unemployment, which is assumed to be the value of matching with the least productive firm type, p_{min} . Employed workers who contact new employers may leave their current job (if the new employer is able to offer more attractive contract terms) or may leverage an outside offer to receive a wage increase (if the incumbent employer is able to offer more attractive contract terms), in either case receiving a share, β , of the match-specific rents above and beyond their relevant threat point. Workers also exogenously separate from their employers at rate $\delta \in [0, 1]$ (and immediately rematch at rate $\kappa \in [0, 1]$), and leave the labor force altogether at exogenous rate $\mu \in [0, 1]$. The discount rate is ρ .

We selected this model as a baseline due to the harmony between the drivers of wage growth in the model and the channels through which NCAs could affect wages that we discussed in Section 2. In the baseline model, workers’ wage growth occurs because of growth in their human capital, h_t , and their ability to search for higher-paying jobs. These two mechanisms for wage growth match well to potential roles for

NCA. First, NCAs are typically justified as a solution to a hold-up problem, where firms are not willing to invest in workers' human capital (e.g., training, imparting trade secrets, client lists, etc.) for fear that the worker will depart the firm and therefore deny the firm its return on investment. Therefore, an NCA in this model should cause h_t to grow at a greater rate, as the firm is more willing to invest in the worker. Second, NCAs prevent workers from changing jobs or threatening to change jobs, meaning that workers will not be able to increase wages by moving to a dominant firm, or by leveraging an outside offer to increase their wages at their current firm. The tradeoff between these two competing mechanisms will partially determine the difference in the rates of wage growth with and without an NCA for the worker.

A.2 Modifications to Bagger et al. (2014)

We operationalize NCAs in the model by assuming that workers exogenously sign enforceable NCAs with probability γ when they commence their first employment relationship (in other words, at the beginning of a worker's career, they randomly become an NCA worker or a non-NCA worker, and this designation will never change). With this inclusion of NCAs, we make three additional primary modifications to the model.

First, we assume that workers with NCAs accumulate human capital at a faster rate. Accumulation of human capital, h_t , is stochastic in Bagger et al. (2014), with the deterministic component of workers' human capital at time t represented by $g(t)$. Here, we define $g^C(t)$ and $g^F(t)$ to be the deterministic component of, respectively, a constrained (i.e., NCA-signing) and free (i.e., non-NCA signing) worker's human capital at time t ⁴⁵. Since human capital evolves faster for those with NCAs, if $g^C(t-1) = g^F(t-1)$, then $g^C(t) > g^F(t)$. This assumption is a natural implication of the argument that NCAs solve a hold-up problem. Firms might be unwilling to invest in human capital of workers who can freely leave, because they do not expect to recoup the returns on their investment. NCAs, by ensuring that workers cannot freely leave, incentivize firms to invest in workers, causing their human capital to develop more rapidly.

The second primary modification is that workers with NCAs are unable to change jobs: the offer arrival rate of new jobs for employed workers with NCAs is zero, or $\lambda_1^C = 0$. In other words, if a worker has an NCA, they will continue to work for the

⁴⁵The superscripts C and F will be used frequently to differentiate functions and parameters that differ between signers and non-signers.

same employer unless they experience an exogenous separation.⁴⁶ Though assuming that NCAs strictly prohibit job changing may seem drastic (because, for example, workers may be able to buy out of NCAs or can move to firms in different industries or geographic locations), this assumption substantially improves tractability and does not change the predictions of the model. For example, the model generates qualitatively similar predictions to a model in which workers must pay a cost to change jobs (e.g. representing an explicit buyout cost); indeed, if the cost is steep enough, in the limit assuming a cost is identical to assuming that the worker is unable to change jobs.

The third modification we make, outlined in Section A.4, is assuming that the offer arrival rate is lower for workers in thinner labor markets (i.e., markets with a lower measure of workers available to match to new firms). Specifically, we allow the offer arrival rate for employed workers in jobs with no NCA, λ_1 , to vary with T . T represents the thickness of the freely mobile labor market (i.e., the measure of workers available to match in non-NCA jobs) defined as $T \equiv N^F + (1 - \gamma)U$ (where N^F is the proportion of workers employed in non-NCA jobs, and $(1 - \gamma)U$ is the proportion of unemployed workers who will randomly match to firms without an NCA). We assume that $\frac{d\lambda_1(T)}{dT} > 0$: the thicker the labor market, the more often workers will be contacted on-the-job. This modification, which generates spillover effects in the model, has sound empirical justification. As shown in Starr et al. (2018), when NCAs are used at a higher rate in a given state-industry combination, workers who do not sign NCAs in that state-industry receive fewer job offers. Furthermore, in thin labor markets (which may result from NCAs since, when a worker signs an NCA, the supply and demand sides of the labor market are diminished as workers and vacancies disappear from the unmatched pool), workers and firms match less often (Bleakley and Lin, 2012; Gan and Li, 2016), resulting in lower contact rates in this model.

Under these modifications, we now generate multiple predictions which relate directly to the empirical work found in this paper.

⁴⁶We make two additional modifications related to this one. First, we assume that, after an exogenous separation, a worker who had previously signed an NCA will continue to work in a job with an NCA. This assumption significantly increases tractability by limiting flows between the two types of jobs. One way to view this assumption is that workers work in industries that use NCAs or in industries that do not; this could occur due to the value of accumulated industry-specific human capital. The second assumption is that workers immediately find new work upon an exogenous separation with their employer. This assumption increases tractability of the model and does not change the qualitative predictions. Furthermore, we view it as reasonable: roughly half of states do not enforce NCAs when employees are fired, leaving such workers able to find other jobs quickly in the event of an involuntary separation.

A.3 Direct Effects of NCAs on Wages of NCA Signers

First, we examine the wages of a worker who signs versus does not sign an NCA, from which we can extrapolate the average wages when the proportion of workers who sign NCAs, γ , changes. Wages depend on human capital (which develops more rapidly when workers have NCAs) and mobility (which is restricted when workers sign NCAs). This tension forms the substance of the direct effects.

The wage of a worker is given by $w_{i,t} = \alpha_i + g^j(t) + \varepsilon_{i,t} + p_{i,t} + r$, where α_i is a worker heterogeneity parameter, $g^j(t)$ is the deterministic component of human capital accumulation of the worker for $j \in \{C, F\}$ for workers who are constrained by an NCA or free to change employers, respectively, and $\varepsilon_{i,t}$ is a stochastic worker human capital shock. Firm productivity, $p_{i,t}$ (where i represents the worker and t represents time), and r (the piece rate of the worker) round out wages.

In order to compare wages across workers, we compare the individual components of wages. By assumption, ε is distributed identically across workers and across time, and α is distributed identically across workers, so in expectation, there are no differences in ε or α for workers with and without NCAs.

By assumption, human capital evolves at a higher rate for those with NCAs: if $g^C(t-1) = g^F(t-1)$, then $g^C(t) > g^F(t)$.

What is left to compare are firm productivities and the piece rates of workers. Intuitively, workers with NCAs will face a worse distribution of firm productivities because they are unable to search for higher-paying jobs—*i.e.* they are unable to climb the job ladder. In fact, since they are immobile and exit occurs independently of firm productivity, the distribution of firm productivities at which NCA-constrained workers are employed ($L^C(p)$) is exactly equal to the exogenous productivity distribution for a worker entering employment: $L^C(p) = F(p)$.

The steady state distribution for those who do not sign NCAs is derived in Bagger et al. (2014) (equation A14): $L^F(p) = \frac{(\mu+\delta)F(p)}{\mu+\delta+\lambda_1 F(p)}$, where $\bar{F}(p) = 1 - F(p)$. Since workers only move *up* the job ladder, $L^F(p)$ first-order stochastically dominates $L^C(p)$.

Finally, we turn to piece rates. Piece rates for nonsigners evolve identically to those in the baseline model of Bagger et al. (2014). However, the piece rate for signers does not evolve over time: lacking the ability to change the piece rate by leveraging outside offers or engaging in job-to-job mobility, the piece rate for a worker with an NCA is determined at the advent of their job spell.

In Bagger et al. (2014), the piece rate (r) is a function of the most recent firm from which the worker was able to, or would have been able to, extract all available surplus (by virtue of having a high enough competing offer)⁴⁷:

⁴⁷Note that the piece rate is negative: wages are given by $w_t = r + p + h_t$, where $p + h_t$ is the

$$r = - \int_{q_{i,t}}^{p_{i,t}} \phi(x) dx$$

where $\phi(x) = (1 - \beta) \frac{\rho + \delta + \mu + \lambda_1 \bar{F}(x)}{\rho + \delta + \mu + \lambda_1 \beta \bar{F}(x)}$, $F(x) = 1 - \bar{F}(x)$ is the exogenous distribution of firm productivities from which workers draw upon matching with a firm, and $q_{i,t}$ represents the productivity of the last firm from which the worker was able to extract all surplus, by virtue of leveraging a competing offer (see Equation 6 in Bagger et al. (2014) for details on the derivation of this equation). The greater is $q_{i,t}$, the greater the worker's wage will be. If $q_{i,t} = p_{i,t}$, then the worker was able to extract all surplus from their current firm and therefore $r = 0$.

In the case of an NCA signer, the last "job" from which the worker was able to extract all surplus was unemployment, since workers cannot leverage outside options or job hop. The piece rate of signers is therefore determined by the worker having outside option p_{min} (the lowest productivity a firm can have), since by assumption, the value of unemployment is equal to the value of employment in the least productive firm. Simplifying (since $\lambda_1^C = 0$ for signers by assumption), the piece rate of NCA signers will be:

$$\begin{aligned} r &= - \int_{p_{min}}^{p_{i,t}} \phi(x) dx \\ &= - \int_{p_{min}}^{p_{i,t}} (1 - \beta) \frac{\rho + \delta + \mu + \lambda_1^C \bar{F}(x)}{\rho + \delta + \mu + \lambda_1^C \beta \bar{F}(x)} dx = -(p_{i,t} - p_{min})(1 - \beta) \end{aligned}$$

The wage processes of signers versus nonsigners are given by:

$$\text{Nonsigners: } w_{i,t}^F = \alpha_i + g^F(t) + \varepsilon_{i,t} + p_{i,t} - \int_{q_{i,t}}^{p_{i,t}} \phi(x) dx$$

$$\text{Signers: } w_{i,t}^C = \alpha_i + g^C(t) + \varepsilon_{i,t} + p_{i,t} - (p_{i,t} - p_{min})(1 - \beta)$$

We now compare expected wages for workers with and without an NCA. First, we examine workers new to the workforce:

Proposition A.1. *In steady state, workers signing NCAs will receive higher initial wages in expectation than workers not signing NCAs: $E_{i,t-1}[w_{i,t}^C] > E_{i,t-1}[w_{i,t}^F]$*

Proof. In the first period in which workers match, the firm productivity distributions are identical (since workers have not had a chance to switch jobs). In expectation, α_i

marginal product of the worker (p is the firm's productivity and h_t is the worker's productivity due to human capital accumulation). Therefore, the piece rate r represents the share of the worker's productivity that is allocated to the firm.

and $\varepsilon_{i,t}$ are identical for those with and without NCAs. By assumption, $E_{t-1}[g^C(t)] > E_{t-1}[g^F(t)]$, so the proposition is proven if

$$E_{i,t}[(p_{i,t} - p_{min})(1 - \beta)] < E_{i,t} \left[\int_{p_{min}}^{p_{i,t}} \phi(x) dx \right],$$

since the worker initially bargains with outside option p_{min} .

Rewriting the right hand side, we must show that

$$E_{i,t} \left[\int_{p_{min}}^{p_{i,t}} (1 - \beta) dx \right] < E_{i,t} \left[\int_{p_{min}}^{p_{i,t}} \phi(x) dx \right],$$

which is true since $\phi(x) > (1 - \beta) > 0$. □

This proposition highlights two reasons for greater (initial) pay under NCAs: first, a greater accumulation of human capital leading to greater productivity, and second, the compensating differential associated with NCAs. Workers who initially match with NCAs are compensated to some extent for their limited future mobility.

However, as workers remain at their jobs longer, three things happen: first, workers with NCAs accumulate more human capital. Second, workers without NCAs climb the job ladder, moving to jobs with greater firm productivities, $p_{i,t}$. Third, when they leverage outside offers, they negotiate better piece rates, r . The first increases earnings by more for those who sign NCAs, while the latter two increase earnings by more for those who do not sign NCAs. The overall comparison, then, is indeterminate: if human capital grows more quickly than mobile workers climb the job ladder and negotiate better piece rates, workers with NCAs will have earnings that grow more quickly than those without, and vice versa. We summarize in Proposition A.2, but first introduce the condition used in the proposition. The condition states that the growth rate of human capital is lower than the growth rate of the lost ability of the worker to bargain for higher wages, and as such, the proposition is just an algebraic simplification. While there is no major intuitive leap contained in the proposition, its goal is to show that there is a direct tradeoff between human capital growth and job mobility which governs wage dynamics.

Condition 1.

$$\begin{aligned}
& E_t[(g^C(t+1) - g^C(t)) - (g^F(t+1) - g^F(t))] \\
& < \left(\int_{q_{j,t}}^{p_{j,t}} \int_{p_{j,t-1}}^p \phi(x) dx dF(p) \right) \\
& + \left(\int_{p_{j,t}}^{p_{max}} p - p_{j,t} - \left(\int_{p_{j,t}}^p \phi(x) dx - \int_{q_{j,t}}^{p_{j,t}} \phi(x) dx \right) dF(p) \right)
\end{aligned}$$

Proposition A.2. *If worker i has an NCA and worker j does not, then conditional on remaining employed and experiencing identical shocks in period t ($\varepsilon_{i,t} = \varepsilon_{j,t}$), $E_t[w_{i,t+1}] - w_{i,t} < E_t[w_{j,t+1}] - w_{j,t}$ whenever Condition 1 holds, and $E_t[w_{i,t+1}] - w_{i,t} > E_t[w_{j,t+1}] - w_{j,t}$ when it does not.*

Proof. The condition is an algebraic simplification of the inequality $E_t[w_{i,t+1}] - w_{i,t} < E_t[w_{j,t+1}] - w_{j,t}$. The left hand side may be rewritten as:

$$E_t[\alpha_i + \varepsilon_{i,t+1} + g^C(t+1) + p_{i,t+1} - (1-\beta)(p_{i,t+1} - p_{min})] - [\alpha_i + \varepsilon_{i,t} + g^C(t) + p_{i,t} - (1-\beta)(p_{i,t} - p_{min})]$$

Since $p_{i,t} = p_{i,t+1}$ for i , who has an NCA, this reduces to $E_t[g^C(t+1) - g^C(t) + \varepsilon_{i,t+1} - \varepsilon_{i,t}]$. The right hand side may be rewritten as

$$\begin{aligned}
& E_t[\alpha_j + \varepsilon_{j,t+1} + g^F(t+1) + p_{j,t+1} - \int_{q_{j,t+1}}^{p_{j,t+1}} \phi(x) dx] - [\alpha_j + \varepsilon_{j,t} + g^F(t) + p_{j,t} - \int_{q_{j,t}}^{p_{j,t}} \phi(x) dx] \\
& = E_t[g^F(t+1) - g^F(t) + \varepsilon_{j,t+1} - \varepsilon_{j,t}] \\
& \quad - \left[\int_{q_{j,t}}^{p_{j,t}} \left(\int_p^{p_{j,t}} \phi(x) dx - \int_{q_{j,t}}^{p_{j,t}} \phi(x) dx \right) dF(p) \right] \\
& \quad + \left[\int_{p_{j,t}}^{p_{max}} p - p_{j,t} - \left(\int_{p_{j,t}}^p \phi(x) dx - \int_{q_{j,t}}^{p_{j,t}} \phi(x) dx \right) dF(p) \right] \\
& = E_t[g^F(t+1) - g^F(t) + \varepsilon_{j,t+1} - \varepsilon_{j,t}] \\
& \quad + \left(\int_{q_{j,t}}^{p_{j,t}} \int_{q_{j,t}}^p \phi(x) dx dF(p) \right) \\
& \quad + \left[\int_{p_{j,t}}^{p_{max}} p - p_{j,t} - \left(\int_{p_{j,t}}^p \phi(x) dx - \int_{q_{j,t}}^{p_{j,t}} \phi(x) dx \right) dF(p) \right]
\end{aligned}$$

We expand the expectation by using the fact that the lowest productivity level a worker will be able to leverage to achieve an increase in earnings is $q_{j,t}$. If the worker contacts a new employer whose productivity is less than $q_{j,t}$, productivity will not

change and the worker will not renegotiate the piece rate. If the worker contacts a new employer with productivity between $q_{j,t}$ and $p_{j,t}$, they will remain employed at productivity $p_{j,t}$ but will renegotiate the piece rate. Finally, if the worker contacts a new employer with productivity above $p_{j,t}$, the worker will change jobs, changing both productivity and the piece rate.

Combination of the reduced right and left hand sides yields the condition stated in the proposition. \square

Proposition A.2 simplifies the condition under which workers have larger wage growth with NCAs versus without. An alternative way of interpreting this proposition is that, when the inequality condition holds, workers without NCAs will see wage increases relative to workers with NCAs.

Let \bar{w}_t^S represent average wages for workers with tenure t (where $t = 0$ represents workers new to the labor market) with or without NCAs ($S \in \{C, F\}$). Since workers initially match using an identical productivity distribution, since α and ε are distributed identically for workers with and without NCAs, and since separation is independent of wage or productivity, induction on Proposition A.2 generates the following corollary immediately:

Corollary A.3. $\bar{w}_0^C - \bar{w}_0^F > \bar{w}_t^C - \bar{w}_t^F$ whenever Condition 1 holds, and $\bar{w}_0^C - \bar{w}_0^F < \bar{w}_t^C - \bar{w}_t^F$ when it does not.

Corollary A.3 generates an indeterminate prediction regarding the relationship between average earnings and NCA enforceability, explored in Section 4. The balance of the tradeoff between human capital growth and climbing the job ladder cannot be assessed theoretically, and we therefore test it empirically.

A.4 Indirect Effects NCAs on Wages of non-NCA Signers

As described above, NCAs thin labor markets, which reduces on-the-job contact rates even for workers without NCAs (Bleakley and Lin, 2012; Gan and Li, 2016). For the remainder of this section, we therefore assume that the on-the-job contact rate (λ_1) is a function of market thickness. As described above, we define market thickness as $T \equiv N^F + (1 - \gamma)U$, where N^F is the proportion of workers employed in jobs with no NCA, and U is the proportion of unemployed workers. Multiplying U by $1 - \gamma$ yields the proportion of workers who are unemployed and will randomly match to firms without an NCA. We assume, then, that $\frac{d\lambda_1(T)}{dT} > 0$: the thicker the labor market, the more often workers will be contacted on-the-job.

To derive conditions under which increased frequency of NCA use yields lower earnings for workers without NCAs, we must first generate expressions representing

the proportion of matched workers, in steady state, without an NCA (N^F), with an NCA (N^C), and who are unemployed ($U \equiv 1 - N^C - N^F$). We do this using flow equations for each type of worker. The flow into unemployment is $(N^C + N^F)(\mu + \delta(1 - \kappa))$, and the flow out is $\lambda_0(1 - N^C - N^F)$, generating the flow balance equation

$$1 - N^C - N^F = \frac{\mu + \delta(1 - \kappa)}{\mu + \delta(1 - \kappa) + \lambda_0},$$

which is effectively identical to equation A7 in Bagger et al. (2014).

The flow balance equations for jobs with and without NCAs, respectively, at productivity p or less are given by

$$\begin{aligned} \text{NCA: } \lambda_0(1 - N^C - N^F)F(p)\gamma + \delta\kappa(1 - L^C(p))N^C F(p) \\ = [\mu + \delta(1 - \kappa) + \delta\kappa\bar{F}(p)]L^C(p)N^C \end{aligned}$$

$$\begin{aligned} \text{No NCA: } \lambda_0(1 - N^C - N^F)F(p)(1 - \gamma) + \delta\kappa(1 - L^F(p))N^F F(p) \\ = [\mu + \delta(1 - \kappa) + (\delta\kappa + \lambda_1(T))\bar{F}(p)]L^F(p)N^F \end{aligned}$$

Setting $p = p_{max}$ and solving generates closed form solutions for N^C and N^F , as well as for $1 - N^C - N^F$:

$$\begin{aligned} N^F &= \frac{\lambda_0(1 - \gamma)}{\mu + \delta(1 - \kappa) + \lambda_0} \\ N^C &= \frac{\lambda_0\gamma}{\mu + \delta(1 - \kappa) + \lambda_0} \\ 1 - N^C - N^F &= \frac{\mu + \delta(1 - \kappa)}{\mu + \delta(1 - \kappa) + \lambda_0} \end{aligned}$$

Therefore, $T = 1 - \gamma$, and $\frac{dT}{d\gamma} = -1$: market thickness is negatively associated with the probability that workers sign an NCA, all else equal. Under the assumption that $\frac{d\lambda_1(T)}{dT} > 0$, it follows that $\frac{d\lambda_1(T)}{d\gamma} < 0$.

The remaining analysis of spillover effects mirrors the analysis of the direct effect of NCAs on wages, to some extent. Workers who take a new job will receive slightly higher initial wages when γ increases. This is because the firm's job offer includes an implicit promise of future mobility, which is cheapened when future mobility is partially compromised due to a thin market. Therefore, the firm must increase the

worker's piece rate to continue to guarantee a β share of that firm's surplus:

Proposition A.4. $\frac{dE_{t=0}[w_{i,1}]}{d\gamma} > 0$

Proof. The workers expected wage before matching with an employer is

$$E_{t=0}[w_{i,1}] = (1 - \lambda_0)V_0(h_0) + \lambda_0 E_{t=0} \left[\alpha_i + g(1) + \varepsilon_{i,1} + p_{i,1} - \int_{p_{min}}^{p_{i,t}} \phi(x) dx \right]$$

Therefore:

$$\begin{aligned} \frac{dE_{t=0}[w_{i,1}]}{d\gamma} &= - \int_{p_{min}}^{p_{i,t}} \frac{d\phi(x)}{d\gamma} dx \\ &= - \int_{p_{min}}^{p_{i,t}} \frac{1 - \beta}{(\rho + \delta + \mu + \lambda_1(T(\gamma))\bar{F}(x))^2} (\bar{F}(x)(1 - \beta)(\rho + \delta + \mu) \frac{d\lambda_1}{d\gamma}) dx \\ &> 0, \end{aligned}$$

where the last inequality follows because $\frac{d\lambda_1}{d\gamma} < 0$. □

Wage growth after a match follows a pattern similar to that found in Proposition A.2, except that workers not bound by NCAs who are subject to spillovers from NCAs do not experience faster than usual growth of human capital. Therefore, workers with NCAs experience slower wage growth than workers without NCAs, which will eventually lead to lower wages for those workers, even in the presence of the compensating differential described in Proposition A.4.

A.5 Empirical Implications of Theoretical Results

Suppose we found that, in the cross section, stricter NCA enforceability led to lower wages on average. How would we rationalize such a result using the theoretical model we have just described? Ignoring any spillover effects, this result would imply, given Corollary A.3, that Condition 1 holds: the lower wage growth due to diminishment of the gains from job search dominates any benefit of faster human capital accumulation (the “late-career effect”), and this negative effect outweighs any initial compensating differential the worker receives (the “early career effect”). Other models could yield alternative explanations. For example, even though the Bagger et al. (2014) does not admit a compensating differential for human capital accumulation, a different model could imply that workers accept a *negative* compensating differential for signing an NCA, since workers anticipate faster future wage growth due to faster human capital accumulation. In this latter explanation, a negative relationship between NCA enforceability and wages on average would not imply that workers are worse off under stricter enforceability.

While we cannot rule out this latter explanation, we present empirical results in Section 6 that support an interpretation that is consistent with the mechanisms implied by our model. In that section, we show evidence that NCA enforceability reduces the benefits from job search by limiting workers' mobility, particularly in strong labor markets. Strong labor markets will positively impact the right hand side of Condition 1, for example by increasing the arrival rate of offers for workers able to accept them. Therefore, strong labor markets make Condition 1 more likely to be satisfied. An unanticipated positive labor market shock enabled a worker not bound by an NCA to leverage his improved outside option to bargain for a higher wage; a similar worker bound by an NCA cannot credibly threaten to leave, and thus is unable to take advantage of the stronger labor market to secure a wage increase. We describe these results more fully in Section 6.

B Appendix Figures & Tables

Table B.1: Dimensions of NCA Enforceability, According to Bishara (2010)

Question Number	Question
Q1	Is there a state statute that governs the enforceability of covenants not to compete?
Q2	What is an employer's protectable interest and how is that defined?
Q3	What must the plaintiff be able to show to prove the existence of an enforceable covenant not to compete?
Q3a	Does the signing of a covenant not to compete at the inception of the employment relationship provide sufficient consideration to support the covenant?
Q3b/c	b) Will a change in the terms and conditions of employment provide sufficient consideration to support a covenant not to compete entered into after the employment relationship has begun? / c) Will continued employment provide sufficient consideration to support a covenant not to compete entered into after the employment relationship has begun?
Q4	If the restrictions in the covenant not to compete are unenforceable because they are overbroad, are the courts permitted to modify the covenant to make the restrictions more narrow and to make the covenant enforceable? If so, under what circumstances will the courts allow reduction and what form of reduction will the courts permit?
Q8	If the employer terminates the employment relationship, is the covenant enforceable?

Source: Bishara (2010).

Table B.2: The Effect of NCA Enforceability on Earnings, by Component of NCA Score

Q1: State Statute	-0.031	(0.024)
Q2: Protectable Interest	-0.045*	(0.025)
Q3: Plaintiff Burden of Proof	0.040	(0.028)
Q3a: Consideration, Start of Employment	-0.055***	(0.014)
Q3bc: Consideration, Continued Employment	-0.031**	(0.012)
Q4: Judicial Modification	-0.023	(0.016)
Q8: Enforceable if Employer Terminates	0.011	(0.039)
NCA Score without Question 1	-0.120***	(0.037)
Observations	1216726	

Each of the first seven rows represents a separate regression (corresponding to Column 1 of Table 3) in which the variable $Enforceability_{st}$ in Equation 1 has been replaced with each component of the NCA Enforceability Score separately. The coefficient on the score component is reported, alongside SEs clustered by state in parentheses. The final row uses as an independent variable a modified NCA Enforceability Score that omits the score for Q1 (whether there exists a state statute that governs NCA enforceability) in the calculation, but is otherwise equivalent to the NCA Enforceability Score used in the main analysis.

***P<.01, **P<.05, *P<.1

Table B.3: The Effect of NCA Enforceability on Earnings: Robustness to Political & Economic Controls

	Log Earnings		Log Hours	Log Wage	Log Average Earnings
	(1)	(2)	(3)	(4)	(5)
NCA Enforceability Score	-0.077**	-0.071***	-0.030**	-0.068***	-0.113***
	(0.033)	(0.024)	(0.013)	(0.023)	(0.020)
Observations	1139890	1139890	1448431	1139890	3431264
R^2	0.274	0.357	0.132	0.346	0.942
Geographic FE	State	State	State	State	County
Time FE	Div x Year	Div x Year	Div x Year	Div x Year	Div x Quarter
Occupation FE	N	Y	Y	Y	N
Sample	ASEC	ASEC	ASEC	ASEC	QWI

This table replicates Table 3, but additionally controls for all variables (except ideology variables) introduced in Table 2.

SEs clustered by state in parentheses. ***P<.01, **P<.05, *P<.1

Table B.4: The Effect of NCA Enforceability on Earnings: Excluding States with Partisan Judicial Elections

	Log Earnings		Log Hours	Log Wage	Log Average Earnings
	(1)	(2)	(3)	(4)	(5)
NCA Enforceability Score	-0.135*** (0.042)	-0.121*** (0.032)	-0.046*** (0.015)	-0.122*** (0.032)	-0.153*** (0.037)
Observations	989854	989854	1262128	989854	2695840
R^2	0.272	0.356	0.130	0.345	0.941
Geographic FE	State	State	State	State	County
Time FE	Div x Year	Div x Year	Div x Year	Div x Year	Div x Year-Quarter
Occupation FE	N	Y	Y	Y	N
Sample	ASEC	ASEC	ASEC	ASEC	QWI

This table replicates Table 3, but drops the 6 states in which judges are selected via partisan election. SEs clustered by state in parentheses. ***P<.01, **P<.05, *P<.1

Table B.5: The Effect of NCA Enforceability on Earnings: Excluding States with Judicial Elections (Partisan or Non-partisan)

	Log Earnings		Log Hours	Log Wage	Log Average Earnings
	(1)	(2)	(3)	(4)	(5)
NCA Enforceability Score	-0.116 (0.092)	-0.112 (0.074)	-0.045** (0.019)	-0.107 (0.074)	-0.125 (0.082)
Observations	699036	699036	890737	699036	1531543
R^2	0.272	0.359	0.128	0.348	0.941
Geographic FE	State	State	State	State	County
Time FE	Div x Year	Div x Year	Div x Year	Div x Year	Div x Year-Quarter
Occupation FE	N	Y	Y	Y	N
Sample	ASEC	ASEC	ASEC	ASEC	QWI

This table replicates Table 3, but drops the 21 states in which judges are selected via election (partisan or non-partisan).

SEs clustered by state in parentheses. ***P<.01, **P<.05, *P<.1

Table B.6: The External Effects of NCA Enforceability on Earnings (Weighted by Employment)

	(1)	(2)	(3)
Own State NCA Score	-0.063 (0.042)	-0.067 (0.045)	-0.050 (0.049)
Donor State NCA Score		-0.014 (0.054)	-0.116* (0.067)
Own Cty Emp/CZ Emp \times Own State NCA Score			-0.078 (0.093)
Own Cty Emp/CZ Emp \times Donor State NCA Score			0.257** (0.109)
Observations	613679	613679	613679
R^2	0.943	0.943	0.943

The dependent variable is log earnings. The sample is the QWI from 1991-2014 and includes individuals between ages 19-64. All regressions include controls for male, age group, as well as division by year by quarter and county fixed effects. Own Cty Emp/CZ Emp is the ratio of sex- and age-group-specific employment in own county divided by sex- and age-group-specific employment in the entire commuting zone. Each regression is weighted by cell-specific employment. Standard errors are clustered by own state in Column (1), and two-way clustered by own state and commuting zone in columns (2) and (3). ***P<.01, **P<.05, *P<.1

Table B.7: The External Effects of NCA Enforceability on Earnings on Counties Far from State Borders

	(1)	(2)	(3)	(4)
Own State NCA Score	-0.200*** (0.060)	-0.205*** (0.056)	-0.184*** (0.055)	-0.206 (0.190)
Nearest Neighboring State's NCA Score	-0.143** (0.060)	-0.063 (0.049)	-0.042 (0.050)	0.009 (0.079)
Observations	615097	2015741	1594870	545696
R^2	0.898	0.889	0.887	0.877
Border Sample	Y	N	N	N
Distance to Nearest State Restriction	None	None	50 miles	100 miles

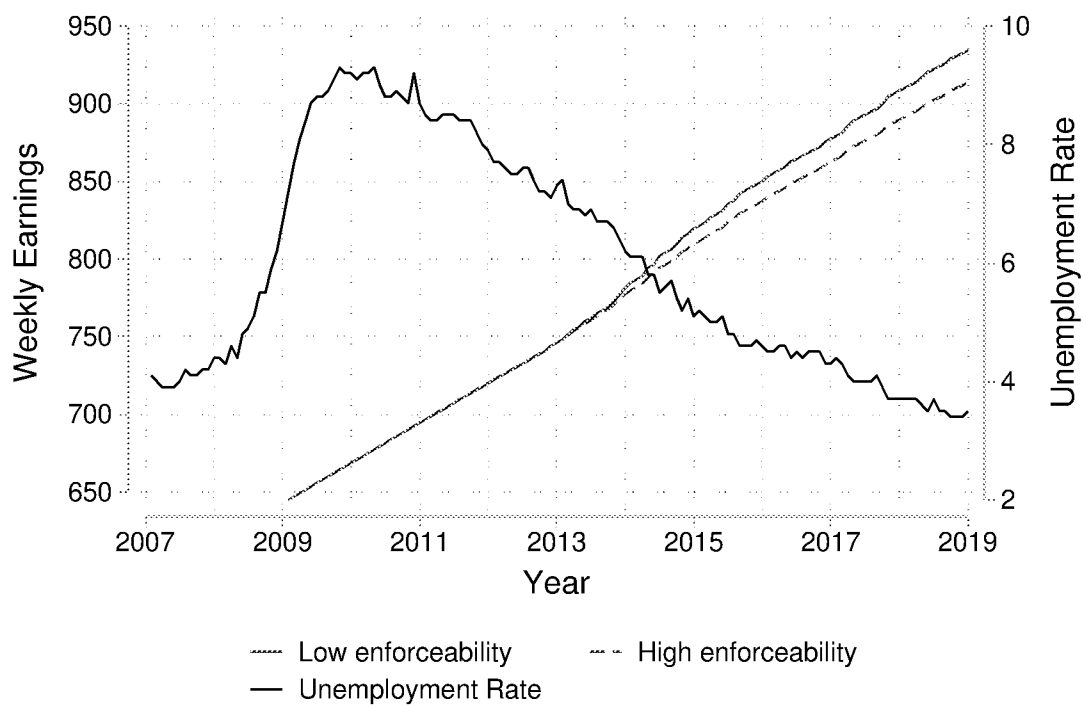
The dependent variable is log earnings. The sample is the QWI from 1991-2014 and includes individuals between ages 19-64. Column 1 uses the sample from Table 6, while Columns 2, 3, and 4 use counties that are neither on state borders nor members of border-straddling commuting zones. Columns 3 and 4 further restrict by the distance from the focal county's centroid to the nearest county centroid in a different state. All regressions include controls for male, age group, as well as division by year by quarter and county fixed effects. Standard errors are clustered by own state. *** $P < .01$, ** $P < .05$, * $P < .1$

Table B.8: The External Effects of NCA Enforceability on Mobility: Hires and Separations

	Hires			Separations		
	(1)	(2)	(3)	(4)	(5)	(6)
Own State NCA Score	-0.286** (0.111)	-0.305** (0.122)	-0.223 (0.144)	-0.269** (0.132)	-0.291** (0.140)	-0.193 (0.167)
Donor State NCA Score		-0.115 (0.165)	-0.182 (0.187)		-0.141 (0.167)	-0.200 (0.190)
Own Cty Emp/CZ Emp × Own State NCA Score			-0.476 (0.557)			-0.568 (0.590)
Own Cty Emp/CZ Emp × Donor State NCA Score			0.413** (0.162)			0.407** (0.163)
Observations	604322	604322	603466	604512	604512	603659
R^2	0.951	0.951	0.951	0.950	0.950	0.950
Sample	Border	Border	Border	Border	Border	Border

The sample is the QWI from 1991-2014 and includes individuals between ages 19-64. All regressions include controls for male, age group, as well as division by year by quarter and county fixed effects. Standard errors are clustered by own state in columns (1) and (4), and two-way clustered by own state and commuting zone in columns (2), (3), (5), and (6). ***P<.01, **P<.05, *P<.1

Figure B.1: Divergence in Weekly Earnings in States with High versus Low NCA Enforceability



The figure depicts simulated earnings for workers who began their jobs in 2009 and held them through 2019. As the labor market improves, workers in low enforceability states receive wage increases according to the tightness of the labor market, while workers in high enforceability states are unable to do so due to the increased costs of mobility imposed by NCAs.

Table B.9: Heterogeneous Effects of NCA Enforceability on Earnings by Race and Sex

	(1)	(2)	(3)	(4)
NCA Score	-0.136*** (0.046)			
Female & White=1	-0.469*** (0.011)	-0.417*** (0.027)	-0.423*** (0.026)	-0.416*** (0.027)
Female & Black=1	-0.572*** (0.011)	-0.519*** (0.026)	-0.526*** (0.024)	-0.513*** (0.030)
Male & Black=1	-0.339*** (0.008)	-0.280*** (0.016)	-0.282*** (0.017)	-0.271*** (0.015)
Female & Not Black or White=1	-0.502*** (0.019)	-0.424*** (0.016)	-0.438*** (0.015)	-0.436*** (0.016)
Male & Not Black or White=1	-0.146*** (0.010)	-0.132*** (0.017)	-0.144*** (0.016)	-0.142*** (0.015)
White Male × NCA Score		-0.091* (0.047)	-0.031 (0.053)	-0.069 (0.047)
Female & White=1 × NCA Score		-0.166*** (0.057)	-0.097* (0.054)	-0.138** (0.054)
Female & Black=1 × NCA Score		-0.166*** (0.053)	-0.097* (0.050)	-0.154*** (0.052)
Male & Black=1 × NCA Score		-0.175*** (0.048)	-0.112* (0.056)	-0.132*** (0.048)
Female & Not Black or White=1 × NCA Score		-0.221*** (0.045)	-0.141*** (0.045)	-0.199*** (0.043)
Male & Not Black or White=1 × NCA Score		-0.106** (0.048)	-0.029 (0.046)	-0.083* (0.044)
College Educated Worker=1 × NCA Score			-0.114*** (0.026)	
High NCA Use Occ=1 × NCA Score				-0.038*** (0.012)
Observations	1537454	1537454	1537454	1537454
R^2	0.275	0.275	0.276	0.289

The dependent variable is log weekly earnings. The sample in all columns is the CPS ASEC from 1991-2014 and includes individuals between ages 18-64 who reported working for wage and salary income at a private employer the prior year. All regressions include fixed effects for state, fixed effects for Census division by year, fixed effects for broad occupational class, and individual controls for male, white, Hispanic, age, age squared, whether the individual completed college, and indicators for the metropolitan city center status of where the individual lives. In Column (4), High NCA Use Occupations are occupations with NCA use greater than the national average, as tabulated by Starr et al. (2018). A separate indicator for High NCA Use Occupation is included in those regressions. SEs clustered by state in parentheses. ***P<.01, **P<.05, *P<.1

University of Chicago Law School

Chicago Unbound

Coase-Sandor Working Paper Series in Law and
Economics


Working Papers

2020

The Economic Basis of the Independent Contractor/Employee Distinction

Eric Posner
dangelolawlib+ericposner1@gmail.com

Follow this and additional works at: https://chicagounbound.uchicago.edu/law_and_economics_wp

 Part of the Law Commons

Chicago Unbound includes both works in progress and final versions of articles. Please be aware that a more recent version of this article may be available on Chicago Unbound, SSRN or elsewhere.

Recommended Citation

Eric Posner, "The Economic Basis of the Independent Contractor/Employee Distinction", Coase-Sandor Working Paper Series in Law and Economics, No. 909 (2020).

This Working Paper is brought to you for free and open access by the Working Papers at Chicago Unbound. It has been accepted for inclusion in Coase-Sandor Working Paper Series in Law and Economics by an authorized administrator of Chicago Unbound. For more information, please contact unbound@law.uchicago.edu.

The Economic Basis of the Independent Contractor / Employee Distinction

Eric A. Posner¹

June 6, 2020

Abstract. In recent years, a controversy has erupted over the distinction between employees and independent contractors. Commentators have argued that in the modern “gig economy,” many people traditionally classified as independent contractors are as vulnerable as employees and should be granted the legal protections that employees alone normally enjoy. However, the distinction between the two categories remains inescapable, and the theoretical basis for it has not been identified. I argue that the distinction is derived from market structure. Employees are workers who, because they must make relationship-specific investments in a single firm, are subject to labor monopsony. Independent contractors do not make such relationship-specific investments, and hence normally operate in a competitive labor market. Employment and labor law may be explained as a method for protecting workers from labor monopsony; because independent contracts are not subject to labor monopsony, they do not require such protection.

Introduction

The law’s distinction between employees and independent contractors (or, merely “contractors”) has sparked intense debate over the last few years. As a result of advances in technology, some workers who have traditionally been classified as employees are now being treated as contractors: they make a living by undertaking a series of “gigs” for different “labor buyers” (as I will call the firms or households that purchase the services of workers) rather than working for a single employer.² Many commentators worry that these gig contractors are being exploited because they are not entitled to the protections of employment law,³ a related view is

¹ Kirkland & Ellis Distinguished Service Professor, The University of Chicago Law School. Thanks to Daniel Hemel, Max Huffman, William Hubbard, Aneil Kovvali, Genevieve Lakier, Mark Lemley, Jonathan Masur, Sarath Sanga, and participants at a workshop at the University of Chicago Law School, for helpful comments, and to Michael Christ and Justin Taleisnik for research assistance.

² Some commentators claimed that these alternative work arrangements (outside the employment relationship) had transformed labor markets throughout the United States, a claim that received a boost from a 2016 study, which found that alternative work arrangements increased from 10.7 percent in 2005 to 15.8 percent in 2015. See Lawrence F. Katz and Alan B. Krueger, *The Rise and Nature of Alternative Work Arrangements in the United States, 1995-2015*, NBER Working Paper No. 22667, September 2016. However, the authors later revised their estimate down to a 1 percent increase. See Lawrence F. Katz and Alan B. Krueger, *Understanding Trends in Alternative Work Arrangements in the United States* (2019), <https://www.nber.org/papers/w25425>.

³ See, e.g., Keith Cunningham-Parmeter, *From Amazon to Uber: Defining Employment in the Modern Economy*, 96 B.U. L. REV. 1673, 1677 (2016) (“[C]urrent judicial pronouncements . . . embrace a cabined vision of employment that shields firms from liability.”); Brishen Rogers, *Employment Rights in the Platform Economy: Getting Back to Basics*, 10 HARV. L. & POL’Y REV. 479, 500, 505–07 (2016) (arguing that underinclusive employment classification tests leave workers at risk of excessive employer domination in violation of fundamental democratic and egalitarian values, suggesting that companies like Uber should be required to treat drivers as employees); Jennifer Pinsof, Note,

that workers who are really employees are being deliberately misclassified as contractors by rapacious employers for the same reason.⁴ Commentators have proposed numerous reforms designed to bring contractors, or a subset of them, under the protection of employment law.⁵

While new technology and employment trends have highlighted these problems, they are not new. The distinction between employees and contractors is deeply entrenched in the law, and

A New Take on an Old Problem: Employee Misclassification in the Modern Gig-Economy, 22 MICH. TELECOMM. & TECH. L. REV. 341, 344 (2016) (suggesting that Uber drivers and similar workers should be classified as employees, which would provide necessary worker protections and benefits); Seth D. Harris & Alan Krueger, *A Proposal for Modernizing Labor Laws for Twenty-First-Century Work: The “Independent Worker”*, BROOKINGS INST. (2015) (proposing protections for non-employees).

⁴ See John A. Pearce II & Jonathan P. Silva, *The Future of Independent Contractors and Their Status as Non-Employees: Moving on from a Common Law Standard*, 14 HASTINGS BUS. L.J. 1, 3 (2018) (“The claim that a business’s economic motivations led to classifying workers as independent contractors has been the basis for numerous legal challenges by gig-economy workers claiming that they were improperly classified as independent contractors and that the nature of their work makes them employees; Pinsof, *supra* note 3, at 349–54 (describing “rampant” misclassification, particularly in the gig economy); Alexia Fernández Campbell, *Companies Often Mislabeled Employees as “Freelancers” to Cut Costs. Workers are Fighting Back.*, VOX (Mar. 20, 2019), <https://www.vox.com/policy-and-politics/2019/3/20/18272918/conde-nast-epicurious-employee-freelancer-contractor> (suggesting that financial incentives to misclassify are powerful and pervasive, potentially affecting millions of workers); Harris & Krueger, *supra* at 7 (Brookings Institute Discussion Paper 2015-10, 2015) (noting that legal uncertainty has created both “intentional and unintentional” misclassification). Allegations of misclassification have led to a flurry of class action lawsuits. See Liya Palagashvili, *Disrupting the Employee and Contractor Laws*, 2017 U. CHI. LEGAL F. 379, 382–83, 405–08 (2017) (describing cases).

⁵ The literature on the question of worker classification has grown massive. For research from the last few years, see, e.g., Rogers, *Employment Rights in the Platform Economy: Getting Back to Basics*, *supra* note 3, at 82–83 (suggesting that courts should not shy away from reliance on normative values to determine employment status, imposing duties on employers when doing so advances the underlying policy goals of employment law such as preventing employee domination); Naomi B. Sunshine, *Employees as Price-Takers*, 22 LEWIS & CLARK L. REV. 105, 110 (2018) (suggesting a rebuttable presumption that workers lacking power to set their rates should be classified as employees); Cunningham-Parmeter, *supra* note 3, at 1677–78 (noting that judicial focus on daily, direct control over working conditions is underinclusive, and instead arguing that courts should use a broader definition of control that is less formalist and considers power imbalances); Brishen Rogers, Am. Constitution Soc’y for Law & Policy, *Redefining Employment for the Modern Economy* 7 (2016), https://www.acslaw.org/wp-content/uploads/2018/04/Redefining_Employment_for_the_Modern_Economy.pdf (proposing expanding the definition of employee to include all workers who are “economically dependent” on the employer); Richard R. Carlson, *Employment by Design: Employees Independent Contractors and the Theory of the Firm*, 71 ARK. L. REV. 127, 130 (2018) (proposing a test based on Ronald Coase’s “make or buy” theory of the firm that would incorporate analysis of the firm’s motives for hiring rather than buying labor); Matthew T. Bodie, *Participation as a Theory of Employment*, 89 NOTRE DAME L. REV. 661, 665–66 (2013) (proposing a definition of employee based on the motive of the employer); Benjamin Means & Joseph A. Steiner, *Navigating the Uber Economy*, 49 U.C.D. L. REV. 1511, 1054 (2016) (suggesting a flexible, case-by-case approach that avoids “sweeping all workers . . . into one category or the other” due to the varied circumstances of workers even within a single company). Another group of commentators have suggested the creation of a third, middle category between employee and contractor, to be granted some but not all of the protections afforded to employees. See, e.g., Harris & Krueger, *supra* note 4, at 5 (suggesting creation of an “independent worker” category); Pearce & Silva, *supra* note 4, at 2 (suggesting creation of a “dependent contractor” to capture workers that are functionally independent contractors but are nevertheless economically vulnerable); Michael L. Nadler, *Independent Employees: A New Category of Workers for the Gig Economy*, 19 N.C. J.L. & TECH. 443, 480 (2018) (suggesting an “independent employee” intermediate category of worker); Miriam A. Cherry & Antonio Aloisi, *Dependent Contractors in the Gig Economy: A Comparative Approach*, 66 AM. U. L. REV. 635, 637 (2017) (observing that an intermediate, “hybrid” category is found in different legal systems around the world).

reflects a basic intuition about the organization of labor markets. Consider, for example, a person trained as an electrician. She might choose to set up her own business. She advertises her services, and spends her days working for various homeowners who pay her to repair the fuse box or install new lighting. The electrician seems like an a business owner, not an employer, and indeed she would be legally classified as an independent contractor. Or she might go to work for, say, a company that manufactures electric turbines. She shows up at a worksite every day at 9 am, leaves at 5 pm, and draws a salary from a single firm. Here, she would be classified as an employee. As an employee, the electrician would be protected by numerous federal and state laws that control wages, working conditions, and benefits. If instead she works as a contractor, she would enjoy none of these legal protections.

Because the same person doing the same type of work might be self-employed or an employee of someone else, the distinction between employee and contractor can be elusive. The distinction is made more complicated still by the administrative requirements that have grown up around it. Because contractors often charge for a job rather than by the hour, it may be difficult to calculate an hourly wage, and thus to apply the minimum wage laws to them. Because contractors often work alone, it might seem that the right to organize a union would do them no good. Because contractors often choose their own tools and control working conditions, it would make little sense to compel those who buy their labor to comply with legal requirements for workplace safety. Contractors often work for homeowners and other consumers who lack the legal sophistication and administrative resources for complying with the huge number of legal restrictions that apply to employers, including the obligation to withhold taxes. And, until recently, it was common to think of contractors as highly trained professionals—electricians, plumbers, lawyers, doctors—who were not as vulnerable to mistreatment by buyers of their labor than ordinary employees were. Contractors did not seem to need employment law protections.

But technology has put pressure on these intuitions. We now see that companies can organize their businesses so that drivers, janitors, and home healthcare workers are classified as contractors rather than employees. Compensation can be structured so that it is hourly or based on the accomplishment of tasks; control over working conditions can be assigned to the worker, retained by an organization, or divided between them; and organizations can match workers with consumers so that consumers, rather than the organization, seem like the employers. Organizations can knit workers together into loose teams, keep them apart from each other, or use them as a conventional workforce; they can assign the price-setting power to workers or keep it for themselves. Or both: Uber normally sets wages for drivers but recently has allowed drivers in some cities to set their own prices up to five times a base price.⁶ With the scrambling of categories, the intuitions have lost their force, and we need to look deeper for the policy reasons behind the distinction between employee and contractor.

I argue that contractor or employee status, properly understood, depends on market structure—whether workers operate in a competitive labor market or not. A real-world labor market falls somewhere along the spectrum from perfect competition to monopsony. When numerous (or, technically, an infinite number of) buyers compete for the labor of a worker, the market is perfectly competitive. When only a single buyer of that labor exists, the market is a

⁶ Preetika Rana, *Uber Tests Feature Allowing Some California Drivers to Set Fares*, WALL ST. J. (Jan. 21, 2020), <https://www.wsj.com/articles/uber-is-testing-a-feature-that-lets-some-california-drivers-set-fares-11579600801>.

monopsony. When the market is perfectly or relatively competitive, the existence of multiple alternative buyers ready to buy work from the worker offers the worker adequate protections from abuse. When the market leans toward monopsony, those alternative buyers do not exist in sufficient quantity to protect the worker. Legal intervention to protect the worker from abuse may be beneficial not only for the worker, but for overall economic health.⁷

Why do some workers work in competitive labor markets and others do not? The assignment of workers to different types of market is likely derived from the nature of the interaction between the worker, assets under the worker's control, and the labor buyer who benefits from the worker's manipulation of those assets. As a rough approximation, I distinguish between what I will call "discrete" and "relational" work. Work is discrete when its value is best exploited by a worker acting alone (in the sense of not being subject to the control of another), in most cases bringing to bear her labor on an asset or various assets that she owns. Work is relational when its value increases with the degree of coordination between the worker in question and other workers, including peers, subordinates, and superiors, all in relationship with assets that the worker uses but does not own.

Because relational workers invest in their firm, they earn higher wages than they could at alternative firms. This is what is meant by labor monopsony. Because such workers cannot earn comparable wages from competing employers, those workers lack a credible threat to quit if their wages stagnate and conditions worsen at the margin. Employment and labor law step in to offer these workers protection: they are effectively a form of price (wage) and conditions regulation that counters the downward pressure of monopsony. By contrast, discrete workers are not subject to labor monopsony. If one buyer of their work refuses to pay them their market price, those workers can exit and find another buyer who will. Employment and labor law will either not help these workers or will help them only by enabling them to extract rents from others—as I will explain in due course. Thus, classification law performs the important task of segregating workers into the class that should receive the protection of employment and labor law, and those who should not. The first group are classified as employees and the second group are classified as contractors.

The legal tests for distinguishing employees and contractors are famously ambiguous. They involve many factors that often point in different directions and that firms can manipulate. The main problem is not the factors themselves but the failure of courts and regulators to understand how those factors relate to the goal of labor market regulation. The discrete/relational distinction provides guidance for identifying workers as contractors (discrete) or employees (relational).

Consider, for example, the controversy over Uber's treatment of its drivers. Uber claims that the drivers are contractors and that Uber merely matches them with customers, like a dating service.⁸ Many drivers argue that they are employees, and thus should be protected by minimum

⁷ From within the economic framework taken by this paper. Many employment laws can be defended for advancing public values and changing attitudes in positive ways; I take no position on these claims. However, these public values do not justify the employee/contractor distinction in the law. If the law seeks to eliminate invidious discrimination, it should not apply different standards to people who buy labor from employees and people who buy labor from contractors.

⁸ Aarian Marshall, *Why Uber Still Thinks It Can Still Call Its Drivers Contractors*, WIRED (Sept. 12, 2019), <https://www.wired.com/story/why-uber-still-call-drivers-contractors/>.

wage and related law, and should be allowed to organize a union.⁹ Litigation and commentary have focused on whether drivers “control” their worksite, and gallons of ink have been spilled on an essentially unanswerable question. I argue that the answer depends on whether Uber enjoys a monopsony over drivers because they engage in relational work for it. I will discuss this issue in Part III.B.

An enormous academic literature has developed in the wake of the controversies over Uber and other gig-economy companies. Most commentators argue that these companies have sought to avoid employment law protections, and advocate an expanded definition of employment so as to sweep in gig-economy workers, or a third category such as the “dependent contractor” who would be entitled to certain protections.¹⁰ However, no one has provided a satisfactory theoretical argument for the distinction. By rooting employment law protections in the problem of monopsony, I can offer a test that is (by the standards of law) clear and stable.¹¹

Thus my thesis. In the absence of law other than basic contract and property law, (1) some workers will end up in competitive labor markets and others will end up in monopsonized labor markets, based on the type of work they do (discrete or relational). (2) The workers in competitive labor markets should not be protected by employment and labor law; the workers in monopsonized labor markets should be. (3) Firms that misclassify employees as contractors cause social harm by evading restrictions on labor monopsony power. Their behavior should be seen as arbitrage, a way to evade laws that have been developed to reduce the harms from monopsonistic labor market competition.¹² (4) The classification test should be based either on whether workers are relational or not, or, more broadly, the extent to which they are subject to labor monopsony. I will discuss this choice in Part II. To be clear, my contribution is not the insight that employment and labor law may counter labor monopsony. This idea has a long history. The purpose of this Article is to bring to bear the literature on labor monopsony on the misclassification debate, and in particular, to criticize and reform the classification test.

My argument can be contrasted with the dominant view about the misclassification controversy in the law review literature. That view sees employment and labor law mainly as devices for protecting low-income workers, and argues that these protections should be extended to contractors who are similarly vulnerable. Some commentators thus argue that the labor exemption in antitrust law—which allows employees to organize—should be extended to

⁹ David Ingram & Diana Dasrath, *More Drivers Sue Uber, Saying They’re Employees, Not Contractors*, NBC NEWS (Dec. 18, 2019), <https://www.nbcnews.com/tech/tech-news/more-drivers-sue-uber-saying-they-re-employees-not-contractors-n1103936>.

¹⁰ See, e.g., Pearce & Silva, *supra* note 4; Nadler, *supra* note 5; Harris & Krueger, *supra* note 4. See also sources cited in notes 3–5, *supra*.

¹¹ Two earlier law review papers draw on the theory of the firm, as I do (see *infra*), but offer different tests and do not derive the distinction between contractor and employee from the problem of labor market monopsony. See Matthew T. Bodie, *Participating as a Theory of Employment*, 89 NOTRE DAME L. REV. 661, 665–66 (2013) (arguing that an employee is a worker who “participates” in a firm); Richard R. Carlson, *Employment by Design: Employees Independent Contractors and the Theory of the Firm*, 71 ARK. L. REV. 127, 130 (2018) (proposing a test that drawn on employer’s economic motives for hiring rather than buying labor).

¹² As I will discuss later, it is not clear that these laws, or all of them, are actually effective; but that is an empirical question that I put to one side.

contractors.¹³ I argue that employment and labor law are not oriented toward wealth distribution but toward efficiency—by countering labor monopsony. These laws thus should not apply to competitive labor markets, and the purpose of classification law is to prevent them from doing so.

I. Discrete and Relational Work

A. Employee and Contractor

At least as far back as Ronald Coase's paper on the "Nature of the Firm," published in 1937, economists have identified the distinguishing feature of employment as the employer's control over the worker.¹⁴ An employee is "in" the firm because the managers of the firm can issue orders to her and expect her obedience. A contractor is "outside" the firm because the managers of the firm can elicit cooperation only through a negotiated bargain. The idea of control has also been central to the common-law definition of employment (or master-servant) relationships for centuries.¹⁵ But the nature of control has turned out to be elusive. Criticizing Coase's reliance on control for the definition of the employment relationship, economists Armen Alchian and Harold Demsetz pointed out that a customer exerts control over a contractor as well.¹⁶ For example, if a grocer refuses to do what a customer wants—stock a certain product, for example—the customer may stop patronizing the business. The fear of losing customers puts the grocer under customers' control—just as the stock clerk's fear of being fired causes her to obey the grocery store owner's order to fill the shelves with one product rather than another.

A moment's reflection, however, reveals the problem with this argument. Suppose a customer tells a grocer to move the candy bars to a shelf where children will not see it, or to sweep the floor because it is too dirty, or to change the window display. Even at risk of losing the customer, the grocer is likely to tell him to get lost. By contrast, a grocer could certainly tell her stock clerks to do any of these things. While the stock clerks could refuse and quit, most likely they would obey the grocer's directions. The control that an employer exerts over an employee is different from the kind of control that a customer exerts over a contractor.

A line of literature has made progress with the notion of control, rooting it in the idea that contracts—including employment contracts—can never fully specify the optimal actions of the parties, and so unavoidably allocate discretion among them.¹⁷ As a rough approximation, contracts

¹³ See Marina Lao, *Workers in the "Gig" Economy: The Case for Extending the Antitrust Labor Exemption*, 51 U.C.D. L. REV. 1543, 1543, 1547 (2018) (arguing that the gig economy's hybrid work relationships and the underlying purpose of the labor exemption suggests that extension to non-employees would be appropriate).

¹⁴ Ronald H. Coase, *The Nature of the Firm*, 4 ECONOMICA 386, 403–04 (1937).

¹⁵ See Scott E. Masten, *Legal Basis for the Firm*, 4 J. LAW, ECON. & ORG. 181 (1988).

¹⁶ Armen A. Alchian & Harold Demsetz, *Production, Information Costs, and Economic Organization*, 62 AMER. ECON. REV. 777, 778 (1972).

¹⁷ The starting point for this literature is a series of articles by Sanford Grossman, Oliver Hart, and John Moore (GHM), who pointed out that because contracts are necessarily incomplete (in the sense of being unable to specify all the optimal actions for both parties), it is important for the parties to specify which party will have discretion within the scope of incompleteness. The parties can assign discretion by allocating property rights: the party with a relevant property right has discretion over the use of that property—"residual" discretion or control in the sense that the discretion is limited by any specified contractual terms. The parties should then assign that discretion (via the assignment of property rights) to whichever party is more likely to use discretion in the joint interest of both parties—and that is roughly the party whose incentive to invest in the joint project is more sensitive to the return on that

between labor buyers and workers that allocate key aspects of discretion to the labor buyers create employment relationships, while those that allocate those aspects of discretion to the workers create contract relationships. In the balance of this section and the next, I draw on this literature. Then I will turn to its implications for the misclassification test.

To understand what “control” means in the employment context, consider a simple example. A person we will call the Rider needs a car and driver in order to get around town. The Rider may choose between two arrangements. Under the first, the Rider owns a car and contracts with a person, the Driver, to chauffeur her for a period of time. Under the second, the Rider does not own the car; the Driver does. The Rider and the Driver enter a contract under which the Driver agrees to drive the Rider around for a period of time.

The two contracts look nearly identical, and it is easy to imagine that in practice the “output”—the routes used by the Driver, the amenities of the service, and so on—is identical under both contracts. If the Driver is capital-constrained, the Rider could lend him the money to buy the car, and now the two contracts seem even more similar. We can further imagine that the parties, in each case, enter into highly detailed contracts that specify numerous attributes of the relationship—when the Driver must show up every day, how far the Driver must drive, even the routes and any chitchat that will take place while the Driver and Rider share the car.

Yet there is an important difference. When the Rider owns the car, the Rider enjoys “residual control”—meaning control over how the car is used where the contract fails to specify the Driver’s obligations. And when the Driver owns the car, the Driver enjoys residual control. As a concrete example, imagine that the parties enter into a 100-page contract that specifies nearly every aspect of the work relationship but omits, say, whether the Driver can talk on the phone with friends while driving the Rider. If the Driver does talk on the phone, and the Rider objects, the assignment of residual control matters. If the Driver owns the car, and the Rider tells him to stop talking to his friends, he can simply refuse because he is not prohibited from doing so under the contract. If the Rider fires him, she breaches the contract. She could offer to pay him more to stop talking to his friends, and he may agree, but this additional bargain is costly to negotiate, and the Driver has the bargaining power. If the Rider does not have good alternative people to hire as the Driver (and is reluctant to approach a stranger who has not acquired what we will call relationship-specific knowledge), the Driver can “hold up” the Rider for a high additional price.

investment. The difference between an employee and a contractor is that the employee does not own the asset in which she applies her labor—that asset is owned by the employer—while a contractor does own the asset. The contractor has greater bargaining power vis-à-vis the firm; and the stronger incentive to maintain the asset but a weaker incentive to use the asset to benefit the firm. See generally Sanford J. Grossman & Oliver D. Hart, *The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration*, 94 J. POLIT. ECON. 691 (1986); Oliver Hart & John Moore, *Property Rights and the Nature of the Firm*, 98 J. POLIT. ECON. 1119 (1990); Philippe Aghion and Richard Holden provide a lucid discussion. See Philippe Aghion & Richard Holden, *Incomplete Contracts and the Theory of the Firm: What Have We Learned over the Past 25 Years?*, 25 J. ECON. PERSPECTIVES 181 (2011). Other authors further developed the special role of the employee. See, e.g., Bengt Holmstrom & John Roberts, *The Boundaries of the Firm Revisited*, 12 J. ECON. PERSPECTIVES 73 (1998); George P. Baker & Thomas N. Hubbard, *Make Versus Buy in Trucking: Asset Ownership, Job Design, and Information*, 93 AMER. ECON. REV. 551 (2003); Eric Van den Steen, *On the Origin of Shared Beliefs (and Corporate Culture)*, 41 RAND J. ECON. 617 (2010); Wouter Dessein, *Incomplete Contracts and Firm Boundaries*, 30 J. L., ECON., & Org. 113 (2014). For an early piece that anticipated some of these arguments, see Alchian & Demsetz, *supra* note 16, at 778, who argue that firms need control over employees because of the difficulty of controlling team production by contract.

If the Rider owns the car, and tells the driver to stop talking to his friends, the Driver is required to obey. He does not have residual control over the car, and cannot use it in ways prohibited by the owner. The Rider can order the Driver to stop the car and leave. Of course, the Driver could quit. But now the Rider has the bargaining power. The Driver might prefer to work, and might be willing to accept a reduction in wages in return for the right to speak, but now he has to pay the Rider for this right. If the Driver has made a relationship-specific investment—preferring to drive *this* Rider rather than other people because he has learned her needs and doesn't want to have to learn the needs and idiosyncrasies of a new boss—he will need to make concessions.

The parties will assign the property interest (that is, car ownership) to the party whose use of residual control is more likely to maximize the surplus of their interaction. Consider first the Driver. If the Driver owns the car, then he will be given strong incentives to maintain the car. He will drive it carefully, take it to the garage frequently, and so on. He will anticipate that in the future the relationship may end, and he will want the car in good working order as he searches for new customers. We will call this behavior “worker care.”

The problem with maximizing worker care is that the worker may have correspondingly weak incentives to follow the Rider's orders. The Rider might order the Driver to take the fastest route, and the Driver might prefer not to because the fastest route is a potholed road that is hard on the car. The Driver might therefore choose another route. The Rider might be willing to tolerate a less-than-perfectly-maintained car in return for greater obedience. We call the Rider's concern “worker coordination”—by which we mean, the value of coordination between the worker's action and assets owned by, or other workers managed by, the labor buyer, here the Rider. To ensure worker coordination, the Rider requires “managerial direction” (or “managerial discretion,” to emphasize the freedom to direct the Driver); specification of optimal Driver behavior by contract is impossible.

Thus, when the parties negotiate the contract, they will trade off these two values. Car maintenance may not be a serious issue because maintenance can be specified by contract, or because the car is well constructed and poor maintenance will not significantly reduce its value. In that case, the parties will assign the car to the Rider. Otherwise, the parties will assign the car to the Driver. In this economic sense, the Driver is an employee in the first case, and a contractor in the second.

A useful real-world example comes from a paper by economists George Baker and Thomas Hubbard, which examines the shift from the owner-operator model of trucking to the company driver model.¹⁸ Owner-operators are independent contractors; they own their trucks and contract with shippers. Company drivers are employees; they work for a single, usually very large, shipper. For truckers, “worker care” is the same as in our example of the Driver-Rider relationship: “Wear and tear on the truck is minimized when drivers drive at a steady and moderate speed, but drivers may prefer to drive fast and then take long breaks because it allows them to rest longer, visit

¹⁸ George P. Baker & Thomas N. Hubbard, *Contractibility and Asset Ownership: On-Board Computers and Governance in U.S. Trucking*, 1443 QUARTERLY J. ECON. 1443, 1443 (2004).

friends, etc., and still arrive on time.”¹⁹ Worker coordination is also important because “hauls vary in their desirability to drivers in ways that are not captured in agreements with carriers. Those that take drivers into congested or dangerous areas are less desirable,” as do those that take drivers to remote areas where there are no desirable “backhauls,” that is, return trips with a new load of cargo.²⁰ The problem faced by the shipper and the trucker is how to optimize care (minimizing wear and tear) while ensuring that the best routes are taken.

When drivers operate as contractors, they have strong worker care incentives: they drive their truck carefully because they own it. But they are reluctant to take less desirable routes, and thus dispatchers—who coordinate routes—must negotiate with them case by case, which is disruptive. When drivers operate as company drivers, they have weak worker care incentives: the company bears the maintenance cost. But the company can simply order drivers to take the route that is optimal for the company. Here, we see the basic tradeoff in action. Baker and Hubbard hypothesize that longer hauls create greater problems for worker effort because there are more opportunities for the driver to engage in the suboptimal speed-and-rest strategy, while the backhaul negotiation problem seems the same for long haul and short haul drivers. The hypothesis finds support in evidence showing that long-haul drivers are more likely to be owner-operators than short-haul drivers are. The authors also examine the effect of the introduction of on-board computers, which monitor driver performance, and thus reduce the worker’s ability to shirk on worker care. The evidence indicates, as one would expect, that the introduction of this monitoring technology substantially reduced the share of owner-operated (contractor) trucks.²¹

To sum up, a labor buyer and a worker maximize their joint surplus by assigning the relevant property right to the worker when care is more important than other actions that require management direction and coordination, and otherwise to the labor buyer. In the first case, the worker is called a contractor; in the second case, the worker is called an employee. These factors in turn depend on how completely the parties can specify by contract the relevant actions of the worker. Where the worker can be given optimal or very good incentives to engage in care through contractual specification or technology, the labor buyer will employ the worker. Where the worker can be given optimal or very good incentives to engage in coordination with the assets and other worker through contractual specification or technology, the labor buyer will treat the worker as a contractor.

B. Relational Work and Discrete Work

A major difference between an employee (who interacts with assets owned by the firm as well as the firm’s other workers) and a contractor (who interacts with assets owned by the contractor) is that the employee makes a relationship-specific investment in the “firm,” which is to say in the assets owned by the firm and the other employees of the firm. The contractor does not make a relationship-specific investment, or much more limited investment. This difference is related to two concepts we discussed earlier: contract specificity and worker coordination.

¹⁹ *Id.* at 1446.

²⁰ *Id.* at 1447.

²¹ *Id.* at 1476.

When a contract can be specified in a relatively substantial sense, the underlying reason is that the task in question is standardized—routine, predictable, the same in different contexts. For example, when a homeowner hires a locksmith to fix a lock, the task is fully specified by the customer (though usually on the basis of a custom that the contract references rather than through detailed contractual specification). The locksmith performs the same task from place to place and does not need to learn anything about the individual interests or needs of different homeowners. Many tasks cannot be specified, however. An elderly homeowner who hires a personal companion to keep her company, maintain her records, deal with outsiders, and so on, will expect the companion to invest a great deal of time and effort in learning the homeowner’s needs, interests, and ways of doing things. As a result, the companion’s work is “relational,” in the sense that abilities that the companion obtains over time as a result of frequent interaction with the homeowner are not easily transferable to other contexts—so if the companion quits and goes to work for someone else, she will start again from the beginning with a new employer.

Thus, I distinguish between “discrete work” like the locksmith’s and “relational work” like the companion’s. The homeowner’s concern is “worker coordination,” the need to ensure that the companion will do just what the homeowner requires. The homeowner does not need, or needs less, worker coordination with respect to the locksmith, who relies on his own, transferable training and experience when he fixes the lock. The locksmith is a contractor; the companion is an employee.

The nature of work—relational or discrete—does not necessarily depend on the profession. A locksmith might be a contractor or employee in different contexts. Let us consider a more complex example involving two lawyers. Andrew is a solo practitioner who drafts wills for clients. He operates out of an office that he rents from a commercial landlord. Most of Andrew’s clients see him only once or a small number of times over their lives. He has drafted wills for thousands of people. It takes him only an hour or so to draft a will in most cases.

Beth is a tax attorney who works full time in the family office of a noted tycoon. She works 9 to 5 in a suite of offices with several other workers—accountants, investment advisors, and other lawyers. She works closely with this team and consults frequently with the tycoon and various family members. She does not work outside the office; and, like most other workers in this office, she has worked there for many years and is vague about when or even whether she might leave for a different job.

While Andrew and Beth may have attended the same law school and received the same legal training, their skills are by now quite different. Andrew’s body of knowledge—a deep understanding of testamentary law and related areas of law—allows him to offer identical services to numerous different people. His skill (or human capital) is general.²² Moreover, while his clients all have different financial resources and needs, along the relative dimension for which they hire Andrew—the disposition of their assets at death—the services that Andrew offers are very similar, one might even say commodified. We can thus think of Andrew’s work as *discrete* in the following sense: the value of the service he offers is independent of his relationship with the client—whether he has known the client for a long time, for example. Andrew’s work is discrete in another sense: its value to the client is independent of Andrew’s relationship with other workers. He works alone.

²² See Gary S. Becker, HUMAN CAPITAL (3d. ed. 1993).

The nature of Andrew’s skills—his knowledge of testamentary law, plus his ability to listen to and understand his clients—enables him to offer discrete services to a range of clients.

Beth, like Andrew, possesses a body of knowledge that she brings to bear as she provides services for her client, the tycoon. When she first goes to work for the family office, she brings the same type of discrete skill to the operations of the office as someone like Andrew would. But after working for a few years for the client, the nature of the skill has changed significantly. The hiring launches a relationship, and it is in the context of this relationship that most of Beth’s value is generated. Over time, she develops an increasingly deep understanding of the client’s resources, needs, and idiosyncrasies. She also develops an understanding of the other workers—how they work together, what they need from her, the appropriate way to behave in the office, and so on. Unlike Andrew, Beth engages in primarily relational work in the following sense: the value of the service she offers is dependent on her relationship with the client and with her coworkers (and, specific assets owned by the employer, for example, office equipment or, more plausibly, bespoke software and records). As a rough approximation, the longer and deeper the relationship, the greater the value that Beth confers on the family office. Beth’s major skills are thus relationship-specific. Aside from her knowledge of the law, the skills she develops at the family office—including her skills in working with others—generate value.

We can now return to the concepts of contract specificity and managerial direction. Because Andrew’s work produces a relatively standard product, the customers do not need control, that is managerial direction, over how Andrew does his work. The customer merely supplies some information in advance that Andrew uses to conform the product at the margin to the customer’s needs. In contrast, because Beth’s product is constantly changing in response to changing conditions, and the idiosyncratic needs of the firm she works for, someone—the manager, or the person who coordinates the joint production of all the workers in the family office—is given managerial direction over Beth’s work.

The distinction between discrete and relational work is an old one. In the middle ages, craft workers typically proceeded through two stages—a relational and a discrete stage.²³ The relational stage was the apprenticeship, during which the worker worked for a single master, and became steadily more value to the master as he gained experience both at the craft and at serving the master’s interests. Once the apprenticeship was over, the worker could sell his services, now discrete work, to the market, by forging swords, cobbling shoes, and building walls—directly for customers. During the industrial revolution, some workers worked in teams at factories, while others did piecework at home. The factory workers did relational work—and would be regarded as employees today. The pieceworkers did discrete work, often at home, and resembled contractors. There have been cases in which a single worksite housed workers who worked in teams and workers who worked independently on components; the latter group of workers thought of themselves as craftsmen with higher status than the first group of workers, who resembled today’s employees.²⁴ In the pre-internet age of the twentieth century, relational work continued in

²³ See Carlson, *supra* note 5 at 146-47.

²⁴ For a detailed illustration based on the Winchester Company, see John Buttrick, *The Inside Contract System*, 12 J. ECON. HIST. 205 (1952), and the very helpful discussion in Carlson, *supra* note 5 at 148-59. Buttrick observes that the “employees,” more than the contractors, were subject to monopsony (because the employees ended up working

factories and other workplaces, while discrete work was often conducted by skilled professionals—plumbers, doctors, artists, writers, lawyers. Thus, the notion that some work lends itself to relationships, and other work does not, is not a new one; it is a thread that runs through the history of labor relations.

C. Market Structure

The distinction between relational and discrete work matter for policy because it roughly maps onto a distinction in market structure. Because the seller of discrete work maximizes the value of the work without entering into relationships with buyers, she can sell that work to many different buyers. The “cost of exit”—the cost to the worker if any specific labor buyer stops buying from her—is low because the worker can simply find another labor buyer. (Think of the locksmith who goes from household to household and will not be significantly harmed if one household stops using the locksmith’s services.) The market for discrete work is thus (relatively) competitive. By contrast, the seller of relational work maximizes the value of that work by working for a single labor buyer and remaining with that buyer for an extended period of time. She must make a relationship-specific investment in assets owned by the labor buyer and in other people who work for the labor buyer. This means that the highest-value buyer of the worker’s labor is the person or firm for whom the worker already works, and thus that the worker’s cost of exit is high. The market is (relatively) monopsonistic.

In making this argument, I use the term “monopsony” in a specialized way familiar to labor economists but mostly unfamiliar to antitrust lawyers and industrial organization economists. In the antitrust world, monopsony means a single buyer (or, sometimes, a small number of buyers, where technically oligopsony prevails); in the context of labor, this means a single employer. Thus, a city like New York with dozens of law firms has a competitive labor market for lawyers because the firms compete among each other to hire the best lawyers, while a small town with, say, a single law firm and no other employers of lawyers has a monopsonized market. For labor economists, however, monopsony simply means that the employer can (and thus presumably does) pay the worker below that worker’s marginal revenue product—roughly, the value that the worker generates for the employer—because the worker’s outside options are limited—exit cost is high. The worker’s outside options might be limited because few other employers exist (a traditional monopsony). But the worker’s outside options might be limited for other reasons—for example, it takes a long time to find similar employers who are hiring (called “search costs”) or nominally similar employers are actually quite different in ways that matter for employees, including the location of the worksite (which affects the commute), the workplace culture, and so on (called “job differentiation”).²⁵ All of this derives from the relational nature of work.

For our purposes, the distinction is unimportant. By definition, a relational worker who works for an employer is subject to monopsony pressure. Because her services are worth more to the current employer than (in normal cases) to another employer, the current employer can pay her more than any outside employer. But by the same token the worker’s outside options are limited. If she earns \$100,000 at the current employer, and outside employers will offer only \$50,000, then

exclusively for Winchester while the contractors sold their services to other companies as well), as does Carlson—my point here. But neither of them recognizes its centrality to the distinction between employer and contractor.

²⁵ Alan Manning, *MONOPSONY IN MOTION: IMPERFECT COMPETITION IN LABOR MARKETS* (2003).

she cannot make a credible threat to leave if the current employer stops giving her raises. The current employer has strong incentives to limit wage increases (or even to reduce wages) or to worsen working conditions (for example, by demanding more work after hours). As labor economists have observed, virtually all employers thus have monopsony power. The usual explanation they give is that search costs and job differentiation gives rise frictions, but the simpler way of putting this is that work for employers is relational. The buyer of discrete work, by contrast, has no such power.

A possible objection to this argument is that people who apply to become relational workers can protect themselves by demanding an employment contract that protects them from such ex post exploitation. When Beth applies for the position with the family office, she is effectively a discrete worker (with a commodified body of knowledge about tax law) who hopes to obtain relationship-specific skills and the higher compensation that comes with them. If she is rational and far-sighted, however, she knows that as she develops these skills, the family office will be able to use its market power over her to pay her less than her marginal revenue product. To protect herself, she could, in principle, demand that her compensation rise indefinitely as her marginal revenue product rises, and insist on a contract to that effect. If the family office breaches, she would be able to sue for damages.

This counterargument brings us back to the issue of contract specificity. The property rights literature on which I draw takes for granted the impossibility of such contracting; that is why residual control is so important. Because the parties cannot specify value-maximizing behavior in advance, they assign residual control to whichever party has the best incentives to use it to generate a surplus.²⁶ Empirically, such bargaining is rarely observed, most likely because the employer does not trust a court to determine the marginal revenue product of a particular employee, and requires flexibility to adjust wages and working conditions in response to external shocks. It may also be that few employees have the sophistication to demand such contracts—perhaps aside from a handful of highly experienced, highly skilled employees who can hire agents and lawyers to protect them. And if few job applicants are sophisticated and wealthy enough to bargain for wage protection, then employers do better by looking for those who lack sophistication. Thus, the competition for workers that takes place when a soon-to-be-relational worker is hired (say, recent law school graduates) does not translate into market protection for those workers years or decades later in their careers.

As a general proposition, discrete workers are protected by competition for their work; relational workers are not. By “protected,” I mean merely that the workers are paid the efficient wage, their marginal revenue product—the wage necessary to maximize production, rather than a lower wage.

The argument is a generalization; exceptions exist. While relational work unavoidably generates monopsony, markets for discrete work are not always competitive. I will return to this issue in Part III.B, below. Moreover, monopsony power can result from other aspects of the employment relationship not captured by relationship-specific investment.²⁷

²⁶ See *supra*.

²⁷ Acemoglu, e.g., employer’s private information about workers.

D. Policy Implications

I have argued that work can be divided into discrete and relational types, and that discrete work tends to yield competitive labor markets while relational work tends to yield monopsonized labor markets. Why does this matter? The answer is that the appropriate legal protections are different for workers in monopsonized labor markets and workers in competitive labor markets. Where markets are monopsonized, they fail: they produce inefficiently low output. Legal regulation is called for, and it takes the form of what is conventionally called “employment law” and “labor law.” The label of the first body of law should come as no surprise given my claim that workers who faced monopsonized conditions are properly called “employees.” Where markets are competitive, they do not fail. Legal regulation is not called for—at least not required for the purpose of correcting a market failure.²⁸ This is why there is no separate body of “contractor” law that confers protections on contractors. The table below summarizes the argument to this point.

Legal Label	Labor Market	Type of Work	Property Right	Investment	Important Incentive	Exit Cost	Regulation
Contractor	Competitive	Discrete	Worker	General	Worker care	Low	None
Employee	Monopsonized	Relational	Labor Buyer	Relationship-specific	Worker coordination	High	Employment and Labor Law

This is why classification matters. When workers engage in discrete work in competitive labor markets, they should be classified as contractors because competition adequately protects them, while employment and labor law can do little good for them and possibly cause harm to them and society at large. When workers engage in relational work in uncompetitive labor markets, they should be classified as employees because both employment and labor law can help them, and in a socially beneficial way.²⁹

By “employment law,” I mean the range of federal and state protections for workers, including minimum wage laws (which will be my focus), maximum hours laws, laws that regulate workplace safety and other conditions of employment, antidiscrimination laws, and so on. By “labor law,” I mean the National Labor Relations Act and its amendments, which give workers the right to organize unions and strike, along with the labor exemption in antitrust law, which protects workers who engage in such organization from the antitrust laws.

Employment law and labor law reflect different approaches to labor relations, and so I will discuss them separately. Among economists, employment law protections have always been a bit of a puzzle. In a perfectly competitive labor market, employment laws cannot benefit workers and will likely harm them.³⁰ Consider the minimum wage. If the labor market is competitive and the

²⁸ If contractors require legal protection or other policy responses, the reasons lie elsewhere. For example, one might favor the extension of antidiscrimination norms in employment law to contractors because one believes that discrimination against contractors is morally reprehensible, or that such an extension will advance public values. These arguments lie outside of the narrow economic perspective that I take in this paper.

²⁹ As noted above, *supra* note __, the relationship between worker and labor buyer may produce other sources of monopsony power, for example, from the buyer’s private information about the worker’s ability. In such cases, the relational work test will be underinclusive. It is possible that it should be expanded to address this problem.

³⁰ For a discussion, see Suresh Naidu & Eric A. Posner, *Labor Monopsony and the Limits of the Law* (Jan. 31, 2019), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=3365374. In the legal literature, Richard Epstein has used this

market wage exceeds the minimum wage, the minimum wage law does not affect wages. If the market wage is below the minimum wage, then employers will fire workers who generate benefits less than their wage cost. While the wage of other workers may rise, in aggregate workers will be harmed.

This analysis applies to “employment mandates” as well, including family leave policies, for example, and tax advantages for retirement savings. The logic is the same. Suppose, for example, that a family leave policy costs an employer \$1,000 per worker. If workers value the family leave policy more than \$1,000, say, at \$2,000, then the employer will offer it without legal compulsion. The employer can thus save \$2,000 in wages by offering a benefit that costs it only \$1,000. If workers value the family leave policy at less than \$1,000, say, at \$500, then legal compulsion can only make the worker worse off—forcing her to accept a benefit of \$500 rather than wages of \$1,000.

Labor law is also a puzzle from this standpoint, although for a different reason. Labor law authorizes workers to organize a union even though the agreement among workers in a union to strike if they are not paid the same wage, or according to the same compensation schedule, is equivalent to a price- (or actually wage-) fixing cartel. If the labor market is competitive, then the employer would be required to pay market wages to the workers. If the workers cartelize the labor market through the union, they can insist on an above-market wage, which would reduce both overall employment in the labor market and output, which would probably raise prices for consumers as well. If labor markets are competitive, there would be no reason for the government to encourage or even tolerate unions.

However, as we have seen, labor markets are not competitive. Search costs and related costs ensure that employers enjoy labor monopsony; labor market concentration, which exists in most employment markets, further enhances employers’ market power.³¹ In conditions of labor monopsony, employment law and labor law have stronger justifications.

The best illustration of this claim is the debate about minimum wage laws, which have been subject to extensive empirical research. As noted above, the traditional view of economists assumed competitive labor markets and held that minimum wage laws cannot benefit workers and can only harm them.

Yet after years of controversy, the emerging consensus is that minimum wage laws do not have this predicted disemployment effect. They usually raise wage levels without reducing employment.³² While this result contradicts the assumption of competitive labor markets, it is

insight to launch a wholesale assault on labor and employment law. *See, e.g.*, Richard A. Epstein, FORBIDDEN GROUNDS: THE CASE AGAINST EMPLOYMENT DISCRIMINATION LAWS (1992); Richard A. Epstein, *Contractual Solutions for Employment Law Problems*, 38 HARV. J.L. & PUB. POL’Y 789 (2015); Richard A. Epstein, *Labor Unions: Saviors or Scourges?*, 41 CAP. U. L. REV. 1 (2013). His argument is, of course, based on the assumption that labor markets are competitive, or nearly so.

³¹ For a recent overview of the empirical literature on labor market concentration, *see* Suresh Naidu, Eric A. Posner & Glen Weyl, *Antitrust Remedies for Labor Market Power*, 132 HARV. L. REV. 536, 560–69 (2018).

³² *See* Arindrajit Dube, T. William Lester & Michael Reich, *Minimum Wage Shocks, Employment Flows, and Labor Market Frictions*, 34 J. LAB. ECON. 663, 664 (2016); José Azar et al., *Minimum Wage Employment Effects and Labor*

compatible with labor monopsony. Monopsonistic employers do not fire workers after a compelled wage increase because a monopsonist will make money off workers even when forced to pay above the monopsony wage as long as the minimum wage is not too high. The minimum wage law pushes wages toward the competitive rate that would prevail in a competitive market. And if the minimum wage is closer to the competitive wage than the monopsony wage is, the minimum wage will result in higher employment—since more workers will work for the higher wage.³³

Similarly, labor law is a straightforward legal response to the problem of monopsonized labor markets. If the employer enjoys a monopsony over workers, then it can pay a wage lower than the marginal revenue product. The main reason that employers can do this is that workers rarely have a credible threat to quit if they are paid below the market wage—again because of search costs, employer concentration, and related frictions. Workers can increase their market power by agreeing among themselves to quit en masse (that is, strike) if the employer refuses to pay them a higher wage. As a result, the wage should be pushed toward the competitive level (though unions could demand wages higher than the competitive level if they are powerful enough). Labor law could be justified for its role in preventing employers from using aggressive tactics to defeat organization so as to preserve their market power over wages.

We can now see why the law goes to such trouble to classify workers as “contractors” and “employees,” and why employers try so hard to reclassify employees as contractors. Employees benefit from legal protection because they are subject to labor monopsony and hence are not protected by market competition. Employment law protections prevent employers from using their monopsony power to push down wages and worsen conditions. Labor law enables workers to counter employer monopsony power with their own aggregated bargaining power. Contractors do not benefit from employment law protections because market competition already protects them, while the right to organize would enable them to form cartels that charge above-market prices.³⁴ The market protects contractors because their discrete skills are valued similarly by numerous labor buyers. And this is why firms that buy work from relational workers have an incentive to misclassify them as contractors; by doing so, they evade employment and labor laws that restrict their ability to exercise monopsony power over their workers.

It is important to see that this justification for employment and labor law—for limiting the law to employees rather than extending it to contractors—is based on market structure, not on income inequality or poverty, which is the focus of the law review literature. From an empirical standpoint, the market structure theory is clearly superior. Employment and labor law protects wealthy employees as well as poor employees; and contractors are deprived of that protection regardless of whether they are rich or poor. From a normative standpoint, the market structure theory is superior as well. Employment law and labor law counter labor monopsony, which should generate wealth. While employment and labor law under this understanding should also redistribute wealth from on-average wealthier investors to on-average poorer employers, it does

Market Concentration 3–4 (Nat’l Bureau of Econ. Research, Working Paper No. 26101, 2019), <http://papers.nber.org/tmp/88516-w26101.pdf>.

³³ However, not all elements of employment law can be justified from a labor monopsony standpoint. When workers are paid above the minimum wage because of market competition, employment mandates and related protections do not benefit them and may harm them. See Naidu & Posner, *supra* note 30.

³⁴ Unless labor buyers violate the antitrust laws. See Part II.3.B, *infra*.

not do so in a targeted way.³⁵ They also do nothing for the very poor, who are often out of work. Traditional antipoverty programs and tax-and-transfers are a much more suitable way to address the problems of poverty and inequality.³⁶

II. Legal Implications

A. Misclassification

1. The Traditional Test

Because employment and labor law restricts firms' ability to exploit their labor monopsony power, firms have an incentive to classify workers as contractors, regardless of whether the workers are employees in an economic sense, that is, workers who are subject to managerial direction and face a high cost of exit because of their relationship-specific investments in the firm. Uber's classification of drivers as contractors have recently attracted public attention, but the problem is a longstanding one.

Courts use various tests to resolve disputes over whether a worker is an employee or an independent contractor. The common-law test turned on "control": the worker is an employee if the labor buyer "controls" her, and an independent contractor otherwise.³⁷ While that test was originally developed to determine whether a labor buyer is liable for torts committed by a worker, it continues to be used for classification issues in the employment context.³⁸ For federal legislation that protects the rights of workers, the relevant test is the "economic reality" or "economic

³⁵ Indeed, even minimum wage laws may not help the poor. The reason is that they are very crude rules that can end up raising wages above the competitive rate even when labor monopsony exists; when they do, the higher costs may be passed on to low-income buyers. For empirical analysis, see Thomas MaCurdy, *How Effective Is the Minimum Wage at Supporting the Poor?*, 123 J. POL. ECON. 497 (2015); Doruk Cengiz, Arindrajit Dube, Attila Lindner, Ben Zipperer, *The Effect of Minimum Wages on Low-Wage Jobs*, 134 Q.J. ECON. 1405 (2019).

³⁶ There remains a significant question how much employment and labor law helps workers even when labor markets are monopsonized. I have abstracted from this question for the purpose of this paper, but a few comments are in order. First, standard employment law protections—including the minimum wage—can probably help only low-skill workers who would otherwise be paid below the minimum wage. The minimum wage does not help higher-skill workers; and it is unlikely that other protections, for example, mandatory family medical leave, can benefit them since employers can reduce wages to offset the cost of mandates. See Lawrence H Summers, *Some Simple Economics of Mandated Benefits*, 79 AMER. ECON. REV. 177 (1989). Second, even the minimum wage may not benefit low-income workers directly if employers offset the cost by raising prices in product markets. The extent of this effect remains empirically ambiguous, and probably varies by location. See MaCurdy, *supra*; Peter Harasztosi and Attila Lindner, *Who Pays for the Minimum Wage?*, 109 AMER. ECON. REV. 2693 (2019). Third, labor laws that encourage union organization probably do benefit workers; union organization is probably the most effective remedy for labor monopsony. Several recent empirical papers have found as much. See Efraim Benmelech, Nittai Bergman & Hyunseob Kim, *Strong Employers and Weak Employees: How Does Employer Concentration Affect Wages?* 4 (Nat'l Bureau of Econ. Research, Working Paper No. 24307, 2018), <http://www.nber.org/papers/w24307.pdf>; Elena Prager & Matthew Schmitt, *Employer Consolidation and Wages: Evidence from Hospitals* 4 (Wash. Ctr. Equitable Growth, 2019), http://papers.ssrn.com/sol3/papers.cfm?abstract_id=3391889.

³⁷ For a statement of the common law test, see *Cnty. for Creative Non-Violence v. Reid*, 490 U.S. 730, 751–52 (1989).

³⁸ Carlson, *supra* note 4, at 159–60. I discuss *respondeat superior* below.

dependence” test: a worker is an employee if she is “dependent” on the labor buyer.³⁹ Related state legislation uses similar tests.

The economic dependence test involves six factors:

- (1) the nature and degree of the alleged employer’s control as to the manner in which the work is to be performed;
- (2) the alleged employee’s opportunity for profit or loss depending upon his managerial skill;
- (3) the alleged employee’s investment in equipment or materials required for his task, or his employment of workers;
- (4) whether the service rendered requires a special skill;
- (5) the degree of permanency and duration of the working relationship;
- (6) the extent to which the service rendered is an integral part of the alleged employer’s business.⁴⁰

Commentators agree that courts apply this test inconsistently.⁴¹ Part of the problem is that all these factors are vague, and relationships between labor buyers and workers are extremely diverse. It is easy to think of examples of employees who have a great deal of control over their work (traveling salesmen); who are paid bonuses based on managerial skill (managers, for example); who have special skills (industrial scientist); and so on. But the main problem is that the normative goal of the test has been forgotten or perhaps never fully understood, and so there is no single principle that disciplines how courts apply the various factors.

2. The Relational Work Test

The solution is a test grounded in the normative goal of employment and labor law, which, as I have argued, is to counter labor monopsony where workers make relationship-specific investments and face high exit costs. The two major factors—economic dependence and control—are unified: a worker is economically dependent on an employer because of the high costs of exit, and because of the high cost of exit the worker is willing to subject herself to the firm’s control. Thus, economic dependence does not mean poverty; it means labor monopsony. Economic

³⁹ See, e.g., *Bartels v. Birmingham*, 332 U.S. 126, 130 (1947) (noting that application of social legislation such as the Social Security Act does not hinge on employer “control,” but rather the extent to which the employees “as a matter of economic reality are dependent upon the business to which they render service”).

⁴⁰ The factors are quoted from *Scantland v. Jeffrey Knight, Inc.*, 721 F.3d 1308, 1312 (11th Cir. 2013).

⁴¹ The Supreme Court recognized the difficulty of applying multi-factor employment tests over 70 years ago. See *NLRB v. Hearst Publications, Inc.*, 322 U.S. 111, 121 (1944) (“Few problems in the law have given greater variety of application and conflict in results than the cases arising in the borderland between what is clearly an employer-employee relationship and what is clearly one of independent entrepreneurial dealing.”). See also Carlson, *supra* note 4, at 171–74 (describing difficulties with the multi-factor test).

dependence and control are not in tension; they are different aspects of the employment relationship.

According to the “relational work” test, as I will call it, a worker is a person who performs activities (“work”) for another person in return for pay. A worker is an employee of a firm (or household) when the worker’s cost of finding alternative work of the same type and at the same level of pay is high (“high exit option”). A worker has a high exit option when the work is “relational,” that is, the work is worth more if performed for a single firm over time than if performed in discrete units. In short: employment is relational work; independent contractor status arises for discrete work.

Seen in this way, the factors used in the economic dependence test can be given greater specificity:

(1) Control. When work is relational, the labor buyer retains “control” over the worker, in the sense of discretionary authority over the worker’s behavior as the worker performs services. Control in this sense is necessary because the labor buyer must constantly coordinate the worker’s work with the work of others. In contrast, discrete work can be largely specified by contract in advance of the work: thus, the labor buyer does not retain control, that is, discretionary authority, over the worker’s behavior as the worker performs tasks.

(2) Opportunity for profit or loss depending upon the workers’ managerial skill. When work is relational, the worker allows herself to be managed by the labor buyer who directly, or through subordinates, coordinates the behavior of multiple workers or adjusts work in response to the buyer’s changing needs. The independent contractor uses managerial skill in the course of contracting herself out to multiple labor buyers, and thus is compensated for that managerial skill as well as taking on the risk of managerial failure that results in the loss of clients and hence profits.

(3) Investment in equipment or materials required for his task, or employment of workers. This factor reflects the importance of the assignment of property rights so as to locate discretion in the party that is most likely to use it to maximize the surplus. When the work is discrete, the property right is assigned to the worker. When the worker is relational, the property right is assigned to the employer.

(4) The service rendered requires a special skill. If we interpret “special” to mean relational, this factor fits the relational test. A plumber’s skill is the same whether the plumber is self-employed or employed by a firm; what is special about the skill in the second case is that plumber adjusts it, through relationship-specific investment, so it satisfies the unique needs of that employer.

(5) The degree of permanency and duration of the working relationship. Because the value of relational work is highest for the labor buyer with whom the worker has a relationship, and low for other potential labor buyers, the relational worker will tend to remain with the

labor buyer in question. Discrete workers, by contrast, may move from labor buyer to labor buyer.

(6) The extent to which the service rendered is an integral part of the alleged employer's business. Relational work, unlike discrete work, is integral, in the sense that it is valuable to the extent that it is used in the operations of the labor buyer with whom the worker has a relationship.

This gloss of the factors in the economic dependence test shows that they can be integrated into the relational work framework. My claim is not, however, that the courts consistently interpret the factors in this way. On the contrary, in the hands of the courts, the factors have become unmoored from any plausible goal of employment and labor legislation. The courts frequently seem more focused on whether workers are low-income than on whether they are economically dependent in the market-structure sense. But, as noted earlier, low-income status is not a reason for classifying workers as employees rather than as independent contractors.

In 2019, the California state legislature passed a law that expanded and simplified its definition of employee. The earlier test resembled the federal economic dependence test. The new law provided that anyone who works for remuneration is an employee unless the labor buyer shows that:

- (A) The person is free from the control and direction of the hiring entity in connection with the performance of the work, both under the contract for the performance of the work and in fact.
- (B) The person performs work that is outside the usual course of the hiring entity's business.
- (C) The person is customarily engaged in an independently established trade, occupation, or business of the same nature as that involved in the work performed.⁴²

The first provision repeats the control test, and the third provision reflects the idea that discrete work can, and often will, be performed independently of a single labor buyer. But the second provision is broader than the federal rule, and seem questionable as well. Under the second provision, a plumbing contractor who establishes contacts with plumbers and contracts them out to households would be deemed an employer. Yet unless the plumbers work in teams on large projects, it does not seem that their plumbing skills are relational to the operations of the plumbing contractor.

3. Alternative Monopsony-Based Tests

Relationship-specific is not the only source of monopsony power. As discussed earlier, labor monopsony may exist simply because few employers compete for workers in labor markets. I have argued that this problem should be left to antitrust law, but an argument could also be made

⁴² Assem. Bill 5, 2019–2020 Reg. Sess. (Cal. 2019), http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201920200AB5.

that workers should be classified as employees in monopsonized labor markets regardless of the source of monopsony.

Another source of monopsony is the labor buyer's private information about a worker's ability. If a worker tries to find work at another firm, that firm may worry that the worker has been forced out by the incumbent firm because of the worker has low ability. Even when this is not true, the phenomenon will tend to raise exit costs for workers.⁴³ Labor economists emphasize search costs and job differentiation as major sources of monopsony.⁴⁴ Thus, one might be worried that the relational test is too narrow.

But private information, search costs, and job differentiation are all closely related to relationship-specificity, even if the latter concept does not capture the whole range of sources of monopsony power. And as one broadens the scope of the inquiry, one will have greater difficulty formulating a test that courts can administer. Perhaps a workable test is cost of exit. Workers whose cost of exit exceeds some threshold—measured as the percentage pay cut the worker would accept before moving to another job (or residual labor market elasticity)—would be classified as employees. Using surveys and empirical tests based on data on labor mobility, courts could evaluate classification claims. The exit cost test would address all sources of labor monopsony, unlike the relational work test. On the other hand, it may be too difficult to administer.

4. The Misclassification Debate

Gig-economy workers float somewhere between the traditional employee and the traditional contractor. Take the case of Uber. Drivers control some aspects of their work (they choose their automobile, within constraints; they choose when to work) and not others (choice of routes); they do not manage anyone but they do invest in equipment, including the automobile itself; they do not seem to exercise a special skill (driving) but they are free to switch among other labor buyers including other ride-sharing companies and their own clients. Drivers sometimes seem economically dependent in the traditional sense because most of them earn a paltry sum for their work, but they are not economically dependent in the sense of being dependent on just one company, or even on the occupation of driving. Other gig-economy firms locate their workers in other places along the spectrum from contractor to employee. Some workers are allowed to negotiate with customers, set their prices, and exert greater control over working conditions, while others are given less control, and prohibited from using other platforms. Critics argue that firms evade the spirit of the employment laws by giving workers minimal freedoms or responsibilities necessary to qualify them for contractor status rather than allowing those conditions to be determined by the nature of the business.

We can use the controversy over Uber's classification of drivers as an illustration. Drivers use "effort" to drive expeditiously and maintain their cars. By allocating ownership of the cars to the drivers, Uber gives them high-powered incentives to maintain their cars. But it is easy to imagine the alternative approach in which Uber owns the cars. While the drivers' care incentives

⁴³ See Daron Acemoglu & Jörn-Steffen Pischke, *Why Do Firms Train? Theory and Evidence*, 113 Q.J. ECON. 79 (1998).

⁴⁴ For a discussion of the literature, see Suresh Naidu, Eric A. Posner, and E. Glen Weyl, *Antitrust Remedies for Labor Market Power*, 132 Harv. L. Rev. 536 (2018).

would be diminished, Uber would gain greater control over other aspects of their work—for example, the level of courtesy they offer passengers, which may be important for Uber’s brand. If Uber owns the cars, then its threat to punish discourteous drivers is strengthened: it can not only kick the drivers off the platform (where drivers can use other platforms); it can take away their cars as well. The analysis is similar to the analysis of trucking where (in my example) the courtesy problem takes the place of the backhaul problem in trucking.

From a legal perspective, the question is whether Uber classifies its drivers as contractors because the care incentives are more important and more responsive to the assignment of property rights than courtesy incentives. If the answer is yes, then Uber’s classification of drivers reflects the underlying economic reality of ride-sharing rather than an attempt by Uber to evade the law. The drivers do not make relationship-specific investments in Uber’s platform, and so the drivers are protected by competition—they can find work with other platforms or independently find clients—and so do not need the protection of the law.

Unfortunately, it is difficult to determine what Uber (or any other ride-sharing or taxi or limousine company) would do in the absence of a legal regime that will push them to classify workers as contractors in order to avoid the costs of complying with employment and labor laws. A court or regulator can at best do a rough analysis. Applying the relational work test, the most striking aspect of the Uber case is that it (like the trucking companies after the introduction of onboard monitoring) imposes a considerable level of control over drivers through contract specificity (thanks to the platform technology). This allows Uber to control their level of courtesy with the star rating system, for example. And this allows Uber to assign the property right over the car to the worker so as to enhance the worker’s motivation to take care. On the other hand, the monitoring technology that enables Uber to control the routes, courtesy, and other elements of the service would also enable it to monitor the driver’s care level (as the trucking example shows). And the empirical evidence suggests that drivers cannot easily abandon the Uber platform once they being using it,⁴⁵ suggesting a nontrivial relationship-specific investment in using and mastering the technology.

While we will not try to resolve the Uber question here, the example illustrates how the relational work test should be applied.

B. Antitrust Law

The gig economy has spawned antitrust suits that allege that firms like Uber have cartelized labor markets. However, while many commentators allege that Uber has exploited drivers, at least one group of plaintiffs argue that Uber coordinated price-fixing by drivers.⁴⁶ Which is it? Are drivers complicit wrongdoers or passive victims?

The problem comes back to the question of whether drivers are employees or contractors. Labor law assumes that employees lack bargaining power—that is, they face an employer-

⁴⁵ See Joshua D. Angrist, Sydnee Caldwell & Jonathan V. Hall, *Uber vs. Taxi: A Driver’s Eye View* 1-2 (Nat’l Bureau of Econ. Research, Working Paper No. 23891, 2017), <https://www.nber.org/papers/w23891.pdf>.

⁴⁶ See *Meyer v Kalanick*, 174 F. Supp. 3d 817, 819–20 (S.D.N.Y. 2016), *vacated*, *Meyer v. Uber Technologies, Inc.*, 868 F.3d 66 (2d Cir. 2017).

monopsonist—and gives them the right to organize, so that they can counter the employer’s bargaining power with their own. Antitrust law, for this reason, recognizes a labor exemption: employees do not violate antitrust law by organizing even though a union is a type of price- (or actually wage-) fixing arrangement.⁴⁷ Labor union organization when employees are subject to monopsony should increase production and (in many cases) lower prices, while raising wages as well. However, because contractors are not employees they do not benefit from the labor exemption, and they are not permitted to organize.⁴⁸ Organization by contractors in a competitive market would result in a cartel, and hence less production and higher prices. That is why drivers who organize violate the law if they are contractors but not if they are employees.

There has been some confusion about this issue. Sanjukta Paul argues that Uber should not be permitted to set the prices for workers if antitrust law prohibits workers from doing it themselves.⁴⁹ She argues that it would be anomalous if Uber were allowed to organize a sellers’ cartel by inviting drivers to its platform, fixing prices, and collecting the rents paid by consumers—when drivers are prohibited by antitrust law from fixing prices.⁵⁰ However, the seeming paradox dissolves upon inspection. Uber is not allowed to organize drivers and set their prices if the result would be a monopoly. That would violate section 1 and 2 of the Sherman Act, and section 7 of the Clayton Act as well. Lawsuits against Uber for cartelizing the market have failed because Uber faces competition from taxis, other rideshare companies like Lyft, public transportation, and other services. And while it is true that if some drivers fixed prices, they would violate the antitrust laws, that is only because of the crude per se ban on price-fixing. The drivers could easily avoid liability by merging into a single firm rather than fixing prices as long as the merger does not encompass an excessive share of the market.

The discussion should make clear how employment law protections fit in with antitrust and labor law to create a general legal structure that governs the relationships between workers and labor buyers. Antitrust law prohibits workers in competitive labor markets from forming cartels, while labor and employment law protects workers in monopsonized labor markets—labor law, by allowing them to organize, and employment law (including the minimum wage law), by regulating prices and conditions. On this view, workers are divided into “employees” who are subject to monopsony and contractors who are not.

Thus, it is important in the antitrust context to get straight the reasons why workers should be classified as employees or contractors. When workers operate in a competitive labor market, they should be classified as contractors, and thus forbidden to organize, because the right to organize would result in cartelization and above-market wages. When workers operate in a monopsonized labor market, they should be classified as employees, and thus allowed to organize, because the right to organize should allow them to counter employer market power and thus raise wages toward the competitive level.

⁴⁷ See *United States v. Hutcheson*, 312 U.S. 219, 232, 235–37 (1941) (holding that the Clayton Act and Norris-LaGuardia Act together provide unions a statutory exemption from antitrust liability).

⁴⁸ See *Columbia River Packers Ass’n v. Hinton*, 315 U.S. 143, 144–45 (1942) (excluding “independent entrepreneurs” from the labor exception); *L.A. Meat and Provision Drivers Union, Local 626 v. United States*, 371 U.S. 94, 96, 99–102 (1962) (same).

⁴⁹ Sanjukta M. Paul, *Uber as For-Profit Hiring Hall: A Price-Fixing Paradox and its Implications*, 38 BERKELEY J. OF EMPLOYMENT & LABOR L. 233, 236–39 (2017).

⁵⁰ *Id.*

We see analogies on the product market side. When firms have natural monopolies, they are generally immune to antitrust challenge, but they may be subject to price regulation and related regulations designed to prevent them from abusing their market power. Given the parallel nature of product and labor markets, it should not be surprising that there are analogous rules on the labor side even if they have rarely been recognized as such because of the different legal terminology.

It is possible to argue that if employment and labor law counter labor monopsony power, the antitrust laws are not needed, and workers should not be permitted to bring antitrust actions against employers. But the two approaches to monopsony are complements. On the product market side, antitrust law coexists with various forms of price regulation, like usury laws, anti-price-gouging laws, and insurance premium limits. Antitrust law focuses on tactics employed by firms to increase their market power or extend monopolies into new markets, while price regulation limits the negative effects of firms that have achieved market power lawfully. Similarly, while antitrust law can be used to prevent firms from increasing their power over labor markets through collusion, mergers, and the like, other areas of law are needed to counter the negative effects of labor market power that is achieved lawfully.⁵¹

A final point is that while I have assumed that discrete-work markets are competitive, they may not be. Discrete-work markets should be more competitive than relational-work markets because the discrete worker does not make a relationship-specific investment in a primary labor buyer and hence does not have high exit costs derived from the relationship. But a discrete worker may still face high exit costs if there are few labor buyers. Concentration arises simply because there are few buyers, and concentration could occur either legally (because there are economies of scale, for example) or illegally (because mergers result in a single dominant labor buyer). In this case, the logical source of legal protection for the discrete worker is antitrust law, which distinguishes between concentration that is regarded as socially harmful and concentration that is regarded as tolerable.⁵²

C. Respondeat Superior

Tort law imposes liability on employers for torts committed by employees under certain conditions. The law distinguishes employers and independent contractors using the control test. If a labor buyer exerts control over a worker, then it will normally be classified as an employer and be held liable. If a labor buyer does not exert control, then it will be classified as a “customer” (with the worker classified as a contractor) and not be held liable. However, there are numerous exceptions to this principle. An employer is not liable for torts committed by employees outside the scope of employment, and a firm may be liable for torts committed by independent contractors, if the tort is partly the result of the labor buyer’s negligence or is the result of inherently dangerous activity.⁵³

⁵¹ Naidu & Posner, *supra* note 30.

⁵² Naidu, Posner & Weyl, *supra* note 31.

⁵³ Restatement (Third) of Agency § 7.05, § 7.07 (2006); Alan O. Sykes *The Economics of Vicarious Liability*, 93 YALE L.J. 1231 (1984).

Alan Sykes influentially argued that respondeat superior (and vicarious liability for labor buyers more generally) is justified when the worker is judgment-proof and the labor buyer can sufficiently control the worker so as to deter the worker from committing torts—for example, through training, inspections, monitoring, and sanctions.⁵⁴ If the worker is not judgment-proof, the law can impose liability directly on the worker in order to deter torts. Respondeat superior is necessary because a firm and a worker might otherwise agree that the worker’s judgment-proofness will be used to shield the firm from liability in return for which the firm compensates the worker. But if the firm has no practical means of controlling the worker’s behavior, then respondeat superior does not improve incentives to take care, but is simply a tax on the enterprise.

One puzzle in Sykes’ analysis is that it does not explain why the law uses the categories of employee and contractor. His analysis suggests instead that the law should always impose liability on person A if person B performs a task at A’s request and for A’s benefit, person B is judgment proof, and person A can control person B. The categories of employment and contractor do no work in this analysis. Indeed, when courts evaluate respondeat superior claims, they do not rely on the firm’s and worker’s own classification of the worker. Courts simply look for evidence of control, and then classify the worker based on how much control the firm exerts over her. Thus the association of respondeat superior with the employment relation seems all the more puzzling.⁵⁵

The solution to this puzzle brings us back to relationship-specificity. A worker who makes relationship specific investments in a firm (its assets and other workers) is subject to the firm’s control because, as a result of these investments, the worker’s cost of exit is high. Because the cost of exit is high, the worker will be concerned that if she commits torts for which the employer is liable, she will be denied bonuses, raises, and promotions—or that she will be fired. By contrast, a discrete worker does not fear such sanctions—because she can work for anyone else. Sykes implicitly acknowledges this distinction, noting that “agents often earn returns in excess of what they can earn in their next-best employment opportunity, and expect those returns to continue into the future. Such agents have an important stake in retaining their current positions.”⁵⁶ Labor monopsony, which ensures that the exit cost is high, thus is important for giving labor buyers leverage for disciplining workers who commit torts. This distinction explains why respondeat superior is associated with employment relationships rather than the broader class of buyer-worker relationships encompassing discrete workers.

D. The Employee/Contractor Distinction in Other Areas of the Law

Congress and state legislatures have built an enormous administrative structure on the distinction between employee and contractor. Employees are entitled to unemployment insurance; contractors are not. Employees benefit from tax-subsidized retirement and health insurance plans operated by employers; contractors do not. And taxes on the wages of employees are withheld by employers; taxes on payments made to contractors are not withheld by the buyers of their services. A question arises why these programs and practices should apply to employees only, and not to workers more broadly.

⁵⁴ Sykes, *supra* note 53, at 1232, 1235–39, 1244.

⁵⁵ As noted by Keith Cunningham-Parmeter, *Gig-Dependence: Finding the Real Independent Contractors of Platform Work*, 39 N. Ill. U. L. Rev. 379, 406-08 (2019).

⁵⁶ Sykes, *supra* note 53, at 1254 (citations omitted).

The existing system may not be optimal but it follows a certain logic. Relationship-specificity creates monopsony, which is another way of saying that workers develop thick relationships with labor buyers and cannot easily exit those relationships. Employees are therefore more vulnerable if laid off than contractors who lose business. A locksmith who loses one customer can find another; but a factory worker whose value to a firm is relationship-specific will have trouble finding a comparable job if laid off. Thus, unemployment insurance may be more appropriate for employees than for contractors, at least in normal times, as opposed to a recession that dries up demand for the contractor's services.

For health insurance, the stickiness of the relationship between employee and employer may help address adverse selection problems in the insurance market and justify group-based insurance. By contrast, contractors who try to create an insurance pool take the risk that the healthiest will opt out. And as for retirement subsidies, there does not seem to be a good reason for offering them to employees but not contractors. But the law does offer various self-employed retirement subsidies that contractors can take advantage of. The law seems to draw a distinction between employee-based retirement subsidies and contractor retirement subsidies so as to enlist employers to manage retirement plans—taking advantage of the administrative capacity of large labor buyers who have continuous relationships with workers.

Finally, tax withholding is a form of monitoring and enforcement and these tasks should be assigned to the party (labor buyer or worker) who is most likely to perform the efficiently and conscientiously. As a rough approximation, labor buyers who employ people will have the infrastructure in place for withholding, while households and even firms that rely on contracts will not.

Conclusion

Let us sum up the analysis. A worker is an employee when she makes a relationship-specific investment in, and applies her labor to, assets owned by, or other workers employed by, another party. A worker is a contractor when the value of her output is independent of such relationships. An employee's output is thus relationship-specific or "relational": the output is valued more by the labor buyer than by outsiders. A contractor's output is "commodified": its value does not depend on the identity of the labor buyer. Because of relationship-specificity, the employer is a monopsonist with respect to the worker's output. The buyer from a contractor is not (normally) a monopsonist. If it is, it is not because of the relationship with the worker but because of market concentration.

Because an employer is a monopsonist, it can (and will) pay the worker less than marginal revenue product, resulting in inefficiency (lower output) and often inequity. Thus, there is a labor market failure. When the worker is a contractor, there is no such labor market failure. The labor buyer is not a monopsonist (normally). Exit is costly for the employee, cheap for the contractor.

In light of these distinctions, employment and labor law should be understood as legal mechanisms for addressing the labor market failure caused by labor monopsony. Employment law is a form of price (wage) and quality (conditions) regulation, understood to raise wages/conditions

toward the competitive rate. Labor law is a form of power aggregation, understood to convert monopsony into bilateral monopoly, with better outcomes closer to the competitive rate. This is not to say that all employment law protections work as intended; support for unions can also produce perverse consequences. But these areas of law are best understood as the legal response to the problem of labor monopsony. Antitrust law provides an additional residual set of protections that are invoked when anticompetitive efforts to cartelize or monopsonize markets result in higher market power.

Firms misclassify workers as contractors in order to avoid employment and labor law protections so that they can exploit their monopsony power and suppress wages. To prevent misclassification, courts and regulators must understand the true economic relationship between a labor buyer and a worker—essentially whether the labor buyer exerts monopsony power because of relationship-specificity. The factors used in the various misclassification tests should be interpreted in this light.

This account should be understood as an attempt to rationalize labor and economic law, and the law of classification—in the sense of offering the best justification for these laws. I do not try to prove that this body of law does what it is supposed to do. There is some evidence that it does, but the evidence is mixed, and varies for different types of rights and protection.

This analysis leaves some questions. One could imagine an alternative legal regime in which no formal distinction is made between employees and contractors and instead employment protections are extended to all workers on the basis of market structure. For example, a minimum wage law could be designed to apply to all labor markets in which the level of monopsony (typically measured as residual labor supply elasticity, which measures the extent to which a worker will tolerate a below-market wage before quitting) exceeds a threshold.⁵⁷ What would be wrong with such a system? One possible answer is that measurement is too complex, and courts and regulators require a simpler even if blunter test. It is also possible that our current system reflects historical contingency. Policymakers have found the contractor/employee distinction useful in a range of contexts, where the consequences of the distinction are different. That may explain both why the distinction is so entrenched and why the modern test has gotten so muddled.

⁵⁷ Recent papers have measured the residual labor supply elasticity in various settings. *e.g.*, David W. Berger, Kyle F. Herkenhoff & Simon Mongey, *Labor Market Power* (Nat'l Bureau of Econ. Research, Working Paper No. 25719, 2019), <https://www.nber.org/papers/w25719>, and related variables, like the workers' outside option or what I have called exit options. See Sydnee Caldwell and Oren Danieli, *Outside Options in the Labor Market* (Nov. 7, 2018), http://scholar.harvard.edu/files/danieli/files/danieli_jmp.pdf.

U.S. PATENT AND TRADEMARK OFFICE
Patent Technology Monitoring Team (PTMT)

PATENT COUNTS BY ORIGIN AND TYPE CALENDAR YEAR 2020

(01-JAN-2020 to 31-DEC-2020)

This report, prepared from the Technology Assessment and Forecast (TAF) database, displays U.S. Patent and Trademark Office (USPTO) patent grants by state/territory and country of origin. Separate counts are provided for utility, design, plant, and reissue patents.

(patent origin is determined by the residence of the first-named inventor)

The following table displays patent counts by origin and patent type -- utility, design, plant, and reissue patents.

US or Foreign	Code	State, Territory, or Country	Utility	Design	Plant	Reissue	Total
US	AL	ALABAMA	537	69	1	0	607
US	AK	ALASKA	59	8	0	0	67
US	AZ	ARIZONA	2870	264	2	4	3140
US	AR	ARKANSAS	453	79	2	1	535
US	CA	CALIFORNIA	46040	4168	208	80	50496
US	CO	COLORADO	3357	380	0	3	3740
US	CT	CONNECTICUT	3259	213	3	3	3478
US	DE	DELAWARE	333	25	0	0	358
US	FL	FLORIDA	4659	801	37	16	5513
US	GA	GEORGIA	2942	398	21	4	3365
US	HI	HAWAII	141	19	1	0	161
US	ID	IDAHO	1161	51	0	1	1213
US	IL	ILLINOIS	5379	781	11	4	6175
US	IN	INDIANA	2242	284	0	3	2529
US	IA	IOWA	1130	72	0	0	1202
US	KS	KANSAS	788	115	2	0	905

US or Foreign	Code	State, Territory, or Country	Utility	Design	Plant	Reissue	Total
US	KY	KENTUCKY	776	92	1	0	869
US	LA	LOUISIANA	454	47	11	0	512
US	ME	MAINE	204	24	0	0	228
US	MD	MARYLAND	2309	167	2	4	2482
US	MA	MASSACHUSETTS	8176	587	9	18	8790
US	MI	MICHIGAN	6493	829	147	9	7478
US	MN	MINNESOTA	4350	372	2	10	4734
US	MS	MISSISSIPPI	165	34	6	0	205
US	MO	MISSOURI	1510	143	1	3	1657
US	MT	MONTANA	148	27	4	1	180
US	NE	NEBRASKA	365	62	0	0	427
US	NV	NEVADA	895	150	0	0	1045
US	NH	NEW HAMPSHIRE	990	110	0	3	1103
US	NJ	NEW JERSEY	4479	530	4	14	5027
US	NM	NEW MEXICO	479	30	0	0	509
US	NY	NEW YORK	9361	1155	3	21	10540
US	NC	NORTH CAROLINA	3603	294	11	9	3917
US	ND	NORTH DAKOTA	128	10	0	0	138
US	OH	OHIO	4583	790	8	5	5386
US	OK	OKLAHOMA	626	56	6	1	689
US	OR	OREGON	3192	877	18	6	4093
US	PA	PENNSYLVANIA	4271	448	13	5	4737
US	RI	RHODE ISLAND	407	95	0	1	503
US	SC	SOUTH CAROLINA	1206	173	3	0	1382
US	SD	SOUTH DAKOTA	145	19	0	1	165
US	TN	TENNESSEE	1227	257	3	7	1494

US or Foreign	Code	State, Territory, or Country	Utility	Design	Plant	Reissue	Total
US	TX	TEXAS	12201	953	23	15	13192
US	UT	UTAH	1699	336	0	1	2036
US	VT	VERMONT	325	34	0	0	359
US	VA	VIRGINIA	2827	165	1	9	3002
US	WA	WASHINGTON	8553	530	22	7	9112
US	WV	WEST VIRGINIA	118	15	1	0	134
US	WI	WISCONSIN	2416	506	9	8	2939
US	WY	WYOMING	147	12	0	0	159
US	DC	DISTRICT OF COLUMBIA	318	21	0	0	339
US	GU	GUAM	2	2	0	0	4
US	PR	PUERTO RICO	52	19	1	0	72
US	VI	U.S. VIRGIN ISLANDS	3	2	0	0	5
US	US	U.S. UNSPECIFIED REGION	19	1	0	0	20
US		-- Subtotal --	164572	17701	597	277	183147
FOREIGN	ALX	ALBANIA	0	1	0	0	1
FOREIGN	ADX	ANDORRA	1	0	0	0	1
FOREIGN	AIX	ANGUILLA	1	0	0	0	1
FOREIGN	ARX	ARGENTINA	96	7	1	0	104
FOREIGN	AMX	ARMENIA	21	0	0	0	21
FOREIGN	AUX	AUSTRALIA	1862	428	30	10	2330
FOREIGN	ATX	AUSTRIA	1511	96	0	1	1608
FOREIGN	AZX	AZERBAIJAN	1	0	0	0	1
FOREIGN	BSX	THE BAHAMAS	6	0	0	0	6
FOREIGN	BHX	BAHRAIN	2	0	0	0	2
FOREIGN	BDX	BANGLADESH	2	0	0	0	2

US or Foreign	Code	State, Territory, or Country	Utility	Design	Plant	Reissue	Total
FOREIGN	BBX	BARBADOS	4	2	0	0	6
FOREIGN	BYX	BELARUS	14	5	0	0	19
FOREIGN	BEX	BELGIUM	1374	88	19	1	1482
FOREIGN	BMX	BERMUDA	2	1	0	0	3
FOREIGN	BOX	BOLIVIA	3	0	0	0	3
FOREIGN	BAX	BOSNIA AND HERZEGOVINA	1	0	0	0	1
FOREIGN	BRX	BRAZIL	494	45	0	0	539
FOREIGN	VGX	BRITISH VIRGIN ISLANDS	2	1	0	0	3
FOREIGN	BNX	BRUNEI DARUSSALAM	8	0	0	0	8
FOREIGN	BGX	BULGARIA	56	3	0	0	59
FOREIGN	BIX	BURUNDI	1	0	0	0	1
FOREIGN	CAX	CANADA	7274	610	9	21	7914
FOREIGN	KYX	CAYMAN ISLANDS	8	7	0	0	15
FOREIGN	TDX	CHAD	1	0	0	0	1
FOREIGN	CLX	CHILE	69	0	5	0	74
FOREIGN	HKX	CHINA, HONG KONG S.A.R.	742	267	0	2	1011
FOREIGN	MOX	CHINA, MACAU S.A.R.	18	0	0	0	18
FOREIGN	CNX	CHINA, PEOPLE'S REPUBLIC OF	21428	5399	2	16	26845
FOREIGN	COX	COLOMBIA	44	7	0	0	51
FOREIGN	CRX	COSTA RICA	38	0	7	0	45
FOREIGN	HRX	CROATIA	17	4	0	0	21
FOREIGN	CUX	CUBA	3	0	0	0	3
FOREIGN	CYX	CYPRUS	13	1	0	0	14
FOREIGN	CZX	CZECH REPUBLIC	325	14	1	0	340
FOREIGN	DKX	DENMARK	1180	149	42	3	1374
FOREIGN	DOX	DOMINICAN REPUBLIC	3	0	0	0	3

US or Foreign	Code	State, Territory, or Country	Utility	Design	Plant	Reissue	Total
FOREIGN	ECX	ECUADOR	2	0	0	0	2
FOREIGN	EGX	EGYPT	36	1	0	0	37
FOREIGN	SVX	EL SALVADOR	1	0	0	0	1
FOREIGN	EEX	ESTONIA	27	9	0	0	36
FOREIGN	FIX	FINLAND	1456	90	0	6	1552
FOREIGN	FRX	FRANCE	7090	482	39	3	7614
FOREIGN	GEX	GEORGIA (REPUBLIC OF)	2	0	0	0	2
FOREIGN	DEX	GERMANY	17785	1282	84	22	19173
FOREIGN	GIX	GIBRALTAR	2	0	0	0	2
FOREIGN	GRX	GREECE	115	11	0	0	126
FOREIGN	GDX	GRENADA	1	0	0	0	1
FOREIGN	GTX	GUATEMALA	6	1	0	0	7
FOREIGN	GGX	GUERNSEY	0	1	0	0	1
FOREIGN	HUX	HUNGARY	155	8	0	1	164
FOREIGN	ISX	ICELAND	52	5	0	0	57
FOREIGN	INX	INDIA	5861	121	0	2	5984
FOREIGN	IDX	INDONESIA	12	5	0	1	18
FOREIGN	IRX	IRAN	121	2	0	0	123
FOREIGN	IQX	IRAQ	1	0	0	0	1
FOREIGN	IEX	IRELAND	916	23	1	2	942
FOREIGN	IMX	ISLE OF MAN	5	0	0	0	5
FOREIGN	ILX	ISRAEL	4661	154	26	3	4844
FOREIGN	ITX	ITALY	3196	569	26	3	3794
FOREIGN	JMX	JAMAICA	9	0	0	0	9
FOREIGN	JPX	JAPAN	51619	2050	24	86	53779
FOREIGN	JEX	JERSEY	4	11	0	0	15

US or Foreign	Code	State, Territory, or Country	Utility	Design	Plant	Reissue	Total
FOREIGN	JOX	JORDAN	14	0	0	0	14
FOREIGN	KZX	KAZAKHSTAN	6	1	0	0	7
FOREIGN	KEX	KENYA	23	0	0	0	23
FOREIGN	KRX	KOREA, SOUTH	21977	1667	0	61	23705
FOREIGN	KVX	KOSOVO	1	0	0	0	1
FOREIGN	KWX	KUWAIT	29	0	0	0	29
FOREIGN	KGX	KYRGYZSTAN	0	1	0	0	1
FOREIGN	LAX	LAOS	1	0	0	0	1
FOREIGN	LVX	LATVIA	14	4	0	0	18
FOREIGN	LBX	LEBANON	12	1	0	0	13
FOREIGN	LIX	LIECHTENSTEIN	33	2	0	0	35
FOREIGN	LTX	LITHUANIA	17	4	0	0	21
FOREIGN	LUX	LUXEMBOURG	63	24	0	0	87
FOREIGN	MKX	MACEDONIA	1	0	0	0	1
FOREIGN	MYX	MALAYSIA	269	23	1	0	293
FOREIGN	MTX	MALTA	19	1	0	0	20
FOREIGN	MUX	MAURITIUS	1	0	0	0	1
FOREIGN	MXX	MEXICO	341	41	2	0	384
FOREIGN	MDX	MOLDOVA	2	0	0	0	2
FOREIGN	MCX	MONACO	13	2	0	0	15
FOREIGN	MNX	MONGOLIA	1	0	0	0	1
FOREIGN	MAX	MOROCCO	4	0	0	0	4
FOREIGN	NAX	NAMIBIA	0	2	0	0	2
FOREIGN	NPX	NEPAL, FED. DEM. REPUBLIC OF	1	0	0	0	1
FOREIGN	NLX	NETHERLANDS	2835	205	366	11	3417
FOREIGN	NZX	NEW ZEALAND	368	108	16	2	494

US or Foreign	Code	State, Territory, or Country	Utility	Design	Plant	Reissue	Total
FOREIGN	NGX	NIGERIA	2	1	0	0	3
FOREIGN	NOX	NORWAY	725	36	0	2	763
FOREIGN	OMX	OMAN	2	0	0	0	2
FOREIGN	PKX	PAKISTAN	18	0	0	0	18
FOREIGN	PAX	PANAMA	2	2	0	1	5
FOREIGN	PYX	PARAGUAY	3	0	0	0	3
FOREIGN	PEX	PERU	9	5	0	0	14
FOREIGN	PHX	PHILIPPINES	69	2	0	1	72
FOREIGN	PLX	POLAND	395	39	2	0	436
FOREIGN	PTX	PORTUGAL	136	24	0	0	160
FOREIGN	QAX	QATAR	25	0	0	0	25
FOREIGN	ROX	ROMANIA	109	8	0	0	117
FOREIGN	RUX	RUSSIAN FEDERATION	677	32	0	0	709
FOREIGN	KNX	SAINT KITTS AND NEVIS	1	0	0	0	1
FOREIGN	LCX	SAINT LUCIA	1	0	0	0	1
FOREIGN	WSX	SAMOA	4	1	0	0	5
FOREIGN	SMX	SAN MARINO	1	0	0	0	1
FOREIGN	SAX	SAUDI ARABIA	974	7	0	0	981
FOREIGN	RSX	SERBIA	28	2	0	0	30
FOREIGN	SCX	SEYCHELLES	3	0	0	0	3
FOREIGN	SGX	SINGAPORE	1045	65	0	2	1112
FOREIGN	SKX	SLOVAKIA	56	6	0	0	62
FOREIGN	SIX	SLOVENIA	76	11	0	0	87
FOREIGN	ZAX	SOUTH AFRICA	155	27	12	1	195
FOREIGN	ESX	SPAIN	1030	128	16	0	1174
FOREIGN	LKX	SRI LANKA	8	0	1	0	9

US or Foreign	Code	State, Territory, or Country	Utility	Design	Plant	Reissue	Total
FOREIGN	RWX	RWANDA	1	0	0	0	1
FOREIGN	SEX	SWEDEN	3018	390	0	3	3411
FOREIGN	CHX	SWITZERLAND	2928	324	0	0	3252
FOREIGN	SYX	SYRIA	1	0	0	0	1
FOREIGN	TWX	TAIWAN	12141	947	0	17	13105
FOREIGN	THX	THAILAND	127	14	13	0	154
FOREIGN	TTX	TRINIDAD AND TOBAGO	3	0	0	0	3
FOREIGN	TNX	TUNISIA	4	0	0	0	4
FOREIGN	TRX	TURKEY	194	30	0	0	224
FOREIGN	TMX	TURKMENISTAN	1	0	0	0	1
FOREIGN	UAX	UKRAINE	78	17	0	0	95
FOREIGN	AEX	UNITED ARAB EMIRATES	113	17	0	0	130
FOREIGN	GBX	UNITED KINGDOM	7418	981	56	15	8470
FOREIGN	UYX	URUGUAY	13	0	0	0	13
FOREIGN	UZX	UZBEKISTAN	1	0	0	0	1
FOREIGN	VEX	VENEZUELA	3	0	0	0	3
FOREIGN	VNX	VIET NAM	32	13	0	0	45
FOREIGN	PSX	WEST BANK/GAZA	1	0	0	0	1
FOREIGN	ZWX	ZIMBABWE	2	1	0	0	3
FOREIGN		-- Subtotal --	187438	17176	801	299	205714
ALL	ALL	TOTAL- All Regions, U.S. and Foreign	352010	34877	1398	576	388861

The following table displays patent counts by origin and patent type -- utility, design, plant, and reissue patents.

Count or Percent	US or Foreign	Code	State, Territory, or Country	Utility	Design	Plant	Reissue	Total

Count or Percent	US or Foreign	Code	State, Territory, or Country	Utility	Design	Plant	Reissue	Total
COUNT	US		TOTAL- All U.S. States/Territories (#)	164572	17701	597	277	183147
COUNT	FOREIGN		TOTAL- All Foreign (#)	187438	17176	801	299	205714
COUNT	ALL		TOTAL- All Regions, U.S. and Foreign (#)	352010	34877	1398	576	388861
PERCENT	US		TOTAL- All U.S. States/Territories (%)	46.8	50.8	42.7	48.1	47.1
PERCENT	FOREIGN		TOTAL- All Foreign (%)	53.2	49.2	57.3	51.9	52.9
PERCENT	ALL		TOTAL- All Regions, U.S. and Foreign (%)	100.0	100.0	100.0	100.0	100.0

PTMT Contacts

Questions regarding these reports should be directed to:

U.S. Patent and Trademark Office
 Electronic Information Products Division - PTMT
 P.O Box 1450
 Alexandria VA 22313-1450

email: oeip@uspto.gov

address of PTMT pages at the USPTO Web Site: <http://www.uspto.gov/web/offices/ac/ido/oeip/taf/reports.htm>
 selected PTMT files available for download at : <http://www.uspto.gov/web/offices/ac/ido/oeip/taf/data/>

[Top](#)

[Home > Listing of Viewable PTMT Reports](#)

Last Modified : 04/15/2024 12:10:20



September 2020

TAX ADMINISTRATION

Opportunities Exist to Improve Oversight of Hospitals' Tax-Exempt Status

Why GAO Did This Study

Slightly more than half of community hospitals in the United States are private, nonprofit organizations. IRS and the Department of the Treasury have recognized the promotion of health as a charitable purpose and have specified that nonprofit hospitals are eligible for a tax exemption. IRS has further stated that these hospitals can demonstrate their charitable purpose by providing services that benefit their communities as a whole.

In 2010, Congress and the President enacted PPACA, which established additional requirements for tax-exempt hospitals to meet to maintain their tax exemption.

GAO was asked to review IRS's implementation of requirements for tax-exempt hospitals. This report assesses IRS's (1) oversight of how tax-exempt hospitals provide community benefits, and (2) enforcement of PPACA requirements related to tax-exempt hospitals.

What GAO Recommends

GAO is making one matter for congressional consideration to specify in the IRC what services and activities Congress considers sufficient community benefit. GAO is also making four recommendations to IRS, including to establish a well-documented process to ensure hospitals' community benefit activities are being reviewed, and to create codes to track audit activity related to hospitals' community benefit activities. IRS agreed with GAO's recommendations.

View GAO-20-679. For more information, contact Jessica Lucas-Judy at (202) 512-9110 or lucasjudj@gao.gov.

TAX ADMINISTRATION

Opportunities Exist to Improve Oversight of Hospitals' Tax-Exempt Status

What GAO Found:

Nonprofit hospitals must satisfy three sets of requirements to obtain and maintain a nonprofit tax exemption (see figure).

Requirements for Nonprofit Hospitals to Obtain and Maintain a Tax-Exemption

ORGANIZATIONAL AND OPERATIONAL REQUIREMENTS

A hospital must be organized and operate to achieve a charitable purpose—the promotion of health for the benefit of the community.

COMMUNITY BENEFITS

Internal Revenue Service has identified 30 factors that demonstrate community benefit:

- Operate an emergency room open to all, regardless of ability to pay
- Maintain a board of directors drawn from the community
- Maintain an open medical staff policy
- Provide care to all patients able to pay, including those who do so through Medicare and Medicaid
- Use surplus funds to improve facilities, equipment, and patient care
- Use surplus funds for education, medical training, research, and research

PATIENT PROTECTION AND AFFORDABLE CARE ACT (PPACA) REQUIREMENTS

Hospitals must:

- Conduct a community health needs assessment
- Set a limit on charges
- Maintain a written financial assistance policy
- Set billing and collection limits

IRS must review each tax-exempt hospital's community benefit activities at least once every 3 years.

Source: GAO review of relevant laws and regulations. | GAO-20-679

While PPACA established requirements to better ensure hospitals are serving their communities, the law is unclear about what community benefit activities hospitals should be engaged in to justify their tax exemption. The Internal Revenue Service (IRS) identified factors that can demonstrate community benefits, but they are not requirements. IRS does not have authority to specify activities hospitals must undertake and makes determinations based on facts and circumstances. This lack of clarity makes IRS's oversight challenging. Congress could help by adding specificity to the Internal Revenue Code (IRC).

While IRS is required to review hospitals' community benefit activities at least once every 3 years, it does not have a well-documented process to ensure that those activities are being reviewed. IRS referred almost 1,000 hospitals to its audit division for potential PPACA violations from 2015 through 2019. However, IRS could not identify if any of these referrals related to community benefits. GAO's analysis of IRS data identified 30 hospitals that reported no spending on community benefits in 2016, indicating potential noncompliance with providing community benefits. A well-documented process, such as clear instructions for addressing community benefits in the PPACA reviews or risk-based methods for selecting cases, would help IRS ensure it is effectively reviewing hospitals' community benefit activities.

Further, according to IRS officials, hospitals with little to no community benefit expenses would indicate potential noncompliance. However, IRS was unable to provide evidence that it conducts reviews related to hospitals' community benefits because it does not have codes to track such audits.

Contents

Letter		1
	Background	3
	Opportunities Exist to Improve Oversight of Tax-Exempt Hospitals	9
	IRS Could Improve Oversight of PPACA Community Benefit Requirements	19
	Conclusions	26
	Matter for Congressional Consideration	27
	Recommendations for Executive Action	27
	Agency Comments	28
Appendix I	Objectives, Scope, and Methodology	29
Appendix II	Comments from the Internal Revenue Service	31
Appendix III	GAO Contact and Staff Acknowledgments	35
Tables		
	Table 1: Percent of Tax-Exempt Hospitals Reporting Compliance with Patient Protection and Affordable Care Act Requirements, by Tax Year	21
	Table 2: Number of Hospital Organizations with Little to No Community Benefit Spending, Tax Years 2014-2016	21
Figures		
	Figure 1: Requirements for Nonprofit Hospitals to Obtain Federal Tax-Exempt Status	4
	Figure 2: Location of Community Benefit Factors on Internal Revenue Service Form 990, Schedule H	14
	Figure 3: Illustration of Community Benefit Data Reported by Hospital Organizations and Independent Hospitals	18
	Figure 4: Results of Community Benefit Activity Reviews (CBAR), Fiscal Years 2015-2019	20

Abbreviations

CBAR	Community Benefit Activity Review
IRC	Internal Revenue Code
IRS	Internal Revenue Service
PPACA	Patient Protection and Affordable Care Act
SOI	Statistics of Income
TE/GE	Tax Exempt and Government Entities

This is a work of the U.S. government and is not subject to copyright protection in the United States. The published product may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.



September 17, 2020

The Honorable Charles E. Grassley
Chairman
Committee on Finance
United States Senate

The Honorable Kevin Brady
Republican Leader
Committee on Ways and Means
House of Representatives

Slightly more than half of community hospitals in the United States are private, nonprofit organizations.¹ Nonprofit organizations can obtain and maintain a federal tax exemption if they are organized for one or more purposes specified in the Internal Revenue Code (IRC) section 501(c)(3). These purposes could include providing charity or education, and continuously operating in pursuit of those purposes.² The public policy underlying the exemption for charitable organizations is based on the concept that the federal government's loss of tax revenue is offset by relief from the need to appropriate public funds and from benefits resulting from the promotion of general welfare.³

Hospitals are on the front line of our national response to Coronavirus Disease 2019 (COVID-19), with hospitals across the country working at maximum capacity to treat the sick. The Internal Revenue Service (IRS) and the Department of the Treasury (Treasury) have recognized the promotion of health as a charitable purpose and have specified that nonprofit hospitals are eligible for a tax exemption.⁴ IRS has further stated that these hospitals can demonstrate their charitable purpose by providing services that benefit their communities as a whole.

¹American Hospital Association, Fast Facts, accessed July 8, 2020, <https://www.aha.org/statistics/fast-facts-us-hospitals>. Community hospitals exclude nonfederal psychiatric hospitals and other hospitals, including long-term care hospitals and those within an institution.

²26 U.S.C. § 501(c)(3).

³H.R. Rep. No. 75-1860, at 19 (1938).

⁴Rev. Rul. 69-545, 1969-2 C.B. 117.

The Joint Committee on Taxation estimated the total revenue loss from the tax exemption of hospitals at \$12.6 billion in 2002.⁵ Hospitals reported that they provided \$76 billion in community benefits in 2016—the most recent data available at the time of our review.⁶

In a 2008 review, we found that IRS guidance allowed hospitals broad latitude in determining what constitutes community benefit activities. That guidance allowed individual hospitals wide discretion in the determination and measurement of those activities as community benefit for federal purposes.⁷ In addition, we reported that prior studies by us and the Congressional Budget Office indicated that tax-exempt hospitals may not have been defining community benefit in a consistent manner that would enable policymakers to hold them accountable for providing benefits commensurate with their tax exemption.

Since then, Congress and the executive branch have taken steps to bolster IRS's oversight of tax-exempt hospitals. In 2008, IRS began to require tax-exempt hospitals to report information about their community benefits on their annual tax returns. Two years later, the Patient Protection and Affordable Care Act (PPACA) was enacted, which established additional requirements for tax-exempt hospitals to meet to maintain their tax exemption.⁸

You asked us to review IRS's implementation of requirements for tax-exempt hospitals. This report assesses IRS's (1) oversight of how tax-

⁵Congressional Budget Office, *Nonprofit Hospitals and the Provision of Community Benefits* (Washington, D.C.: December 2006) reports the Joint Committee on Taxation estimate. Rosenbaum et. al., "The Value of The Nonprofit Hospital Tax Exemption Was \$24.6 Billion In 2011," *Health Affairs*, vol. 35, no. 7 (2015) estimate the total revenue loss at \$24.6 billion for 2011. Most recently, Ernst and Young estimated the federal cost of the tax exemption in a report for the American Hospitals Association at \$9 billion in 2016. EY, *Estimates of the Federal Revenue Forgone Due to the Tax-Exemption of Non-Profit Hospitals Compared to the Community Benefit They Provide*, 2016 (2019). All of these estimates were based on federal tax rates prior to the implementation of the Tax Cuts and Jobs Act, which reduced corporate tax rates and would likely result in lower estimates of the revenue loss.

⁶For the purposes of this report, we use the term "tax-exempt hospitals" to refer to nongovernmental, nonprofit, and tax-exempt hospitals. Government hospitals—including those at the federal, state, local, and tribal levels—are also exempt from federal taxation.

⁷GAO, *Nonprofit Hospitals: Variation in Standards and Guidance Limits Comparison of How Hospitals Meet Community Benefit Requirements*, GAO-08-880 (Washington, D.C.: Sept. 12, 2008).

⁸26 U.S.C. § 501(r).

exempt hospitals provide community benefits, and (2) enforcement of PPACA requirements related to tax-exempt hospitals.

For both objectives, we reviewed relevant provisions of the Internal Revenue Code, Treasury regulations, revenue rulings, and guidance. We also reviewed IRS policies, procedures, audit plans, and determining factors for reviewing tax-exempt hospitals and interviewed IRS officials. We examined the most recent data available (tax year 2016) from forms hospitals are required to file with IRS documenting the community benefits they provide and their compliance with PPACA. To assess data reliability, we analyzed the content of those data for discrepancies and interviewed IRS officials about their procedures for preparing the data. We determined the data were sufficiently reliable for the purposes of our objectives.

In addition, we interviewed selected stakeholder groups—interest groups representing both tax-exempt and for-profit hospitals, patient advocacy groups, and research organizations—to obtain their views on the clarity and enforcement of the community benefit standard and requirements included in PPACA. We identified these groups using our prior reports on tax-exempt hospitals and the recommendations of representatives of the organizations we interviewed. For more information on our methodology, see appendix I.

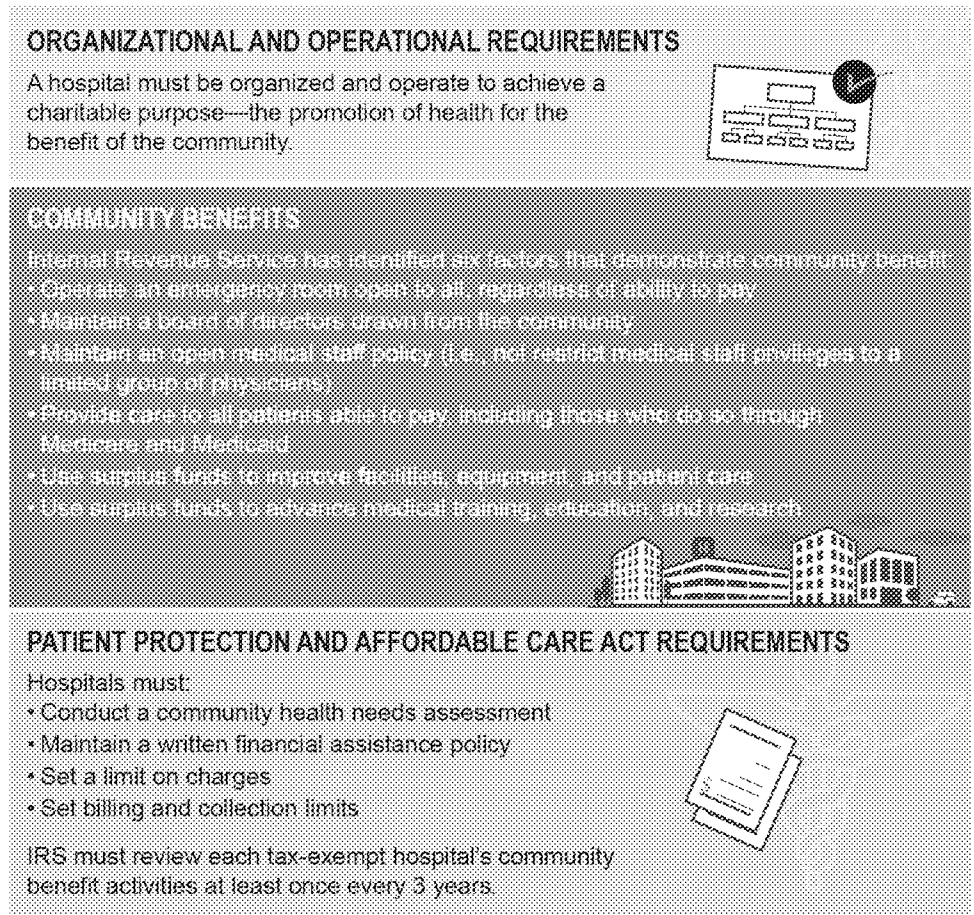
We conducted this performance audit from March 2019 to September 2020 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Background

Requirements for Hospitals to Obtain and Maintain a Federal Tax Exemption

Nonprofit hospitals must satisfy three sets of requirements to obtain and maintain a federal tax exemption (see figure 1).

Figure 1: Requirements for Nonprofit Hospitals to Obtain Federal Tax-Exempt Status



Source: GAO review of relevant laws and regulations. | GAO-20-679

Organizational and Operational Requirements

The IRC requires that all organizations seeking a tax exemption under section 501(c)(3) be organized and operated for one or more purposes, which can be charitable, religious, or educational, among others.⁹ The IRC does not specifically identify hospitals as being eligible for a tax

⁹Section 501 of the IRC covers the majority of these organizations, which include public charities, social welfare organizations, business leagues, and private foundations. Other types of organizations, such as education-oriented programs, farmers' cooperatives, and political organizations, are also wholly or partially tax exempt. 26 U.S.C. §§ 501(c)(3), 521, 527, 529-530.

exemption. However, IRS and federal courts have recognized that the promotion of health for a community's benefit is a charitable purpose.¹⁰

In addition, Treasury regulation provides that an organization will be considered as operating exclusively for one or more exempt purposes if it engages primarily in activities that accomplish those purposes.¹¹ The IRC does not clarify what activities can demonstrate a charitable purpose, but IRS has issued revenue rulings with this type of information.¹² For example, in a 1956 revenue ruling, IRS required tax-exempt hospitals to provide charity care to the extent of their financial abilities.¹³

IRS determined in the ruling that only hospitals that operated for the benefit of those not able to pay, and not exclusively for the benefit of those who were able and expected to pay, could qualify for a tax exemption. Then, in 1959, Treasury updated its regulations to establish that organizations can receive tax-exempt status by demonstrating a charitable purpose such as the promotion of health.¹⁴

Community Benefits

In 1969, 4 years after Congress and the President created Medicare and Medicaid, IRS removed the requirement for tax-exempt hospitals to provide charity care—patient care without charge or at rates below cost—when it issued Revenue Ruling 69-545.¹⁵ The ruling compares the extent to which two hypothetical hospitals satisfy the IRC's requirements for a tax exemption. In making that comparison, the ruling identifies six factors that distinguish how one hospital satisfies the requirements and how the second does not. There is no specific definition of community benefit. These six factors currently serve as the primary examples of community benefits that hospitals can provide to obtain and maintain a tax

¹⁰See *Geisinger Health Plan v. Comm'r*, 985 F.2d 1210, 1216 (3d Cir. 1993) (discussing IRS policy and cases construing exemption provisions for hospitals).

¹¹26 C.F.R. § 1.501(c)(3)-1(c)(1).

¹²A revenue ruling is an official interpretation of the IRC, related statutes, tax treaties or regulations as applied to a specific set of facts.

¹³Rev. Rul. 56-185, 1956-1 C.B. 202. Charity care is generally defined as care provided to patients whom the hospital deems unable to pay all or a portion of their bills.

¹⁴26 C.F.R. § 1.501(c)(3)-1(d)(2).

¹⁵Rev. Rul. 69-545, 1969-2 C.B. 117.

exemption. The factors are commonly referred to as the community benefit standard. IRS describes the six factors on its website:

- **Operate an emergency room open to all, regardless of ability to pay.** A hospital that does not operate a full-time emergency room may not be fulfilling the community's need for emergency health care. If that emergency room is not open to everyone regardless of ability to pay, the hospital may not be serving a significant segment of the community.¹⁶
- **Maintain a board of directors drawn from the community.** A hospital board of directors comprised of independent civic leaders helps to ensure that the hospital serves public, rather than private, interests, and therefore operates for the benefit of the community.
- **Maintain an open medical staff policy (i.e., not restrict medical staff privileges to a limited group of physicians).** A hospital that restricts its medical staff privileges to a limited group of physicians is likely to be operating for the private benefit of the staff physicians rather than for the public interest.
- **Provide care to all patients able to pay, including those who do so through Medicare and Medicaid.** A hospital that restricts admissions to patients of staff members, or otherwise discriminates against patients with the ability to pay for nonemergency services, is not operating for the benefit of the community.
- **Use surplus funds to (1) improve facilities, equipment, and patient care; and (2) advance medical training, education, and research.** The use of surplus funds for these purposes demonstrates that a hospital is promoting the health of the community.¹⁷

The standard makes clear that the factors are examples of ways in which hospitals can demonstrate community benefits. The standard states that a hospital need not meet all of the factors to qualify for a tax exemption.

¹⁶IRS Revenue Ruling 83-157 established that if a state health planning agency determined that additional emergency facilities would be unnecessary and duplicative, or if the hospital offers medical care limited to special conditions unlikely to necessitate emergency care, such as eye or cancer hospitals, then the fact that a hospital organization does not operate an emergency room will not, by itself, disqualify it from a tax exemption. Rev. Rul. 83-157, 1983-2 C.B. 94.

¹⁷IRS, *Charitable Hospitals – General Requirements for Tax-Exemption Under Section 501(c)(3)*, accessed April 30, 2020.
<https://www.irs.gov/charities-non-profits/charitable-hospitals-general-requirements-for-tax-exemption-under-section-501c3>.

Patient Protection and
Affordable Care Act
Requirements

The absence of any one factor, or the presence of others, may not necessarily be conclusive of the hospital's community benefits. IRS says that though a hospital is no longer required to provide charity care it considers doing so to be a significant factor indicating community benefit. Furthermore, IRS considers all of a hospital's facts and circumstances relevant when determining whether a hospital's community benefits are sufficient to warrant a tax exemption.

The Patient Protection and Affordable Care Act (PPACA) established four additional requirements that tax-exempt hospitals must meet to maintain a tax exemption.¹⁸

- **Conduct a community health needs assessment.** Every 3 years, each tax-exempt hospital must identify the community's health needs and develop an implementation plan for how it will address those needs.¹⁹
- **Maintain a written financial assistance policy.** Each tax-exempt hospital must publish a written policy that identifies who can qualify for financial assistance for medical services, how the hospital calculates costs for those services, and the actions the hospital will take in the event of nonpayment.
- **Set a limit on charges.** A tax-exempt hospital cannot charge individuals eligible for financial assistance more for medical services than they do patients with insurance.
- **Set billing and collection limits.** A tax-exempt hospital may not take extraordinary collection actions against an individual, such as filing a lawsuit, before the hospital determines whether that individual is eligible for financial assistance.

In addition, the law established a new requirement for IRS to review the community benefit activities of each tax-exempt hospital at least once every 3 years.²⁰

IRS Oversight of Tax-
Exempt Hospitals

The Tax Exempt and Government Entities (TE/GE) division within IRS oversees tax-exempt hospitals. TE/GE officials said they use two primary

¹⁸26 U.S.C. § 501(r).

¹⁹PPACA establishes that a tax-exempt hospital that does not meet the community health needs assessment requirement must pay an excise tax. 26 U.S.C. § 4959.

²⁰PPACA, Pub. L. No. 111-148, title IX, subtitle A, § 9007(c), 129 Stat. 119, 857 (2010).

Form 990, Schedule H

methods to collect information about tax-exempt hospitals and enforce their compliance with applicable law and guidance: (1) annual tax forms and (2) regular reviews of hospitals' community benefit activities.

Certain tax-exempt nonprofit organizations, including hospitals, are required to file Form 990 annually.²¹ The form requires organizations to report information including

- employees, governance, and compensation;
- revenue and expenses;
- assets and liabilities;
- employment tax compliance; and
- specific organizational issues, such as lobbying by charities and private foundations.

In addition, a tax-exempt hospital is required to file Schedule H, titled "Hospitals," with its Form 990 annually.²² TE/GE uses Schedule H to collect information on the activities and policies of tax-exempt hospital organizations and the hospital facilities and other nonhospital health care facilities they operated during the tax year.²³ Specifically, the schedule requires tax-exempt hospitals to report information on their

- community benefits provided;
- community building activities;
- bad debt, Medicare costs, and collection practices;
- management structure;
- facilities, including how they met PPACA requirements for each; and

²¹26 U.S.C. § 6033.

²²In certain circumstances, tax-exempt hospitals may be required to file other schedules with their Form 990, in addition to Schedule H.

²³IRS defines a hospital organization as an entity that operated at least one hospital facility during a tax year. A hospital facility is an entity that is required to be licensed, registered, or similarly recognized by a state as a hospital. Nonhospital health care facilities may include, but are not limited to, rehabilitation and other outpatient clinics, mobile clinics, and skilled nursing facilities.

Community Benefit Activity Reviews

- supplemental information, including how they promote health.²⁴

TE/GE conducts Community Benefit Activity Reviews (CBAR) to meet the PPACA requirement that it review each tax-exempt hospital's community benefit activities at least once every 3 years. TE/GE states that these reviews determine if tax-exempt hospitals are in compliance with both the community benefit standard and PPACA requirements. TE/GE revenue agents, who conduct the CBARs, may refer a hospital for audit if they determine there is potential noncompliance with either the community benefit standard or PPACA requirements.²⁵

Opportunities Exist to Improve Oversight of Tax-Exempt Hospitals

Congress Could Clarify the IRC to Improve Oversight of Tax-Exempt Hospitals

Congress has taken actions that convey an expectation that hospitals, in exchange for a tax exemption, should provide services and activities that benefit the immediate communities in which they operate. Specifically, in PPACA Congress required tax-exempt hospitals to identify each hospital's community's health needs indicating an expectation that hospitals provide benefits to the immediate community.²⁶ Hospitals that fail to do so must pay an excise tax. However, a broad range of activities fall within the IRC's requirement for a tax exemption for charitable organizations, making it challenging to effectively ensure that the community benefits hospitals provide justify their tax exemption.

Ensuring hospitals are able to meet community health needs is especially important at this time, as the COVID-19 response strains many of the nation's public resources. Congress has provided billions in direct funding to hospitals to help prevent, prepare for, and respond to the coronavirus.

²⁴IRS defines community building activities as activities that improve or protect a community's health and safety.

²⁵An IRS audit is an examination of an organization's or individual's accounts and financial information to ensure information is reported correctly according to the tax laws and to verify the reported amount of tax is correct.

²⁶26 U.S.C. § 501(r).

In doing so, however, Congress set explicit expectations for the use of that funding.²⁷

By contrast, IRS does not have authority to define specific types of services and activities that a hospital must undertake to qualify for a tax exemption. Instead, it provides guidance on the types of activities that can demonstrate community benefits. Some health care industry stakeholders we spoke to told us that IRS's community benefit standard does not ensure that the community benefits that tax-exempt hospitals provide justify their tax exemptions. They explained that the standard only provides examples and does not establish requirements or expectations of services and activities that can demonstrate a hospital's community benefits. They also identified the following specific issues:

- **Hospitals could address some of the standard's factors in ways that do not benefit surrounding communities.** Patient advocate organizations and researchers we spoke to told us, for example, that a hospital could use its surplus funds to conduct research that does not specifically benefit the individual community in which the hospital is located—though it may benefit the healthcare industry as a whole. Similarly, a hospital could use surplus funds to build a new facility, such as a state-of-the-art cancer treatment center, that primarily benefits affluent, insured patients. Such an activity would not necessarily benefit individuals across the community in which the hospital is located.
- **Some of the standard's factors may have lost relevance.** Some factors in the community benefit standard may no longer be relevant for distinguishing between nonprofit and for-profit hospitals. In 2005, the Commissioner of Internal Revenue told Congress that some community benefit factors, such as maintaining an open medical staff policy and accepting patients on Medicare and Medicaid, are now common features of all hospitals.²⁸ Additionally, the Emergency Medical Treatment and Active Labor Act, signed into law in 1986,

²⁷Specifically, the funds were to be used for certain types of activities, such as building temporary structures and emergency operation centers, buying medical supplies and equipment including personal protective equipment and testing supplies, and retrofitting facilities. See, for example, Coronavirus Aid Relief, and Economic Security Act, Pub. L. No. 116-136, division B, title VIII, 134 Stat. 281 (2020).

²⁸*The Tax-exempt Hospitals Sector before the Committee on Ways and Means U.S. House of Representatives*, 109th Cong. 8-18, (2005) (statement of Mark W. Everson, Commissioner of Internal Revenue).

requires that all hospitals that operate emergency rooms provide emergency treatment to all, regardless of ability to pay.²⁹ As a result, these standards may be a less useful gauge for measuring community benefit than they once were.

- **The standard does not identify some factors that can demonstrate substantial community benefit.** Stakeholders told us, for example, that the standard does not clearly identify that a hospital's spending on social determinants of health is an example of community benefits. Social determinants of health are economic or social conditions, such as the quality of one's housing, that influence health outcomes within groups or individuals. IRS's instructions to hospitals for completing their annual tax returns state that some spending in this area can be claimed as community benefit. However, stakeholders told us these instructions are not clear. As a result, tax-exempt hospitals might underinvest in such activities, which reduces the benefit to communities.

Given this ambiguity, a hospital could, in theory, maintain a tax exemption by operating an emergency room open to all and accepting patients on Medicare or Medicaid, which are common among hospitals, while spending little to no money on charity care or other community benefit activities.

However, other stakeholders, such as representatives of tax-exempt hospitals, told us that current law and the community benefit standard offer hospitals needed flexibility in demonstrating community benefits. They said community health needs vary substantially across the country. Therefore, community benefits can vary substantially from place to place. For example, a hospital located in a remote rural community may be the only hospital within hundreds of miles. Its primary benefit may be the fact that it exists to serve the community. Such a reason could be sufficient to justify its tax exemption.

IRS states that it reviews hospitals' services and activities to ensure that they are providing community benefits that justify their tax exemptions. For example, one of the purposes of IRS's Community Benefit Activity Reviews is to enforce tax-exempt hospitals' compliance with the requirement that they provide community benefits. However, IRS officials told us that they could not identify whether any tax-exempt hospitals were referred to its audit division during the period from fiscal years 2015

²⁹Emergency Medical Treatment and Active Labor Act, Pub. L. No. 99-272, title IX, § 9121(b), 100 Stat 164 (1986).

through 2019 for potentially providing insufficient community benefits because, as discussed later, the agency does not track this information. Furthermore, IRS officials told us that the agency has not revoked a hospital's tax-exempt status for failing to provide sufficient community benefits in the last 10 years.

We have previously reported that criteria for a good tax system include transparency and administrability.³⁰ A transparent tax system is one that taxpayers are able to understand. Administrable tax systems allow the government to collect taxes as cost effectively as possible. The way the tax system is structured by Congress can affect how it is administered, and this can affect compliance. A nontransparent tax system is challenging to administer because tax administrators will have difficulty consistently applying the law to taxpayers in similar situations.

As we have previously reported, the IRC, and IRS's implementation of it, provides tax-exempt hospitals with broad latitude to determine the nature and amount of community benefits they provide.³¹ However, these broad requirements create challenges for IRS in administering tax law. IRS, in its 2009 report on hospital compliance, stated that the community benefit standard is difficult to administer.³² The lack of clarity makes it difficult for IRS to ensure that hospitals receiving a tax exemption undertake services and activities that provide benefits to the communities in which they operate.

Additional clarity in the IRC about specific services and activities Congress believes would provide sufficient community benefits could improve IRS's ability to oversee tax-exempt hospitals.

³⁰GAO, *Understanding the Tax Reform Debate: Background, Criteria, and Questions*, GAO-05-1009SP (Washington, D.C.: September 2005).

³¹GAO, *Nonprofit Hospitals: Variation in Standards and Guidance Limits Comparison of How Hospitals Meet Community Benefit Requirements*, GAO-08-880 (Washington, D.C.: Sept. 12, 2008).

³²IRS, "Tax Exempt and Government Entities," *IRS Exempt Organizations Hospital Study Executive Summary of Final Report* (February 2009) https://www.irs.gov/pub/irs-tege/execsum_hospprojrept.pdf.

IRS Could Improve the Transparency of Community Benefit Data

IRS Does Not Collect Consistent Data on Community Benefits

IRS requires a tax-exempt hospital to file a Schedule H with its Form 990 annually to provide the public with information on their policies, activities, and the community benefits that their facilities provide. IRS has stated that a tax-exempt organization's Form 990, along with its schedules, can be the primary or sole source of information the public uses to understand a tax-exempt organization's operations, such as the community benefits a hospital provides. The publicly available data are also intended to enable researchers and the broader public to better understand the level of community benefits that these hospitals provide.

However, Form 990, Schedule H solicits information inconsistently, resulting in a lack of clarity about the community benefits hospitals provide. The schedule includes questions intended to capture information on each of the six factors of the community benefit standard. These questions are located on different parts of the schedule and hospitals are instructed to address them in different ways (see figure 2).

Figure 2: Location of Community Benefit Factors on Internal Revenue Service Form 990, Schedule H

LOCATION ON FORM 990, SCHEDULE H

**COMMUNITY
BENEFIT
FACTORS**

Part I (aggregated)

7 Financial Assistance and Certain Other Community Benefits at Cost						
Financial Assistance and Means-Tested Government Programs	(a) Number of activities or programs (optional)	(b) Persons served (optional)	(c) Total community benefit expense	(d) Direct offsetting revenue	(e) Net community benefit expense	(f) Percent of total expense
a Financial Assistance at cost (from Worksheet 1)						
b Medicaid (from Worksheet 3, column a)						
c Costs of other means-tested government programs (from Worksheet 3, column b)						
d Total. Financial Assistance and Means-Tested Government Programs:						
Other Benefits						
e Community health improvement services and community benefit operations (from Worksheet 4)						
f Health professions education (from Worksheet 5)						
g Subsidized health services (from Worksheet 6)						
h Research (from Worksheet 7)						
i Cash and in-kind contributions for community benefit (from Worksheet 8)						
j Total. Other Benefits:						
k Total. Add lines 7d and 7j:						

Provide care to all patients able to pay, including those who do so through Medicare and Medicaid.

Use surplus funds to advance medical training, education, and research.

Part V (by facility)

Policy Relating to Emergency Medical Care				
21	Did the hospital facility have in place during the tax year a written policy relating to emergency medical care that required the hospital facility to provide, without discrimination, care for emergency medical conditions to individuals regardless of their eligibility under the hospital facility's financial assistance policy?	21		
If "No," indicate why:				
a	<input type="checkbox"/> The hospital facility did not provide care for any emergency medical conditions			
b	<input type="checkbox"/> The hospital facility's policy was not in writing			
c	<input type="checkbox"/> The hospital facility limited who was eligible to receive care for emergency medical conditions (describe in Section C)			
d	<input type="checkbox"/> Other (describe in Section C)			

Operate an emergency room open to all regardless of ability to pay.

Part VI (aggregated, open question)

5	Promotion of community health. Provide any other information important to describing how the organization's hospital facilities or other health care facilities further its exempt purpose by promoting the health of the community (e.g., open medical staff, community board, use of surplus funds, etc.).			
---	--	--	--	--

Not restrict medical staff privileges to a limited group of physicians (e.g., maintain an open medical staff policy).
 Maintain a board of directors drawn from the community.
 Use surplus funds to improve facilities, equipment, and patient care.

Source: GAO. | GAO-20-679

For three of the factors in the community benefit standard—provide care to all patients able to pay; use surplus funds to advance medical training, education, and research; and operate an emergency room open to all—IRS explicitly directs tax-exempt hospitals to report on Form 990, Schedule H the extent to which they have addressed them. For example, regarding the factors on surplus funds, IRS directs hospitals to identify the specific costs they incur by providing health education and medical research.

However, IRS does not explicitly direct tax-exempt hospitals to report on the other three community benefit factors on Form 990, Schedule H—maintain an open medical staff policy; have a board of directors drawn from the community; and use surplus funds to improve facilities, equipment, and patient care. Rather, IRS asks hospitals to describe narratively additional information important to understanding the full scope of the community benefits they provide. In doing so, those three factors are suggested only as examples they could use in their description.

This reporting structure can affect the comprehensiveness and consistency with which a tax-exempt hospital presents its community benefits:

- It creates uncertainty about where information on certain types of activities should be reported. For example, hospital representatives we spoke to said some of their members are confused about whether they should report information under community benefits or community building activities.
- It generally precludes tax-exempt hospitals from specifically identifying the amount of surplus funds used to improve facilities, equipment, and patient care. According to IRS officials, those costs are located on the main Form 990. However, the information to which IRS referred—functional expenses and the hospital’s overall balance sheet—do not break out costs attributable to the improvement of facilities, equipment, or patient care.
- It could result in potentially incomplete information on how hospitals are providing community benefits. For example, our analysis of hospitals’ Form 990, Schedule H filings for tax years 2015 through 2018 identified hospitals that described their open medical staff policy and how their board of directors was made up of members from the community. These hospitals also provided numerous examples of how they used surplus funds to improve their facilities and patient

care. Conversely, we identified other hospitals that did not address whether they had an open medical staff policy, a board of directors drawn from the community, or the use of surplus funds to improve patient care, facilities, or equipment.

- It limits the comprehensiveness of publicly available data. IRS releases quantitative, machine-readable files on the community benefits reported by each tax-exempt hospital on Form 990, Schedule H.³³ However, the files do not contain any community benefit information that hospitals describe narratively.

We have previously reported on the importance of transparently reporting open government data.³⁴ Specifically, we reported that data designed to be open to the public should be provided in useful formats, such as ensuring users have detailed and disaggregated data. In addition, agencies should facilitate data discovery for all users, such as utilizing central data repositories and catalogues to help users easily find the data they seek.

IRS officials told us that the current Form 990, Schedule H is sufficient, as it allows tax-exempt hospitals to fully identify and describe their community benefit activities throughout the form. IRS officials also stated that Form 990, Schedule H does not specifically ask about surplus costs for facilities, equipment, and patient care because that information is already reported by hospitals in another part of Form 990. IRS officials said they believe that revising the form to include more specific information would put undue burden on hospitals. According to IRS officials, hospitals often use prior forms to update their information, and changes to a form could make it challenging to do so. Additionally, officials said that new forms take time to design and vet.

However, the Form 990, Schedule H, as it is currently structured, does not enable tax-exempt hospitals to demonstrate clearly for the public the extent to which they provide community benefits. Furthermore, the reliance on an optional narrative answer for some factors results in data that potentially provide an incomplete picture of a hospital's community benefits. A revised Form 990, Schedule H that enables tax-exempt hospitals to present community benefit information clearly, consistently, and comprehensively could help IRS, Congress, and the broader public

³³Forms 990 are disclosable to the public and can be requested by submitting Form 4506-A.

³⁴GAO, *Treasury Could Better Align USAspending.gov with Key Practices and Search Requirements*, GAO-19-72 (Washington, D.C.: Dec. 13, 2018).

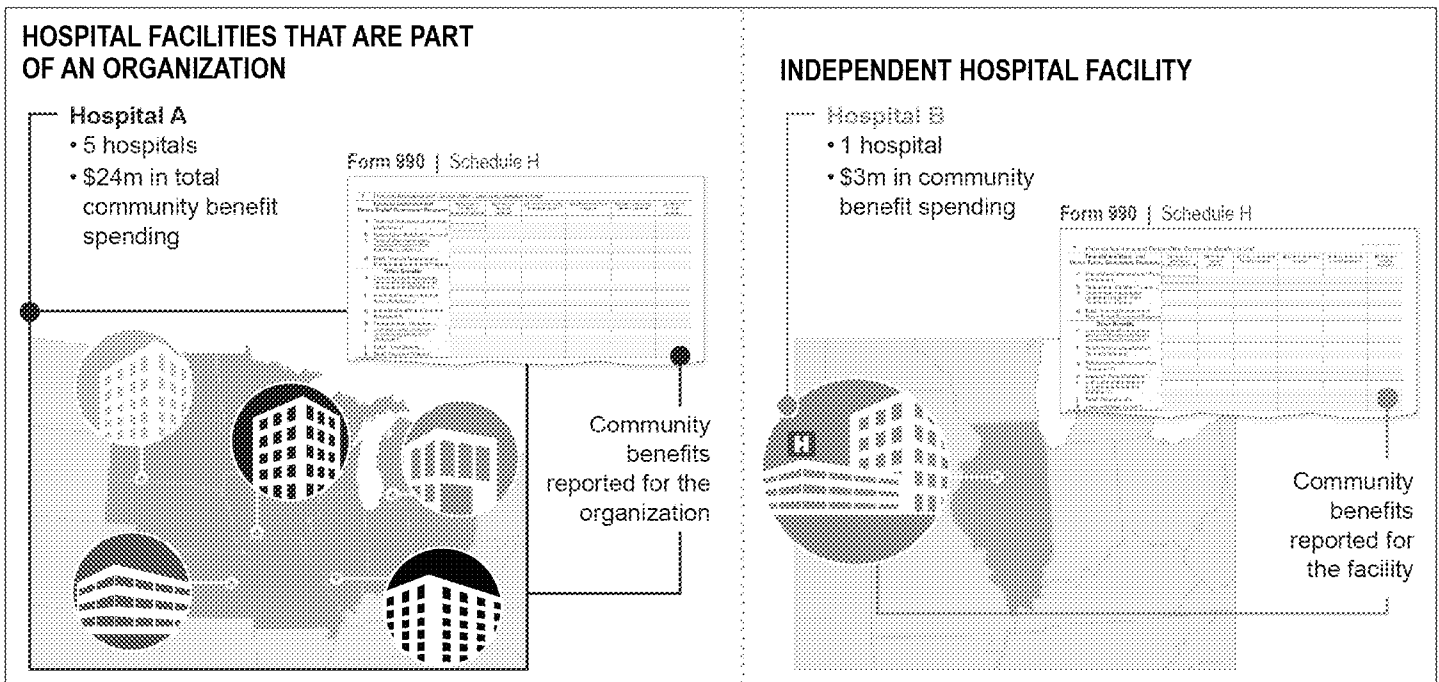
IRS Does Not Require
Hospitals to Report Community
Benefit Expenses by Facility

better understand the full scope of the community benefits a hospital provides and whether the benefits sufficiently justify a tax exemption.

Form 990, Schedule H directs tax-exempt hospitals to report their community benefit expenses at the hospital organization level rather than at the facility level. Therefore, hospital organizations report community benefits in the aggregate for all of their facilities. For example, a hospital organization reports the amount of charity care it provides and its costs for medical training, education, and research for all of its facilities as a whole, not for each facility.

In tax year 2016, 46 percent of hospital facilities were part of a hospital organization, and therefore those facilities' community benefit expenses were reported as part of the organization as a whole. For example, a hospital organization with five facilities could report \$24 million in community benefit expenses on its Form 990, Schedule H. In that case, it would not be transparent how much each facility contributed to the total. Two of the facilities could contribute \$12 million each in community benefit expenses, while the other three contribute none. Alternatively, the community benefit expenses reported by independent hospital facilities would be transparent (see figure 3).

Figure 3. Illustration of Community Benefit Data Reported by Hospital Organizations and Independent Hospitals



Source: GAO. | GAO-20-679

Congress and the executive branch have previously signaled the importance of tax-exempt hospitals reporting information on hospital facilities by including in the Patient Protection and Affordable Care Act (PPACA) requirements for those hospitals to create community health needs assessments and associated implementation plans for each of their facilities. Furthermore, *Standards for Internal Control in the Federal Government* states that an organization should use quality information to achieve its objectives and communicate that information externally. These standards also call for management to design and implement internal controls within programs based on the related benefits and costs.³⁵

IRS officials told us that the agency requires a tax-exempt hospital to report its community benefit expenses on their Form 990, Schedule H at the organization level because the IRC provides a tax exemption at the organizational level, not the facility level. IRS directs hospital organizations to report information about PPACA requirements at the

³⁵GAO, *Standards for Internal Control in the Federal Government*, GAO-14-704G (Washington, D.C.: September 2014).

facility level on the form because it is legally required.³⁶ Because hospitals are now required by PPACA to provide information at the facility level, reporting all information at the facility level could potentially reduce the burden of aggregating data to the organization level. However, IRS has not assessed the benefits and costs of requiring hospitals to report their community benefit expenses at the facility level. Without doing so, IRS may be missing an opportunity to collect information that would more clearly and transparently demonstrate the benefits tax-exempt hospitals provide to the communities in which they operate. This information, in turn, would allow Congress, IRS, and the public to weigh the costs and benefits of the hospital's tax exemption.

IRS Could Improve Oversight of PPACA Community Benefit Requirements

IRS Verifies Hospitals' Self-Reported Compliance with the Four PPACA Requirements

IRS requires hospitals to self-report compliance with all four PPACA requirements on Form 990, Schedule H, Part V.³⁷ Hospitals must answer a series of yes or no questions for each of the four PPACA requirements. For example, they must answer the question, "During the tax year or either of the two immediately preceding tax years did the hospital facility conduct a community health needs assessment?"

The Tax Exempt and Government Entities (TE/GE) division of IRS verifies many aspects of the hospitals' reports during its triennial Community Benefit Activity Reviews (CBAR). Using guidance called the ACA Desk Guide, TE/GE revenue agents answer a list of questions to review how hospitals comply with the PPACA requirements. These questions are primarily tied to the questions on the Form 990, Schedule H, Part V, but also include follow-up questions that depend on hospitals' responses. For example, the guide asks revenue agents to verify that the hospital

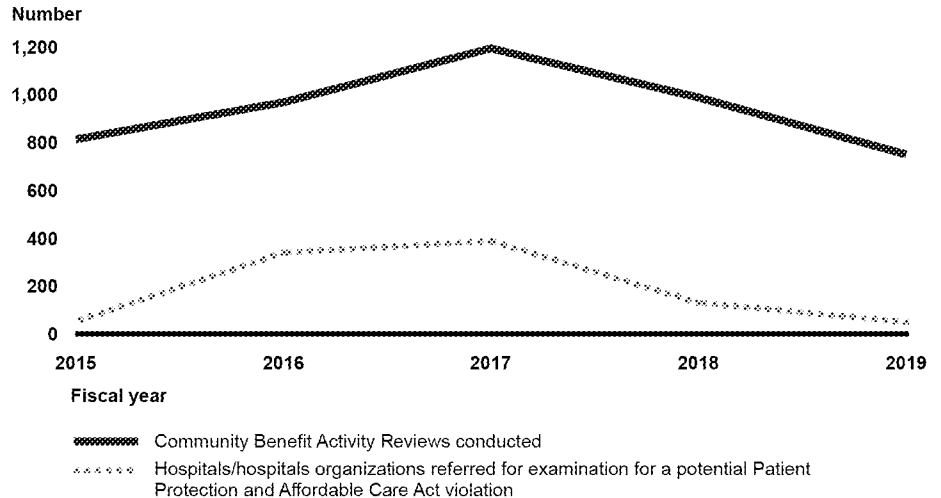
³⁶26 U.S.C. § 501(r).

³⁷Those requirements are to (1) conduct a community health needs assessment every 3 years and develop an implementation plan for how it will address those needs; (2) publish a written financial assistance policy; (3) not charge individuals eligible for financial assistance more for medical services than they do patients with insurance; and (4) not take extraordinary collection actions against an individual before the hospital determines whether that individual is eligible for financial assistance.

conducted a community health needs assessment, but also includes detailed questions on how it was conducted.

TE/GE revenue agents determine a hospital organization’s compliance by collecting information from a range of public sources, but they do not contact the organization itself when conducting the CBAR.³⁸ If the revenue agent cannot verify a hospital’s compliance with all the PPACA requirements, the hospital is to be referred for audit.³⁹ Over the period from fiscal years 2015 through 2019, TE/GE conducted more than 4,700 CBARs and referred almost 1,000 tax-exempt hospitals for audit because they identified a potential PPACA violation (see figure 4). Figure 4 shows that there was an initial rise in the number of CBARs conducted. According to TE/GE officials, changes in the number of CBARs conducted are due to several factors, including openings, closings, and mergers, as well as refinements in selection criteria.

Figure 4: Results of Community Benefit Activity Reviews (CBAR), Fiscal Years 2015-2019



Source: GAO analysis of Internal Revenue Service data. | GAO-20-679

³⁸Public sources revenue agents are instructed to use include Google, Medicare, and state and hospital websites.

³⁹Revenue agents may also refer a hospital for a compliance check. A compliance check, while less detailed than an audit, consists of a more thorough review than the CBAR and is used to determine if the hospital is adhering to recordkeeping and reporting requirements and if its activities are consistent with its stated tax-exempt purpose. It does not relate to determining tax liability or verifying the hospital’s responses coincide with its records.

Note: Hospitals can demonstrate their charitable purpose by providing services that benefit their communities.

Both TE/GE enforcement of the PPACA requirements and an increase in self-reported compliance may have contributed to the recent decline in referrals. As part of that enforcement, TE/GE has issued more than 300 written advisories and levied excise taxes on 40 noncompliant hospitals from CBARs and resulting audits from fiscal years 2015 through 2018. In addition, representatives from tax-exempt hospitals told us that TE/GE's reviews have been thorough. For example, they cited instances in which IRS required their hospitals to provide additional documentation to support information they reported on the Form 990, Schedule H. Our analysis of Form 990, Schedule H data shows that hospitals' self-reported compliance increased each year from fiscal years 2014 to 2016 (see table 1).

Table 1: Percent of Tax-Exempt Hospitals Reporting Compliance with Patient Protection and Affordable Care Act Requirements, by Tax Year

	2014	2015	2016
Conduct Community Health Needs Assessments Every 3 Years	85	89	90
Maintain a Financial Assistance Policy	16	33	62
Set a Limit on Charges	91	92	94
Set Billing and Collections Limits	56	70	77

Source: GAO analysis of Internal Revenue Service data. | GAO-20-679

TE/GE Could Improve Its Review of Hospitals' Community Benefit Activities

IRS referred almost 1,000 hospitals to its audit division for potential PPACA violations from fiscal years 2015 through 2019. However, IRS could not identify whether any of these referrals related to community benefits. Our analysis of schedule H data indicates there were hospitals that could have been at risk for noncompliance with the community benefit standard during a similar period (see table 2).

Table 2: Number of Hospital Organizations with Little to No Community Benefit Spending, Tax Years 2014-2016

	2014	2015	2016
No financial assistance	64	68	48
No community benefit spending	48	45	30
Less than 1 percent community benefit spending	142	137	108

Source: GAO analysis of Internal Revenue Service data. | GAO-20-679

Note: Financial assistance includes financial aid (i.e., charity care), Medicaid, and other means-tested government programs. The calculation of community benefit corrects for hospitals that reported negative spending values due to excess off-setting revenues, such as grants or Medicaid reimbursements.

Specifically, we identified 30 hospitals in 2016 that reported no spending on community benefits, as shown in table 2.⁴⁰ TE/GE states that it sends back forms that are materially incomplete and requests that hospitals complete the missing information; however, we found that some of the hospitals left the required community benefit section of Form 990, Schedule H blank. These hospitals may have actually spent funds on community benefit activities, but did not complete the form. Other hospitals reported spending amounts that were approximately 0 percent of expenses.⁴¹

In addition to requiring IRS to review hospitals' compliance with the PPACA requirements, PPACA required IRS to review information about hospitals' community benefit activities at least every 3 years. However, TE/GE does not have a well-documented process, such as clear instructions on referring hospitals for audit during its triennial reviews or automated queries to identify hospitals at risk for noncompliance with the community benefit standard. TE/GE also does not have a way to track audits related to tax-exempt hospitals' community benefit activities.

Referring Hospitals for Audit

TE/GE revenue agents are to follow the ACA Desk Guide to conduct their triennial reviews of hospitals' community benefit activities. The ACA Desk Guide states that the purpose of the CBAR is to determine if hospitals are compliant with the community benefit standard as well as the additional requirements for tax-exempt hospitals in PPACA. It also states that the CBAR survey addresses "501(r)" questions, which relate specifically to PPACA requirements. However, the instructions for conducting CBARs do not identify how or when a revenue agent should refer a hospital at risk of providing insufficient community benefits for audit.

The instructions for conducting CBARs provide general guidance for the types of issues that a revenue agent may identify. However, this guidance does not specifically include potential noncompliance with the community

⁴⁰The calculation of community benefit spending included direct offsetting revenue, such as grants or Medicaid reimbursements that could result in zero net spending.

⁴¹IRS agents in the Statistics of Income group in the Research Applied Analytics and Statistics Division correct some of the Form 990, Schedule H data for obvious errors before posting the public files onto IRS's website. However, those changes do not extend to the forms themselves that TE/GE officials would review in a CBAR.

benefit standard. While there are specific questions that address the community benefit factors, there is no direction on when a hospital should be referred for audit if the revenue agent is unable to verify the factor. According to TE/GE officials, during the CBAR, revenue agents may refer a hospital for audit because it provides insufficient community benefits, but the agents do not have guidance that would enable them to make this determination consistently.

The ACA Desk Guide has specific questions related to community benefits. These include questions concerning whether the hospital:

- maintains a financial assistance policy;
- budgets amounts for free or discounted care;
- operates an emergency room, open to all without regard for ability to pay;
- maintains medical staff at each facility that is open or if it is restricted to a certain group of physicians;
- maintains a governing body that is composed of a majority of members of the community; or
- uses surplus funds to improve patient care, expand facilities, or advance medical education and research.

TE/GE officials stated that these questions ask the revenue agent to verify that the organization demonstrates these community benefit factors.

However, those questions simply ask revenue agents if they can verify that the hospitals reported in the affirmative; during the CBAR review, the revenue agent is not able to request any more detailed information on the amount of care provided or activities conducted. During an audit, an examiner can request additional information. Further, the guide does not indicate which or how many of those questions that the revenue agent is unable to verify justifies referring the hospital for audit.

According to IRS officials, the decision to refer the hospital for an audit is based on the facts and circumstances of each review conducted. Updating the instructions for conducting the CBARs to more clearly direct revenue agents on how to identify and recommend for audit tax-exempt hospitals at risk of providing insufficient community benefits could help ensure that revenue agents are identifying hospitals that are potentially noncompliant with the community benefit standard.

According to TE/GE officials, the CBAR reviews are IRS's primary method used for overseeing hospitals' tax-exempt status. TE/GE will accept referrals for audit independent of the CBAR, such as those resulting from audit selections based on the Form 990 that did not relate specifically to hospitals. However, it does not have a system outside the CBAR for identifying hospitals at risk for providing insufficient community benefits. According to TE/GE officials, TE/GE only reviews Form 990 data during the CBAR.

TE/GE uses models to analyze data from the Form 990 that most tax-exempt organizations file. Those models include questions to identify responses on returns that may indicate noncompliance because they do not meet certain criteria or expected values. Although a hospital may be selected for an audit as a result of these automated queries, none of the queries address the community benefit standard or apply to the schedule H that a hospital includes with its Form 990. According to TE/GE officials, hospitals with little to no community benefit expenses would be indicative of potential noncompliance and may warrant an audit. However, as discussed below, TE/GE was unable to provide us evidence that it conducts reviews specifically related to hospitals' community benefits.

TE/GE officials stated that of 37 hospitals that reported zero or negative community benefit spending in tax year 2016, 21 were referred for examination or compliance check as a result of their CBAR reviews.⁴² Six of the hospitals were referred for audit based on CBAR review of the 2016 Form 990. The other 15 referrals were made based on other tax years. In all these cases, the referrals were made as a result of possible issues with the financial assistance policy or community health needs assessment. TE/GE officials said that the other 16 hospitals that reported no spending on community benefits were not referred for audit because they met the PPACA requirements. According to TE/GE officials, these requirements address community benefit issues as they relate to financial assistance policies and community health needs assessments, but they do not necessarily address the facts and circumstances determination of meeting the community benefit standard.

TE/GE officials also said automated queries for selecting hospitals to audit related to the community benefit standard are not needed because

⁴²We provided IRS with a list of 37 hospitals that, based on our review of Form 990, Schedule H data, reported zero or negative net community benefit spending for tax year 2016. This number is larger than the amount reported in table 2, because the values in table 2 correct for the cases for which hospitals reported negative spending in Medicaid.

Tracking Review of Hospitals’
Community Benefit Activities

the entire population of hospitals is reviewed at least every 3 years during the CBAR. TE/GE officials stated that, through these reviews, IRS collects information necessary for its role in administering the tax law. Risk-based audit selection methods, such as automated queries to flag values below certain percentages, could help better identify a hospital at risk for noncompliance rather than relying solely on the CBAR reviews.

PPACA requires IRS to review, at least once every 3 years, the community benefit activities of each hospital organization subject to the PPACA provisions.⁴³ A well-documented and consistently implemented process for identifying hospitals at risk for noncompliance with the community benefit standard would help IRS ensure it is effectively reviewing hospitals’ community benefit activities.

TE/GE also does not have a way to track audits related to tax-exempt hospitals’ community benefit activities. Specifically, it does not have a way to determine if hospitals are being selected for audit for potential noncompliance related to community benefits during a CBAR. Hospitals can also be referred for audit through channels outside of the CBAR and not all cases referred for audit are accepted for audit. IRS also does not have a method to track how many hospitals have actually been audited based on potential noncompliance with community benefits.

According to TE/GE, it uses audit issue codes that differentiate between PPACA-related noncompliance and other noncompliance. Revenue agents and examiners use the same set of codes for PPACA-related issues—revenue agents use them to indicate a deficiency with PPACA found during a CBAR review and examiners use them to identify issues found while conducting an audit. TE/GE designates these codes “ACA,” and they include seven specific categories related to the individual PPACA requirements. According to TE/GE, issues unrelated to the PPACA requirements, such as having unrelated business income, are labeled “non-ACA.” However, there are no codes related to potential noncompliance with the community benefit standard.

According to IRS, from 2016 through 2019, fewer than 10 cases each year were referred to its audit division during the CBAR for an issue not related to PPACA. As stated above, IRS officials told us that they could not identify whether any tax-exempt hospitals were referred to its audit

⁴³PPACA, Pub. L. No. 111-148, title IX, subtitle A, § 9007(c), 129 Stat. 119, 857 (2010).

division from 2016 through 2019 for potentially providing insufficient community benefits.

According to TE/GE officials, if a hospital was referred for audit due to concerns about the community benefits it was providing, a revenue agent would probably use one of the non-ACA codes, such as “operational requirements” or possibly “disqualifying operations.” Officials further stated TE/GE does not have an issue code specifically related to “community benefit” because the community benefit standard relates to qualification for exemption, which has existed since 1969. According to TE/GE officials, the codes they currently use in the CBAR were implemented specifically for tracking issues related to the PPACA provisions identified during CBARs and not for issues with qualification for exemption. They also said it would be inefficient to have codes that identify every type of exemption or operational issue that could be encountered in a case.

Standards for Internal Control in the Federal Government states that agencies should design information systems to obtain and process information to meet the agency’s objectives.⁴⁴ This design allows the agency to effectively monitor its operations and ensure it is meeting its goals.

Similar to how IRS established seven issue codes for the PPACA provisions, establishing codes for the six community benefit factors would allow IRS to track audit activity related to community benefits. These codes for community benefit issues would enable revenue agents to indicate in their respective review summaries that a hospital is being referred for audit based on potential noncompliance with the community benefit standard. In addition, these codes would enable TE/GE to systematically track whether examiners found deficiencies in how hospitals provide community benefits during their audits. This information would enable IRS to demonstrate to Congress and the public that hospitals’ community benefit activities are being reviewed and whether hospitals are complying with the community benefit standard.

Conclusions

Hospitals have been able to receive a tax exemption since their inception. The reasons for that exemption have changed over time from a focus on charity to a focus on providing community benefits. Congress and the

⁴⁴GAO, *Standards for Internal Control in the Federal Government*, GAO-14-704G (Washington, D.C.: September 2014).

executive branch have taken steps to bolster IRS's oversight of tax-exempt hospitals, including establishing additional requirements to better ensure that hospitals are adequately serving their communities. However, the lack of clarity in the law regarding what types of activities hospitals should be engaged in to justify that tax exemption make it challenging for IRS to ensure effective oversight.

IRS has identified factors that demonstrate community benefits, but the agency does not require any one factor to be satisfied and it audits hospitals on a facts and circumstances basis. These factors vary significantly from activities that demonstrate direct benefits, such as using surplus funds for patient care, to broader concepts of community benefit, such as general medical research. The data IRS collects to provide the public with information on the policies, activities, and community benefits that hospital facilities provide are not clear or detailed, reducing both the transparency of hospitals' activities and IRS's ability to effectively verify the benefits hospitals provide to their communities. IRS actively reviews hospitals' compliance with PPACA requirements. However, IRS does not have a well-documented process to ensure or demonstrate it is consistently reviewing the community benefits hospitals provide. By taking steps to improve its oversight of hospitals' tax-exempt status, IRS could provide Congress and the public with confidence that these hospitals are adequately serving their communities.

Matter for Congressional Consideration

Congress should consider specifying in the IRC what services and activities it considers sufficient community benefit. (Matter for Consideration 1)

Recommendations for Executive Action

We are making the following four recommendations to IRS:

The Commissioner of Internal Revenue should update Form 990, including Schedule H and instructions where appropriate to ensure that the information demonstrating the community benefits a hospital is providing is clear and can be easily identified by Congress and the public, including the community benefit factors. (Recommendation 1)

The Commissioner of Internal Revenue should assess the benefits and costs, including the tax law implications, of requiring tax-exempt hospital organizations to report community benefit expenses on Schedule H by individual facility rather than by collective organization and take action, as appropriate. (Recommendation 2)

The Commissioner of Internal Revenue should establish a well-documented process to identify hospitals at risk for noncompliance with the community benefit standard that would ensure hospitals' community benefit activities are being consistently reviewed. (Recommendation 3)

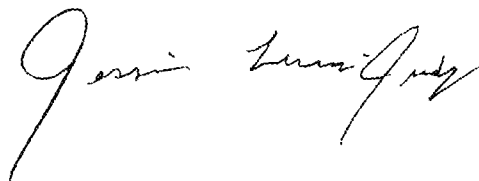
The Commissioner of Internal Revenue should establish specific audit codes for identifying potential noncompliance with the community benefit standard. (Recommendation 4)

Agency Comments

We provided a draft of this report to the Commissioner of Internal Revenue for review and comment. IRS provided written comments, reproduced in appendix II, stating that it agreed with GAO's recommendations. In addition, after a discussion with the IRS Commissioner for Tax Exempt and Government Entities and other senior program officials on August 20, 2020, we clarified two recommendations. First, we clarified Recommendation 1 to include updating instructions where appropriate and to ensure the information addresses the community benefit factors. Second, we clarified Recommendation 2 to include tax law implications as part of assessing the benefits and costs.

As agreed with your offices, unless you publicly announce the contents of this report earlier, we plan no further distribution until 30 days from the report date. At that time, we will send copies of this report to the appropriate congressional committees, Commissioner of Internal Revenue, and other interested parties. In addition, the report will be available at no charge on the GAO website at <http://www.gao.gov>.

If you or your staff have any questions about this report, please contact me at (202) 512-9110 or lucasjudyj@gao.gov. Contact points for our offices of Congressional Relations and Public Affairs are on the last page of this report. GAO staff who made key contributions to this report are listed in appendix III.



Jessica Lucas-Judy
Director, Strategic Issues

Appendix I: Objectives, Scope, and Methodology

This report assesses the Internal Revenue Service's (IRS) (1) oversight of how tax-exempt hospitals provide community benefits, and (2) enforcement of requirements related to tax-exempt hospitals included in the Patient Protection and Affordable Care Act (PPACA).¹

To assess IRS's oversight of how tax-exempt hospitals provide community benefits we reviewed relevant provisions of the Internal Revenue Code, Department of the Treasury regulations, revenue rulings, and guidance. We assessed IRS's oversight efforts against relevant federal internal control standards and our criteria for a good tax system. The relevant internal control principles focus on using and communicating quality information and designing information systems to achieve objectives and respond to risks. The criteria for a good tax system describe principles of transparency and administrability.² We also reviewed IRS policies, procedures, audit plans, and determining factors for reviewing tax-exempt hospitals.

We examined data on community benefit information hospitals report from Forms 990, Schedule H, which hospitals are required to file with IRS documenting the community benefits they provide. Those data were obtained from IRS Statistics of Income (SOI) public microdata files that cover the entire population of tax-exempt hospitals for tax year up to 2016, the most recent year available at the time of our review. While we found instances in which some sections of the form H were incomplete, we determined that the data were sufficiently reliable for addressing our objectives. We made this determination by performing detailed tests for errors or discrepancies and interviewing SOI officials on their procedures for preparing the data.

To assess IRS's enforcement of requirements related to tax-exempt hospitals included in PPACA, we reviewed its primary guidance for Tax Exempt and Government Entities (TE/GE) reviewers—the ACA Desk Guide—to determine what topics their triennial reviews of hospitals cover. We also analyzed TE/GE's audit referral system to determine the steps IRS has in place to enforce hospitals' compliance with the PPACA provisions and the community benefit standard. We compared data on

¹26 U.S.C. § 501(r).

²GAO, *Standards for Internal Control in the Federal Government*, GAO-14-704G (Washington, D.C.: September 2014), and *Understanding the Tax Reform Debate: Background, Criteria, and Questions*, GAO-05-1009SP (Washington, D.C.: September 2005).

Forms 990, Schedule H on hospitals' self-reported compliance with PPACA and their reported community benefits with the triennial review guidance.

In addition, we interviewed selected interest groups representing both tax-exempt and for-profit hospitals to obtain their views on the clarity and enforcement of the community benefit standard and requirements included in PPACA. To identify these groups, we reviewed our past reports on tax-exempt hospitals that yielded an initial list of seven groups. Two of those groups declined to participate. The remaining groups that we spoke to were: American Hospital Association, Association of American Medical Colleges, Catholic Health Association of the United States, Federation of American Hospitals, and Healthcare Financial Management Association. In each of those interviews, we solicited suggestions for other groups to interview that yielded one group not already included. That group also declined to be interviewed.

We also spoke to national patient advocacy groups that represent broad patient interests. We initially identified five groups by soliciting suggestions from our own health specialists and methodologists. Two declined to be interviewed and we solicited suggestions for other groups to contact during our interviews with the three remaining groups identified. Those groups were Families USA, American Public Health Association, and Community Catalyst. They suggested additional groups to interview that were either part of our hospital group interviews, policy group interviews, or outside our scope.

Lastly we conducted a detailed literature review to identify groups that have written on the policy implications related to tax-exempt hospitals. In addition to numerous academic, government, and trade publications, it also yielded nine articles from think tanks. Based on the results of the think tank articles we interviewed the following organizations: Hilltop Institute, Commonwealth Fund, Baker Institute for Public Policy, Brookings Institute, and Urban Institute. The results of our interviews with interest groups, advocacy groups, and think tanks may not represent the views of all groups involved in or with an interest in tax-exempt and for-profit hospitals. However, they illustrate a range of perspectives on these topics.

Appendix II: Comments from the Internal Revenue Service



Deputy Commissioner

DEPARTMENT OF THE TREASURY
INTERNAL REVENUE SERVICE
WASHINGTON, DC 20224

September 2, 2020

Jessica K. Lucas-Judy
Director, Strategic Issues
United States Government Accountability Office
441 G Street, NW
Washington, DC 20548

Dear Ms. Lucas-Judy:

Thank you for the opportunity to review the draft report of the Government Accountability Office, GAO-20-679, entitled ***Tax Administration: Opportunities Exist to Improve Oversight of Hospitals' Tax-Exempt Status*** (job code 103426). We appreciate the analysis in the report, which observes that administration of the tax law in this area presents challenges.

Over the past years, the Internal Revenue Service (IRS) has worked toward continuous improvement of tax administration for tax-exempt hospitals. In 2008, the IRS concluded a study of tax-exempt hospitals and used this information to develop the Form 990, Schedule H. This Schedule was designed, with stakeholder input and consultation, to collect uniform information from tax-exempt hospitals with respect to the community benefit standard described in Revenue Ruling 69-545. In 2010, we implemented key provisions of the Affordable Care Act (ACA) addressing additional, statutorily mandated requirements that must be met by hospitals that are tax-exempt under Section 501(c)(3) of the Internal Revenue Code. As noted in the report, the ACA also requires the IRS to conduct a review of the community benefit activities of every tax-exempt hospital at least every three years. We quickly instituted a program to conduct those reviews, and we continue to review these organizations. We have also engaged in education with respect to the various requirements, and we have seen the positive results of this outreach in the form of increased compliance in the sector.

It is true that there are challenges with respect to administration of the community benefit standard, which requires consideration of "all of the relevant facts and circumstances in each case" and under which "the absence of particular factors [set out in Revenue Ruling 69-545] or the presence of other factors will not necessarily be determinative." See Rev. Rul. 69-545, 1969-2 C.B. 117. Those challenges are a product of the inherent flexibility required under the existing legal guidance to ensure that hospitals' activities are considered in the context of their own communities. This is something the IRS continues to address, and to that end, we have fully considered each of your recommendations. There are inherent advantages and disadvantages of

Appendix II: Comments from the Internal Revenue Service

2

imposing either a standard or a rule, a question of policy on which the IRS must defer to the Department of the Treasury and the Congress.

We appreciated the opportunity to review and comment on the draft report. Responses to your specific recommendations to the IRS are enclosed. If you have questions, please contact me, or a member of your staff may contact Maria D. Hooke, Director, Compliance Planning and Classification, at 214-413-5500.

Sincerely,

Sunita B. Lough

Digitally signed by Sunita B. Lough
Date: 2020.09.02 20:56:49 -0400

Sunita Lough
Deputy Commissioner for
Services and Enforcement

Enclosure

Enclosure

Matter for Congressional Consideration:

Congress should consider specifying in the IRC what services and activities it considers sufficient community benefit.

Comment:

No comment from the IRS.

Recommendation 1: The Commissioner of Internal Revenue should update Form 990, including Schedule H and Instructions where appropriate to ensure that the information demonstrating the community benefits a hospital is providing is clear and can be easily identified by Congress and the public, including the community benefit factors.

Comment:

The IRS agrees with this recommendation. The current Form 990, including Schedule H allows for, but is not limited to, reporting on the six factors identified in Revenue Ruling 69-545. Those six factors were relevant in the application of the law to the particular set of facts in that revenue ruling, but as indicated in the revenue ruling itself, no factor is determinative in every instance, and other factors may be relevant. The current Form 990, including Schedule H thus appropriately provides flexibility to report information on “all of the relevant facts and circumstances” regarding the demonstration of community benefit. Nevertheless, the IRS will review and revise forms and/or instructions to further clarify and allow for clear identification of information that demonstrates community benefit as appropriate.

Recommendation 2:

The Commissioner of Internal Revenue should assess the benefits and costs, including the tax law implications, of requiring tax-exempt hospital organizations to report community benefit expenses on Schedule H by individual facility rather than by collective organization and take action, as appropriate.

Comment:

The IRS agrees with this recommendation. The IRS will assess the benefits and costs of requiring community benefit reporting on a facility-by-facility basis and will take action as appropriate. As part of the assessment, IRS will take into account that under pre-existing law, tax exemption is granted and the community benefit standard is therefore applicable at the organizational level. On the other hand, as enacted by Congress in 2010, the ACA imposes additional requirements codified in Section 501(r) that must be met, specifically, on a facility-by-facility basis.

Recommendation 3:

The Commissioner of Internal Revenue should establish a well-documented process to identify hospitals at risk for non-compliance with the community benefit standard that would ensure hospitals' community benefit activities are being constantly reviewed.

Comment:

The IRS agrees with this recommendation. Where compliance with the existing community benefit standard is based on "all of the relevant facts and circumstances," it would not be feasible to identify uniformly applicable indicators of noncompliance with the standard. However, IRS will update instructions and procedures to improve documentation of the relevant community benefit facts and circumstances considered during a review.

Recommendation 4:

The Commissioner of Internal Revenue should establish specific audit codes for identifying potential non-compliance with the community benefit standard.

Comment:

While potential noncompliance with the community benefit standard is an "operational issue" for which there are existing codes, the IRS continually seeks to improve its processes. Accordingly, the IRS agrees with this recommendation and will establish an Exam Issue Code to more specifically identify potential non-compliance with the community benefit standard.

Appendix III: GAO Contact and Staff Acknowledgments

GAO Contact

Jessica Lucas-Judy, (202) 512-9110 or LucasJudyJ@gao.gov

Staff Acknowledgments

In addition to the contact named above, Sonya Phillips (Assistant Director), Jennifer G. Stratton (Analyst-in-Charge), William R. Chatlos, Steven Flint, Robert Gebhart, James A. Howard, Matthew Levie, Ed Nannenhorn, Cynthia Saunders, Sonya Vartivarian, Daniel Webb, and Alicia White made key contributions to this report.

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through our website. Each weekday afternoon, GAO posts on its website newly released reports, testimony, and correspondence. You can also subscribe to GAO's email updates to receive notification of newly posted products.

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, <https://www.gao.gov/ordering.htm>.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

Connect with GAO

Connect with GAO on Facebook, Flickr, Twitter, and YouTube. Subscribe to our RSS Feeds or Email Updates. Listen to our Podcasts. Visit GAO on the web at <https://www.gao.gov>.

To Report Fraud, Waste, and Abuse in Federal Programs

Contact FraudNet:

Website: <https://www.gao.gov/fraudnet/fraudnet.htm>

Automated answering system: (800) 424-5454 or (202) 512-7700

Congressional Relations

Orice Williams Brown, Managing Director, WilliamsO@gao.gov, (202) 512-4400, U.S. Government Accountability Office, 441 G Street NW, Room 7125, Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800, U.S. Government Accountability Office, 441 G Street NW, Room 7149, Washington, DC 20548

Strategic Planning and External Liaison

James-Christian Blockwood, Managing Director, spel@gao.gov, (202) 512-4707, U.S. Government Accountability Office, 441 G Street NW, Room 7814, Washington, DC 20548



Please Print on Recycled Paper.



Coming Forward

Key Trends and Data
from the TIME'S UP
Legal Defense Fund





The National Women's Law Center

The National Women's Law Center (NWLC) fights for gender justice—in the courts, in public policy, and in our society—working across the issues that are central to the lives of women and girls. NWLC uses the law in all its forms to change culture and drive solutions to the gender inequity that shapes our society and to break down the barriers that harm all of us—especially those who face multiple forms of discrimination, including women of color, LGBTQIA+ people, and low-income women and families. For more than 45 years, NWLC has been on the leading edge of every major legal and policy victory for women.



The TIME'S UP Legal Defense Fund

The TIME'S UP Legal Defense Fund is housed at and administered by the National Women's Law Center Fund. The TIME'S UP Legal Defense Fund was created after women in Hollywood spoke out about their experiences of sex harassment. Farmworker women responded in an open letter, recognizing the women in Hollywood as their "sisters," and committing to work in solidarity to tackle the issue of sex harassment in both their workplaces.¹ This alliance, and the related groundswell of public support, led to the creation of the TIME'S UP Legal Defense Fund, a national fund—the first of its kind—dedicated to helping individuals, particularly workers in low-paid jobs, challenge workplace sex harassment and retaliation. The TIME'S UP Legal Defense Fund connects attorneys with individuals who experience sexual misconduct including assault, harassment, abuse, and related retaliation in the workplace or in trying to advance their careers. The TIME'S UP Legal Defense Fund also provides financial support to defray the costs of legal and public relations assistance in selected cases.

Authors

Jasmine Tucker, Director of Research

Jennifer Mondino, Director of Legal Programs, TIME'S UP Legal Defense Fund

Acknowledgements

The authors would like to thank Fatima Goss Graves, Emily Martin, Sharyn Tejani, Ellie Driscoll, Amanda Barroso, Nicole Rinier, Vivian Alvarado, Sky Sealey, Hilary Woodward, Beth Stever, and Tom Gamertsfelder for their substantive contributions to and review and design of the report, and for all that they have done and continue to do to further the work of the TIME'S UP Legal Defense Fund.

The authors also thank Jae Aron, Rakeen Mabud, and Amanda Harrington for their review of the report and the TIME'S UP Foundation for their support of the TIME'S UP Legal Defense Fund.

Executive Summary

Since its launch in January 2018, the TIME'S UP Legal Defense Fund has heard from **thousands of people, in every state, industry, and type of job,** about their experiences of being harassed at work.

There is power, truth, and pain in each and every one of these stories, to be sure, and a core part of the mission of the TIME'S UP Legal Defense Fund is to connect each of these individuals with legal help and ultimately, we hope, a measure of justice. But when we take a step back and consider these requests for legal help as a collective, and see the commonalities in what these individuals have experienced and how it has affected them, there is a different sort of power and truth: it shines a

light on certain realities about the dynamics and impact of workplace sex harassment. For example, that more often than not, speaking out about sex harassment will be met with retaliation. That sex harassment can take a serious toll on the financial stability and mental health of survivors. And that, all too often, supervisors and others in positions of power are the ones perpetrating sex harassment, and employers are failing to take action to address this abuse.

A WORD ABOUT HOW THIS DATA WAS GATHERED:

To receive attorney contact information from the TIME'S UP Legal Defense Fund, an individual must complete an online request form. In completing that form, they provide their e-mail address, state, and a brief description of their situation, enough to make clear that the situation involved workplace sex harassment. There is no requirement that the individual give any other specific details about themselves or their experiences, and all the other information solicited in the online request form, such as demographic information, is provided voluntarily. In addition, as is noted on the online request form itself, the TIME'S UP Legal Defense Fund keeps the information provided confidential to the extent allowed by law. Because connecting with attorneys is the individual's primary goal in completing the TIME'S UP Legal Defense Fund online request form, the amount of information and level of detail in each request for legal assistance varies dramatically. As a result, the data here is based solely on what individuals chose to share with us. Therefore, in interpreting the results and trends described in this report, it is important to keep in mind that if a person did not mention something (such as the sex of the perpetrator, how the harassment affected their mental health, or whether they were retaliated against) it does not mean it was not a part of their experience. Rather, it simply means that the person did not volunteer that information. It should be understood, therefore, that the figures provided throughout this report refer to the percentage or number of people who volunteered this information.

“

Standing up to my assaulter has been one of the worst experiences of my life—I was punished for it.

—Colorado

No single analysis will ever be able to capture the full scope and impact of sex harassment in the workplace. This is both because many people never report the harassment they experience and because the long-term impacts of sex harassment on survivors' careers, financial situation, mental health, and more are difficult to wholly measure. This report does not attempt to draw conclusions about how a larger population of people is experiencing sex harassment in the United States. People who come to the TIME'S UP Legal Defense Fund for legal help are inherently unique from the larger population who have experienced sex harassment: the individuals who reach out to the TIME'S UP Legal Defense Fund not only have reached a point where they want or need justice, but they are sufficiently well informed and able to try to seek out legal help—in contrast to most survivors of workplace sex harassment. Surveys

indicate the majority of those who experience workplace sex harassment will never report their experiences, much less seek help from an attorney. Still, in considering the trends in the characteristics and experiences of the individuals who sought help from the TIME'S UP Legal Defense Fund, we can gain an understanding of the dynamics and impacts of workplace sex harassment more broadly. This snapshot into survivors' lived experience with workplace sex harassment can and should be useful to survivors, policymakers, employers, and the public in thinking about how to prevent and address sex harassment in the workplace. This report provides an overview of the demographics and main trends found in the experiences of people seeking legal help from the TIME'S UP Legal Defense Fund to address workplace sex harassment. Between January 1, 2018, and April 30, 2020, the TIME'S UP Legal Defense Fund received 3,317 requests from individuals seeking legal help for situations involving their own experiences of workplace sex harassment. Most people contacted the TIME'S UP Legal Defense Fund by submitting an online request form. The information they provided in the online request form was solicited for the purpose of connecting these individuals with attorneys in the TIME'S UP Legal Defense Fund/Legal Network for Gender Equity. See this report's full methodology in Appendix B for more information.



Topline Findings

» **More than seven in 10 survivors who experienced workplace sex harassment faced some form of retaliation, including termination, being sued for defamation, and denial of promotions.**

› More than seven in 10 people (72 percent) said they experienced some form of retaliation when they complained about harassment. Of those who experienced retaliation, the most common form mentioned was being fired (36 percent), followed by 19 percent who said they received poor performance evaluations, had their work products or behavior scrutinized, or were otherwise treated poorly at work.



» **Workplace sex harassment had a severe impact on individuals' economic, physical, and mental health well-being.**

› More than one in five people (22 percent) volunteered² information about their experience of workplace sex harassment negatively impacting their economic or financial well-being.

› Nearly one in five people (19 percent) volunteered that the harassment had a damaging impact on their mental health.

› More than one in four people (28 percent) volunteered that the harassment they experienced was not an isolated incident.

› More than one in five people (21 percent) volunteered information about their perpetrator harassing multiple victims.

› More than one in three people (36 percent) said they experienced sexual assault, assault, rape, or other physical harassment.

› Nearly one in nine people (11 percent) said that they had reported the harassment to the police.

» **Individuals frequently experienced intersecting forms of discrimination.**

› Nearly one in five people (18 percent) said that they had experienced discrimination or harassment based on sex and other aspects of their identities; for example, they were harassed because they were a woman with a disability, a woman of color, or a woman born outside of the United States.

› Nearly one in nine people (11 percent) said that they had experienced both sex and race discrimination at work.



» **Even when individuals reported harassment, their perpetrators were not held accountable.**

- › Seven in 10 people (70 percent) said they reported the harassment they experienced.
- › Of the people who identified a perpetrator of their harassment, nearly two in five people (37 percent) said that nothing happened to the perpetrator.

» **Harassment frequently comes from someone who has power over the worker.**

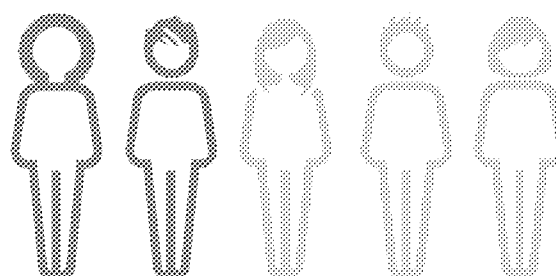
- › More than half of the survivors (56 percent) who identified their perpetrator in the online request form said they were harassed by someone they reported to at work, including a supervisor, superior, owner, or top executive.

» **Individuals are turning first to their employers to report harassment, but employers are not taking action.**

- › Of those people who said they reported the harassment (i.e., reported the harassment to an employer, a government agency, a court, or law enforcement), nearly two in three people (64 percent) reported the harassment to their employer.

- › Of people who reported harassment, nearly three in 10 (29 percent) said nothing was done about it.

These trends were almost entirely consistent regardless of people's race or national origin. The differences across race or national origin were only rarely statistically significant and thus are generally not detailed in this report.



Of the people who identified a perpetrator of their harassment, nearly two in five people (37 percent) said that nothing happened to the perpetrator.

Introduction

Since its launch in January 2018, the TIME'S UP Legal Defense Fund has heard from **thousands of people** facing workplace sex harassment in every state, industry, and type of job.

It has heard from individuals—mostly women—who have been assaulted, grabbed, and groped. Who have been told that their job or their promotion hinged on having sex with a supervisor. Who have been propositioned. Who regularly hear comments about how they look, their bodies, and how their harasser wants to assault them. We also hear from individuals who are belittled constantly at work—for no reason other than their gender—and whose work is ignored or criticized—in direct response to their speaking out about harassment.

Too often, people who experience sex harassment feel alone and blame themselves for the harassment. This report provides a unique window into how people approaching the TIME'S UP Legal Defense Fund for help experienced sex harassment and how it affected them, and it should reassure those facing harassment in their own workplaces: they are not alone, and the harassment is not their fault. This is a nationwide problem in all types of jobs and industries and happens to all types of people.

Further, this report shows the need for the systemic changes catalyzed by the #MeToo and TIME'S UP movements—changes in state policy, commitments by employers, and more willingness from survivors to come forward. Finally, the report also shows that there is a great deal of work left to do, especially to stop retaliation, ensure appropriate employer responses, and address the effects of harassment on the mental health and economic well-being of workers so that everyone can experience a safe and respectful workplace.

“

I feel like a sex object in the eyes of my...male co-workers.

—Hawaii



Sex harassment is a form of sex discrimination and includes a range of unwelcome behavior motivated by the sex or gender of the person targeted for the harassment.³ Throughout this report, the term sex harassment is used as an umbrella term to include two things:

SEX HARASSMENT,

or unwelcome sexual behavior. For example, it may include unwanted sexual advances, requests for sexual favors, unwanted touching, sexual assault, abusive sexual language, or demands to engage in sex as a condition of getting or keeping a job.

GENDER-BASED HARASSMENT,

which includes a broad range of verbal and non-verbal behaviors that convey insulting, hostile, and degrading attitudes based on gender. Although it may not include sexual overtures, this type of harassment communicates hostility to someone because of their gender or based on gender stereotypes. For example, such harassment may include hostile comments about a woman not conforming to stereotypes about how she should dress or behave or about women not being suited to a particular job. This type of harassment also may include hostile comments or other unwelcome behavior based on the individual's sexual orientation, gender identity, or pregnancy.

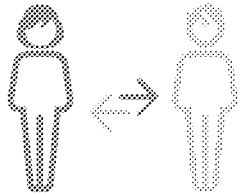
“

I'm tired of the good old boys club...I'm a[n] opinionated girl and I don't fit in. [They] let me know it.

—Illinois

The TIME'S UP Legal Defense Fund

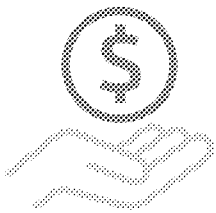
The TIME'S UP Legal Defense Fund has four major projects:



Connecting Individuals with Attorneys

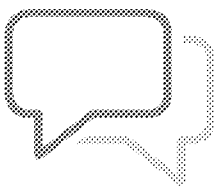
The TIME'S UP Legal Defense Fund connects individuals facing workplace sex harassment and retaliation with attorneys in the Legal Network for Gender Equity and provides individuals with information about the laws that prohibit sex harassment at work. Attorneys in the Legal Network provide free initial legal consultations to individuals who are connected to them through this process.

Since January 2018, the TIME'S UP Legal Defense Fund has provided this information in response to over 4,700 requests, helping scores of people understand their legal rights and obtain legal representation. Currently, there are 689 attorneys in our Legal Network for Gender Equity.



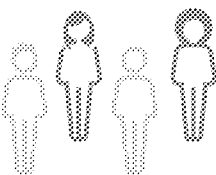
Funding Cases

The TIME'S UP Legal Defense Fund provides funding to cover attorney's fees and costs in selected cases of workplace sex harassment and related retaliation. These cases are selected based on a review of attorneys' funding applications, which include information about financial need and how the case fits the priorities of the TIME'S UP Legal Defense Fund. Decisions about cases, eligibility, and funding are made by TIME'S UP Legal Defense Fund staff, based on a rigorous evaluation against a consistent set of guidelines free from influence by any outside agenda, person, organization, or entity.



Public Relations Assistance

The TIME'S UP Legal Defense Fund provides funding to cover professional public relations firms to assist in selected matters in which individuals are speaking out about their experience with workplace sex harassment. These public relations specialists help people prepare to talk to the press, find outlets to speak with, and respond to press inquiries.



Outreach Grants

The TIME'S UP Legal Defense Fund has funded 18 organizations focused on the rights of low-wage workers and/or survivors of sexual violence through outreach grants. The grants help support programs to inform people about their rights regarding workplace sexual harassment and how to report it and access resources. The groups focus on populations such as farmworkers, restaurant workers, poultry workers, nail salon workers, LGBTQIA+ workers, and recent immigrants.

Overview of Methodology

This report provides an overview analysis of the **3,317 requests** for legal help submitted between January 1, 2018 and April 30, 2020.

The cases discussed here relate specifically to workplace sex harassment and were submitted by the individual who experienced the workplace sex harassment described in the request. See this report's full methodology in Appendix B for a more detailed explanation of the nature of the requests included in the analysis.

All the requests for legal help included in the analysis involved either sex harassment or gender-based harassment or both.

Nearly nine in 10 people (88 percent) who requested legal help said they experienced sex harassment while more than one in four (26 percent) said they experienced harassment based

on their gender. Many people said they experienced both types of harassment, often saying they were subjected to unwanted sexual advances and then called sexist slurs for refusing those advances.

Demographic data of people requesting legal help

Most individuals seeking legal help from the TIME'S UP Legal Defense Fund identify as women. Of the 3,317 people requesting legal assistance for workplace sex harassment, 91 percent of people (3,020) self-identified their gender.⁴ Of those who identified their gender, 96 percent identified as women, three percent identified as men, and less than one percent identified as nonbinary or another gender identity.

About 40 percent of people requesting legal help from the TIME'S UP Legal Defense Fund identify as people of color.

Of the 3,317 requests, 84 percent of people (2,796) self-identified their race/ethnicity.⁵ Of those:

- 59%** identified as white
- 17%** identified as Black
- 6%** identified as Latinx
- 5%** identified as Asian American or Pacific Islander (AAPI)
- 1%** identified as Native American
- 11%** identified as two or more races⁶
- 2%** identified as another race

These proportions closely mirror the makeup of the U.S. workforce in 2018.⁷

Most people requesting legal help from the TIME'S UP Legal Defense Fund identify as low-income.

Nearly six in 10 people (59 percent) indicated on their online request form that they may qualify for legal services for low-income legal clients. However, it is important to note that people could have responded yes to this question for a variety of reasons. For example, they may have experienced harassment while working in a low-paid job and cannot afford to pursue legal action. Or, they may have lost a higher-paid job due to harassment and thus have had limited resources at the time they completed the online request form. Additionally, people likely interpreted "low-income" in different ways.

The TIME'S UP Legal Defense Fund receives requests for legal help from all across the United States.

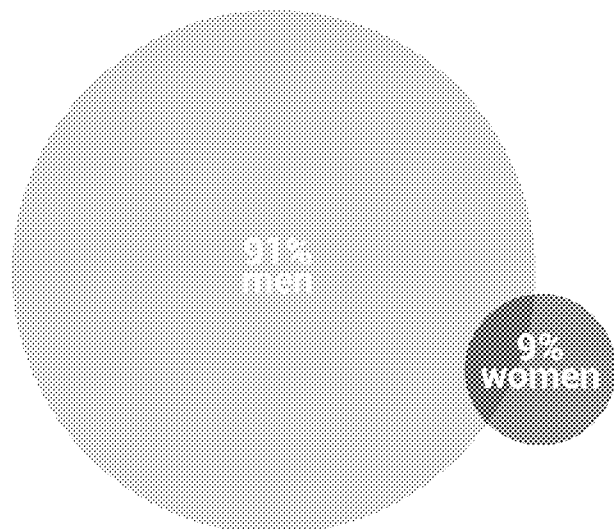
Nearly every person requesting legal help (99 percent) indicated the state where their harassment occurred. Of those 3,270 people requesting assistance that specified a state:⁸

» Thirty-four percent were in the South (including South Atlantic, East South Central, and West South Central regions; i.e., Alabama, Arkansas, District of Columbia, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, West Virginia).

- » Thirty percent were in the West (including Mountain and Pacific regions; i.e., Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming).
- » Twenty-one percent were in the Northeast (including New England and Middle Atlantic regions; i.e., Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont).
- » Fifteen percent were in the Midwest (including East North Central and West North Central regions; i.e., Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin).

The online request form never prompts survivors for information about the perpetrator of harassment. Despite that, more than eight in 10 people (81 percent) volunteered information about their perpetrator.

Of those who volunteered the gender of their harasser, 91 percent identified perpetrators as men and nine percent identified perpetrators as women. Of people who identified the gender of their harasser, the majority of both men and women reported being harassed by a man. However, women (93 percent) were more likely than men (54 percent) to identify their perpetrator as a man, while men (46 percent) were more likely than women (seven percent) to identify their perpetrator as a woman.



Of the 2,208 people who identified the gender of their harasser, more than 9 in 10 said they were harassed by a man.

Key Trends and Data Points

Reporting Harassment

Because of low rates of reporting, the prevalence of sex harassment in the workforce has long been and continues to be difficult to measure. According to the EEOC, some studies found as many as 85 percent of women report experiencing sex harassment in the workplace.⁹ However, given barriers such as fear of not being believed or fear of retaliation, many people do not report their experience: it is estimated that anywhere from 87 percent to 94 percent of people who experience workplace harassment never file a formal complaint.¹⁰

Of the 3,317 requests for legal assistance to the TIME'S UP Legal Defense Fund, seven in 10 people (70 percent or 2,330 requests) said that they reported the harassment they experienced to someone else such as their employer, the police, or the Equal Employment Opportunity Commission (EEOC). Given that these people were seeking legal assistance, it is unsurprising that they might be more likely than the wider workforce to come forward with a complaint about harassment.

Of those who did report the harassment, nearly two in three (64 percent) said they reported it to their

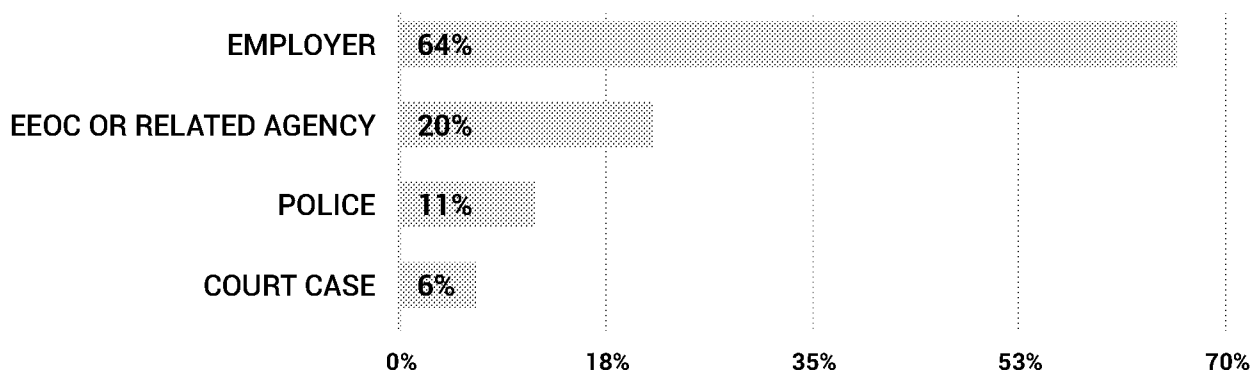


When I complained about my boss[’s] treatment...the acting HR individual told me I was “too sensitive”, and that my boss[’s] behavior was just brusque.

—California

employer (either to a supervisor or other higher-up, human resources, or Title IX office), and in many instances, people reported the harassment to multiple people in their workplace. Of those who reported harassment, one in five people (20 percent) reported to the EEOC or other government agency; those that identified as Black (30 percent) were nearly two times more likely than those who identified as white (16 percent) to indicate they reported the harassment to the EEOC or a related office. Others reported the harassment to the police (11 percent) or filed a case in court against their perpetrator or employer (six percent).

Where people reported sex harassment



Nearly three in 10 people (29 percent) who reported harassment said that nothing was ever done about it. Instead, many people said they were not believed or were ridiculed for reporting the harassment. That type of response can have a chilling effect on reporting and, in turn, exacerbate a toxic workplace culture; if workers see that nothing is done when a co-worker complains, for example, they may be less likely to come forward themselves, and harassment may continue and worsen over time.

Retaliation

Retaliation occurs when a worker faces adverse consequences or worse treatment as a result of reporting sex harassment or otherwise trying to stop sex harassment. Retaliation can take many forms, such as being transferred to less desirable or more burdensome work assignments or locations, receiving harsher treatment from supervisors or worse performance reviews, or even being fired outright or sued. Of the people the TIME'S UP Legal Defense Fund heard from, the vast majority faced retaliation.

» **More than seven in 10 people (72 percent) requesting legal assistance said they experienced retaliation in some form after reporting or trying to stop the harassment, and many experienced multiple forms of retaliation.** An analysis of charges made to the Equal Employment Opportunity Commission (EEOC) suggests this pattern is consistent with the experiences of survivors of workplace sex harassment more broadly; in fiscal years 2016 and 2017, 72 percent

of sex harassment charges filed with the EEOC included allegations of both sex harassment and retaliation.¹¹

» **More than one in three people (36 percent) who experienced retaliation said they were fired from their job.**

› Of those who experienced retaliation, people who identified as men (53 percent) were more likely than people who identified as women (34 percent) to say they were fired.

» **Nearly one in five people (19 percent) said they were given poor performance reviews, had their work products or behavior scrutinized, or were otherwise treated poorly.**

› Of those who experienced retaliation, people who identified as Black (26 percent) were more likely than those who identified as white (18 percent) to say they were given negative reviews or had their behavior scrutinized.

› Of people who experienced retaliation, those who identified as women (20 percent) were nearly three times more likely than those who identified as men (7 percent) to say they were given negative reviews.

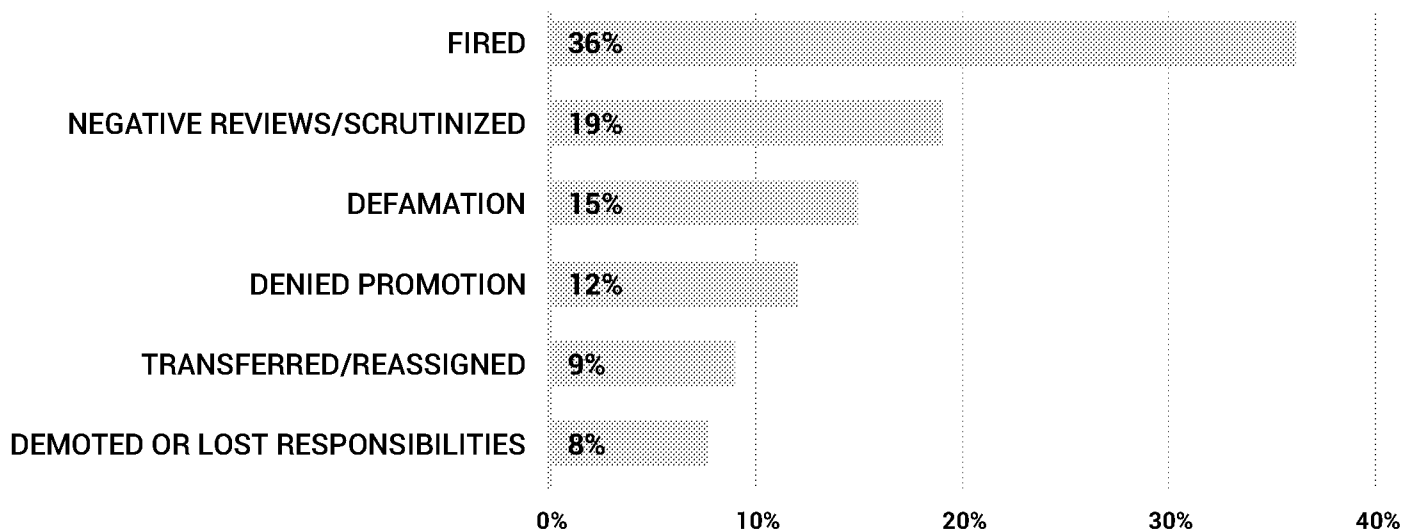
» **Additionally, more than one in seven people (15 percent) who experienced retaliation said they were slandered or had their reputation damaged in some way by their perpetrator or employer.** Further, more than 100 of these people specifically stated that the slander led to them being blacklisted in their company or field.

“

They said they were sorry if I was offended and that I took it the wrong way, it “wasn’t like that.” I responded that I had not been offended I had been assaulted.

—Ohio

More than 1 in 3 people who experienced retaliation were fired



“

The men who harassed me advised teammates to beware of interacting with me, my department colleagues limit their interactions with me and/ or ignore me including management.

—Arizona

Similarly, for a considerable number of people, the harassment and retaliation were so severe that they felt forced to quit their jobs.

» **One in 10 people (10 percent) requesting legal assistance volunteered that they quit their job as a result of the harassment, often citing an ongoing hostile work environment or concerns about their safety.** Although these people were not terminated outright, leaving the job was not voluntary; many of these people said they were pushed out of their workplace or felt they had no other options because they needed to put distance between themselves and their harassers. Others said they could not afford to lose their job and had to endure the harassment in order to keep working.

Individuals also reported retaliation taking the form of detrimental changes to people's work assignments and opportunities.

» **Nearly one in eight people (12 percent) who experienced retaliation said they were denied a promotion or other career advancement.** Nine percent were transferred, reassigned, or displaced from their department or a project.

» Of those who experienced retaliation, people who identified as women (10 percent) were 10 times more likely than people who identified as men (one percent) to say they were transferred, displaced, or reassigned.

- » Nearly one in 12 people (8 percent) said they were demoted or otherwise lost work responsibilities.
 - ↳ Of those who experienced retaliation, people who identified as women (nine percent) were much more likely than people who identified as men (one percent) to say they were demoted.



This employer repeatedly failed to protect me and even told me that they would demote me if I continued to report allegations.

—North Dakota

More than 140 people said they experienced isolation or exclusion as part of their harassment or in retaliation and more than 120 people said rumors were spread about them throughout their workplace. Others volunteered that their perpetrators “gaslit” them—in other words, perpetrators responded as if the harassment never happened. In many cases, women that had accused men of sex harassment or retaliation were labeled as irrational or “crazy.” This type of “gaslighting” behavior is particularly harmful because it can make targets of harassment and retaliation both more vulnerable to further workplace abuse and less likely to rely on workplace supports.¹²

Individuals reported a variety of efforts by employers to discourage or punish the reporting of harassment or other attempts to seek justice; for example, people reported being discouraged from pursuing legal action, being pressured to sign nondisclosure agreements (NDAs), and facing complaints and lawsuits filed against them.

- » **Nearly one in five people (19 percent) said that people in the workplace either discouraged them from pursuing legal action regarding their experience or explicitly told them to keep the harassment quiet.**
- » People who identified as men (28 percent) were more likely than people who identified as women (19 percent) to be discouraged from sharing their story or pursuing legal action, with many men saying they were discouraged for reasons entrenched in toxic masculinity: they were warned they would look weak or be embarrassed for reporting harassment, especially if their perpetrator was a woman. In many of these cases the male workers were told they should be flattered by the attention or that the harassment was harmless flirtation.

Many people said there was more than just discouragement—there were outright threats.

- » **More than one in seven people (15 percent) said that they were threatened with legal action, with losing their job, or even physical harm if they told anyone about their experience.** Sometimes the discouragement and threats came from the perpetrators themselves, but other times it came from supervisors or other higher-ups, or even from someone in their workplace’s human resources department, the very place many people are told to go if they ever experience sex harassment. In some cases, people said that they were explicitly told by someone they reported the harassment to that nothing would happen with their complaint because the perpetrator was too senior to experience repercussions or because the perpetrator was too valuable to the company’s bottom line.
- » **Six percent of people seeking legal help said that they were presented with or signed a nondisclosure agreement.** Most of them said it was presented to them after the harassment or as part of a settlement agreement or severance

“

I reported to HR immediately. The next day one of the company owners called me and said I would have to decide [i]f I could continue working with him because he was an integral part of their business ventures.

—Arizona

payment from their employer. Many others said they were asked to sign a release of claims, which means they are unable to pursue legal action against their employer, and still others said they were given something to sign but did not know exactly what the document was.

Harassment by supervisors and people in positions of power

More than half (56 percent) of the survivors who identified their harasser in their online request form said it involved someone they reported to at work, including a supervisor, superior, owner, or executive.

Relatedly, many people specifically said they experienced quid pro quo harassment. For these individuals, they had to endure harassment—verbal or physical—from a supervisor or superior in order to keep their job.

“

My employer forced me to perform oral sex on him before he would give me my paycheck.

—California

Accountability and employers' response to harassment

Of those who reported their harassment, nearly two-thirds (64 percent) said they reported it to their employer (either to a supervisor or other superior, human resources department, or a Title IX office). Indeed, in many instances, people reported the harassment to multiple people in their workplace. This points to the fact that when employers have an internal process in place for making complaints about harassment, some workers are aware of and following this process, at least as a first step.

At the same time, some people (one percent) volunteered that their workplace did not have a formal policy or any training on sex harassment. Of those people who volunteered information about their harasser's identity in their online request form, most people did not specify what happened to their harasser after they reported the harassment. Yet only five percent of people who identified their harasser in their online request form noted that their harasser suffered any consequences, such as termination, demotion, or being pushed into early retirement.

In many instances, it seems that the harassment individuals experienced was widely known throughout the workplace.

“

I was chosen by my former supervisor as a target of bullying, gaslighting, and other general intimidation behaviors for a year after disputing a decision he made...He has a history of targeting and bullying female employees at his last three places of employment and was simply moved around in the district...

—Oregon

» **More than one in five people (21 percent) said that they were not the only person being harassed in their workplace—other people were targeted, too.** Further, people who identified as women (22 percent) were more likely than those who identified as men (13 percent) to say that there were other victims of harassment. **One in 11 people (nine percent) seeking legal help mentioned that co-workers, customers, or others were witnesses or bystanders to their harassment.** Many of these people said having witnesses did not help their case, however, and others said that witnesses or bystanders were unwilling to come forward in reporting the harassment they saw for fear of becoming a target themselves or for fear of retaliation.

“

I dreaded coming into work on many days...I finally was so depressed because of the harassment, that I just didn't report anymore.

—Nevada

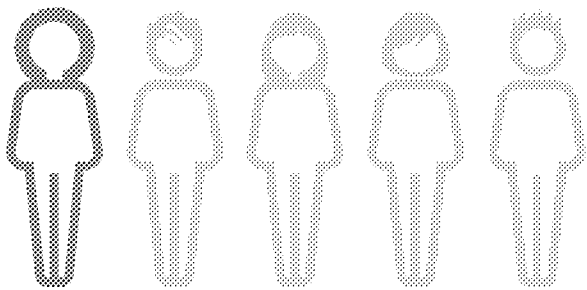
Intersections of sex harassment and other forms of discrimination

All too often, sex harassment is paired with, and exacerbated by, discrimination motivated by hostility against survivors' intersecting identities, such as their race, national origin, sexual orientation, and disability status. In an analysis of sex harassment charge data filed by women with the EEOC between 2012 and 2016, for example, NWLC found that Black women and Latinas filed charges at a higher rate than their white peers, suggesting that they experience harassment at higher rates.

» Of people seeking legal assistance from the TIME'S UP Legal Defense Fund, those who identified as Black (26 percent), AAPI (23 percent), Latinx (15 percent), two or more races (16 percent), or another race (18 percent) were all significantly more likely than those who identified as white (three percent) to say they experienced race-based harassment along with sex-based harassment.

» Further, these numbers likely undercount the actual prevalence of race- or national origin-based harassment experienced by these survivors, as the online request form prompts are focused on sex harassment, and thus it is likely that people would not have volunteered information about race- or national origin-based harassment even when it was part of their experience.

In addition to race-based harassment, many people requesting legal help cited harassment based on another identity: more than 200 people said they experienced harassment because of a disability, and nearly 100 people said they experienced harassment because of their sexual orientation or gender identity. Many others reported harassment based on religion, age, pregnancy, nationality, or immigration status in addition to sex.



Nearly one in five people (19 percent) requesting legal assistance from the TULDF volunteered that the harassment they experienced had a negative impact on their mental health.

“

As a gay female with a disability in the male-dominated technology industry and the only female on my team in my company, I was subjected to ongoing harassment from my manager...[He subjected] me to unsafe and uncomfortable situations in order to provoke fear and anxiety in me and intentionally intensifying the effects of my mental disability.

—Florida

Severity of risk of harm from and impact of sex harassment

Mental health impact

For many survivors of workplace sex harassment, the experience of harassment results in mental health consequences such as depression, anxiety, substance abuse, and eating disorders.¹³

» **Nearly one in five people (19 percent) requesting legal assistance from the TIME'S UP Legal Defense Fund volunteered that the harassment they experienced had a negative impact on their mental health.**

» These results mirror those of a recent survey showing that a common outcome for people who have experienced sex harassment and assault is feeling anxiety or depression.¹⁴

» The mental health impact was even more pronounced for Native Americans: two in five people seeking legal help who identified as Native American (40 percent) said the harassment had an adverse impact on their mental health.

“

The work environment became so hostile that I dreaded going to work, suffered from lack of sleep, questioned my own sanity, had a panic attack, and even considered taking my own life to escape from the pressure of continuing to work with [my perpetrator].

—District of Columbia

Of those people seeking legal assistance who volunteered information about a damaging impact on their mental health:

- » More than one in three (34 percent) said they felt emotional distress or increased stress, more than one in four (28 percent) said they experienced anxiety or panic attacks, nearly one in four (24 percent) said they experienced symptoms of or were diagnosed with post-traumatic stress disorder (PTSD), and more than one in five (22 percent) reported being depressed after their experience.
- » Dozens of people said they considered or attempted suicide after their experience and others reported low self-esteem or a loss of confidence. Of those who reported experiencing an impact on their mental health, less than one

in five people (18 percent) mentioned seeking assistance from a mental health professional.

Severity and breadth of harm

Many people said the sex harassment they experienced included physical harm, multiple victims or perpetrators, or harassment occurring repeatedly and through multiple means, including through social media and other technology.

- » More than one in three people (36 percent) said they experienced sexual assault, assault, rape, or other physical harassment by their perpetrator, often in conjunction with verbal or other harassment.
- » And nearly one in nine (11 percent) people who reported the harassment made a report to the police.
- » **More than one in four people (28 percent) requesting legal help said the harassment they experienced was not an isolated incident—it happened more than once or continuously.** People who identified as women (29 percent) were more likely than those who identified as men (19 percent) to say the harassment was not an isolated incident. More than 160 people volunteered information about the harassment they experienced getting worse over time. Many of those people said the harassment started with verbal comments that later escalated to physical harassment, such as sexual assault.

With the ubiquity of e-mail, social media, and smartphones, today's technology means perpetrators have new ways of harassing their targets. More than 250 people seeking legal help said they experienced online or electronic harassment. Many said they were sent unwanted nude photos of their perpetrators or other pornographic material. Seventy people said they experienced image-based sexual abuse, where a private, often explicit, photo of them was shared without their consent.¹⁵



“

I have been harassed, belittled and [discriminated] against...when I was the only female District Service Manager...I can no longer find work in [my] area close to home because the Slander and Defamation of my character continues...I went from making six figures to unemployed and homeless.

--Illinois

In many cases, the harassment involved multiple people, as well, both in terms of the perpetrators and the victims.

- » More than one in five people (21 percent) volunteered that there were other victims of harassment.
- » In addition, nearly one in five people (18 percent) who identified a perpetrator in their online request form mentioned that there were multiple perpetrators of harassment.

Financial impact

Sex harassment can set off a chain of events with effects not only on the job where the harassment occurred, but on a survivor's entire career. For example, being forced out of a job as a result of harassment (whether someone was fired or felt forced to quit to escape a hostile working



environment) can lead to a substantial loss of earnings. Still others may be blacklisted in their field: they may not find a comparable job and may be forced to start over in an entirely different field of work. For people in rural areas, there may not be another workplace matching their skills within a reasonable commuting distance. And more than just lost earnings in the short term, harassment has implications for survivors' lifetime earnings, future career growth, and even retirement. For these reasons and more, sex harassment has been identified as one driver of the gender wage gap.¹⁶

More than one in five people (22 percent) volunteered information about the devastating impact their experience with sex harassment had economically or financially, some even saying that the experience left them destitute and struggling to pay their bills.

Of those people who volunteered a negative economic or financial impact as a result of harassment, nearly one in three people (31 percent) said that their hours, shifts, or pay had been cut. This can be especially harmful for someone who is already living paycheck to paycheck, such as those working in a low-paid job or in one that relies on tips.

Of those who volunteered information about an economic or financial impact, nearly one in four (23 percent) said they had money troubles because of a difficulty finding another job. Conversely, nearly one in nine people (11 percent) said they endured the harassment because they needed the money or otherwise could not afford to lose their job.

Seven percent said they lost health insurance or other benefits as a result of the harassment, and dozens of others mentioned that the harassment impacted their retirement. Some said this because they were forced into an early retirement; others mentioned that they had to draw down retirement savings while they looked for other work, which incurs large penalties for workers under 59½.¹⁷ Still others said they had to leave a job and settle for another one that didn't offer an employer-sponsored retirement plan.

“

The[y] have cut my hours, so I am not forced to work with him, [but], essentially, I am the one being punished for his bad behavior.

—Missouri

Conclusion

The strength and bravery of the thousands of people who have come forward to the TIME'S UP Legal Defense Fund to share their stories and **seek justice for themselves and others** is astounding. Despite obstacles such as discouragement from reporting, retaliation, and negative impacts on their mental and financial well-being, people requesting legal help frequently said they were coming forward because others were too afraid to do so or because they wanted to prevent anyone else from experiencing a similar situation.

This analysis of their requests for legal help provides a window into their experience, into a better understanding of workplace sex harassment and its effects, in order to build safer and more equitable workplaces.

Policymakers, advocates, employers, legislators, and others should use this information to identify and implement needed changes in how to support survivors and respond to harassment at work. For recommendations and additional information about such changes, please visit the National Women's Law Center [here](#).¹⁸

Appendix A: Variables Coded from Online Request Forms

Codebook	Number of sex harassment requests with code	Percent of sex harassment requests with code
Bullying mentioned	292	9%
Criminal behavior by employer/perpetrator	87	3%
Discouraged from legal action or sharing story	640	19%
Domestic violence mentioned	45	1%
Economic impact*	714	22%
Employment relationship mentioned	80	2%
Forced arbitration	38	1%
Formal complaint was made*	2330	70%
Formal complaint was made but nothing was done	952	29%
Harassment (h.)	3317	100%
h. Electronic	259	8%
h. Exclusion/isolation	142	4%
h. Image-based sexual abuse	70	2%
h. Intimidation/aggression	388	12%
h. Physical	648	20%
h. Pornography	100	3%
h. Rumors	123	4%
h. Sexual	2923	88%
h. Verbal	502	15%
Harassment continued	934	28%
Harassment got worse	163	5%
Health impact	123	4%
Hostile environment	474	14%
Impact on family/relationships	94	3%
Industry (i.)	1911	58%
i. Arts/writing/design/architecture	42	1%

Appendix A: Variables Coded from Online Request Forms (continued)

Codebook	Number of sex harassment requests with code	Percent of sex harassment requests with code
i. Athletics/gym	11	0%
i. Automotive	10	0%
i. Computer/IT/tech	60	2%
i. Corrections	31	1%
i. Cosmetology/massage	11	0%
i. Delivery business /shipping/gov (UPS/USPS)	59	2%
i. Domestic work	21	1%
i. Education/school	234	7%
i. Entertainment/modeling/music	134	4%
i. Faith based org or institution/church	12	0%
i. Farming/food processing	13	0%
i. Film/photography	44	1%
i. Finance/banking	50	2%
i. Government	195	6%
i. Health	157	5%
i. Hospitality	44	1%
i. Native American reservation	3	0%
i. Journalism/media/advertising	50	2%
i. Law enforcement	45	1%
i. Legal field	47	1%
i. Manufacturing/construction/factory	72	2%
i. Military	66	2%
i. Nonprofit/social service/civil service	80	2%
i. Real estate	20	1%
i. Restaurant/fast food	100	3%
i. Retail	83	3%

Appendix A: Variables Coded from Online Request Forms (continued)

Codebook	Number of sex harassment requests with code	Percent of sex harassment requests with code
i. Sales/business	96	3%
i. Sex industry	9	0%
i. Sports	9	0%
i. Trade	6	0%
i. Transportation	49	1%
Intersectional (sex harassment and harassment based on race, religion, etc.)	584	18%
Intersectionality/discrimination (in.)	3317	100%
in. Age	120	4%
in. Disability	216	7%
in. Gender	861	26%
in. Immigration status	21	1%
in. Nationality	26	1%
in. Pay discrimination	164	5%
in. Pregnancy/maternity/mother	41	1%
in. Race/ethnicity	349	11%
in. Religion	52	2%
in. Sexual	2922	88%
in. Sexuality	94	3%
LGBTQ identified	239	7%
Low income	1961	59%
Male victim	57	2%
Male-dominated workplace	208	6%
Mental health impact*	614	19%
Metoo or TIME'S UP mentioned	134	4%
Nondisclosure agreement/other secrecy agreements (nda.)	192	6%
nda. For investigation	6	0%

Appendix A: Variables Coded from Online Request Forms (continued)

Codebook	Number of sex harassment requests with code	Percent of sex harassment requests with code
nda. Not sure if pre or post dispute	31	1%
nda. Post incident/part of settlement or severance	106	3%
nda. Pre incident	16	0%
nda. Refused to sign	36	1%
Other victims mentioned	694	21%
Perpetrator mentioned*	2687	81%
Quid pro quo harassment	243	7%
Race/ethnicity	2796	84%
Release of claims	9	0%
Resulted in unemployment/quitting	336	10%
Retaliation*	2394	72%
Sexual assault/assault/rape/physical harassment	1184	36%
Statute of limitations (s.l.)	599	18%
s.l. 1-3 years	197	6%
s.l. 4 or more years	231	7%
s.l. under 1 year	164	5%
Threats and retaliation threats	508	15%
Union involvement/membership	191	6%
Victim not believed	100	3%
Victim was blamed	125	4%
Want to prevent future harassment/help other women	76	2%
Whistleblower	46	1%
Witnesses to harassment	284	9%
Workplace had no policy or training on sex harassment	31	1%

*Denotes a separate table below provides more detail in these broad codes

Appendix A: Variables Coded from Online Request Forms (continued)

Of 2,394 Workplace Sex Harassment Requests Reporting Retaliation*	
Were terminated	36%
Reported having their behavior scrutinized, given bad reviews, or treated poorly	19%
Reported fearing retaliation	16%
Were slandered	15%
Were denied promotion or other advancement of their careers	12%
Were transferred, reassigned, or displaced from a project	9%
Were demoted or who lost responsibilities	9%

*Figures do not total 100% because some people reported multiple types of retaliation; others mentioned retaliation broadly

Of 2,330 Workplace Sex Harassment Requests Reporting a Formal Complaint*	
Reported the harassment to their employer (via a supervisor, higher-up, human resources, or title ix office)	64%
Reported the harassment to the EEOC or related agency	20%
Reported the harassment to the police	11%
Filed a court case in response to the harassment	6%

*Figures do not total 100%; some people reported the harassment in multiple ways; others did not specify to whom they reported the harassment

Of 614 Workplace Sex Harassment Requests Reporting a Mental Health Impact*	
Reported feeling emotional distress or stress	34%
Reported experiencing anxiety or panic attacks	28%
Reported experiencing PTSD	24%
Reported experiencing depression	22%
Reported seeking assistance from a mental health professional	18%
Reported feeling suicidal	6%
Reported low self-esteem	5%

*Figures do not total 100%; some people reported experiencing more than one mental health impact; others did not specify one of the above impacts

Appendix A: Variables Coded from Online Request Forms (continued)

Of 2,687 Workplace Sex Harassment Requests Identifying a Perpetrator*	
Identified their perpetrator as male**	75%
Identified their perpetrator as female**	7%
Identified their perpetrator as a supervisor/higher-up/owner/top executive	56%
Reported nothing happened to their perpetrator	37%
Identified multiple perpetrators of harassment	22%
Identified their perpetrator to be a co-worker or peer	21%
Reported their perpetrator suffered a consequence	5%

*Figures do not total 100%; some people identified more than one of the above regarding their perpetrator

**Of requests that identified the sex of their perpetrator, 91% identified them as male, 9% identified them as female

Of 714 Workplace Sex Harassment Requests Reporting an Economic or Financial Impact*	
Reported having their hours/shifts/pay cut	31%
Reported financial trouble as a result of not being able to find a new job	23%
Said they could not afford to lose their job	11%
Reported a loss of health insurance or other benefits	7%
Reported an impact on retirement	6%

*Figures do not total 100%; some people reported experiencing more than one economic impact above; others did not specify one of the above impacts or mentioned other economic impacts

Appendix B: Methodology

This research analyzes information and stories provided in requests for legal assistance related to workplace sex discrimination through the Legal Network for Gender Equity and TIME'S UP Legal Defense Fund, which are operated and administered by the National Women's Law Center Fund LLC.

The Legal Network for Gender Equity provides attorney information to individuals facing sex discrimination (including harassment) at work, in education, and as patients receiving health care. The TIME'S UP Legal Defense Fund helps people facing workplace sex harassment, including by connecting people to attorneys in the Legal Network. Between January 1, 2018, and February 16, 2018, people requesting legal assistance from

the TIME'S UP Legal Defense Fund could select whether their request was related to workplace sex discrimination. Between February 16, 2018, and July 3, 2019, people requesting legal assistance self-selected the category that best fit their situation from the following options: anti-LGBTQ discrimination in schools, biased school discipline, defense against defamation claims, discrimination against pregnant and parenting students, discrimination in health care, discrimination in school athletics, employment issues for abortion providers, military sex harassment/sex discrimination, other sex discrimination in education, other workplace sex discrimination, pay discrimination, retaliation (education), retaliation (health care), sex harassment/assault of students, workplace breastfeeding or pregnancy



Appendix B: Methodology (continued)

discrimination, workplace sex harassment. After July 3, 2019, people self-selected whether their situation occurred in education, in health care, or in the workplace and a TIME'S UP Legal Defense Fund staff person or volunteer selected the appropriate category from the list above after reviewing a request. Depending on the date of the request, then, people either self-selected a category or a category was selected on their behalf. Either way, workplace related requests were those in the following categories: defense against defamation claims, military sex harassment/sex discrimination, other workplace sex discrimination, pay discrimination, retaliation (workplace), workplace breastfeeding or pregnancy discrimination, and workplace sex harassment. As explained below, this study only used those intakes that involved workplace sex harassment.

In filling out the online request form, people were required to provide the following information, while answers to all other questions were optional: last name, email address, state where the incident occurred, the general context in which the situation occurred (i.e., in seeking health care, in attending school, or in a workplace), a brief summary of what occurred, and whether the TIME'S UP Legal Defense Fund could share their name with attorneys to alert them that the individual may contact them. Note that people were not required to share any other information; for example, people were not required to identify their employer, industry, or harasser. In addition, the TIME'S UP Legal Defense Fund does not track and is not aware of whether individuals seeking legal help ultimately take legal action.

Nonetheless, people frequently volunteered information beyond what was required to complete the form. Throughout this report, we have referred to what people "said" to describe information

provided in response to specific questions on the online request form; by contrast, we refer to what people "volunteered" to describe information provided without being specifically prompted to do so. Between July 3, 2019, and April 30, 2020, a small number of requests were taken by phone and a TIME'S UP Legal Defense Fund staff person or volunteer filled out the online request form on someone's behalf. In some instances where there was not enough information in a request, a TIME'S UP Legal Defense Fund staff person or volunteer called or emailed the person requesting legal assistance for additional information about their experience. Information captured in this way by a TIME'S UP Legal Defense Fund staff person or volunteer was excluded from this analysis. Because people filling out the online request form were only seeking an initial connection with attorneys and because the required brief summary of what occurred was an open-ended question, the level of detail in each summary varies dramatically; some individuals provided a one sentence description while others outlined their experience over several pages. In this report, where we have provided quotations from online request forms, the quoted individuals had indicated in their online request form that the TIME'S UP Legal Defense Fund could use general or anonymized information about their experiences to support its education or advocacy efforts.

Between January 1, 2018, and April 30, 2020, the Legal Network for Gender Equity received 5,072 requests that were related to workplace sex discrimination using the parameters above. After eliminating duplicate requests and requests that were misidentified as workplace-related but were in fact education—or health care-related, 4,552 requests remained for this analysis. Duplicate requests were identified as more than one request for assistance by the same email address or the same name, location, and similar description. For

Appendix B: Methodology (continued)

requests between January 1, 2018, and October 31, 2019, duplicate requests were merged by email address or name before they were analyzed in this research. For requests after October 31, 2019, only the first request from an email address or the same name, location, and similar description was analyzed.

This research deals only with those workplace requests that are workplace sex harassment. Through the coding process for this report, 3,616 workplace related requests were identified as workplace sex harassment, 3,317 of which were in first person and 299 of which were in third person—meaning they were either submitted by someone other than the person who experienced

the harassment (such as a bystander, witness, or friend or family member) or through notes recorded by a TIME'S UP Legal Defense Fund staff person or volunteer after a phone call. Information captured in third person by a TIME'S UP Legal Defense Fund staff person or volunteer was excluded from this analysis.

Requests were read and coded using Atlas.ti and the codebook in Appendix A. The data was then exported to SPSS to determine frequencies and means as well as statistical significance tests between groups by gender, race, and industry. Only differences that were statistically significant at the $p < .05$ level are reported throughout this analysis.

Appendix C: Glossary of Terms

Sex Discrimination

Sex discrimination occurs when an individual is treated less favorably due to their sex, which includes sexual orientation, gender identity or expression, pregnancy or pregnancy-related condition (including lactation), or a sex stereotype.

Sex Harassment

Sex harassment is a form of sex discrimination and includes a range of unwelcome behavior motivated by the sex or gender of the person targeted for the harassment. Sex harassment may include hostile verbal or physical conduct, whether or not sexual overtures are involved. Throughout this report, the term sex harassment is used as an umbrella term to include sex harassment and gender-based harassment, which are separately defined in this appendix.

For example, sex harassment may occur when:

- » A person's submission to or rejection of sexual advances is used as the basis for employment decisions, or submission to sexual advances is made as a condition of employment (quid pro quo harassment).
- » Unwelcome sexual conduct or gender-based harassment is sufficiently severe or pervasive that it creates an intimidating, hostile, or offensive work environment.
- » A person is subject to unwelcome communication of a sexual nature, for example through emails, text messages, messages through social media, or audible jokes or remarks.
- » An individual is forced or caused to view unwelcome images of a sexual nature, such as pornography, lewd graffiti, or sexual gestures.

Gender-Based Harassment

Gender-based harassment is unwelcome behavior that is motivated by the sex or gender of the person targeted for the harassment and that does not necessarily involve sexual overtures; rather, this type of harassment includes a broad range of verbal and nonverbal behaviors that convey insulting, hostile, and degrading attitudes based on gender. For example, gender-based harassment may include slurs, taunts, or hostile comments, or physical threats or attacks, based on the individual's actual or perceived sexual orientation, gender identity, or pregnancy.

Sex Harassment

Sex harassment is a pattern of unwelcome sexual behavior. For example, it may include unwelcome sexual advances, requests for sexual favors, unwanted touching, sexual assault, abusive sexual language, or demands to engage in sex as a condition of getting or keeping a job.

Appendix C: Glossary of Terms (continued)

Quid Pro Quo Harassment

Quid pro quo is a form of sex harassment that occurs when a supervisor, manager, or other person with positional authority or power conditions an employment benefit or decision on an individual's submission to or rejection of unwelcome sexual advances. For example, quid pro quo harassment may involve a supervisor requesting sexual favors as a condition to hire or promote someone. It also may involve a supervisor threatening to fire or demote someone if the individual denies them sexual favors.

Retaliation

Retaliation occurs when an employer punishes an employee for reporting an experience of sex harassment or sex discrimination in the workplace. Retaliation may take the form of any action that has an adverse impact on the individual's employment, such as demotion, discipline, firing, salary reduction, or job or shift reassignment. It may also be more subtle, such as leaving the individual who reported the harassment out of meetings or email threads, denying them opportunities to receive training or be considered for promotions, or giving them poor performance evaluations.

Appendix D: Original Online Request Form

Online Form for Legal Assistance

1. Name

- a. First Name:
- b. Middle
- b. Last Name:*
- c. Suffix

2. Contact Information

- a. Zip Code:
- b. City and State of Residence:*
- c. Home Phone
 - 1. Note
 - 2. Safe?
- d. Mobile Phone
 - 1. Note
 - 2. Safe?
- e. Email:

3. Do you speak a primary language other than English and would it be helpful to have language assistance services during your intake process or in consulting with a network attorney?

- a. Yes
- b. No
- c. If yes, what language:

4. Please list any necessary disability related modifications or accommodations that would be helpful during your intake process on in consulting with a network attorney.

5. How did you learn about us?

6. Is there other information you would like to provide in connection with this request?

Demographic Information

The National Women's Law Center will use the demographic data to better connect you to an attorney from the Legal Network for Gender Equity. For example, some of the attorneys only take cases from people with incomes below a certain level. We also seek to ensure that the network is reaching individuals from multiple communities and backgrounds.

* red asterisks are required questions

Appendix D: Original Online Request Form (continued)

7. What is your age?

1. Under 18
2. 18–39 years old
3. 40–64 years old
4. 65 and older

8. Gender

1. Nonbinary
2. Prefer not to say
3. Female
4. Male

9. Do you wish to identify as a member of the LGBTQ community?

1. Yes
2. No
3. If yes, how do you identify?:

10. What is your race and/or ethnicity? (select all that apply)

1. Asian/Pacific Islander
2. Black/African American
3. Hispanic/Latinx
4. Native American
5. White (Not Hispanic)

11. Additional race or ethnicity?

1. Asian/Pacific Islander
2. Black/African American
3. Hispanic/Latinx
4. Native American
5. White (Not Hispanic)

12. Some of our legal network attorneys serve only low income clients. If you believe you may qualify for legal services for low income clients and would like to be connected to an attorney who serves such clients, please check 'yes'

1. Yes
2. No

13. Sharing our stories is an important way to educate the public about the issues that come up in women's lives. Could we share your general story to help with our advocacy and education efforts without using your name or other specific identifying information? (This is optional and not required to receive information from the network.)

1. Yes
2. No

Appendix D: Original Online Request Form (continued)

Legal Issue: Please note which of the following topics your issue relates to, so that we can better direct your inquiry.

13. Legal Problem Category: Education, Healthcare, Workplace (please choose one)

14. Legal Problem Code (please choose one): Defense against defamation claims, military sex harassment/sex discrimination, other workplace sex discrimination, pay discrimination, retaliation (workplace), workplace pregnancy and breastfeeding discrimination, workplace sex harassment.

15. If Healthcare/Education:

1. Briefly describe the situation below.*
2. Is it okay for us to send your name to attorneys to alert them that you might be contacting them through the Legal Network?*

16. If Retaliation (healthcare/education)

1. Briefly describe the situation below.*
2. I would like to have my information on this intake form, including my contact information, shared with a public relations (PR) firm to be evaluated for possible free PR support when it comes to telling my story. (Please note: Checking "yes" here will not guarantee that you will receive PR help.)
3. Is it okay for us to send your name to attorneys to alert them that you might be contacting them through the Legal Network?*

17. If Workplace—defense against defamation claims, other workplace sex discrimination, pay discrimination

1. What is your industry or type of employer?
2. What is the name of your employer (This question is entirely optional.)
3. Are you a member of a union?
4. Briefly describe the situation below.*
5. I would like to have my information on this intake form, including my contact information, shared with a public relations (PR) firm to be evaluated for possible free PR support when it comes to telling my story. (Please note: Checking "yes" here will not guarantee that you will receive PR help.)
6. Is it okay for us to send your name to attorneys to alert them that you might be contacting them through the Legal Network?*

* red asterisks are required questions

Appendix D: Original Online Request Form (continued)

18. If Workplace—military sex harassment/sex discrimination, workplace sex harassment

1. What is your industry or type of employer?
2. What is the name of your employer (This question is entirely optional.)
3. Are you a member of a union?
4. Briefly describe the situation below.*
5. Have you reported the sex harassment to anyone? If so, what happened?
6. Did you experience retaliation? If so, what happened?
7. Did your employer discourage you from sharing your story or pursuing legal action—either by citing a nondisclosure agreement, insisting on mediation, or some other way? If so, what happened?
8. Did you experience other harassment in addition to sex harassment, or connected to the sex harassment—for instance, harassment because of your race, LGBTQ status, immigration status, or disability?
9. I would like to have my information on this intake form, including my contact information, shared with a public relations (PR) firm to be evaluated for possible free PR support when it comes to telling my story. (Please note: Checking “yes” here will not guarantee that you will receive PR help.)
10. Is it okay for us to send your name to attorneys to alert them that you might be contacting them through the Legal Network?*

19. If Workplace—pregnancy and breastfeeding discrimination

1. What is your industry or type of employer?
2. What is the name of your employer? (This question is entirely optional.)
3. Are you a member of a union?
4. Briefly describe the situation below.*
5. Who did you request accommodations from (supervisor, HR, both, other)? What did they say?
Was HR involved in this process?
6. Is it okay for us to send your name to attorneys to alert them that you might be contacting them through the Legal Network?*

Please click the submit button below to send us your application.

Thank you for completing our online application. Please click the submit button below to send us your application.

We will contact you within 10 days to let you know if we can help you or if we need additional information. If you have any questions, please review our FAQs page or contact us at legalnetwork@nwlc.org.

To ensure that you receive that email please add legalnetwork@nwlc.org to your list of email contacts. Please also check your junk email for emails from legalnetwork@nwlc.org.

* red asterisks are required questions

Appendix E: Current Online Request Form

Legal Help

The Legal Network for Gender Equity and sex discrimination, including sex harassment, at work or in your career, at school, or when getting healthcare.

We can also help connect you to attorneys for legal help or questions about accessing the new leave benefits for workers affected by COVID-19: paid sick leave and paid leave to care for a child whose school or child care provider is closed because of COVID-19.

To start the process, please fill in this form. When you submit it, you will receive an email from legalnetwork@nwlc.org with additional information and steps about how the process works. We will keep the information you send us private to the extent allowed by law.

If you have questions about the form, if you need an accommodation to complete this form because of a disability, or if you need assistance in a language other than English or Spanish, please email legalnetwork@nwlc.org or call 202-319-3053. Please do not come to our office to request assistance; we do not conduct in-person meetings with individuals seeking legal help.

Required information is indicated by a * after the field name.

Language

Select whether you would like the intake form to be in English or Spanish. If you need assistance in another language, please call us at 202-319-3053.

Seleccione si desea que el formulario de admisión sea en inglés o español. Si necesita ayuda en otro idioma, llámenos al 202-319-3053.

- English
 Español

Contact Information

First Name ("Jane/John Doe" is fine) _____

Last Name* _____

Email Address* _____

Is it safe to use this email to send you information?

- No
 Yes

Phone Number _____

* red asterisks are required questions

Appendix E: Current Online Request Form (continued)

Tell Us Why You Reached Out Today

We can help people to connect to lawyers to challenge sex discrimination and harassment that happened in the workplace, in education, or in healthcare.

We can also help connect you to attorneys for legal help or questions about accessing the new leave benefits for workers affected by COVID-19: paid sick leave and paid leave to care for a child whose school or child care provider is closed because of COVID-19.

Did you experience this situation:* _____

Where did this occur? _____

City _____

State* _____

Thank you for telling us a bit about your situation. Please take a couple minutes to tell us just a little more about yourself using the form below. The TIME'S UP Legal Defense Fund/Legal Network for Gender Equity uses this information to make sure that the Network is reaching the individuals we want to ensure to include through this program. Sharing this information is optional and does not affect whether you will receive help from us.

Tell Us About You

What is your race or ethnicity?

1. Asian/Pacific Islander
2. Black/African American
3. Hispanic/Latinx
4. Native American
5. White (Not Hispanic)

What is your gender?

1. Nonbinary
2. Prefer not to say
3. Female
4. Male

Do you identify as LGBTQI+?

1. Yes
2. No
3. If yes, how do you identify?:

* red asterisks are required questions

Appendix E: Current Online Request Form (continued)

Do you identify as a person with a disability?

1. Yes
2. No

Demographic Information for Workplace Discrimination Cases

If you experienced discrimination at work or connected to your career, we would like to know little more about your workplace and income. Sharing this information is optional and does not affect your ability to receive help from us.

When the discrimination/harassment happened, what industry did you work in?

When the discrimination/harassment happened, what was the name of your employer?

When the discrimination/harassment happened, approximately how much were you earning? Report by hourly wage or salary

Some of our legal network attorneys serve only low income clients. If you believe you may qualify for legal services for low income clients and would like to be connected to an attorney who serves such clients, please check 'yes'

- Yes
 No

Story Sharing

Sharing our stories is an important way to educate the public about sex discrimination and harassment in our lives. May we share your anonymized story in our advocacy and education efforts? We will not share your name, contact information, or the names of any organizations or individuals involved.

With these protections, can we share your story?

- Yes
 No

Please click the submit button below to send us your application.

Thank you for completing our online application. Please click the submit button below to send us your application.

We will contact you within 10 days to let you know if we can help you or if we need additional information. If you have any questions, please review our FAQs page or contact us at legalnetwork@nwlc.org.

To ensure that you receive that email please add legalnetwork@nwlc.org to your list of email contacts. Please also check your junk email for emails from legalnetwork@nwlc.org.

Appendix E: Current Online Request Form (continued)

Important information:

The Legal Network for Gender Equity and the TIME'S UP Legal Defense Fund are housed at and administered by the National Women's Law Center Fund LLC (NWLCF).

The NWLC has been fighting for gender justice for over 45 years. In 2016, the NWLC created the NWLCF and the Legal Network for Gender Equity to help people facing sex discrimination and harassment in education, the workplace, and health care connect with attorneys.

By providing this information to you, the National Women's Law Center (NWLC) and the National Women's Law Center Fund (NWLCF) are not becoming your attorneys. Filling in the intake form also does not mean that NWLC or NWLCF are becoming your attorneys. Providing information to NWLC or NWLCF will not create an attorney-client relationship unless NWLC/NWLCF expressly agrees to represent you.

The Legal Network for Gender Equity provides the names of lawyers for informational purposes only. By providing contact information of attorneys, the Legal Network for Gender Equity is not endorsing, approving, vouching for, or recommending the lawyers or groups listed. We cannot guarantee that any of the lawyers will agree to represent you or that if they do, you will have a positive result in your case. And it is important to note that just because a lawyer was successful in other cases.

You are responsible for meeting any filing deadlines associated with your legal claims.

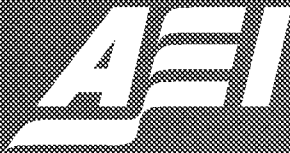
Endnotes

- ¹ Alianza Nacional de Campesinas "Dear Sisters" letter to Hollywood actors (November 10, 2017), available at <https://time.com/5018812/farmworkers-solidarity-hollywood-sexual-assault/> (last visited Aug. 30, 2020); TIME'S UP "Dear Sisters" letter to farmworkers (Dec. 21, 2017), available at <https://www.nytimes.com/interactive/2018/01/01/farts/02women-letter.html?mtrref=www.google.com&assetType=FIGWALL> (last visited Aug. 30, 2020).
- ² In the executive summary and throughout the report, we use the term "volunteered" to describe information that individuals provided without being prompted for information by a question on the online request form. By contrast, we use the term "said" to describe information that an individual provided in response to a specific question on the online request form.
- ³ NWLC, Frequently Asked Questions about Sexual Harassment in the Workplace (Nov. 2016), available at <https://nwlc.org/wp-content/uploads/2016/11/Sexual-Harassment-FAQ.pdf>. See more definitions in Appendix C.
- ⁴ Although people were not required to provide this information, many people self-identified their gender and race. An analysis was conducted on each of the variables coded for this report (see Appendix A for the full list of codes) to determine if statistically significant differences existed by gender or race. In many cases, because experiences were common for people across groups, differences between genders and racial groups were often not statistically significant; only those that were statistically significantly different at the $p < .05$ level are reported throughout this report.
- ⁵ See n. 4, *supra*.
- ⁶ Anyone requesting legal help who self-selected more than one race/ethnicity was categorized as "two or more races." For example, people who identified as Black and AAPI, as well as people who identified as white and Latino, were placed in this category. Therefore, it is important to note that the "two or more races" category encompasses people with a wide range of experiences, the nuances of which will not be captured by the analysis in this report.
- ⁷ Those who self-identified their race/ethnicity in their online request form closely match the U.S. workforce in 2018. According to annual Census data, 62% of workers were white, non-Hispanic, 18% were Latinx, 12% were Black, 7% were AAPI, and 1% were Native American. NWLC calculations of U.S. Census Bureau, 2019 Current Population Survey using IPUMS. In the annual Current Population Survey, respondents self-identify their race and separately self-identify whether they are of Hispanic, Latino, or Spanish origin. For more information, see <https://www2.census.gov/programs-surveys/cps/techdocs/questionnaires/Demographics.pdf>. Race/ethnicity data in the TIME'S UP Legal Defense Fund online form, however, was collected in one question and people were able to select as many options as applied to them.
- ⁸ States were divided into four regions based on U.S. Census Bureau divisions. See U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau, Census Regions and Divisions of the United States, available at https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf. In 2018, 37% of the U.S. workforce resided in the South, 24% in the West, 22% in the Midwest, and 17% in the North. NWLC calculations of U.S. Census Bureau, 2019 Current Population Survey using IPUMS.
- ⁹ U.S. Equal Employment Opportunity Commission, Select Task Force on the Study of Harassment in the Workplace 16 (June 2016), available at <https://www.eeoc.gov/select-task-force-study-harassment-workplace>.
- ¹⁰ U.S. Equal Employment Opportunity Commission, Select Task Force on the Study of Harassment in the Workplace 16 (June 2016), available at <https://www.eeoc.gov/select-task-force-study-harassment-workplace>.
- ¹¹ Jocelyn Frye, Center for American Progress, Not Just the Rich and Famous, available at <https://www.americanprogress.org/issues/women/news/2017/11/20/449139/not-just-rich-famous/>.
- ¹² Paige L. Sweet, The Sociology of Gaslighting, *American Sociological Review*, Vol. 85, No. 5 (2019), pp.851-875.
- ¹³ C. Garcia Moreno, A. Riechler-Rosser, Violence Against Women and Mental Health, *Indian Journal of Medical Research*, Vol 142, No. 5 (Nov 2015), 629-630.
- ¹⁴ Holly Kearn, Stop Street Harassment, The Facts Behind the #MeToo Movement: A National Study on Sexual Harassment and Assault (Feb. 2018), available at <http://www.stopstreetharassment.org/wp-content/uploads/2018/01/Full-Report-2018-National-Study-on-Sexual-Harassment-and-Assault.pdf>.
- ¹⁵ Clare McGlynn and Erika Rackley, Image-Based Sexual Abuse, *Oxford Journal of Legal Studies*, Vol. 37, No. 3 (2017), pp. 534-561.
- ¹⁶ Heather McLaughlin, Christopher Uggen, Amy Blackstone, The Economic and Career Effects of Sexual Harassment on Working Women, *Gender and Society*, Vol. 31, No. 3 (June 2017), 333-358.
- ¹⁷ People who withdraw money from retirement accounts such as 401ks or Individual Retirement Accounts (IRAs) before age 59½ are subject to a 10% penalty plus income tax. However, this 10% penalty was temporarily waived for many people in tax year 2020 as part of the CARES Act. See <https://www.irs.gov/newsroom/major-changes-to-retirement-plans-due-to-covid-19>.
- ¹⁸ For information about NWLC's policy advocacy on workplace sex harassment and related resources, visit: <https://nwlc.org/issue/sexual-harassment-in-the-workplace/>



TIME'S[™]
LEGAL DEFENSE FUND
UP

The TIME'S UP Legal Defense Fund is housed and administered by the National Women's Law Center Fund, LLC. Questions? Contact us at legalnetwork@nwlcf.org



A Better Bargain

HOW NONCOMPETE REFORM CAN
BENEFIT WORKERS AND BOOST
ECONOMIC DYNAMISM

John W. Lettieri

DECEMBER 2020

A M E R I C A N E N T E R P R I S E I N S T I T U T E

Executive Summary

Noncompete agreements hinder the mobility of roughly 20 percent of the American workforce and reduce overall dynamism in the economy. Once reserved for senior executives and those possessing valuable trade secrets, noncompetes have become a pervasive feature of the US labor market, including low-wage fast-food workers, janitors, and warehouse workers.

Recent years have seen growing bipartisan momentum for noncompete reform in state legislatures, Congress, and even President-elect Joe Biden's policy platform. Indeed, 2019 and 2020 were banner years for the noncompete reform movement, leading to several major state-level reforms and the first tangible signs of

bipartisan federal interest in using noncompete reform as an avenue to pursue higher wages, higher rates of entrepreneurship, and a stronger innovation sector.

This report surveys recent state and federal efforts to rein in noncompetes, including a close examination of noteworthy successes in states such as Maine, Virginia, and Washington. Additionally, it outlines key principles and a menu of legislative options to guide policymakers as they consider ways to design noncompete reform to achieve their desired economic benefits. It also includes a comprehensive appendix of state-level noncompete legislation for 2019–2020.

A Better Bargain

HOW NONCOMPETE REFORM CAN BENEFIT WORKERS AND BOOST ECONOMIC DYNAMISM

By **John W. Lettieri**

Today, roughly 20 percent of the American workforce cannot take a better job in the field of their choice—regardless of higher pay, better benefits, improved job satisfaction, or the many other factors that influence an employee’s career decisions.¹ The reason: They are bound by a covenant not to compete, or a “noncompete” agreement. These agreements take various forms but generally prohibit a former employee from starting or working for a “competing” business in the same industry as their former employer for a certain period and in a certain geographic area.²

Once reserved for senior executives and those possessing valuable trade secrets, noncompete agreements are now pervasive across a broad range of sectors and a wide variety of workers, including low-wage fast-food workers, janitors, and warehouse workers. In the absence of many legal restraints, noncompetes have become a reliable tool for employers looking to retain talent, stifle would-be competitors, and keep labor costs down.

However, a mounting body of research points to the harms noncompetes impose on individual workers, labor markets, and the economy overall. As a result, interest in noncompete reform is now building in state legislatures and Congress. Indeed, 2019 and 2020 were banner years for the noncompete reform movement and could foreshadow an even greater acceleration of state and federal legislative activity in the years ahead. One reason for this is that, despite deep political divisions on many economic issues, there is growing bipartisan consensus on the benefits of reining in noncompetes.

In this report, I survey recent noncompete reform activity in Congress and state legislatures nationwide and detail the various approaches being pursued, with particular emphasis on successful legislative efforts in key states in 2019 and 2020. I then outline the core principles that should guide policymakers in evaluating how to limit noncompetes at the state or federal levels and close by distilling the range of potential policy features into a three-tier menu of legislative options.

How Noncompete Agreements Affect the Economy and Labor Market

Healthy labor markets depend on firms competing vigorously for workers and workers’ ability to market their skills to employers freely. Likewise, a dynamic economy depends on the productivity-boosting exchanges that happen when individuals collaborate or apply their collection of ideas and experiences in new contexts. The simple process of workers switching firms is crucial to facilitating innovation and helping know-how proliferate throughout the economy. And this process appears to be under threat. Recent research from economists at the University of Chicago finds that up to 70 percent of the decline in measures of economic dynamism can be attributed to a dramatic slowdown in the rate at which knowledge diffuses across the economy.³

Noncompete agreements, much like occupational-licensing laws and “no-poach” agreements, erect barriers to worker mobility and dampen the US

economy. Research finds they reduce job switching,⁴ lessen wage growth,⁵ and prevent competition from new entrants in the marketplace.⁶ Perhaps for these reasons, employers increasingly incorporate noncompetes into their employment agreements and more frequently take legal action to enforce them.⁷

An estimated one in five American workers is currently covered by a noncompete, and nearly twice as many have signed one at some point in their careers.⁸ Noncompetes are relatively more common among higher-wage workers but numerically more common among lower-wage employees. According to one study, 33 percent of workers earning under \$40,000 report signing such an agreement during their careers.⁹ Much of the recent interest in noncompetes has been driven by the growing awareness that employers are requiring them of fast-food workers, janitors, camp counselors, and other low-wage workers. Such examples indicate that employers' justifications for noncompetes, such as protecting valuable trade secrets, are often little more than a pretext to suppress wages and limit the mobility of already vulnerable workers.

Why is limiting worker mobility problematic? One key reason is that job-hopping—especially early in a worker's career—is associated with stronger lifetime earnings. Research finds that strict enforcement of noncompetes is associated with reduced job-to-job mobility, lower wages, and weaker rates of firm formation. Noncompetes also appear to exacerbate racial and gender wage gaps by exerting much larger wage effects on female and black employees than on white men.¹⁰ Additional research identifies noncompetes as one explanatory factor in the gender gap in entrepreneurship.¹¹ Moreover, noncompetes not only affect the individual workers bound by them but also have a chilling effect on the entire labor market.¹² If human capital is the lifeblood of an economy, noncompetes are a clog in the arteries.

Noncompetes also harm employers. By diminishing the healthy churn of workers throughout the labor market, restrictive covenants reduce the supply of available workers for businesses seeking to grow. This could be especially detrimental to younger and smaller enterprises that already face

disadvantages against larger incumbent businesses. The detrimental effect of noncompetes on new businesses—stifling would-be entrepreneurs and limiting the pool of much-needed talent for startups—should be of particular concern to policymakers given that the US business startup rate remains mired near historically low levels.¹³

The vast majority of noncompete agreements are not subject to any negotiation between the employer and employee, suggesting that the employee is unlikely to receive any benefits in return for their signature. A large share of these agreements are presented for signature only after the employee has already accepted the job offer—often on the first day of work.¹⁴ Employers frequently exploit workers' lack of knowledge and resources when crafting noncompetes. For example, employers commonly request that workers sign noncompetes even in states where they are completely unenforceable—and workers nevertheless sign the agreements assuming they are valid. Likewise, employers often craft extremely broad provisions knowing that employees generally lack both an understanding of what is enforceable and the wherewithal to challenge the terms in court.¹⁵

Guideposts for Reforming the Use of Noncompetes

As discussed later in the report, mounting evidence of the harm caused by noncompetes has fueled a bevy of legislation aimed at limiting their use in states throughout the country. Such efforts are rarely broad-based, however, and tend to lack the full variety of features necessary to mitigate the negative effects of noncompetes on workers and the broader economy. A more standardized and holistic approach is needed.

Before examining specific legislative options in greater detail, it is important to establish the core objectives that should guide state and federal reform. Every reform effort, no matter how far-reaching, should follow these guideposts.

Presume the Right to Compete. The law should presume a worker's right to compete freely in the job

market. From there, lawmakers can evaluate which—if any—narrow circumstances should justify restrictions on worker mobility. Today, we largely find the opposite in most states: the presumption that former employers should be allowed right of refusal over a worker’s future job options, with some narrow exceptions.

Ensure Transparency. Many of the negative effects of noncompetes can be reduced simply by ensuring greater transparency and improving workers’ awareness of their bargaining position. As noted earlier, employers easily and routinely exploit the current lack of transparency associated with noncompetes. For example, without prior-notice requirements, employers will often wait to present the noncompete until an employee’s first day of work—once alternative employment options have been foreclosed. Rules governing noncompetes should be clear and easy to administer, and employees should be given adequate notice and explanation before being asked to sign away future job opportunities.

Establish Disincentives for Overuse. There are few reasons for an employer *not* to require noncompetes of its employees—even overly broad and unenforceable ones and ones that cover employees who have no specialized skills or trade secrets. This is a simple matter of incentives. Therefore, an obvious way to limit the overuse of noncompetes is to make them cost something to the employer. Examples include “garden-leave” laws, which require employers to provide continued compensation while the agreement is enforced and levy stiff fines for employers that knowingly request signature in states where noncompetes are unenforceable.

Limit the Pool of Eligible Workers. Most states have no restrictions on the kinds of workers who can be bound by a noncompete. Policymakers have many options for narrowing the pool of eligible workers by industry, wage level, or educational attainment so their use is reserved only for senior executives and other top talent in strong bargaining positions.

Limit the Scope of Agreements. Even when policymakers see a valid use for noncompetes under certain circumstances, most agree the scope of such agreements should be limited in various ways (e.g., duration or geography). At a minimum, employers should be required to draft noncompetes to be enforceable in the state where they are signed. While this may sound obvious, several states do the opposite and instead require their courts to redraft an unenforceable noncompete so it can be enforced in the employer’s interest.

Policy Recommendations: Organizing Legislative Options from Broad Consensus to Comprehensive Reform

With the above guideposts in mind, I organize a spectrum of potential legislative options into three tiers, starting with the most basic and graduating to the most comprehensive. The options below are generally applicable at both the state and federal levels.

Broad Consensus Reform. There is broad agreement that policymakers should, at minimum, place reasonable limits on how and when noncompete covenants can be used. The following recommendations build on the areas of strongest consensus regarding what kind of limitations are necessary, drawing from various state-level policy initiatives and the strongest empirical evidence. If implemented together, they would set basic parameters around the use of noncompetes, with particular emphasis on protections for lower-wage and more vulnerable workers.

Exempt Low- and Moderate-Wage Workers. The first and most obvious reform should be to exempt workers with the least natural bargaining power—as evidenced by educational attainment and income level—from ever being asked to sign a noncompete. Noncompetes should therefore be disallowed for any worker without a bachelor’s degree or making less than the state or national median income, whichever is greater. Lawmakers must be careful to avoid ambiguity when defining which workers are exempted and

which forms of compensation count toward the earnings threshold.

Ban Unenforceable Signatures. Reform efforts often focus on limiting the enforcement of noncompetes. But simply restricting enforcement is not enough to mitigate their harmful effects; the evidence suggests that, once signed, even *unenforceable* noncompetes have a chilling effect on worker mobility.¹⁶ For example, research finds that workers in states that do not enforce noncompetes, such as California, sign them at roughly similar rates to workers in states that do enforce them.¹⁷ Employers should therefore be prohibited from requesting signature of a noncompete containing terms that are clearly unenforceable, in states where they are unenforceable or if the worker in question is part of an exempted class (e.g., low-wage workers). Failure to comply should be accompanied by fines and other penalties.

Discourage Unreasonable Scope. Policymakers should create meaningful disincentives for the use of overly broad or unenforceable noncompete provisions. To start, any agreement that includes such provisions should be rendered completely void—as opposed to rewritten by the court to make it enforceable—even if it contains otherwise enforceable provisions. Employers that are found by a court to have drafted overly broad provisions should be required to cover the cost of the employee’s litigation, any lost wages, and any litigation expenses incurred by the new employer. These rules should apply to existing agreements and any agreements signed after the legislation’s effective date.

Require Prior Notice. Employers should be required to notify job candidates of their intention to require a noncompete during the interview process and present the agreement when the formal job offer is made—not after the employee has accepted the job. Employers should disclose that a noncompete will be required when advertising for a position. Additionally, employers requesting signature of a noncompete agreement should inform the prospective employee of applicable state and federal laws and allow adequate time for

the candidate to consider the terms of the agreement before deciding.

Limit Duration. The enforcement period for noncompetes should be limited by statute to no more than one year.

Disallow in the Event of Termination. If an employee is terminated, the noncompete should immediately be rendered void. This rule should apply to all existing and future agreements.

Moderate Reform. In addition to or in place of the modest reforms noted above, states or federal lawmakers can go further in limiting noncompetes’ applicability and broadening the pool of workers exempted from them.

Limit Noncompetes to Top Wage Earners. Since the mobility of workers with greater skills and education is especially important to the economy’s health, policymakers may wish to exempt a wider share of the workforce from the restrictions imposed by noncompetes. Accordingly, noncompetes could be disallowed for all workers except those whose income is in the top 5 percent of the state or country, whichever is greater. This would ensure that only those employees with the strongest bargaining power and the wherewithal to negotiate the terms of the agreement with their potential employers are presented with a noncompete.¹⁸

Limit Eligibility Within Firms. Since some firms have many high-wage earners in knowledge-intensive activities, policymakers may also want to limit how pervasively noncompetes can be used *within* a given firm. For example, some have suggested that firms over a certain size—say, 50 employees—be allowed to cover no more than 5 percent of their workforce with a noncompete.¹⁹ This approach, combined with the limit to only high-wage earners, would help ensure that firms carefully consider the circumstances under which to require noncompetes and that the workers covered are truly the ones most valuable to the firm’s success and therefore most capable of negotiating the terms of an agreement.

Nullify Existing Agreements on Certain Workers. Instead of merely limiting the *future* use of noncompete agreements, reformers should consider the more immediate economic benefits of nullifying existing noncompete agreements for low- and moderate-wage workers. Specifically, any noncompete covering a worker who lacks a bachelor's degree or makes below the state or national median income could be immediately nullified, in addition to other prospective policy changes.

Require Garden Leave. Firms could be required to provide compensation to any former employee abiding by the terms of a noncompete, a concept known as “garden leave.” Specifically, former employees should at minimum receive the equivalent of 50 percent of their highest previous annual salary while the agreement is being enforced. Garden leave assigns a tangible cost to an employer seeking to enforce a noncompete, which, in turn, reduces the likelihood of frivolous use. States with garden-leave laws include Massachusetts and Oregon.

Ban by Industry. States could restrict noncompetes on an industry basis to improve wages and mobility in high-value sectors such as technology,²⁰ in sectors where there is no plausible trade secret to protect, such as cosmetology, lawn service, or hospitality; or in sectors such as health care, in which limiting the availability of specialists and service providers has serious implications for consumer health and safety—especially in the wake of the COVID-19 pandemic.

Exempt Employees of New and Small Businesses. Reform efforts could be targeted to consider the types of firms involved. For example, given the widely recognized need to boost entrepreneurship and business dynamism, noncompetes could be rendered void for covered workers who find employment at a startup (a firm that started less than three years ago, for example). Newer firms typically lack the resources to fight a noncompete and may be particularly disadvantaged by the reduction in the supply of available skilled workers induced by noncompetes.

Comprehensive Reform. In light of the significant economic harms and market distortions caused by noncompetes, many policymakers believe the best approach is to disallow the use of noncompetes in all but the most narrow circumstances. Comprehensive reform—in addition to increasing the likely upsides for employees, entrepreneurs, and the economy as a whole—would have the added value of regulatory simplicity for workers and firms alike, making the law easier to interpret and enforce. This is especially true of federal reform efforts, which would eliminate the regulatory uncertainty inherent in the patchwork state-by-state approach.

Generally Prohibit Noncompetes with Few Exceptions. Rather than narrow limitations tied to wage, industry, or skill level, noncompetes could be completely disallowed except when the parties are on level footing, such when a business owner sells to an acquirer and agrees not to start a competing business as part of the terms of sale. States that already take this approach include California, whose innovation ecosystem demonstrates the profound benefits of allowing knowledgeable workers and would-be entrepreneurs to deploy their skills freely in the labor market.

Blanket Retroactive Nullification. All existing noncompetes that do not conform to the exceptions noted above (e.g., sale of a business) would become immediately voided. (This was the approach taken in the recent federal legislation, as discussed later in this report.)

Further Considerations. As is clear from these options, noncompete reform can take many shapes and include an array of features. Short of an outright ban on noncompetes, reform efforts should rely on a combination of core ingredients that work together to mitigate the harmful effects of noncompetes throughout the labor market and economy. For example, the benefits of exempting low-wage workers alone are likely to be far weaker than an exemption paired with strong transparency and prior-notice requirements.

With so many potential legislative options, policymakers must be mindful of why reform is necessary and what it can accomplish if properly crafted. There is perhaps a natural temptation to focus solely on the lowest-wage workers for whom the use of a noncompete seems most abusive and frivolous. Achieving the full promise of noncompete reform, however, also requires enabling skilled and knowledgeable workers to better deploy their talents and ideas throughout the economy, including by starting new firms and bringing innovations to market. Exemptions for low-wage workers alone will fall short of reaching this crucial goal.

The Current Landscape of Noncompete Reform

The years 2019 and 2020 were unusually active periods for noncompete legislation at the state and federal levels. At least 10 states passed laws to limit the use of noncompetes in some form. Meanwhile, two Republican Senators, Marco Rubio (R-FL) and Todd Young (R-IN), introduced legislation that cemented noncompete reform as an issue of bipartisan interest at the federal level. Noncompetes even factored into the policy platforms of most of the major 2020 Democratic presidential candidates, including the nominee, President-elect Joe Biden. What is the state of reform nationwide, and what are the common themes that link reform efforts across the country?

State Legislation. The treatment of noncompetes can differ dramatically among states. Only three states—California, North Dakota, and Oklahoma—broadly prohibit enforcement of noncompete agreements. Most others have permissive rules that give employers broad latitude in how they use noncompetes.

But this is beginning to change. At least 29 state legislatures considered legislation related to noncompetes in the 2019 legislative session. (See Appendix A.) Of the at least 40 bills introduced in 2019, all but two were intended to limit the scope or enforceability of noncompetes. The reform bills run the gamut

from full statewide bans to protections for low-wage workers or workers in certain industries and quite often enjoyed bipartisan support. Texas moved in the other direction by passing a law making it more difficult for individuals covered by a noncompete to seek legal recourse. North Dakota, meanwhile, modestly expanded existing exceptions to the state's general prohibition of noncompete provisions.

Eight states enacted some kind of new limitation on noncompetes in 2019: Florida, Maine, Maryland, New Hampshire, Oregon, Rhode Island, Utah, and Washington. Of those, Maine and Washington passed the most far-reaching legislation. Meanwhile, Florida took a sector-specific approach with reforms to improve the accessibility and affordability of physician specialists in areas controlled by a single health care provider.

In Maine, a state grappling with population loss and a dearth of available workers, lawmakers enacted a broad set of provisions aimed at protecting the most vulnerable workers and improving transparency related to noncompetes.²¹ Under the new law, employees making at or below 300 percent of the federal poverty level would be prohibited from signing a noncompete agreement. The law imposes several prior-notice requirements, including requiring employers to disclose in job postings when a noncompete will be required.

Maine's law also enacts a ban on so-called no-poach agreements among employers under which the parties, such as two fast-food establishments from the same restaurant chain, agree not to recruit or hire each other's employees or former employees. No-poach agreements, which disproportionately affect low-wage workers and are particularly common in the franchise sector,²² have come under greater scrutiny in recent years.²³

Washington state's new law sets a number of preconditions for the enforceability of noncompetes in the state.²⁴ First, a noncompete will only be enforceable if the employee earns over \$100,000 annually—a far higher threshold than any state aside from those that ban enforcement of noncompetes entirely. (For independent contractors, the threshold is \$250,000 annually.) Employers must provide notice of the

terms of the agreement when or before a prospective employee accepts a job offer. Furthermore, employers must compensate any terminated worker subject to the enforcement of a noncompete, and agreements longer in duration than 18 months are generally presumed to be unreasonable and unenforceable.

By combining broad earnings-based exemptions, prior notice, garden leave, strict time limitations, and more, Washington's law is arguably the most comprehensive example of recent state-level reform. Nevertheless, it provides limited benefits to the higher-wage workers who are essential to the state's knowledge-driven economy.

Attempts to limit physician noncompetes have become one of the most common varieties of reform legislation in recent years—a trend that could accelerate even further in response to the strain placed by the COVID-19 crisis on the health care system. While some initiatives aim to enact sweeping prohibitions on physician noncompetes, Florida's 2019 law specifically seeks to foster more competition among health care providers in certain geographic areas. It made noncompetes void and unenforceable for specialists practicing in a county wherein all physicians practicing that same specialty are employed by a single entity—something of particular concern in rural areas with fewer health care providers. The state explicitly did so on the grounds that, in such cases, noncompetes harm consumers by making health care less accessible and more expensive.²⁵ The noncompetes in question must remain unenforceable for three years after a competing employer enters the area offering the same services.

Reform momentum continued into 2020, with noncompete-related legislation introduced in at least 18 states. While legislative success was more limited than in the previous year, Indiana and Virginia enacted laws to rein in noncompete agreements. Louisiana, meanwhile, bucked the national trend by expanding the scope of noncompete enforcement.

Virginia became the latest state to exempt low- and moderate-wage workers for exemption from noncompetes. As of July 1, 2020, employers in the commonwealth may not “enter into, enforce, or threaten to enforce a covenant not to compete with any low-wage

employee.”²⁶ The law employs a fairly broad definition of low-wage, which applies to any worker whose average weekly earnings over the year before termination were less than the average weekly wage statewide or any independent contractor paid an hourly rate less than the previous year's median hourly wage in the state. Unfortunately, like other state-level noncompete statutes, Virginia's new law suffers from ambiguities that risk undermining its overall effectiveness. For example, the law is unclear on what kinds of compensation count toward its earnings threshold.

Indiana, meanwhile, became the latest state to place new conditions on physician noncompetes. Noncompete agreements made on or after July 1, 2020, must include provisions ensuring a physician is given access to medical records of recent patients and any notices sent to such patients regarding the physician's departure. Former employers are also required to provide inquiring patients with the physician's updated contact information and location. Perhaps most importantly, the Indiana law gives physicians the right to “purchase a complete and final release” from the noncompete agreement “at a reasonable price.”²⁷ The law, however, provides no details on how “a reasonable price” is to be determined.

Taken together, one-fifth of all states enacted new limitations on noncompete agreements in 2019–20, and a clear majority of states saw at least one noncompete reform bill put forward for consideration. This flurry of recent legislative activity builds on a series of noteworthy reforms stretching back more than a decade. In particular, Oregon's noncompete law, passed in 2007 and taking effect in 2008, became a model for much of what was to follow. Oregon's far-reaching statute was well ahead of its time in enacting protections for low-wage workers alongside a range of measures aimed at improving transparency and narrowing what is enforceable under state law.²⁸

The centerpiece of the law was a provision to void all new noncompetes for workers earning less than the median income for a household of four. (It also covered hourly workers and employees in certain occupations.) The law also made Oregon the first state to codify the concept of garden leave. Other

important elements included prior-notice standards and a rule limiting enforceable agreements to 18 months or less.

Policymakers at the time were motivated by a belief that low-wage workers were simply not in a position to bargain over the terms of a noncompete and were unlikely to understand what terms were enforceable.²⁹ Noncompetes therefore, in theory, constrained low-wage workers' job mobility without rewarding them for trading off future job opportunities. Recent research on noncompete bans for hourly workers in Oregon finds significant evidence confirming this.³⁰ Researchers Michael Lipsitz and Evan Starr find an average 2–3 percent increase in hourly wages, a 12–18 percent increase in monthly job mobility, an increase in the proportion of salaried workers in the state labor force, and a decrease in the likelihood of being unemployed for affected workers 10 years after passage.

A decade after Oregon's pathbreaking reform, Illinois passed a modest revision to its noncompete laws in 2017.³¹ The state targeted the lowest end of its income distribution by prohibiting noncompetes for any worker earning less than the equivalent of \$13 per hour and immediately voiding any prior agreement in effect covering such an employee.

Another noteworthy legislative success occurred in Massachusetts in 2018.³² Massachusetts is particularly interesting given the strength of its knowledge economy, which has made it a center of venture capital and entrepreneurship. Many observers point to noncompetes as a key reason that Silicon Valley has outpaced Boston as an innovation hub. (California does not enforce them; Massachusetts has historically been a strict enforcer.) The Massachusetts statute, which applies only to agreements signed after October 1, 2018, bans noncompetes entirely for low-wage workers who are classified as nonexempt under the Fair Labor Standards Act, employees terminated without cause, and undergraduate or graduate students in internships.

Perhaps most notably, the law made Massachusetts only the second state to institute garden leave, requiring employers to compensate former employees during the restricted period on a pro rata basis

equal to 50 percent of the employee's highest annualized base salary during the final two years of employment. As an alternative to garden leave, employers can provide "other mutually-agreed upon consideration," a term left conspicuously undefined in the statute.³³ The law also includes prior-notice standards and imposes a one-year limitation on the duration of noncompetes, but maintains the employer-favorable standard of allowing courts to rewrite an overly broad agreement to make it enforceable. It also provides exceptions related to the sale of a business or the dissolution of a partnership.

While most states have targeted reforms to low- and moderate-wage workers, one state explicitly focused on entrepreneurship and innovation as the driving factors behind its reform efforts. In 2015, Hawaii banned noncompetes for workers in its technology industry to attract and retain highly skilled workers and compensate for its unique geographical limitations.³⁴ Research found significant increases in mobility and wages for Hawaii's newly hired tech workers following the ban.³⁵ While other states have yet to follow its lead in targeting innovation-intensive industries for reform, Hawaii's experience holds potentially crucial lessons for other states looking to attract coveted human capital and establish themselves as "rise of the rest" technology hubs. It also undermines the notion that highly skilled workers, such as those found in the high-tech sector, derive significant benefits from signing noncompetes.

Federal Action. Noncompete reform offers a relatively simple and inexpensive place to start for policymakers struggling to find consensus on ways to improve worker mobility, increase business formation, strengthen innovation, and boost wages.

But why should the federal government take action on an issue that has until now been left to states? Noncompetes are clearly an issue of employment law, for which the federal government commonly sets basic standards governing the relationship between employers and employees. In addition, even though employers exercise little self-restraint in how they deploy noncompete agreements, states have generally set few limitations of their own. But none of this

would be of much consequence without the growing evidence of noncompetes' negative effects on wages, job mobility, entrepreneurship, the gender and racial income gaps, and more—all of which underscores the clear and compelling need for Congress to establish guardrails. Without federal policy, a confusing patchwork of policy harms workers and employers and fails to serve our national economic interests.

Unlike state action on noncompete reform, interest among federal policymakers was until recently limited to Democrats. That changed in 2019 when Rubio introduced the Freedom to Compete Act.³⁶ It targeted only the small share of workers who are nonexempt under the 1938 Fair Labor Standards Act, those making less than \$23,600 a year. The legislation would prevent employers from entering into or enforcing a noncompete with such employees, but includes few other reforms. It includes no penalties for employers that violate the law, no requirement of prior notice or transparency for employers that intend to request a noncompete, and no limitations on the duration or geographic scope of noncompetes. While the bill is exceptionally narrow in scope, its significance should not be underestimated: it marked the first tangible sign that bipartisan federal reform to noncompete law is possible.

That first sign of bipartisan interest was soon followed by another. On March 7, 2019, a bipartisan group of six senators—Tim Kaine (D-VA), Chris Murphy (D-CT), Rubio, Elizabeth Warren (D-MA), Ron Wyden (D-OR), and Todd Young (R-IN)—formally requested that the US Government Accountability Office (GAO) initiate a review of noncompete agreements' effects “on workers and the economy as a whole.”³⁷ After noting growing concerns about how the extensive use of noncompetes could be harming the economy's growth and vitality, they asked the GAO to assess three specific questions to inform future congressional activity.

1. What is known about the prevalence of non-compete agreements in particular fields, including low-wage occupations?

2. What is known about the effects of non-compete agreements on the workforce and the economy, including employment, wages and benefits, innovation, and entrepreneurship?
3. What steps have selected states taken to limit the use of these agreements, and what is known about the effect these actions have had on employees and employers?³⁸

After leading the letter to the GAO, Young and Murphy introduced a sweeping reform bill in October 2019—the first bipartisan noncompete legislation on record.³⁹ Their Workforce Mobility Act would prohibit the signature, enforcement, and threat of a noncompete—and render all existing agreements unenforceable.⁴⁰ Exceptions would be made for the sale of a business and the dissolution of a partnership, with the understanding that the parties involved have equal bargaining power and full awareness of what they are negotiating. The legislation also includes public-awareness and transparency provisions aimed at ensuring workers are aware of the law's effects. Notably, the bill includes “teeth” in the form of enforcement authorities to the Federal Trade Commission and the Department of Labor (including civil fines of \$5,000 per week for those in violation) and a private right of action for individuals to pursue damages.

A companion version of the Workforce Mobility Act was introduced in the House in January 2020. Like its Senate counterpart, the House bill was introduced on a bipartisan basis, with Reps. Scott Peters (D-CA) and Mike Gallagher (R-WI) as its lead sponsors.

While neither bill gained traction in the 116th Congress, they represent important milestones on the path to federal reform, providing a bipartisan basis for future legislative efforts to curtail restrictive covenants. Such efforts could soon arrive. President-elect Biden has promised to work with Congress to completely eliminate noncompete agreements except for those “absolutely necessary to protect a narrowly defined category of trade secrets” and to “outright ban all no-poaching agreements.”⁴¹

Conclusion

Workers should be free to seek better jobs and compete in the labor market without permission from their former employers. Employers should be rewarded for winning the competition for talent—not for holding workers hostage. And policymakers should be relentlessly focused on encouraging competition, healthy risk-taking, and worker mobility.

The pervasive use of noncompetes is proof the United States has fallen short on all three fronts. Fortunately, this is not a problem that requires an elaborate new program or endless taxpayer dollars to solve. Policymakers simply need to rediscover the benefits of allowing workers and employers to compete in an open and transparent labor market.

The idea of Congress enacting national restrictions on the use of noncompete covenants would have sounded far-fetched only a few years ago. But with bipartisan legislation in both chambers of Congress, the support of an incoming president, and the interest of prominent Republicans in building a more worker-centric Party,⁴² the prospects of federal legislation—while by no means certain—have never looked better. Meanwhile, what began as a trickle of state-level action over a decade ago has evolved into a flood of bills in dozens of state legislatures. The momentum is unmistakable—and likely irreversible, as each new legislative success makes

the next one easier to achieve. The challenge now is to evolve to a more coherent and comprehensive approach to reform that delivers stronger benefits to workers, entrepreneurs, and the broader economy. In any event, the rising tide of reform means this is one area of policy that is almost certain to become friendlier to workers, more embracing of competition, and more conducive to economic dynamism in the years ahead.

Acknowledgments

The report was made possible through a generous grant by the Rockefeller Foundation. The author wishes to thank David Wilde and Kawit Promrat for their outstanding work as research assistants on this project. Special thanks to Kenan Fikri for reviewing early drafts and providing his usual keen insights.

About the Author

John W. Lettieri is the cofounder, president, and CEO of the Economic Innovation Group, a bipartisan research and public policy organization based in Washington, DC. He was a visiting fellow at AEI from 2018 to 2020.

Appendix A. 2019 Legislation

State	Bill Number	Title	Sponsor(s)	Summary	Type	Status
Arkansas	HB 1068	An Act to Repeal the Ability to Enforce a Covenant Not to Compete Agreement; and for Other Purposes	Reps. Brian Evans (R) and Ricky Hill (R)	Bans noncompetes in the state by repealing Arkansas Code § 4-75-101, which was passed in 2015 to strengthen noncompete enforcement.	Limit	Not enacted
Connecticut	SB 377	An Act Prohibiting the Use of Noncompete Clauses in Physician Employment Contracts	Sen. Heather Somers (R)	Bans noncompetes in physician employment contracts.	Limit	Not enacted
Connecticut	HB 6914	An Act Concerning Non-compete Agreements	Rep. Josh Elliott (D)	Sets an income threshold below which companies are not allowed to force employees to sign a noncompete.	Limit	Not enacted
Connecticut	SB 1033	An Act Concerning Non-compete Agreements in the Blockchain Technology Industry	Commerce Committee	Prohibits the use of noncompete agreements in the blockchain technology industry.	Limit	Not enacted
Florida	SB 882 HB 443	An Act Relating to Restrictive Covenants; Creating s.542.336, F.S.; Providing That Certain Restrictive Covenants Are Void and Unenforceable for a Specified Period; Providing an Effective Date	Committee on Commerce and Tourism and Sen. Joe Gruters (R)	Makes physician noncompetes void and unenforceable for specialists in a county wherein all physicians practicing that specialty are employed by a single entity. Such noncompetes will remain void for three years after the date on which a competing entity begins offering similar specialty services in the county.	Limit	Date signed: June 26, 2019 Date effective: July 1, 2019
Georgia	HB 81	A Bill to Be Entitled an Act	Rep. Todd Jones (R)	Prohibits the use of noncompete agreements with information technology employees.	Limit	Not enacted
Hawaii	HB 1059 SB 328	Relating to Fair Employment Practices	Sens. Stanley Chang (D), Mike Gabbard (D), Gil Keith-Agaran (D), Karl Rhoads (D), and Russell Ruderman (D) Rep. Joy San Buenaventura (D)	Prohibits noncompete agreements for low-wage workers earning less than \$15 per hour.	Limit	Not enacted
Illinois	HB 2565	An Act Concerning Employment	Rep. Ann Stava-Murray (D) with 16 cosponsors (16 Democrats)	Bans noncompete agreements entirely by amending the law that bans noncompetes for low-wage employees to extend to all employees.	Limit	Not enacted

Illinois	HB 2328	An Act Concerning Regulation	Rep. André Thapedi (D)	Provides that noncompetes between hospitals and physicians may not contain any provision to restrict a physician's ability to leave employment with the hospital or hospital affiliate and immediately continue to practice in the same field of medicine in the same geographic area.	Limit	Not enacted
Indiana	SB 348	A Bill for an Act to Amend the Indiana Code Concerning Labor and Safety	Sen. J. D. Ford (D)	Prohibits employers from requiring or enforcing noncompetes against workers earning no more than \$15 per hour.	Limit	Not enacted
Indiana	HB 1357	A Bill for an Act to Amend the Indiana Code Concerning Professions and Occupations	Reps. Robert Morris (R), Chris Judy (R), and Doug Miller (R)	Prohibits hospitals from imposing or enforcing noncompetes against their employee physicians.	Limit	Not enacted
Louisiana	SB 177	Physicians: Provides Relative to Prohibited Restraint of Certain Professions	Sens. Bodi White Jr. (R) and Dan Claitor (R)	Prohibits most noncompetes for physicians and nurses.	Limit	Not enacted
Maine	HB 733	An Act to Keep Workers in Maine	Rep. John Schneck (D) with 9 cosponsors (9 Democrats)	Enacts a range of reforms, including: <ul style="list-style-type: none"> • Prohibiting noncompetes for an employee earning wages at or below 300% of the federal poverty level • Prior notice requirements to ensure prospective employees are aware that a noncompete will be required and that adequate time is given to review and negotiate the terms • Civil penalties of a minimum of \$5,000 • A waiting period before the non-compete takes effect • A ban on no-poach agreements between employers under which the parties agree not to recruit or hire each other's employees or former employees 	Limit	Date signed: June 28, 2019 Date effective: September 18, 2019
Maryland	SB 328 HB 38	An Act Concerning Labor and Employment—Non-compete and Conflict of Interest Clauses	Sen. William Smith Jr. (D) with 14 cosponsors (14 Democrats) Delegate Alfred Carr (D)	Bans noncompetes for any employee earning less than or equal to \$15 per hour or \$31,200 annually.	Limit	Date signed: May 25, 2019 Date effective: October 1, 2019

Massachusetts	SB 1083	An Act Relative to Banning Noncompetition Agreements in the Commonwealth	Sen. Patricia Jehlen (D), Reps. Elizabeth Malia (D), Denise Provost (D), and Tommy Vitolo (D)	Bans all noncompete agreements in the state.	Limit	Not enacted
Michigan	HB 4874 SB 483	Michigan Antitrust Reform Act	Rep. Marie Manoogian (D) with 32 cosponsors (32 Democrats)	Requires that employees be given written notice of a noncompete requirement and bans the use of noncompetes for low-wage employees.	Limit	Not enacted
Minnesota	HF 557 SF 350	A Bill for an Act Relating to Health; Protecting Physician-Patient Relationship by Prohibiting Noncompete Agreements; Proposing Coding for New Law in Minnesota Statutes, Chapter 145	Reps. Alice Mann (D), Steve Elkins (D), and Liz Olson (D) Sens. Scott Jensen (R), Jim Abeler (R), Mike Goggin (R), Rich Draheim (R), and Matt Klein (D)	Prohibits hospitals from imposing or enforcing noncompetes against their employee physicians.	Limit	Not enacted
Missouri	HB 331	An Act to Amend Chapter 431, RSMo, by Adding Thereto One New Section Relating to Covenants Not to Compete	Rep. Doug Beck (D)	Bans noncompetes for hourly employees.	Limit	Not enacted
Nevada	AB 419	An Act Relating to Labor; Revising Provisions Governing Noncompetition Covenants in Employment Practices; and Providing Other Matters Properly Relating Thereto	Assembly Committee on Judiciary	Places time and geographic limitations on noncompete enforcement in the state.	Limit	Not enacted
New Hampshire	SB 197 HB 346	An Act Relative to Non-compete Agreements for Low-Wage Employees	Sen. David Watters (D) with 11 cosponsors (10 Democrats and 1 Republican)	Prohibits an employer from requiring a low-wage employee to enter into a noncompete agreement. "Low-wage employee" is defined as "an employee who earns . . . an hourly rate less than or equal to 200 percent of the federal minimum wage; or an hourly rate less than or equal to 200 percent of the tipped minimum wage."	Limit	Date signed: July 10, 2019 Date effective: September 8, 2019
New Jersey	HB 346 SB 2872	An Act Relative to Non-compete Agreements	Sen. Joseph Cryan (D)	Prohibits noncompetes for undergraduate and graduate students, apprentices, seasonal or temporary employees, employees laid off without cause, independent contractors, minors, and low-wage employees. Also contains a garden-leave provision requiring employers to pay employees 100 percent of the pay they would have earned while the noncompete is in effect.	Limit	Not enacted

New York	A 2504	An Act to Amend the Labor Law, in Relation to Prohibiting Employers from Requiring Low-Wage Employees to Enter into Covenants Not to Compete and Requiring Employers to Notify Potential Employees of Any Requirement to Enter into a Covenant Not to Compete	Assemblyman Jeffery Dinowitz (D) with 18 cosponsors (18 Democrats)	Prohibits the use of noncompetes for low-wage employees.	Limit	Not enacted
North Dakota	HB 1351	A Bill for an Act to Amend and Reenact Section 9-08-06 of the North Dakota Century Code, Relating to Contractual Noncompete Provisions	Rep. Mary Johnson (R) with 7 cosponsors (2 Republicans and 5 Democrats)	Modestly expands existing exceptions to the state's general prohibition on noncompetes.	Expand	Date signed: April 1, 2019 Date effective: August 1, 2019
Ohio	SB 141	To Enact Section 4113.66 of the Revised Code to Prohibit the Use of Noncompete Provisions in Physician Employment Contracts	Sen. Sandra Williams (D)	Prohibits the use of noncompetes for physicians.	Limit	Not enacted
Ohio	SB 75	To Enact Section 4113.66 of the Revised Code to Prohibit the Use of Noncompete Provisions in Employment Contracts in the Broadcasting Industry	Sen. Sandra Williams (D)	Prohibits the use of noncompete provisions in employment contracts in the broadcasting industry.	Limit	Not enacted
Oregon	HB 2992	Relating to Noncompetition Agreements	Rep. Ron Noble (R)	Requires employers to provide a signed, written copy of noncompete agreement within 30 days of termination for the agreement to be enforceable.	Limit	Date signed: May 14, 2019 Date effective: January 1, 2020
Pennsylvania	HB 171	An Act Prohibiting Enforcement of Covenants Not to Compete in Employment Agreements	Rep. Thomas Caltagirone (D) with 5 cosponsors (5 Democrats)	Prohibits the use of noncompete agreements in employment contracts, limiting the use of future noncompetes to the sale of a business or the dissolution of a partnership.	Limit	Not enacted
Pennsylvania	HB 563	An Act Prohibiting Certain Covenants Not to Compete; Conferring Powers and Duties on the Department of Labor and Industry; and Imposing Penalties	Rep. Donna Bullock (D) with 27 cosponsors (24 Democrats and 3 Republicans)	Prohibits the use of noncompetes for low-wage employees.	Limit	Not enacted

Pennsylvania	HB 601	Limiting Restrictive Covenants in Health Care Practitioner Employment Agreements	Rep. Anthony DeLuca (D) with 6 cosponsors (4 Democrats and 2 Republicans)	Prohibits the use of noncompetes for health care practitioners with some exceptions.	Limit	Not enacted
Rhode Island	HB 6019 SB 698	An Act Related to Labor and Labor Relations—Rhode Island Noncompetition Agreement Act	Rep. Christopher Blazewski (D) Sen. Maryellen Goodwin (D)	Prohibits noncompetes for undergraduate and graduate students, minors, and low-wage employees.	Limit	Date signed: July 15, 2019 Date effective: January 15, 2020
Rhode Island	SB 345	An Act Relating to Labor and Labor Relations—Noncompete Agreements—Broadcast Employees	Sens. Frank Lombardi (D), Frank Ciccone (D), Stephen Archambault (D), Michael McCaffrey (D), and Elaine Morgan (R)	Prohibits the inclusion of non-compete agreements in broadcast industry employment contracts that are entered into after January 1, 2020.	Limit	Not enacted
South Dakota	SB 120	An Act to Modify the Time Period Allowable for Certain Covenants Not to Compete	Sen. Brock Greenfield (R) with 8 cosponsors (7 Republicans and 1 Democrat)	Limits the duration of noncompetes to one year.	Limit	Not enacted
Texas	HB 1522	An Act Relating to a Prohibition Against Covenants Not to Compete for Independent Contractors with Oil and Gas Operations	Rep. Chris Paddie (R) with 7 cosponsors (3 Republicans and 4 Democrats)	Largely prohibits noncompetes in the oil and gas industries.	Limit	Not enacted
Texas	SB 2162 HB 2730	An Act Relating to Civil Actions Involving the Exercise of Certain Constitutional Rights	Rep. Jeff Leach (R) with 6 cosponsors (4 Republicans and 2 Democrats) Sen. Angela Paxton (R)	Modifies the Texas Citizens Participation Act to (among other things) exempt from anti-Strategic Lawsuits Against Public Participation laws "legal action[s] to enforce: (A) a noncompete agreement; (B) a nondisclosure agreement; or (C) a non-disparagement agreement." This strengthens Texas's noncompete enforcement, making it harder for people under noncompetes to seek legal recourse.	Expand	Date signed: June 2, 2019 Date effective: September 1, 2019
Texas	HB 2960	An Act Relating to a Prohibition Against Covenants Not to Compete for Certain Low-Wage Employees	Rep. Gina Hinojosa (D)	Prohibits the use of noncompete agreements against low-wage employees.	Limit	Not enacted

Utah	HB 199	Contract Amendments	Rep. Mike Schultz (R) and Sen. Dan Hemmert (R)	Replaces the four-year maximum term for a contract in the broadcasting industry of which a noncompete is a part to a term “of reasonable duration, based on industry standards, and other factors. The maximum duration of the noncompete remains one year.”	Limit	Date signed: March 22, 2019 Date effective: May 14, 2019
Vermont	HB 1	An Act Relating to Agreements Not to Compete	Reps. Martin LaLonde (D) and Annemarie Christensen (D)	Bans noncompetes statewide, except in connection with the sale of a business or dissolution of a partnership or limited liability company.	Limit	Not enacted
Virginia	HB 1792	A Bill to Amend the Code of Virginia by Adding in Article 1 of Chapter 3 of Title 40.1 a Section Numbered 40.1-28.7:7, Relating to Covenants Not to Compete; Low-Wage Employees	Delegate Schuyler VanValkenburg (D) with 14 cosponsors (14 Democrats)	Bans noncompetes for low-wage employees.	Limit	Not enacted
Virginia	SB 1387	A Bill to Amend the Code of Virginia by Adding in Article 1 of Chapter 3 of Title 40.1 a Section Numbered 40.1-28.7:7, Relating to Covenants Not to Compete; Low-Wage Employees; Civil Penalty	Sens. Frank Wagner (R) and Scott Surovell (D)	Bans the use of noncompetes for low-wage employees. SB 1387 is essentially HB 1792 with several additional elements. Specifically, SB 1387 expressly adds “interns, students, apprentices, or trainees employed, with or without pay, at a trade or occupation to gain work or educational experience” to the definition of low-wage employees.	Limit	Not enacted
Washington	HB 1450 SB 5478	An Act Relating to Restraints, Including Non-competition Covenants, on Persons Engaging in Lawful Professions, Trades, or Businesses	Sen. Derek Stanford (D) with 6 cosponsors (6 Democrats) Sen. Marko Lias (D) with 5 cosponsors (5 Democrats)	Under the new rules, noncompete agreements will only be enforceable if an employee earns more than \$100,000 a year, an independent contractor earns \$250,000 a year from the employer proposing a noncompete, the employer discloses terms of the noncompete when making an offer or earlier, the employer compensates employees who are laid off but still subject to noncompete agreements, or the noncompete agreement doesn't cover a period longer than 18 months.	Limit	Date signed: May 8, 2019 Date effective: January 1, 2020

Source: An Act to Repeal the Ability to Enforce a Covenant Not to Compete Agreement; and for Other Purposes, H. 1068, Arkansas Legislature, 92nd sess. (2019), <https://www.arkleg.state.ar.us/Bills/Detail?ddBienniumSession=2019%2F2019R&measureno=HB1068>; An Act Prohibiting the Use of Noncompete Clauses in Physician Employment Contracts, S. 377, Connecticut Legislature, (2019), https://www.cga.ct.gov/asp/cgabillstatus/cgabillstatus.asp?selBillType=Bill&bill_num=SB00377&which_year=2019; An Act Concerning Noncompete Agreements, H. 6914, Connecticut Legislature, (2019), https://www.cga.ct.gov/asp/cgabillstatus/cgabillstatus.asp?selBillType=Bill&bill_num=HB06914&which_year=2019; An Act Concerning Noncompete Agreements in the Blockchain Technology Industry, S. 1033, Connecticut Legislature, (2019), https://www.cga.ct.gov/asp/cgabillstatus/cgabillstatus.asp?selBillType=Bill&bill_num=SB01033&which_year=2019; An Act Relating to Restrictive Covenants; Creating s.542.336, F.S.; Providing That Certain Restrictive Covenants Are Void and Unenforceable for a Specified Period; Providing an Effective Date, S. 882, Florida

Legislature, (2019), <https://www.ilseenate.gov/Session/Bill/2019/882>; A Bill to Be Entitled an Act, H. 81, Georgia Legislature, (2019), <http://www.legis.ga.gov/legislation/en-us/display/20192020/hb/81>; Relating to Fair Employment Practices, H. 1059, Hawaii Legislature, 30th sess. (2019), https://www.capitol.hawaii.gov/Archives/measure_indiv_Archives.aspx?billtype=HB&billnumber=1059&year=2019; An Act Concerning Employment, H. 2565, Illinois Legislature, 101st sess. (2019), <https://www.ilga.gov/legislation/BillStatus.asp?DocNum=2565&GAID=15&DocTypeID=HB&SessionID=108&GA=101>; An Act Concerning Regulation, H. 2328, Illinois Legislature, 101st sess. (2019), <https://www.ilga.gov/legislation/BillStatus.asp?DocNum=2328&GAID=15&DocTypeID=HB&SessionID=108&GA=101>; A Bill for an Act to Amend the Indiana Code Concerning Labor and Safety, S. 348, Indiana Legislature, (2019), <http://iga.in.gov/legislative/2019/bills/senate/348/#document-a1ac3708>; A Bill for an Act to Amend the Indiana Code Concerning Professions and Occupations, H. 1357, Indiana Legislature, (2019), <http://iga.in.gov/legislative/2019/bills/house/1357/#document-ef72e7fd>; Physicians: Provides Relative to Prohibited Restraint of Certain Professions, S. 177, Louisiana Legislature, (2019), <http://www.legis.la.gov/legis/BillInfo.aspx?s=19rs&b=SB177&sbi=y>; An Act to Keep Workers in Maine, H. 733, Maine Legislature, 129th sess. (2019), http://legislature.maine.gov/legis/bills/display_ps.asp?LD=733&snm=129; An Act Concerning Labor and Employment—Noncompete and Conflict of Interest Clauses, H. 38, Maryland Legislature, (2019), <http://mgaleg.maryland.gov/mgawebsite/legislation/details/hb0038?ys=2019rs>; An Act Concerning Labor and Employment—Noncompete and Conflict of Interest Clauses, S. 328, Maryland Legislature, (2019), <http://mgaleg.maryland.gov/mgawebsite/legislation/details/sb0328?ys=2019rs>; An Act Relative to Banning Noncompetition Agreements in the Commonwealth, S. 108, 191st sess. (2019), <https://legiscan.com/MA/bill/S1083/2019>; Michigan Antitrust Reform Act, H. 4874, Michigan Legislature, (2019), <http://www.legislature.mi.gov/documents/2019-2020/billintroduced/House/pdf/2019-HIB-4874.pdf>; Michigan Antitrust Reform Act, S. 483, Michigan Legislature, (2019), <https://legiscan.com/MI/text/SB0483/2019>; A Bill for an Act Relating to Health; Protecting Physician-Patient Relationship by Prohibiting Noncompete Agreements; Proposing Coding for New Law in Minnesota Statutes, Chapter 145, H. 557, 91st sess. (2019), <https://www.revisor.mn.gov/bills/bill.php?b=House&f=HF0557&ssn=0&y=2019>; A Bill for an Act Relating to Health; Protecting Physician-Patient Relationship by Prohibiting Noncompete Agreements; Proposing Coding for New Law in Minnesota Statutes, Chapter 145, S. 350, 91st sess. (2019), <https://www.revisor.mn.gov/bills/bill.php?f=SF350&y=2019&ssn=0&b=senate>; An Act to Amend Chapter 431, RSMo, by Adding Thereto One New Section Relating to Covenants Not to Compete, H. 331, 100th sess. (2019), <https://www.house.mo.gov/Bill.aspx?bill=HB331&year=2019&code=R>; An Act Relating to Labor; Revising Provisions Governing Noncompetition Covenants in Employment Practices; and Providing Other Matters Properly Relating Thereto, A. 419, Nevada Legislature, (2019), <https://www.leg.state.nv.us/App/NELIS/REL/80th2019/Bill/6797/Overview>; An Act Relative to Noncompete Agreements for Low-Wage Employees, S. 197, New Hampshire Legislature, (2019), http://gencourt.state.nh.us/bill_status/bill_status.aspx?lsr=486&sy=2019&sortoption=&txtsessionyear=2019&txtbillnumber=SB197; An Act Relative to Noncompete Agreements, H. 346, New Hampshire Legislature, (2019), [https://legiscan.com/NH/text/HB346/id/1834381#:~:text=New%20Hampshire%20House%20Bill%20346%20\(Prior%20Session%20Legislation\)&text=Bill%20Title%3A%20Relative%20to%20noncompete%20agreements.&text=AN%20ACT%20relative%20to%20noncompete%20agreements.&text=This%20bill%20governs%20noncompete%20agreements%20between%20employers%20and%20employees](https://legiscan.com/NH/text/HB346/id/1834381#:~:text=New%20Hampshire%20House%20Bill%20346%20(Prior%20Session%20Legislation)&text=Bill%20Title%3A%20Relative%20to%20noncompete%20agreements.&text=AN%20ACT%20relative%20to%20noncompete%20agreements.&text=This%20bill%20governs%20noncompete%20agreements%20between%20employers%20and%20employees); An Act Relative to Noncompete Agreements, H. 346, New Jersey Legislature, (2019), <https://legiscan.com/NH/text/HB346/id/1834381>; An Act Limiting Certain Provisions in Restrictive Covenants and Supplementing Title 34 of the Revised Statutes, S. 2872, New Jersey Legislature, 218th sess. (2018), <https://legiscan.com/NJ/bill/S2872/2018>; An Act to Amend the Labor Law, in Relation to Prohibiting Employers from Requiring Low-Wage Employees to Enter into Covenants Not to Compete and Requiring Employers to Notify Potential Employees of Any Requirement to Enter into a Covenant Not to Compete, A. 2504, New York Legislature, (2019), https://assembly.state.ny.us/leg/?default_fld=&bn=AC2504&term=2019&Summary=Y&Actions=Y&Text=Y&Committee%26nbspVotes=Y&Floor%26nbspVotes=Y; A Bill for an Act to Amend and Reenact Section 9-08-06 of the North Dakota Century Code, Relating to Contractual Noncompete Provisions, H. 1341, North Dakota Legislature, (2019), <https://www.legis.nd.gov/assembly/66-2019/bill-actions/ba1351.html>; To Enact Section 4113.66 of the Revised Code to Prohibit the Use of Noncompete Provisions in Physician Employment Contracts, S. 141, Ohio Legislature, 133rd sess. (2019), <https://www.legislature.ohio.gov/legislation/legislation-summary?id=GA133-SB-141>; To Enact Section 4113.66 of the Revised Code to Prohibit the Use of Noncompete Provisions in Employment Contracts in the Broadcasting Industry, S. 75, Ohio Legislature, 133rd sess. (2019), <https://www.legislature.ohio.gov/legislation/legislation-summary?id=GA133-SB-75>; Relating to Noncompetition Agreements, H. 2992, Oregon Legislature, 80th sess. (2019), <https://olis.leg.state.or.us/liz/2019R1/Measures/Overview/HB2992>; An Act Prohibiting Enforcement of Covenants Not to Compete in Employment Agreements, H. 171, Pennsylvania Legislature, (2019), <https://www.legis.state.pa.us/cfdocs/billinfo/BillInfo.cfm?year=2019&ind=0&body=H&type=B&bn=171>; An Act Prohibiting Certain Covenants Not to Compete; Conferring Powers and Duties on the Department of Labor and Industry; and Imposing Penalties, H. 563, Pennsylvania Legislature, (2019), <https://www.legis.state.pa.us/cfdocs/billinfo/BillInfo.cfm?year=2019&ind=0&body=H&type=B&bn=563>; Limiting Restrictive Covenants in Health Care Practitioner Employment Agreements, H. 601, Pennsylvania Legislature, (2019), <https://www.legis.state.pa.us/cfdocs/billinfo/BillInfo.cfm?year=2019&ind=0&body=H&type=B&bn=601>; An Act Related to Labor and Labor Relations—Rhode Island Noncompetition Agreement Act, H. 6019, Rhode Island Legislature, (2019), <https://legiscan.com/RI/bill/H6019/2019>; An Act Related to Labor and Labor Relations—Rhode Island Noncompetition Agreement Act, S. 698, Rhode Island Legislature, (2019), <https://legiscan.com/RI/bill/S0698/2019>; An Act Relating to Labor and Labor Relations—Noncompete Agreements—Broadcast Employees, S. 345, Rhode Island Legislature, (2019), <https://legiscan.com/RI/text/S0345/2019>; An Act to Modify the Time Period Allowable for Certain Covenants Not to Compete, S. 120, South Dakota Legislature, 94th sess. (2019), <https://sdlegislature.gov/#/Session/Bill/9651>; An Act Relating to a Prohibition Against Covenants Not to Compete for Independent Contractors with Oil and Gas Operations, H. 1522, Texas Legislature, 86th sess. (2019), <https://capitol.texas.gov/BillLookup/History.aspx?>

LegSess=86R&Bill=HB1522; An Act Relating to Civil Actions Involving the Exercise of Certain Constitutional Rights, S. 2162, Texas Legislature, 86th sess. (2019), <https://capitol.texas.gov/BillLookup/History.aspx?LegSess=86R&Bill=SB2162>; An Act Relating to Civil Actions Involving the Exercise of Certain Constitutional Rights, H. 2730, Texas Legislature, 86th sess. (2019), <https://capitol.texas.gov/BillLookup/History.aspx?LegSess=86R&Bill=HB2730>; An Act Relating to a Prohibition Against Covenants Not to Compete for Certain Low-Wage Employees, H. 2960, Texas Legislature, 86th sess. (2019), <https://capitol.texas.gov/BillLookup/History.aspx?LegSess=86R&Bill=HB2960>; Contract Amendments, H. 199, Utah Legislature, (2019), <https://le.utah.gov/~2019/bills/static/HB0199.html>; An Act Relating to Agreements Not to Compete, H. 1, Vermont Legislature, (2019), <https://legislature.vermont.gov/bill/status/2020/H.1>; A Bill to Amend the Code of Virginia by Adding in Article 1 of Chapter 3 of Title 40.1 a Section Numbered 40.1-28.7:7, Relating to Covenants Not to Compete; Low-Wage Employees, H. 1792, Virginia Legislature, (2019), <https://lis.virginia.gov/cgi-bin/legp604.exe?191+sum+HB1792&191+sum+HB1792>; A Bill to Amend the Code of Virginia by Adding in Article 1 of Chapter 3 of Title 40.1 a Section Numbered 40.1-28.7:7, Relating to Covenants Not to Compete; Low-Wage Employees; Civil Penalty, S. 1387, Virginia Legislature, (2019), <https://lis.virginia.gov/cgi-bin/legp604.exe?191+sum+SB1387&191+sum+SB1387>; An Act Relating to Restraints, Including Noncompetition Covenants, on Persons Engaging in Lawful Professions, Trades, or Businesses, H. 1450, Washington Legislature, 66th sess. (2019), <https://apps.leg.wa.gov/billsummary/?BillNumber=1450&Year=2019&Initiative=false>; An Act Relating to Restraints, Including Noncompetition Covenants, on Persons Engaging in Lawful Professions, Trades, or Businesses; Adding a New Chapter to Title 49 RCW; and Providing an Effective Date, S. 5478, Washington Legislature, 66th sess. (2019), <https://apps.leg.wa.gov/billsummary/?BillNumber=5478&Chamber=Senate&Year=2019>.

Appendix B. 2020 Legislation

State	Bill Number	Title	Sponsor(s)	Summary	Type	Status
Connecticut	HB 5210	An Act Concerning Non-compete Agreements in the Blockchain Technology Industry	House Commerce Committee	Prohibits the use of noncompete agreements in employment contracts for employees in the blockchain technology industry.	Limit	Not enacted
Connecticut	SB 143	An Act Prohibiting Covenants Not to Compete Involving Physicians	Sen. M. Saud Anwar (D)	Prohibits noncompete agreements involving physicians.	Limit	Not enacted
Connecticut	SB 409	An Act Concerning Home-maker and Companion Services	Sen. M. Saud Anwar (D)	Modifies noncompete agreements between an individual and a home-maker-companion agency, registry, or home health aide agency.	Limit	Not enacted
Hawaii	HB 1841	An Act Relating to Fair Employment Practices	Rep. Joy San Buenaventura (D) with 8 cosponsors (7 Democrats and 1 Republican)	Prohibits noncompete agreements for low-wage workers whose earnings do not exceed the greater of the minimum wage required by applicable federal or state law or \$15 per hour.	Limit	Not enacted
Iowa	SF 2332	An Act Prohibiting Employers from Entering into Non-compete Agreements with Low-Wage Employees	Senate Committee on Labor and Business Relations	Prohibits employers from requiring a noncompete of low-wage employees, defined as any employee earning an hourly wage of 200 percent of the federal minimum wage or less.	Limit	Not enacted
Iowa	SB 3159	An Act Relating to Non-competition Agreements and Mental Health and Disability Services Contracts with a State Board of Regents Institution	Committee on Human Resources	Places restrictions on the enforceability of noncompete provisions for mental health and disability workers.	Limit	Not Enacted

Illinois	SB 3430	An Act Concerning Employment	Sen. Heather Steans (D)	Amends the Illinois Freedom to Work Act to limit noncompetes in several ways, including voiding such agreements unless the employer provides a copy of the agreement in advance and advises the employee in writing to seek the advice of an attorney. For a noncompete to be valid, an employee must have worked for the employer for two years from the date of signature or received some other benefit "specifically bargained for" in exchange for signature. The terms of the noncompete must be limited to what is reasonably required to protect a legitimate business interest without posing undue burden on the employee or harm to the public.	Limit	Not enacted
Indiana	HB 1004	An Act to Amend the Indiana Code Concerning Health	Rep. Ben Smaltz (R) with 8 cosponsors (6 Republicans and 2 Democrats)	Requires physician noncompete agreements to contain certain provisions to be enforceable.	Limit	Date signed: March 18, 2020 Date effective: July 1, 2020
Indiana	SB 337	An Act to Amend the Indiana Code Concerning Insurance	Sen. Victoria Spartz (R) with 5 cosponsors (5 Republicans)	Establishes limitations on non-compete agreements concerning physicians.	Limit	Not enacted
Kentucky	SB 41	An Act Relating to Non-Compete Agreements with Health-Care Providers	Sen. Ralph Alvarado (R)	Prohibits noncompete agreements for certain health care providers.	Limit	Not enacted
Kentucky	HB 310	An Act Relating to Health Care Provider Employment Contracts	Rep. Rob Wiederstein (D)	Prohibits the use of noncompete agreements against nurse practitioners, physicians, and osteopaths.	Limit	Not enacted
Kentucky	HB 86	An Act Relating to Physician Employment Contracts	Reps. Rob Wiederstein (D), George Brown (D), and Maria Sorolis (D)	Prohibits the use of noncompete agreements against physicians and osteopaths.	Limit	Not enacted
Louisiana	SB 345	Provides Relative to Non-compete Agreements	Sen. Ronnie Johns (R)	Expands noncompete enforcement by allowing them to be used against business partners, franchisees, and different business entities	Expand	Date signed: June 9, 2020 Date effective: August 1, 2020

Minnesota	HF 4376	A Bill for an Act Relating to Employment; Providing That Covenants Not to Compete Are Void and Unenforceable; Providing for the Protection of Substantive Provisions of Minnesota Law to Apply to Matters Arising in Minnesota; Proposing Coding for New Law in Minnesota Statutes, Chapter 181	Reps. Jennifer Schultz (D) and Liz Olson (D)	Establishes noncompete agreements as void and unenforceable.	Limit	Not enacted
Minnesota	HF 3673	An Act Relating to Labor Standards; Prohibiting Covenants Not to Compete; Imposing Penalties; Amending Minnesota Statutes 2018, Section 177.27, Subdivision 4; Proposing Coding for New Law in Minnesota Statutes, Chapter 181	Rep. Alice Mann (D) with 29 cosponsors (29 Democrats)	Prohibits all noncompete agreements.	Limit	Not enacted
Missouri	HB 2326	An Act to Amend Chapter 191, RSMo, by Adding Thereto One New Section Relating to Noncompete Agreements for Certain Health Care Providers	Rep. Steve Helms (R)	Prohibits most noncompete agreements against health care providers.	Limit	Not enacted
Missouri	HB 2230	An Act to Amend Chapter 431, RSMo, by Adding Thereto One New Section Relating to Covenants Not to Compete	Rep. Herman Morse (R)	Prohibits the use of noncompete agreements against hourly wage earners.	Limit	Not enacted
New Hampshire	HB 1106	An Act Relative to Noncompete Agreements for Certain Mental Health Professionals	Rep. Pat Abrami (R) and Sen. Tom Sherman (D)	Prohibits noncompete agreements for certain mental health professionals.	Limit	Not enacted
New Jersey	SB 899	An Act Limiting Certain Provisions in Restrictive Covenants and Supplementing Title 34 of the Revised Statutes	Sen. Joe Cryan (D)	Limits certain provisions in and enforceability of restrictive covenants.	Limit	Not enacted
New York	SB 8274	An Act to Amend the Labor Law, in Relation to Prohibiting Non-Compete Agreements	Sen. Alessandra Biaggi (D)	Prohibits enforcement of a non-compete agreement unless it is reasonable in scope, necessary to protect a legitimate business interest, and does not impose undue harm on the employee or the public.	Limit	Not enacted

New York	SB 5790	An Act to Amend the Labor Law, in Relation to Prohibiting Non-Compete Agreements and Certain Restrictive Covenants	Sens. Jessica Ramos (D), Alessandra Biaggi (D), and Andrew Gouardes (D)	Prohibits noncompete agreements against workers earning less than \$75,000 a year (increases annually for inflation).	Limit	Not enacted
Oregon	SB 1527	An Act Relating to Non-competition Agreements; Creating New Provisions; and Amending ORS 63.295	Senate Interim Committee on Labor and Business	Prohibits noncompete agreements against workers making less than \$97,311 a year (increases annually for inflation). Also limits valid non-compete agreements to one year.	Limit	Not enacted
Pennsylvania	HB 2366	An Act Prohibiting Enforcement of Covenants Not to Compete in Broadcast Employment Agreements	Rep. Thomas Murt (R) with 6 cosponsors (2 Republicans and 4 Democrats)	Prohibits noncompete agreements in broadcast employment agreements.	Limit	Not enacted
Rhode Island	HB 7530	An Act Relating to Businesses and Professions—Physician Assistants	Reps. David Bennett (D), Dennis Canario (D), John Edwards (D), Joseph McNamara (D), and Joseph Shekarchi (D)	Limits the use of noncompete agreements against physician assistants.	Limit	Not enacted
Rhode Island	SB 2087	An Act Relating to Labor and Labor Relations—Noncompetition Agreements—Broadcast Employees	Sens. Frank Lombardi (D), Donna Nesselbush (D), Frank Ciccone (D), William Conley (D), and Michael McCaffrey (D)	Prohibits the inclusion of noncompete agreements in broadcast industry employment contracts.	Limit	Not enacted
South Carolina	HB 5242	A Bill to Amend the Code of Laws in South Carolina by Adding Section 41-1-130 so as to Provide That Any Covenant Not to Compete That a Nonprofit Corporation with an Annual Gross Revenue Exceeding One Billion Dollars Has with Current and Former Employees Is Null, Void, and Unenforceable, and to Provide That No Such Covenant May Be Entered into with Future Employees	Rep. Todd Rutherford (D)	Prohibits nonprofit corporations with annual gross revenue exceeding \$1 billion from using noncompete agreements on their employees.	Limit	Not enacted
South Dakota	SB 108	An Act to Revise the Time Period Allowable for Certain Covenants Not to Compete	Sens. Brock Greenfield (F) and Jeff Partridge (R)	Limits the duration of noncompete agreements to one year.	Limit	Not enacted

United States	HR 5710	Workforce Mobility Act of 2020	Reps. Scott Peters (D-CA), Mike Gallagher (R-WI), and Anna Eshoo (D-CA)	Prohibits nearly all noncompete agreements.	Limit	Not enacted
Virginia	HB 330	An Act Relating to Covenants Not to Compete, Definition of Low-Wage Employees, Civil Penalty	Delegate Schuyler VanValkenburg (D)	Prohibits noncompete agreements for low-wage employees, defined as employees with average weekly earnings less than the statewide average weekly wage or independent contractors compensated at an hourly rate that is less than the statewide median hourly wage.	Limit	Date signed: April 9, 2020 Date effective: July 1, 2020

Source: An Act Concerning Noncompete Agreements in the Blockchain Technology Industry, H. 5210, Connecticut Legislature, (2020), https://www.cga.ct.gov/asp/cgabillstatus/cgabillstatus.asp?selBillType=Bill&bill_num=HB05210&which_year=2020; An Act Prohibiting Covenants Not to Compete Involving Physicians, S. 143, Connecticut Legislature, (2020), https://www.cga.ct.gov/asp/cgabillstatus/cgabillstatus.asp?selBillType=Bill&bill_num=SB00143&which_year=2020; An Act Concerning Homemaker and Companion Services, S. 409, Connecticut Legislature, (2020), https://www.cga.ct.gov/asp/cgabillstatus/cgabillstatus.asp?selBillType=Bill&bill_num=SB00409&which_year=2020; An Act Relating to Fair Employment Practices, H. 1841, Hawaii Legislature, (2020), https://www.capitol.hawaii.gov/measure_indiv.aspx?billtype=HB&billnumber=1841&year=2020; An Act Prohibiting Employers from Entering into Noncompete Agreements with Low-Wage Employees, S. 2332, Iowa Legislature, 88th sess. (2020), <https://www.legis.iowa.gov/legislation/BillBook?ga=88&ba=SF2332>; An Act Relating to Noncompetition Agreements and Mental Health and Disability Services Contracts with a State Board of Regents Institution, S. 3159, Iowa Legislature, 88th sess. (2020), <https://www.legis.iowa.gov/legislation/BillBook?ga=88&ba=SSB3159>; An Act Concerning Employment, S. 3430, Illinois Legislature, 101st sess. (2020), <https://www.ilga.gov/legislation/BillStatus.asp?DocNum=3430&GAID=15&DocTypeID=SB&SessionID=108&GA=101>; An Act to Amend the Indiana Code Concerning Health, H. 1004, Indiana Legislature, 121st sess. (2020), <http://iga.in.gov/legislative/2020/bills/house/1004/>; An Act to Amend the Indiana Code Concerning Insurance, S. 337, Indiana Legislature, 121st sess. (2020), <http://iga.in.gov/legislative/2020/bills/senate/337/#document-a2a0fbbc>; An Act Relating to Non-Compete Agreements with Health-Care Providers, S. 41, Kentucky Legislature, (2020), <https://apps.legislature.ky.gov/record/20RS/sb41.html>; An Act Relating to Health Care Provider Employment Contracts, H. 310, Kentucky Legislature, (2020), <https://apps.legislature.ky.gov/record/20RS/hb310.html>; An Act Relating to Health Care Provider Employment Contracts, H. 86, Kentucky Legislature, (2020), <https://apps.legislature.ky.gov/record/20RS/hb86.html>; Provides Relative to Noncompete Agreements, S. 345, Louisiana Legislature, (2020), <http://www.legis.la.gov/legis/BillInfo.aspx?s=20rs&b=SB345&sbi=y>; A Bill for an Act Relating to Employment; Providing That Covenants Not to Compete Are Void and Unenforceable; Providing for the Protection of Substantive Provisions of Minnesota Law to Apply to Matters Arising in Minnesota; Proposing Coding for New Law in Minnesota Statutes, Chapter 181, H. 4376, Minnesota Legislature, 91st sess. (2020), <https://www.revisor.mn.gov/bills/bill.php?b=House&f=HF4376&ssn=0&y=2019>; An Act Relating to Labor Relating to Labor Standards; Prohibiting Covenants Not to Compete; Imposing Penalties; Amending Minnesota Statutes 2018, Section 177.27, Subdivision 4; Proposing Coding for New Law in Minnesota Statutes, Chapter 181, H. 3673, Minnesota Legislature, 91st sess. (2020), <https://www.revisor.mn.gov/bills/bill.php?b=House&f=HF3673&ssn=0&y=2019>; An Act to Amend Chapter 191, RSMo, by Adding Thereto One New Section Relating to Noncompete Agreements for Certain Health Care Providers, H. 2326, Missouri Legislature, 110th sess. (2020), <https://www.house.mo.gov/Bill.aspx?bill=HB2326&year=2020&code=R>; An Act to Amend Chapter 431, RSMo, by Adding Thereto One New Section Relating to Covenants Not to Compete, H. 2230, Missouri Legislature, 100th sess. (2020), <https://www.house.mo.gov/Bill.aspx?bill=HB2230&year=2020&code=R>; An Act Relative to Noncompete Agreements for Certain Mental Health Professionals, H. 1106, New Hampshire Legislature, (2020), http://gencourt.state.nh.us/bill_status/bill_status.aspx?lstr=2028&sy=2020&sortoption=&txtsessionyear=2020&txbillnumber=HB1106; An Act Limiting Certain Provisions in Restrictive Covenants and Supplementing Title 34 of the Revised Statutes, S. 899, New Jersey Legislature, 219th sess. (2020), <https://legiscan.com/NJ/bill/S899/2020>; An Act to Amend the Labor Law, in Relation to Prohibiting Non-Compete Agreements, S. 8274, New York Legislature, (2020), <https://www.nysenate.gov/legislation/bills/2019/S8274>; An Act to Amend the Labor Law, in Relation to Prohibiting Non-Compete Agreements and Certain Restrictive Covenants, S. 5790, New York Legislature, (2019), https://assembly.state.ny.us/leg/?default_fld=&bn=S05790&term=2019&Summary=Y&Actions=Y&Text=Y&Committee%26nbspVotes=Y&Floor%26nbspVotes=Y; An Act Relating to Noncompetition Agreements; Creating New Provisions; and Amending ORS 63.295, S. 1527, Oregon Legislature, 80th sess. (2020), <https://olis.leg.state.or.us/liz/2020R1/Measures/Overview/SB1527>; An Act Prohibiting Enforcement of Covenants Not to Compete in Broadcast Employment Agreements, H. 2366, Pennsylvania Legislature, (2020), https://www.legis.state.pa.us/cfdocs/billinfo/bill_history.cfm?year=2019&sind=0&body=H&type=B&bn=2366; An Act Relating to Businesses and Professions—Physician Assistants, H. 7530, Rhode Island Legislature, (2020), <https://legiscan.com/RI/bill/H7530/2020>; An Act Relating to Labor and Labor Relations—Noncompetition Agreements—Broadcast Employees, S. 2087,

Rhode Island Legislature, (2020), <https://legiscan.com/RI/bill/S2087/2020>; A Bill to Amend the Code of Laws in South Carolina by Adding Section 41-1-130 so as to Provide That Any Covenant Not to Compete That a Nonprofit Corporation with an Annual Gross Revenue Exceeding One Billion Dollars Has with Current and Former Employees Is Null, Void, and Unenforceable, and to Provide That No Such Covenant May Be Entered into with Future Employees, H. 5242, South Carolina Legislature, 123rd sess. (2020), <https://www.scstatehouse.gov/billsearch.php?billnumbers=5242&session=123&summary=B>; An Act to Revise the Time Period Allowable for Certain Covenants Not to Compete, S. 108, South Dakota Legislature, 95th sess. (2020), <https://sdlegislature.gov/#/Session/Bill/12004>; Workforce Mobility Act of 2020, H. 5710, 116th Cong., 2nd sess. (2020), <https://www.congress.gov/bill/116th-congress/house-bill/5710/all-info>; and An Act Relating to Covenants Not to Compete, Definition of Low-Wage Employees, Civil Penalty, H. 330, Virginia Legislature, (2020), <https://lis.virginia.gov/cgi-bin/legp604.exe?201+sum+HB330&201+sum+HB330>.

Notes

1. J. J. Prescott, Norman D. Bishara, and Evan Starr, “Understanding Noncompetition Agreements: The 2014 Noncompete Survey Project,” *Michigan State Law Review* 2016, no. 2 (2016): 369–464, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2799961.
2. Stephen L. Brodsky, “Restrictive Covenants in Employment and Related Contracts: Key Considerations You Should Know,” American Bar Association, February 8, 2019, <https://www.americanbar.org/groups/litigation/committees/commercial-business/practice/2019/restrictive-covenants-employment-related-contracts>.
3. Ufuk Akcigit and Sina T. Ates, “What Happened to U.S. Business Dynamism” (working paper, Becker Friedman Institute, Chicago, IL, April 15, 2019), <https://bfi.uchicago.edu/working-paper/what-happened-to-u-s-business-dynamism/>.
4. Matt Marx, Deborah Strumsky, and Lee Fleming, “Mobility, Skills, and the Michigan Non-Compete Experiment,” *Management Science* 55, no. 6 (2009): 875–89, <https://www.jstor.org/stable/40539267?seq=1>.
5. Evan Starr, “Consider This: Training, Wages, and the Enforceability of Covenants Not to Compete,” *ILR Review* 72, no. 4 (2019): 783–817, <https://doi.org/10.1177/0019793919826060>.
6. Hyo Kang and Lee Fleming, “Non-Competes, Business Dynamism, and Concentration: Evidence from a Florida Case Study,” *Journal of Economics & Management Strategy* 29, no. 3 (2020): 663–85, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3172477.
7. Koby Levin, “As Non-Compete Agreements Proliferate, So Do Lawsuits,” AP News, March 23, 2018, <https://apnews.com/70f0855282de4329908957fa7b1e278d/As-non-compete-agreements-proliferate,-so-do-lawsuits>.
8. Evan Starr, J. J. Prescott, and Norman Bishara, “Noncompete Agreements in the U.S. Labor Force,” *Journal of Law and Economics* (forthcoming), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2625714.
9. Evan Starr, *The Use, Abuse, and Enforceability of Non-Compete and No-Poach Agreements*, Economic Innovation Group, February 2019, <https://eig.org/wp-content/uploads/2019/02/Non-Competes-Brief.pdf>.
10. Matthew S. Johnson, Kurt Lavetti, and Michael Lipsitz, “The Labor Market Effects of Legal Restrictions on Worker Mobility,” September 18, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3455381.
11. Matt Marx, “Punctuated Entrepreneurship (Among Women)” (working paper, US Census Bureau, Center for Economic Studies, Washington, DC, 2018), <https://ideas.repec.org/p/cen/wpaper/18-26.html>.
12. Starr, Prescott, and Bishara, “Noncompete Agreements in the U.S. Labor Force”; Evan Starr, Justin Frake, and Rajshree Agarwal, “Mobility Constraint Externalities,” *Organization Science* 30, no. 5 (September–October 2019): 869–1123, <https://pubsonline.informs.org/doi/10.1287/orsc.2018.1252>.
13. Kenan Fikri, “Dynamism Diagnostics: Five Observations from the Latest Data,” Economic Innovation Group, October 1, 2020, <https://eig.org/news/dynamism-diagnostics-five-observations-from-the-latest-data>.
14. Starr, Prescott, and Bishara, “Noncompete Agreements in the U.S. Labor Force.”
15. Evan Starr and J. J. Prescott, “Subjective Beliefs About Contract Enforceability” (working paper, 2019), <https://sites.google.com/site/starrean/research>.
16. Matt Marx, *Reforming Non-Competes to Support Workers*, Brookings Institution, Hamilton Project, February 2018, https://www.brookings.edu/wp-content/uploads/2018/02/es_2272018_reforming_noncompetes_support_workers_marx_policy_proposal.pdf.
17. Starr, Prescott, and Bishara, “Noncompete Agreements in the U.S. Labor Force.”
18. Laws that prohibit noncompetes for workers earning under a certain amount often suffer from ambiguity in how key terms are defined. Lawmakers must therefore be clear about what kinds of compensation count toward the earnings threshold (e.g., bonuses and commissions) and which circumstances (e.g., a reduction in income after the noncompete is signed) would render an agreement unenforceable.
19. Karla Walter, “The Freedom to Leave: Curbing Noncompete Agreements to Protect Workers and Support Entrepreneurship,” Center for American Progress, January 9, 2019, <https://www.americanprogress.org/issues/economy/reports/2019/01/09/464831/the-freedom-to-leave/>.

20. Natarajan Balasubramanian et al., “Locked In? The Enforceability of Covenants Not to Compete and the Careers of High-Tech Workers” (working paper, US Census Bureau, Center for Economic Studies, Washington, DC, March 2020), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2905782.
21. An Act to Promote Keeping Workers in Maine, H. 733, Maine Legislature, 129th sess. (2019), http://legislature.maine.gov/legis/bills/display_ps.asp?LD=733&snum=129.
22. Alan B. Krueger and Orley Ashenfelter, “Theory and Evidence on Employer Collusion in the Franchise Sector” (working paper, National Bureau of Economic Research, Cambridge, MA, July 2018), <https://www.nber.org/papers/w24831?sy=831>.
23. US Department of Justice, “No-Poach Approach,” September 30, 2019, <https://www.justice.gov/atr/division-operations/division-update-spring-2019/no-poach-approach>.
24. An Act Relating to Restraints, Including Noncompetition Covenants, on Persons Engaging in Lawful Professions, Trades, or Businesses, H. 1450, Washington Legislature, 66th sess. (2019), <https://apps.leg.wa.gov/billssummary/?BillNumber=1450&Year=2019&Initiative=false>.
25. An Act Relating to Restrictive Covenants; Creating s.542.336, F.S.; Providing That Certain Restrictive Covenants Are Void and Unenforceable for a Specified Period; Providing an Effective Date, H. 843, Florida Legislature, (2019).
26. Va. Code Ann. § 40.1-28.7:8.
27. An Act to Amend the Indiana Code Concerning Health, H. 1004, Indiana Legislature, 121st sess. (2020), <https://legiscan.com/IN/text/HB1004/id/2169617/Indiana-2020-HB1004-Enrolled.pdf>.
28. Or. Rev. Stat. § 653.295.
29. Melissa Ilyse Rassas, “Explaining the Outliers: Oregon’s New Non-Compete Agreement Law & the Broadcasting Industry,” *University of Pennsylvania Journal of Business Law* 11, no. 2 (2008): 447–73, <https://scholarship.law.upenn.edu/jbl/vol11/iss2/5/>.
30. Michael Lipsitz and Evan Starr, “Low-Wage Workers and the Enforceability of Non-Compete Agreements,” *Management Science* (forthcoming), <https://ssrn.com/abstract=3452240>.
31. Illinois Freedom to Work Act, S. 3163, Illinois Legislature, 99th sess. (2016), <https://www.ilga.gov/legislation/publicActs/fulltext.asp?Name=099-0860&GA=99>.
32. An Act Relative to the Judicial Enforcement of Noncompetition Agreements, S. 988, Massachusetts Legislature, 190th sess. (2018), <https://malegislature.gov/bills/190/sd1578>.
33. James P. Flynn, “Massachusetts Establishes Garden-Leave-Type Non-Compete Requirements and Limits,” *Trade Secrets & Employee Mobility*, November 30, 2020, <https://www.tradesecretsandemployeemobility.com/2018/08/articles/non-competes-agreements/massachusetts-establishes-garden-leave-type-non-competes-requirements-and-limits/>.
34. An Act Relating to Employment Agreements, H. 1090, Hawaii Legislature, 28th sess. (2015), https://www.capitol.hawaii.gov/session2015/bills/HB1090_CD1_.HTM.
35. Balasubramanian et al., “Locked In?”
36. Freedom to Compete Act, S. 124, 116th Cong., 1st sess. (2019), [https://www.congress.gov/bill/116th-congress/senate-bill/124#:~:text=Introduced%20in%20Senate%20\(01%2F15%2F2019\)&text=This%20bill%20amends%20the%20Fair,entry%20level%2C%20lower%20wage%20workers](https://www.congress.gov/bill/116th-congress/senate-bill/124#:~:text=Introduced%20in%20Senate%20(01%2F15%2F2019)&text=This%20bill%20amends%20the%20Fair,entry%20level%2C%20lower%20wage%20workers).
37. Christopher Murphy et al., letter to Gene Dodaro, Office of Sen. Marco Rubio, March 7, 2019, https://www.rubio.senate.gov/public/_cache/files/3d73589c-7939-497a-8119-7b7b959d43a2/3E7179209663946F9B343FF3A6882412.letter-to-gao-on-non-competes-agreements-from-murphy-young-warren-rubio-wyden-kaine--signed.pdf. The author advised the senators on the content of this legislation in his capacity as chief executive of the Economic Innovation Group. See Economic Innovation Group, “EIG Applauds Introduction of Bipartisan Legislation to Limit Non-Compete Agreements, Boost Worker Mobility,” press release, October 17, 2019, <https://eig.org/news/eig-applauds-introduction-of-workforce-mobility-act>.
38. Murphy et al., letter to Gene Dodaro.
39. Office of Sen. Todd Young, “Young and Murphy Introduce Bill to Limit Non-Compete Agreements, Protect Workers,” press release, October 17, 2019, <https://www.young.senate.gov/newsroom/press-releases/young-and-murphy-introduce-bill-to-limit-non-competes-agreements-protect-workers>.

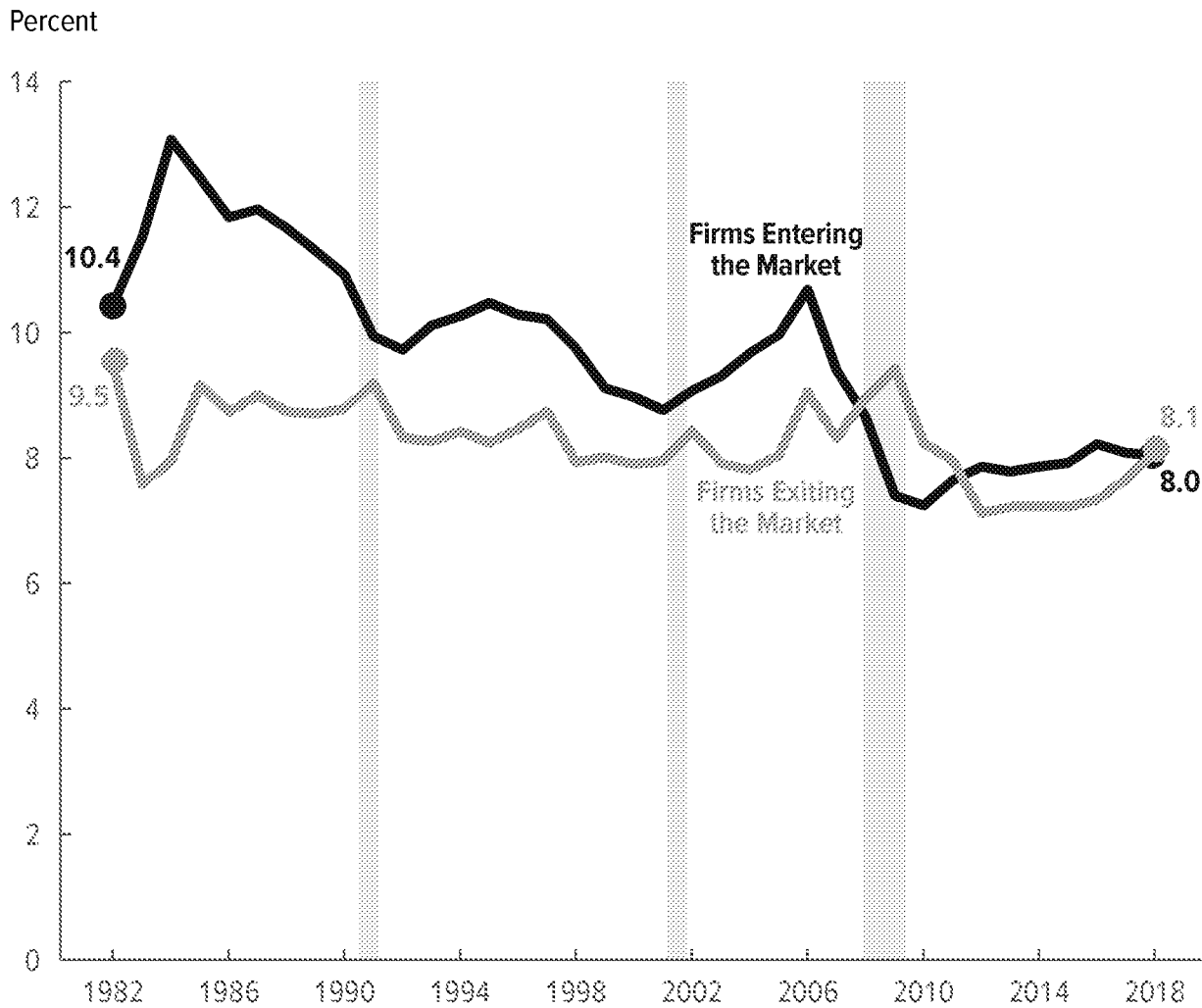
40. Workforce Mobility Act of 2019, S. 2614, 116th Cong., 1st sess. (2019).
41. JoeBiden.com, “The Biden Plan for Strengthening Worker Organizing, Collective Bargaining, and Unions,” <https://joebiden.com/empowerworkers/#>.
42. Alayna Treene, “Rubio Says the GOP Needs to Reset After 2020,” Axios, November 11, 2020, <https://www.axios.com/rubio-gop-reset-trump-872340a7-4c75-4c2b-9261-9612c590ee14.html>.

© 2020 by the American Enterprise Institute for Public Policy Research. All rights reserved.

The American Enterprise Institute (AEI) is a nonpartisan, nonprofit, 501(c)(3) educational organization and does not take institutional positions on any issues. The views expressed here are those of the author(s).

CBO

Federal Policies in Response to Declining Entrepreneurship



DECEMBER 2020

At a Glance

Entrepreneurship in the economy has declined significantly over the past four decades. The rate at which firms were created decreased from 10 percent of all businesses in 1982 to 8 percent in 2018, and the share of employment belonging to new firms (those less than five years old) fell from 14 percent to 9 percent over that same period.

New firms contribute to economic growth through the important role they play in allocating the economy's resources more efficiently. New firms provide innovative products and services, improve the productivity of the workforce, and ensure competition in the marketplace. The decline in entrepreneurship has been associated with a decrease in annual productivity growth whose cumulative effect has made the economy 3 percent to 4 percent less productive than if entrepreneurship had remained unchanged since the early 1980s, in CBO's assessment. The decline is frequently attributed to three types of factors, two of which are supported by strong evidence, whereas the third is uncertain:

- **Financial.** Much of the decrease in the formation of new businesses occurred during recessions, particularly the 2007–2009 recession, as firms faced restricted access to financing and a weaker economy.
- **Demographic.** The slower growth rate of the labor force after 1980 contributed to the decrease in entrepreneurship, as did the decline in the share of the workforce after 2000 of people ages 35 to 54, who are most likely to be successful entrepreneurs.
- **Regulatory.** Regulation affects entrepreneurship. Thus, changes in the regulatory environment might have contributed to the falloff in entrepreneurship, but that link has not been clearly established in the research literature.

If policymakers wish to spur entrepreneurship, they could put in place measures to give new firms greater access to financing or provide more financial support for the development of new technologies to those small businesses that are likely to be innovative. They could also facilitate the immigration of highly skilled workers and entrepreneurs to the United States, or they could alter the regulatory environment to assist new businesses. One challenge to such policies is that identifying potentially successful new firms can be difficult because about as many of them fail as succeed within their first five years. Still another challenge is that existing federal policies aimed at supporting entrepreneurship often focus on small businesses, but even though new businesses generally start small, most small businesses are not new. Different approaches would have their own advantages and disadvantages.



Contents

Summary	1
How Much Has Entrepreneurship Declined Since the 1980s?	1
How Has the Decline in Entrepreneurship Affected Productivity Growth?	1
What Factors Have Caused a Decline in Entrepreneurship?	1
What Federal Policies Would Support Entrepreneurship?	2
The Decline in Entrepreneurship	3
Changes in the Rates at Which Firms Enter and Exit the Market	3
Changes in New Firms and Their Share of Employment	3
Differences by Economic Sector	3
The Effect of the Coronavirus Pandemic on Entrepreneurship	4
BOX 1. THE GIG ECONOMY AND ITS EFFECT ON ENTREPRENEURSHIP	6
The Impact of the Decline in Entrepreneurship on Economic Productivity	7
The Contribution of Entrepreneurship to Greater Efficiency	8
Estimates of the Impact of the Decline in Entrepreneurship on Productivity	10
Factors Affecting Entrepreneurship	12
BOX 2. SOURCES OF ECONOMIC EFFICIENCY THAT FAVOR INCUMBENT FIRMS	13
Financing Constraints and Economic Conditions	14
Demographics and Immigration	15
The Regulatory Environment	17
Federal Policies to Support Entrepreneurship	20
Financing and Financial Support for New Firms	20
Immigration of Highly Skilled Workers and Entrepreneurs	26
The Regulatory Environment's Effects on Competition	28
Appendix: Alternative Measures of Entrepreneurship	29
List of Tables and Figures	31
About This Document	32

Notes

Numbers in the text and figures may not add up to totals because of rounding.

Some of the figures in this report use shaded vertical bars to indicate periods of recession. (A recession extends from the peak of a business cycle to its trough.)

The terms “firm,” “business,” and “company” are used interchangeably throughout the report. The term “establishment” denotes a single physical location where business is conducted. Accordingly, a firm can own one or more establishments.



Federal Policies in Response to Declining Entrepreneurship

Summary

Entrepreneurship is beneficial to the economy in many ways. Policymakers may therefore be concerned about its continued decline over the past four decades and the implications for economic growth. In this report, the Congressional Budget Office examines the falloff in entrepreneurship, its potential economic consequences, factors that have contributed to it, and ways that federal policies could be changed to reverse the trend.

How Much Has Entrepreneurship Declined Since the 1980s?

Several measures of entrepreneurship have declined since the early 1980s. The rate at which new businesses formed decreased from 10 percent of all firms that existed in 1982 to 8 percent in 2018. New firms (defined here as those less than five years old with at least one employee on the payroll) constituted 38 percent of all businesses in 1982 but were only 29 percent of them in 2018. During that period, new firms' share of employment fell by a third, from 14 percent to 9 percent. The decline in new firms' share of employment was fairly consistent in both the retail and services sectors throughout the period, whereas the share of employment attributed to new businesses in the information and high-tech sectors rose in the 1990s, falling thereafter. Although an early indicator of entrepreneurship—applications for an employer identification number submitted by likely employers to the Internal Revenue Service—dropped precipitously after the start of the 2020 coronavirus pandemic, it subsequently rebounded strongly.

How Has the Decline in Entrepreneurship Affected Productivity Growth?

Entrepreneurship plays an important role in allocating resources more efficiently throughout the economy, thereby making it more productive. Innovative new firms can be a source of technological change, often introducing new products and services. New companies can also increase the productivity of workers by improving methods of production, and they can bring competitive

discipline to markets, forcing other companies to become more efficient to maintain business. Even the potential for new firms to enter a market can influence the behavior of existing firms.

The decline in entrepreneurship over roughly the past 40 years appears to have had a moderate impact on the overall growth of productivity. In particular, the decline was related to a falloff in labor productivity of at least 3 percent to 4 percent by the mid-2010s, in CBO's assessment, compared with what it would have been otherwise. In the 1990s, new firms in the information and high-tech sectors supplied products that were useful to a wide range of industries, and the growth of those firms was accompanied by greater productivity growth in the economy. When the growth of new businesses in those sectors later declined, so did the growth rate of productivity.

The effects of a reduction in entrepreneurship in a given sector of the economy may also depend on the cause underlying the reduction. In some cases, economic forces that led to increases in productivity, such as economies of scale and scope, hampered the formation of successful new firms. For example, technological advances commercialized during the last half of the 1990s enabled large incumbent firms in the retail sector to implement new, more efficient business practices. As a result, the decline in new firms and their employment share in that sector was associated with an *increase* in productivity growth, as smaller, local retailers could not compete with the large incumbent firms.

What Factors Have Caused a Decline in Entrepreneurship?

Financing constraints and broader economic conditions have played a significant role in the decline of entrepreneurship, particularly in the aftermath of the 2007–2009 recession. New firms are especially vulnerable to economic downturns and the concomitant adverse effects on revenues and bank lending. Although economic

uncertainty appeared to hinder entrepreneurship during the first few months of the coronavirus pandemic, the subsequent upswing in applications for employer identification numbers from potential new employers suggests many new firms may have launched in the latter part of 2020.

Demographic trends have also affected entrepreneurship. A decline in the growth of the working-age population, from 3 percent at the end of the 1970s to 1 percent at the end of the 2010s, has been linked to the decrease in new businesses. In addition, people between the ages of 35 and 54 are more likely to be entrepreneurial—and successful in their new businesses—than those of other ages, and their share of the workforce has fallen since 2000. Despite those overall demographic declines, the proportion of foreign-born people among the U.S. population grew from 10 percent in 1998 to 14 percent in 2018. Highly educated, foreign-born workers add to the pool of qualified employees for new firms, especially in the high-tech sector. Immigrants have also been more likely than native-born Americans to start new businesses. Looking forward, CBO expects a substantial drop in net immigration to the United States from 2020 through 2023 because of the coronavirus pandemic.

Regulation affects businesses, but the research literature provides mixed evidence regarding the quantitative impact of the regulatory environment on the formation and growth of new businesses. Some individual regulations discourage the entry of new firms, and some studies show that increases in the overall stringency of federal regulations hamper entrepreneurship, but others do not. The details of the regulations and of the industry setting affect the impact. The patent system generally encourages entrepreneurship by providing a legal framework that protects intellectual property, although large concentrations of patents held by incumbent firms can be a barrier to new businesses. Some economists are becoming concerned about diminished competition among firms—especially those firms involved in Internet-based commerce—and its harm to entrepreneurship, and about the impact of noncompete clauses. Changes in health care policy over the past 40 years have had varying effects. The continuation of group coverage upon separation from employment and an increase in the federal income tax deduction for health insurance provided to the self-employed have made entrepreneurship more attractive for many people than it was at the start of the 1980s. The passage of the Affordable Care Act (ACA) in

2010 made health insurance coverage cheaper and more accessible for some entrepreneurs but more expensive for others.

What Federal Policies Would Support Entrepreneurship?

Federal policies can address many of the factors that inhibit entrepreneurship. Policymakers could create a program to give new firms access to credit. A challenge is that new firms lack a track record demonstrating their ability to repay loans, and roughly half of them will fail within their first five years, on average. Providing them with greater access to credit would entail costs to the federal government and would run the risk of failures; it could also provide an incentive for firms to modify their operations solely to qualify for the subsidized credit. Policymakers could expand the Small Business Administration's (SBA's) existing credit programs. The SBA charges borrowers fees that are intended to be large enough to offset losses from loans to small businesses that are not fully repaid, which means that the agency has a limited ability to take risks on new firms. Expanded SBA lending would also potentially provide benefits to small businesses that are not new ones. Although virtually all new firms are small (say, with fewer than 100 employees), only 30 percent of such small businesses are new.

Policymakers could increase financial support for entrepreneurship in several other ways. They could direct a share of the federal government's spending on research and development (R&D) to be set aside for new companies or increase the existing share provided for small firms. Policymakers could take steps to subsidize products that are more likely to be supplied by cutting-edge new firms, or they could make tax preferences used primarily by small businesses more generous. A drawback to such approaches is that financial support directed to small businesses without regard to their age may not reach many innovative new companies. Moreover, it can be challenging for the government to predict which firms will be innovative and to determine which sectors of the economy to target with support for innovation.

Highly skilled immigrants have founded many companies in the United States that have increased innovation, job creation, and economic growth. Policymakers could support entrepreneurship by expanding programs that provide visas for highly skilled workers and entrepreneurs immigrating to the United States. A challenge here is

that expanding the pool of qualified workers through immigration would also entail a broader set of effects on businesses and wages throughout the economy. And programs that attempt to identify immigrant entrepreneurs are costly to administer and susceptible to abuse.

The federal government could make regulatory policies less burdensome for new firms in particular. Policymakers could also increase the scrutiny of incumbent firms' potentially anticompetitive actions directed toward rival start-ups. Finally, concerns about the impact of diminished competition on the formation and growth of new firms could be addressed by restricting the use of noncompete contracts. The effects of such policy changes on entrepreneurship would depend largely on the details of the policies.

The Decline in Entrepreneurship

Long before the coronavirus pandemic affected the economy, entrepreneurship had significantly diminished over the past four decades. The rate at which new firms entered the economy (relative to the total number of firms in operation) fell by about a quarter between 1982 and 2018, and the rate at which firms left the economy also fell below its long-term average.¹ As a result, new businesses account for both a smaller share of firms and a smaller share of employment. Those trends do, however, mask differences between different sectors of the economy at certain times. New businesses' share of employment fell in both the retail and services sectors, whereas it rose in the information and high-tech sectors from the mid-1990s through the early 2000s, falling thereafter.

Changes in the Rates at Which Firms Enter and Exit the Market

The annual rate at which new firms were created decreased from about 10 percent of all businesses in 1982 to about 8 percent in 2018, the most recent year for which such data are available (see Figure 1). Much of that decline occurred during recessions, including a drop of 2 percentage points over the period encompassing and immediately after the 2007–2009 recession. The rate at which firms exited the market typically ranged between 8 percent and 9 percent from 1982 to 2018, though that

rate spiked to a high of approximately 10 percent during the 2007–2009 recession. After the recession, it fell to 7 percent, not returning to its typical range until 2017.

Because this report examines the impact of entrepreneurship on economic productivity, CBO has only included firms with at least one employee on the payroll (sometimes called employer firms) in its measures of entrepreneurship. Several studies have linked such firms to trends in labor productivity. Some people start a business activity with no intent to grow it beyond their own involvement—and those activities are less likely to affect productivity and economic growth than those with employees. If those people eventually hire workers, their businesses will be included in the measure of employer firms used in this report. A person who starts a solo business contributes to the economy as well. But measures of entrepreneurship that include firms without employees, or other businesses that are not expected to grow their employment, are less useful for the purposes of the analysis presented here than employer firms, whose growth and decline have a larger impact on productivity and economic growth (see the appendix).

Changes in New Firms and Their Share of Employment

As the rates at which businesses enter and exit the market have declined over the past several decades, new firms—that is, firms started within the past five years with at least one employee—have shrunk considerably, both as a share of all businesses and in terms of their share of employment (see Figure 2). In 1982, new businesses constituted 38 percent of all employer firms; in 2018, they accounted for only 29 percent. Over that same period, the percentage of workers in those firms as a share of all workers fell by about a third, from about 14 percent to approximately 9 percent.

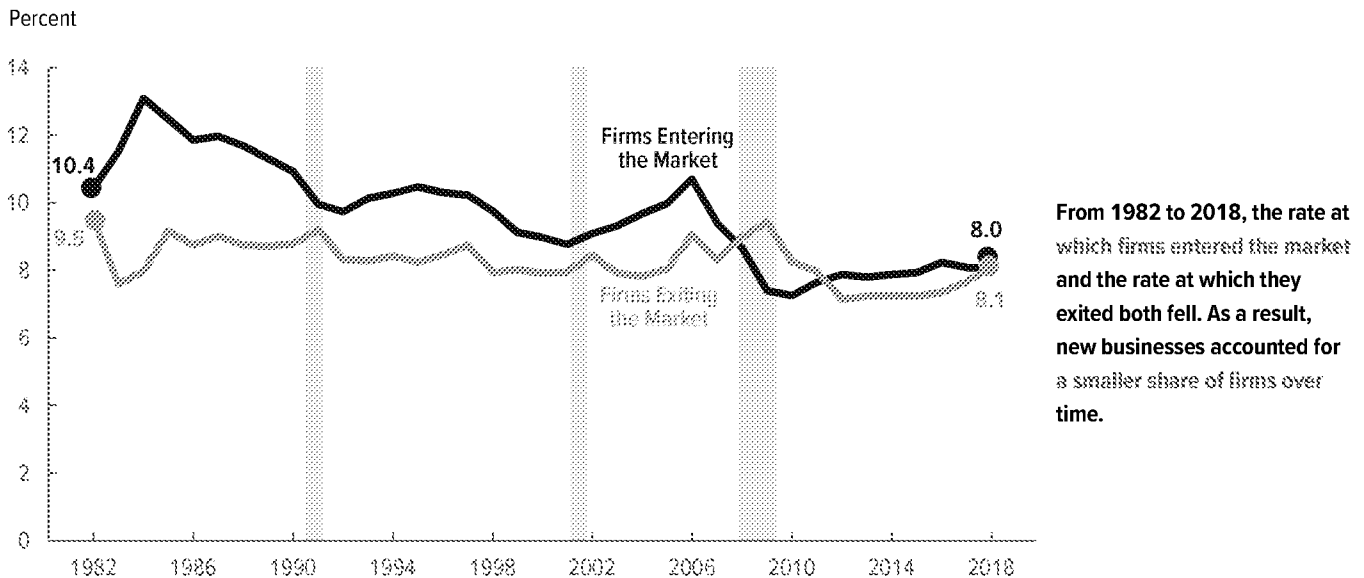
With the advent of the online “gig” economy, a growing number of people now freelance and have a different relationship with the firms that administer their work than with a traditional employer. Although the increased presence of gig workers may have reduced measures of employment at new firms, such workers account for only a small part of the decline in those measures (see Box 1).

Differences by Economic Sector

At certain times, the decrease in new firms' share of employment has varied depending on the sector of the economy in which they operate (see Figure 3 on

1. The rates at which firms enter and exit the market are measured relative to the average number of firms in operation throughout a given year and the preceding one. That approach mitigates the short-term effects of year-to-year fluctuations in the total number of firms on the reported rates of firms entering and exiting the market.

Figure 1.

Rates at Which Firms Entered and Exited the Market, 1982 to 2018

Data source: Congressional Budget Office, using data from the Census Bureau's Business Dynamics Statistics. See www.cbo.gov/publication/56906#data.

page 8). For example, the share of employment at new firms fell fairly steadily in both the retail and services sectors between 1982 and 2018 but was more stable in manufacturing. (Because services constitute such a large portion of economic activity, trends in the overall economy mirror trends in that sector.) And in the information sector, as well as the related high-tech and high-tech manufacturing sectors, the share of employment at new firms rose from the mid-1990s through the early 2000s, falling thereafter.²

The construction sector also saw greater-than-average declines in new firms' share of employment. Like segments of the services sector (such as repair services and personal services), the construction sector includes a larger share of "middle-skilled" business owners than other sectors. Middle-skilled workers are those with some additional education or training after high school, including an associate's degree or training at a technical or trade school, but without a four-year college degree. Entrepreneurship provides a source of economic opportunity for people without a college degree or high-tech

skills. A person with some technical skills can start a construction company, for example, with a moderate amount of capital investment.

Those sector-specific differences in the share of employment have had important implications for productivity growth. As is discussed in more detail below, when the employment share of new firms in the information sector grew, productivity growth increased. And when the employment share in that sector declined, productivity growth declined. But at times, the employment share of new firms in the retail sector has had the opposite relationship with productivity growth, when large retailers were the primary drivers of productivity-enhancing changes in that sector.

The Effect of the Coronavirus Pandemic on Entrepreneurship

The initial spread of the coronavirus significantly curtailed entrepreneurship, but it strongly rebounded. A real-time indicator of early-stage entrepreneurial activity is the number of applications for an employer identification number submitted to the Internal Revenue Service by businesses that are especially likely to become employers. That number fell sharply after the declaration of the public health emergency in mid-March 2020 (see

2. For the definition of high-tech industries, see Daniel E. Hecker, "High-Technology Employment: A NAICS-Based Update," *Monthly Labor Review* (Bureau of Labor Statistics, July 2005), pp. 57–72, <https://go.usa.gov/xGWfb>.

Figure 2.

New Firms as a Share of Total Firms and Total Employment, 1982 to 2018

Percent



Firms less than five years old accounted for a much smaller percentage of all firms and employment in 2018 than they did in 1982.

Data source: Congressional Budget Office, using data from the Census Bureau’s Business Dynamics Statistics. See www.cbo.gov/publication/56906#data.

Figure 4 on page 9). But beginning in early June, the number of applications returned to and then substantially exceeded its prepandemic level. (The moderate falloff in applications during the final quarter of 2020 is consistent with seasonal trends.)

Many factors could be behind the rebound. Some applications may have been deferred from March and April, and others may represent acquisitions of businesses that were especially hard hit in the downturn, reflecting changes in ownership rather than the creation of new

firms. The increase in applications may also reflect a rise in early-stage entrepreneurial activity. For example, layoffs and social distancing may have prompted some people to start businesses offering goods and services that they had previously provided as an employee of another business (although this measure excludes those who become self-employed without hiring employees). Fewer opportunities for employment may also have pushed people to try new ideas for businesses. Starting a new business may have become easier because fewer firms were competing for workers, equipment, and investment.

Box 1.

The Gig Economy and Its Effect on Entrepreneurship

Developments in technology have allowed more of the activities of freelance workers to be marketed and managed online (in some cases through smartphone-based “apps,” or software applications). Many workers participating in what is called the online gig economy are treated as independent contractors—rather than employees—by the firms that manage their work.

The growing popularity of services provided by such firms could affect measures of employment attributed to new businesses. Ride-sharing services, for example, match passengers to participating drivers (based on proximity and availability) who supply their own vehicles and are treated as contractors working for themselves rather than as employees. Drivers working in some food-delivery services operate in much the same way. As such start-ups expand, they might not register employment growth in the same way as other businesses with workers on their payrolls.

Despite its rapid growth, the online gig economy constitutes only about 1 percent of the workforce—still too small a percentage to account for much of the declining share of employment attributed to new firms.¹ The number of people reporting earnings from labor in the online gig economy rose from roughly 22,000 in 2012 to about 2,000,000 in 2016.² Even if all those people were considered to be employees of new firms (with app-based drivers, for example, thought of as employees), that 1 percent of the workforce would not significantly affect the decline in such firms’ share of employment from approximately 14 percent in 1982 to 9 percent in 2018. The growth of the gig

economy reflects its appeal to those workers for whom it provides both income and flexibility, but so far that growth does not explain the decline in the employment share of new firms.

Nor does the decline appear to have been caused by changes in the shares of independent contractors and other self-employed workers who do not rely on an Internet-based company for their job assignments. Those shares have generally remained consistent over the past few decades.³

This analysis counts people working in the online gig economy in the same way that the companies that administer their work report them because classifying them as individual new businesses would overstate the rate at which new firms are created. Those workers generally are not starting independent businesses intended to grow beyond their own involvement. Some people rely on gig-economy work as their primary source of income, whether on an ongoing basis or in between full-time work.⁴ Many other participants in the online gig economy do not earn much income from it (in 2016, most participants received a gross income of \$2,500 or less), and they use the work to supplement income earned elsewhere. Gig work shares characteristics with entrepreneurship in the sense that both gig workers and entrepreneurs have a good deal of control over their working hours.

Several recent developments may slow the growth of the online gig economy or even reduce its size. First, the classification of the workers as independent contractors rather than employees may be changing, which would affect the regulations governing gig workers and companies. A law passed in 2019 in California—a state with a large presence of online gig-economy workers—made gig workers and other nonsalaried workers more like salaried employees for the purpose of labor regulations, such as those governing eligibility for the minimum wage and overtime pay, unemployment insurance and family leave, and bargaining rights. Those changes

1. Some analysts estimate the share of gig workers in the economy to be as high as 10 percent or more. See Laura Schultz, *Defining and Measuring Gig Work* (Rockefeller Institute of Government, March 2020), <https://tinyurl.com/y7g522pf>; and Board of Governors of the Federal Reserve, *Report on the Economic Well-Being of U.S. Households in 2019* (May 2020), <https://go.usa.gov/xGGcn>. Those estimates include work in jobs outside of the online gig economy or include arrangements that have traditionally been informal or taken the form of self-employment. Other estimates that produce larger numbers include people who work in gig jobs only occasionally or for small amounts of time each week. The measure used here is intended to capture gig economy workers who might otherwise have been considered as employees if not for the online gig arrangement.

2. See Brett Collins and others, *Is Gig Work Replacing Traditional Employment? Evidence From Two Decades of Tax Returns* (Internal Revenue Service, Statistics of Income, March 2019), p. 3, <https://go.usa.gov/xGWaG> (PDF, 2.8 MB); and Bureau of Labor Statistics, *Electronically Mediated Work: New Questions in the Contingent Worker Supplement* (September 2018), <https://go.usa.gov/xGWAs>.

3. See Brett Collins and others, *Is Gig Work Replacing Traditional Employment? Evidence From Two Decades of Tax Returns* (Internal Revenue Service, Statistics of Income, March 2019), p. 3, <https://go.usa.gov/xGWaG> (PDF, 2.8 MB).

4. Other characteristics of workers in the online gig economy suggest that its labor force is quite distinct from both independent contractors and other self-employed individuals elsewhere, as well as workers in the economy overall. For example, participants in the online gig economy are much more likely to be single males. They are also heavily concentrated in large cities and effectively absent from many rural areas.

Continued

Box 1.

Continued

The Gig Economy and Its Impact on Entrepreneurship

raised the cost of employing such workers and limited their ability to work flexible hours or part time—a hallmark of the gig economy. To the extent that the associated cost is passed on to customers, it would reduce the demand for the services the workers provide and thus tend to limit growth of the online gig economy. A ballot initiative in the November 2020 election reversed part of the law, effectively exempting

drivers for app-based transportation and delivery services from it. An ongoing challenge for workers in the gig economy is that the spread of the coronavirus has led to much weaker demand for many types of services they provide, such as ride-sharing and home-sharing (although partially offset by an increased demand for food delivery), and it is unclear when it will increase.

New firms may have been formed in response to business opportunities as a result of the conditions created by the pandemic, such as an increase in demand for delivering health- and education-related services remotely. Some of the increase in new business applications may be the result of fraudulent submissions made to obtain federal funds through the SBA's Economic Injury Disaster Loans program.³ In any case, the marked upswing in applications stands in stark contrast to the trend in entrepreneurship during the 2007–2009 recession, when the rate of firms entering the market fell sharply without subsequently returning to its prerecession level.

The economic contraction that initially chilled the formation of firms at the beginning of the pandemic also threatened the viability of businesses started in the past few years. Although weekly payroll data are not specifically reported for new firms, the data for small firms show that, from mid-February to late April, employment fell among them by more than it did among large firms.⁴ Increases in employment since then have made up for some, but not all, of those losses, in absolute terms. The remaining percentage declines are similar among firms of different sizes, suggesting that the net effects of the pandemic (and policies put in place to respond to it) may

not ultimately be much different for small businesses than for large ones.

The Impact of the Decline in Entrepreneurship on Economic Productivity

Entrepreneurs play an important role in allocating the economy's resources more efficiently, thereby increasing economic productivity. Start-ups often commercialize new products, and new firms that supply the same goods and services more efficiently than incumbents can force their competitors to perform better or lose business. Entrepreneurship has also long been viewed in America as a path to upward mobility for people of all skill levels and educational backgrounds, and not all new businesses are oriented toward future expansion. Some “mom-and-pop” stores, for example, are established to serve customers in their local neighborhood. But for those businesses that are geared toward future growth, the ability to expand is widely considered an important indicator of the economy's dynamism.

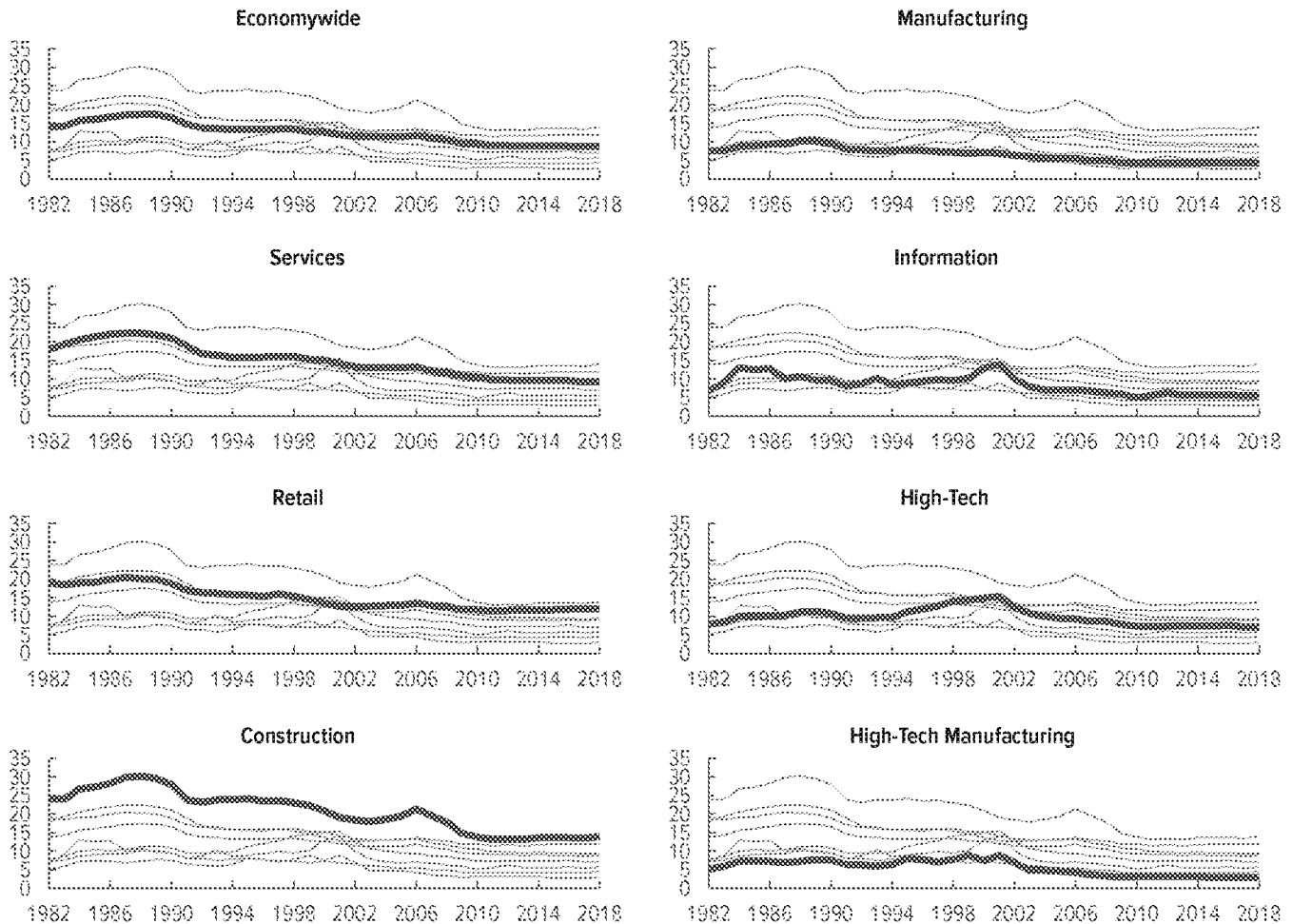
The decline in entrepreneurship—in terms of either the falloff in the rate of business formation or the decline in new firms' share of employment—has been linked to a modest decrease in productivity growth. In particular, evidence suggests that labor productivity was 3 percent to 4 percent lower in the mid-2010s because of the decline. In some cases, alternatively, a decrease in entrepreneurship could be the consequence of the economy's weak productivity performance rather than the cause. Additionally, in certain periods and industries, economies of scale and scope and incentives to perform R&D boosted productivity while favoring incumbent firms over new ones.

3. See Office of the Inspector General, “Serious Concerns of Potential Fraud in EIDL Program Pertaining to the Response to COVID-19” (Small Business Administration, July 2020), <https://go.usa.gov/x7hpz>.
4. See Tomaz Cajner and others, *The U.S. Labor Market During the Beginning of the Pandemic Recession*, Working Paper 2020-58 (Becker Friedman Institute for Economics at the University of Chicago, July 2020), <https://tinyurl.com/y6445pt7> (PDF, 1 MB).

Figure 3.

New Firms' Share of Employment, by Sector, 1982 to 2018

Percent



Between 1982 and 2018, new firms' share of employment fell in all sectors except the information and high-tech manufacturing sectors. In those sectors, employment rose from the mid-1990s to 2001, falling thereafter.

Data source: Congressional Budget Office, using data from Ryan A. Decker and others, *Changing Business Dynamism and Productivity: Shocks vs. Responsiveness*, Working Paper 24236 (National Bureau of Economic Research, January 2018) www.nber.org/papers/w24236. See www.cbo.gov/publication/56906#data.

The Contribution of Entrepreneurship to Greater Efficiency

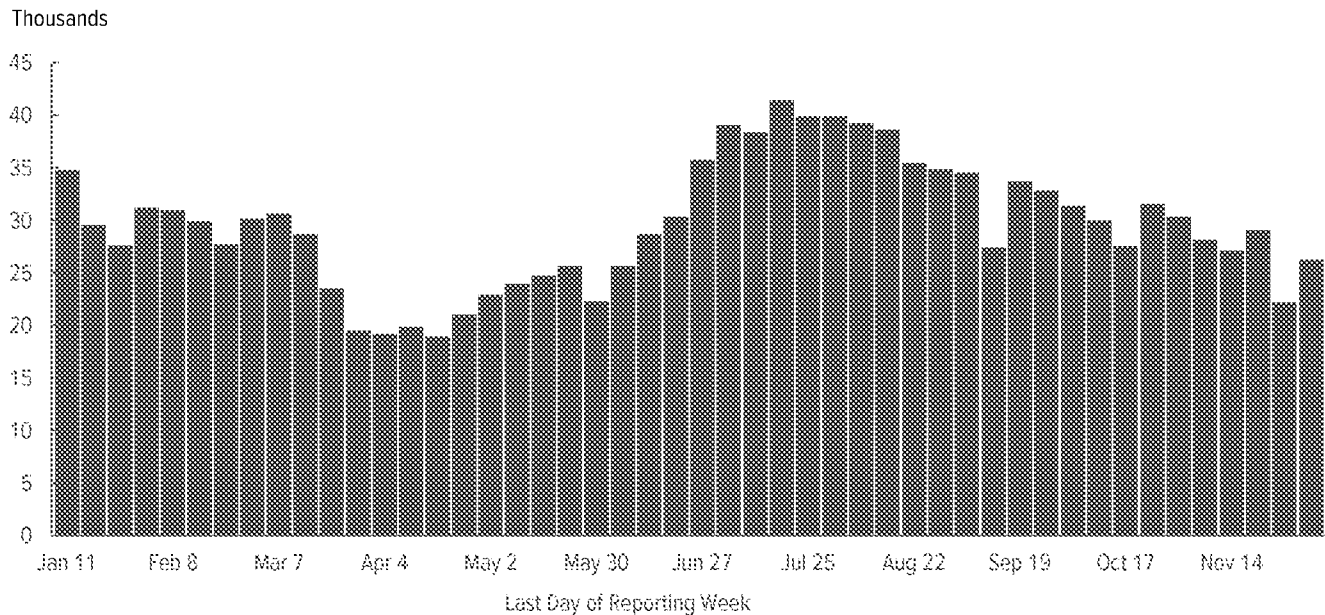
Entrepreneurship brings greater efficiency to the economy in several ways, one of which is by introducing new products and technologies that increase productivity.⁵ Although relatively few new businesses perform R&D, those that do are more R&D-intensive and more likely than older businesses to test new technologies. Indeed,

5. Frederic M. Scherer, *Industrial Market Structure and Economic Performance*, 2nd ed. (Rand McNally & Co., 1980), pp. 437–438.

most recently created manufacturing firms that have grown to become very large carried out R&D when they were new.⁶

6. See Daron Acemoglu and others, "Innovation, Reallocation, and Growth," *American Economic Review*, vol. 108, no. 11 (November 2018), p. 3468 and p. 3471, <https://tinyurl.com/y73vatcs>; and Nikolas Zolas and others, "Measuring Technology Adoption in Enterprise-Level Surveys: The Annual Business Survey" (paper presented at the 2020 American Economic Association meeting, San Diego, Calif., January 3, 2020), <https://tinyurl.com/y7qwjtpc> (PDE 3.9 MB).

Figure 4.

Number of Weekly High-Propensity EIN Applications in 2020

Data source: Congressional Budget Office, using data from the Census Bureau's Business Formation Statistics. See www.cbo.gov/publication/56906#data.

Data for the first week in 2020 are excluded because substantially fewer EIN applications are usually submitted at the beginning (and at the end) of a calendar year, and their number is not informative of the trend in subsequent weeks.

On the basis of characteristics reported on Internal Revenue Service (IRS) Form SS-4, the IRS identifies applicants that have a high propensity of becoming businesses with payrolls. Those characteristics include being a corporate entity; anticipating hiring employees, purchasing a business, or changing organizational type; expecting to pay wages by a specific date; or being in the manufacturing, retail, health care, or restaurant/food-service industry.

EIN = Employer Identification Number.

Entrepreneurship also helps the economy when productive new firms increase their employment, thus promoting greater productivity as workers move from older, less efficient businesses to newer, more efficient ones. Those new businesses may provide new products and technologies, or they may do a better job of providing existing products and services, including middle-skill services such as construction. Entrepreneurship, to a certain extent, involves a variety of skills (not just technical skills) because the entrepreneur has to maintain a financially viable business.⁷ Although workers move among businesses of all ages, new firms typically display an “up-or-out” dynamic—grow and survive or shrink and eventually go out of business. Job gains at growing firms account for much of the employment lost at unsuccessful

ones. When new businesses do not emerge and grow, productivity languishes.⁸

As discussed later in the report, several federal programs established to bolster entrepreneurship target their support to firms of a certain size rather than to new firms. But it turns out that new businesses are more closely associated with increased employment than small ones.⁹ And while nearly all new firms begin small, the rapid

7. See Edward P. Lazear, “Balanced Skills and Entrepreneurship,” *American Economic Review*, vol. 94, no. 2 (May 2004), pp. 208–211, www.jstor.org/stable/3592884.

8. See Ryan A. Decker and others, “The Role of Entrepreneurship in U.S. Job Creation and Economic Dynamism,” *Journal of Economic Perspectives*, vol. 28, no. 3 (Summer 2014), pp. 7–8, www.aeaweb.org/articles?id=10.1257/jep.28.3.3.

9. See John Haltiwanger, Ron S. Jarmin, and Javier Miranda, “Who Creates Jobs? Small Versus Large Versus Young,” *Review of Economics and Statistics*, vol. 95, no. 2 (May 2013), pp. 347–361, https://doi.org/10.1162/REST_a_00288; and Congressional Budget Office, *Small Firms, Employment, and Federal Policy* (March 2012), www.cbo.gov/publication/43029.

growth of highly productive new firms distinguishes them from other small firms (of any age).¹⁰

A third way that entrepreneurship makes the economy more productive is by providing market discipline. New companies can compel established ones to improve their performance to maintain business, or they can even force them from the market. Such market discipline also prevents firms from exercising market power and raising prices to an extent that would be inefficient for the economy. If businesses that supply goods and services to contestable markets raise the prices of their products too high, new firms will have an incentive to enter the market and take customers.

Of course, not all new businesses contribute to the economy's growth. A large share fails within a few years of forming, and some that survive were established for reasons that may limit their eventual economic impact.¹¹

Estimates of the Impact of the Decline in Entrepreneurship on Productivity

The decline in entrepreneurship appears to be related to a recent decrease in productivity growth. Measures of productivity account for the amount of output produced by the economy using specific quantities of labor (or for some measures, labor and capital). The decline in entrepreneurship is associated with a falloff in labor productivity from at least 3 percent to 4 percent in the mid-2010s, in CBO's assessment.

Entrepreneurship has increased productivity through the innovations that new firms provided and the growth of those new firms, although it is difficult to know precisely how much each of those factors contributed to the increase. The shares of employment attributed to new firms in the high-tech sector (including the high-tech

manufacturing and service industries) and the information sector (which also includes a number of high-tech industries) rose from the mid-1990s through 2001. Firms that are more productive than their competitors tend to grow. Newly formed high-tech companies that were more productive than their competitors increased their employment more rapidly during that period than such firms had previously. The advances in information technology and communications brought about by those firms also allowed businesses in other sectors to increase their productivity, which in turn corresponded to an uptick in productivity in the broader economy: At its high point in 2001, the growth rate of labor productivity was more than double what it was in 1995 (see Figure 5).¹² After 2001, the shift in employment from low-productivity firms to high-productivity firms in the high-tech sector was much less pronounced.

In fact, what was true for the high-tech sector after 2001 was also true for the economy overall: More-productive firms tended to add fewer workers than before, while less-productive firms tended to shed fewer employees. Estimates indicate that labor productivity would have been 4 percent higher by 2013 had such an economy-wide decline in responsiveness not occurred. A similar pattern characterized changes in overall labor and capital productivity.¹³ As new firms innovated or otherwise improved on the performance of existing companies, the decline in new firms' share of employment after the early 2000s meant that they contributed less to overall productivity growth than they had in earlier periods. As a result, that measure also indicates that productivity would have been about 4 percent higher by 2013 if new firms had maintained their earlier level of contribution to productivity.¹⁴

10. See John Haltiwanger and others, "High Growth Young Firms: Contribution to Job, Output, and Productivity Growth," in John Haltiwanger and others, eds., *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges* (University of Chicago Press, September 2017), pp. 11–62, www.nber.org/chapters/c13492.

11. See Erik Hurst and Benjamin Wild Pugsley, *What Do Small Businesses Do?* (Brookings Papers on Economic Activity, September 2011), <https://tinyurl.com/y8kvf7ep>; and Antoinette Schoar, "The Divide Between Subsistence and Transformational Entrepreneurship," in Josh Lerner and Scott Stern, eds., *Innovation Policy and the Economy*, Volume 10 (National Bureau of Economic Research, 2010), pp. 57–81, <https://tinyurl.com/y99v89h5>.

12. For a discussion of the nature of those advances in information technology and communications, see Congressional Budget Office, *The Role of Computer Technology in the Growth of Productivity* (May 2002), www.cbo.gov/publication/13675.

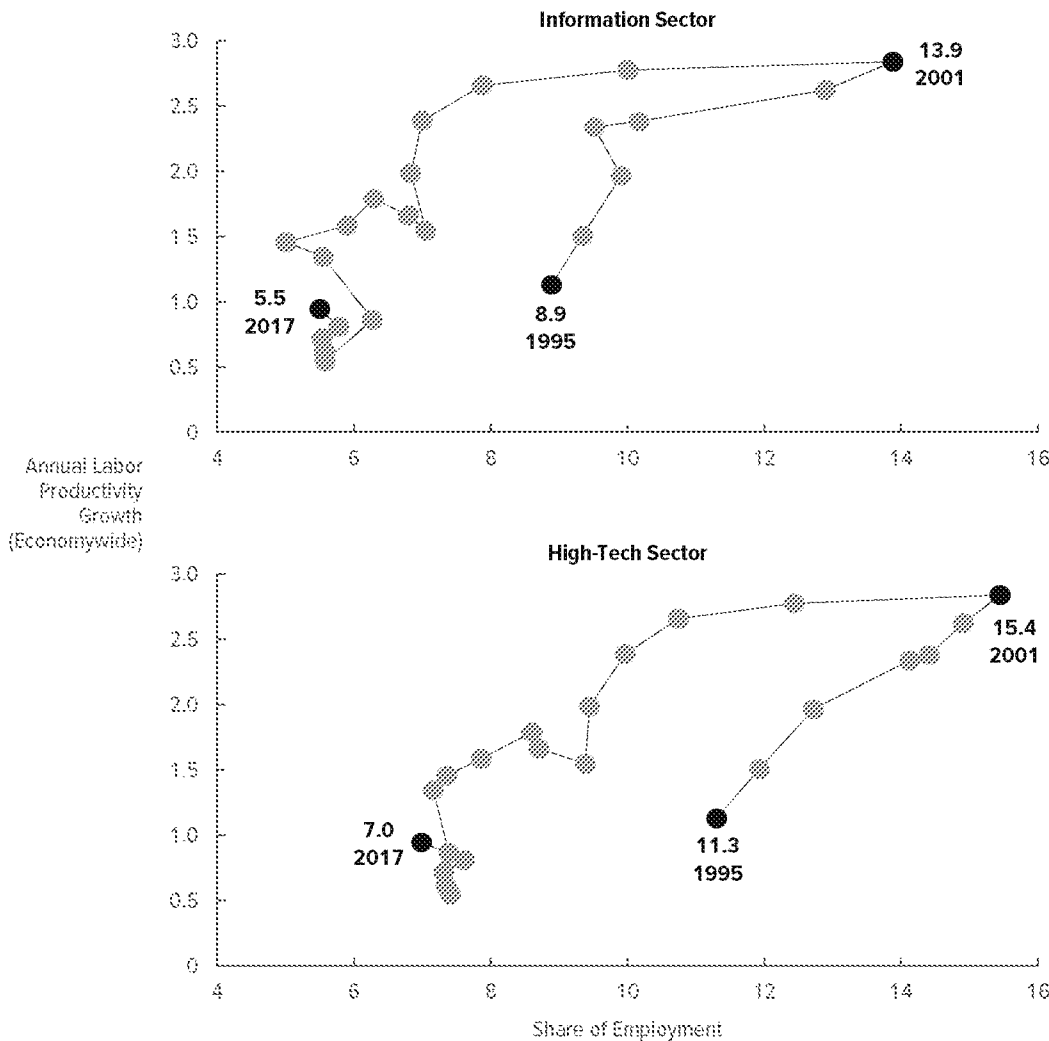
13. See Ryan A. Decker and others, "Changing Business Dynamism and Productivity: Shocks Versus Responsiveness," *American Economic Review*, vol. 110, no. 9 (September 2020), pp. 2859–2898, <https://tinyurl.com/y4ber8kk>.

14. See Peter J. Klenow and Huiyu Li, "Innovative Growth Accounting," in Martin Eichenbaum and Erik Hurst, eds., *NBER Macroeconomics Annual 2020*, Volume 35 (National Bureau of Economic Research, forthcoming), <https://tinyurl.com/yaphjsba>; and Daniel Garcia-Macia, Chang-Tai Hsieh, and Peter J. Klenow, "How Destructive Is Innovation?" *Econometrica*, vol. 87, no. 5 (September 2019), pp. 1507–1541, <https://doi.org/10.3982/ECTA14930>.

Figure 5.

Trends in New Firms' Share of Employment, by Sector, and Labor Productivity Growth in the Economy, 1995 to 2017

Percent



New firms in the high-tech and information sectors that were more productive than their competitors increased their employment more rapidly from the mid-1990s through 2001 than they had previously. The technological advances brought about in those sectors allowed firms throughout the economy to increase their productivity. After 2001, the shift in employment from low-productivity firms to high-productivity firms in the high-tech sector—and the broader economy—ebbed.

Data source: Congressional Budget Office, using data from the Bureau of Economic Analysis, the Bureau of Labor Statistics, and from Ryan A. Decker and others, *Changing Business Dynamism and Productivity: Shocks vs. Responsiveness*, Working Paper 24236 (National Bureau of Economic Research, January 2018), www.nber.org/papers/w24236. See www.cbo.gov/publication/56906#data.

The percentage changes in labor productivity represent five-year moving averages.

Deriving such estimates of productivity gains is complicated because measures of productivity may tend to understate the economy's growth from entrepreneurial innovation. When a new product drives an existing one from the market, the improvements in the new product's quality may not be captured in statistical data if the new product is not directly comparable to the old one. When cell phones were introduced, for example, it was

clear that they offered greater functionality than landline phones, but the transformative nature of their improved capabilities made it difficult to estimate the change in quality. That is because statistical agencies' surveys that do not include the defunct product (because it no longer exists) will fail to account for the new one (because it is

not categorically identical to what it replaced).¹⁵ It can therefore be difficult to determine the improvement that the new product represents—at least in terms of the additional amount of output it creates in the economy.¹⁶

Regardless of whether entrepreneurs create innovative products, newly established businesses are typically more productive than the firms that preceded them. As the rate of the formation of new firms declined, successively smaller shares of more productive new firms joined the economy. As a result, the productivity boost from consecutive waves of new firms—each wave containing, on average, more productive firms than the one before it—diminished over time. After 1980, that decline in new firms reduced productivity growth by an average of one-tenth of one percentage point each year. By 2014, the cumulative effect left labor productivity 3 percentage points lower than it would otherwise have been.¹⁷

The closing of less-productive businesses in conjunction with forming new ones appears to have had a positive effect on productivity as well. Between one-fifth and one-quarter of all total factor productivity growth (a measure that accounts for both labor and capital use) between 1977 and 1997 in a set of manufacturing industries was attributable to both the entry of new plants and the exit of older, less productive ones.¹⁸ The extent to which those measured effects can be combined or overlap each other, though, is ultimately unclear.

15. See Philippe Aghion and others, “Missing Growth From Creative Destruction,” *American Economic Review*, vol. 109, no. 8 (August 2019), pp. 2795–2822, <https://doi.org/10.1257/aer.20171745>.

16. See Jerry Hausman, “Cellular Telephone, New Products, and the CPI,” *Journal of Business and Economic Statistics*, vol. 17, no. 2 (April 1999), pp. 188–194, <https://economics.mit.edu/files/1023> (PDF, 693 KB); and Erica L. Groshen and others, “How Government Statistics Adjust for Potential Biases From Quality Change and New Goods in an Age of Digital Technologies: A View From the Trenches,” *Journal of Economic Perspectives*, vol. 31, no. 2 (Spring 2017), pp. 187–210, <https://tinyurl.com/y3zfp8pa>.

17. See Titan Alon and others, “Older and Slower: The Startup Deficit’s Lasting Effects on Aggregate Productivity Growth,” *Journal of Monetary Economics*, vol. 93 (January 2018), pp. 68–85, <https://tinyurl.com/y6y2f1i6>.

18. A firm may have one or more plants. See Lucia Foster, John Haltiwanger, and Chad Syverson, “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?” *American Economic Review*, vol. 98, no. 1 (March 2008), pp. 394–425, <https://tinyurl.com/y5xgkaf6>.

The correspondence between the declines in entrepreneurship and productivity growth might also have been influenced by other independent factors. For instance, having fewer opportunities to exploit new technology could affect both productivity growth and the formation of new firms. Some scholars believe that there has been less potential for technological advances (or fewer novel ideas to be exploited) since the mid-2000s than during much of the 20th century. In that case, the decline in productivity growth over the past decade would reflect the beginning of an era in which economy-boosting innovation occurred much less often because the potential for it had decreased. Fewer prospects for innovating could have then indirectly resulted in the declining rate at which new businesses were created.¹⁹

Not all declines in entrepreneurship correspond to declines in productivity growth. In at least one case, that of retail trade in the 1990s, a decrease in entrepreneurship was associated with *faster* productivity growth, as large retailers used innovations in information technology and logistics to undercut the prices of smaller stores. Economic forces that favor incumbent firms, such as economies of scale, have at times made the economy more efficient while working against the formation of successful new businesses (see Box 2). Ultimately, having some diversity in the size and structure of businesses within an industry may enhance the prospects for long-term economic growth.²⁰

Factors Affecting Entrepreneurship

Certain factors have contributed to the falloff in entrepreneurship since the early 1980s. New firms are especially vulnerable to economic downturns and reductions in bank lending and were thus greatly harmed by the 2007–2009 recession and its aftermath. Demographic trends also appear to have played a role in the long-term decline of entrepreneurship. The growth rate of the labor force is smaller today than it was in the 1980s, and the share of the workforce most likely to consist of successful entrepreneurs has fallen since 2000. Regulation affects

19. See Robert J. Gordon, *The Rise and Fall of American Growth*, (Princeton University Press, 2018); and Nicolas Bloom and others, “Are Ideas Getting Harder to Find?” *American Economic Review*, vol. 110, no. 4 (April 2020), pp. 1104–1144, <https://tinyurl.com/y9vy42lk> (PDF, 1,014 KB).

20. See Wesley M. Cohen and Steven Klepper, “The Tradeoff Between Firm Size and Diversity in the Pursuit of Technological Progress,” *Small Business Economics*, vol. 4, no. 1 (March 1992), pp. 1–14, www.jstor.org/stable/40228763.

Box 2.

Sources of Economic Efficiency That Favor Incumbent Firms

Productivity growth does not necessarily require the constant presence of robust new firms in all sectors of the economy. Some sources of efficiency—such as economies of scale, economies of scope, and incentives to undertake research and development (R&D)—are associated with larger and more established firms rather than smaller and newer ones. Incumbent firms create many jobs, and competition among those firms over the past two decades has produced a form of economic dynamism reflected in shifts in industry profitability.¹

Economies of Scale and Vertical Integration. In many instances, firms can operate more efficiently as they become larger; changes in technology have sometimes reinforced the advantages that incumbent firms have over new—and typically much smaller—competitors. For example, at the turn of the 20th century, large firms exploited economies of scale by vertically integrating (that is, owning and controlling) input supply, manufacturing, and distribution.²

In the 1990s, the dynamism of the retail sector came primarily at the expense of new businesses, as incumbent firms created new, more efficient establishments. Older firms (such as Walmart) opened a multitude of “big-box” retail stores by exploiting advances in information technology to operate very efficiently at a large scale. As a result, the decline in the share of employment belonging to new firms in the retail sector during the 1990s was accompanied by a rise in productivity.³ A trend toward vertical integration and increased efficiency from operating at a greater scale may also be under way today. For example, some online merchants have begun operating their own delivery services.

In addition, globalization—in the form of increased foreign trade and offshoring—may require a sufficient scale of

operation that new firms lack, making it more difficult for them to compete. Successful businesses may locate their production abroad, reducing their employment growth in the United States. Sectors of the economy that are particularly exposed to foreign trade are thus more likely to experience a decline in new firms.⁴

Economies of Scope. Just as firms can operate more efficiently at larger sizes, they may also do so by offering a wider range of services or a broader set of products. Several established Internet-based companies have used their expertise with information technologies to launch new business lines. Examples of such entrepreneurial ventures by incumbent firms include cloud computing services, autonomous vehicles, and the creation of media content. The benefits of exploiting economies of scope are not, of course, exclusive to the Internet age: Many decades ago, large catalog companies used their expertise in distribution to venture into new business lines, such as selling prefabricated houses.

Incentives to Pursue R&D. Although some new companies carry out R&D more intensively than large established ones do, the latter account for the greatest amount of R&D spending and patenting activity.⁵ Some economists believe that the most promising advanced technologies today, such as the development and application of machine learning to improve production, require large operations well beyond the small size typical of new firms.⁶ Others believe that large firms with market power may have incentives to pursue R&D to help preserve that market power—although to the extent that the result is inefficiently high prices, doing so would inhibit rather than enhance efficiency.⁷

1. See Richard L. Clayton and others, “High-Employment-Growth Firms: Defining Them and Counting Them,” *Monthly Labor Review* (June 2013), pp. 1–14, <https://tinyurl.com/yc2tbp8>; and James Manyika and others, *Superstars: The Dynamics of Firms, Sectors, and Cities Leading the Global Economy* (McKinsey Global Institute, October 2018), <https://tinyurl.com/y5bpqb8q> (PDF, 935 KB).

2. See Alfred D. Chandler Jr., *Scale and Scope: The Dynamics of Industrial Capitalism* (Belknap Press, 1994), and *The Visible Hand: The Managerial Revolution in American Business* (Belknap Press, 1993).

3. See Lucia Foster, John Haltiwanger, and C. J. Krizan, “Market Selection, Reallocation, and Restructuring in the U.S. Retail Trade Sector in the 1990s,” *The Review of Economics and Statistics*, vol. 88, no. 4 (November 2006), pp. 748–758, <https://tinyurl.com/yyc676r>.

4. See Benjamin Wild Pugsley and Ayşegül Şahin, “Grown-up Business Cycles,” *The Review of Financial Studies*, vol. 32, no. 3 (March 2019), pp. 1102–1147, <https://doi.org/10.1093/rfs/hhy063>.

5. See Robert D. Atkinson and Michael Lind, *Big Is Beautiful: Debunking the Myth of Small Business* (MIT Press, 2018).

6. See Ajay K. Agrawal, Joshua S. Gans, and Avi Goldfarb, “Economic Policy for Artificial Intelligence,” in Josh Lerner and Scott Stern, eds., *Innovation Policy and the Economy*, Volume 19 (National Bureau of Economic Research, 2019), pp. 139–159, <https://tinyurl.com/y72nzyzj>; and Ufuk Akcigit and Sina T. Ates, *Ten Facts on Declining Business Dynamism and Lessons From Endogenous Growth Theory*, Working Paper 25755 (National Bureau of Economic Research, April 2019), www.nber.org/papers/w25755.

7. See Richard Gilbert and David M. G. Newbery, “Preemptive Patenting and the Persistence of Monopoly,” *American Economic Review*, vol. 72, no. 3 (June 1982), pp. 514–526, <https://tinyurl.com/yjrx4385>.

businesses and business formation, but the available evidence is not clear about whether changes in the regulatory environment contributed to the decline in entrepreneurship.

Financing Constraints and Economic Conditions

New firms often face financing constraints and, as a result, are especially susceptible to economic downturns. Those factors appear to have been particularly consequential during and after the 2007–2009 recession. Bank lending to start-ups (and other small businesses) fell sharply because of concerns about borrowers' creditworthiness amid the difficult economic conditions. Between 2006 and 2010, the rate at which new businesses were created fell by roughly 30 percent, from nearly 11 percent to just over 7 percent. During that time, the rate at which businesses exited the market spiked at close to 10 percent, and new firms displayed a very high rate of exit (their financial condition tends to be much less resilient than that of their older counterparts).²¹ The difficulty in obtaining financing persisted for several years after the recession: The value of commercial and industrial loans outstanding for small businesses did not return to its 2008 level until 2017.²²

During the 2007–2009 recession, the severe drop in the housing market contributed to entrepreneurs' restricted access to financing because they often use the equity value of a house as collateral for business loans. Housing values also constitute a large portion of people's wealth, changes in which affect consumer spending. The more housing prices dropped in a given locality in the wake of the recession, the more the rate at which businesses were created fell in that area. Conversely, the run-up in housing prices

in various localities before 2007 was associated with increasing rates of new businesses in those areas.²³

Financing constraints may especially affect entrepreneurial ventures that seek to achieve strong growth because such companies often need access to more than one round of financing as their business develops. In particular, companies that receive equity investments from venture capital firms represent a very select group of new businesses that are more likely to grow to become publicly traded and economically consequential. Those businesses typically receive infusions of cash at regular intervals to support their development and expansion so that investors may review the company's progress before committing additional funds. In that way, venture capital firms can manage a portfolio of investments and fully develop the most promising ones, but that process can also leave new ventures vulnerable to downturns in access to equity. An analysis of firms created between 1988 and 2014 found that their growth is correlated with both the general availability of capital and overall economic conditions.²⁴

Although the impact of the 2007–2009 recession was felt for several years after the economy stabilized, the financial environment for entrepreneurship had improved before the onset of the coronavirus pandemic. Housing prices rebounded significantly (but remained below their peak in 2005, after adjusting for inflation).²⁵ Venture capital financing increased substantially after 2008, exceeding \$100 billion in 2018 and 2019, its highest levels since 2000.²⁶

21. For a discussion of factors affecting firms' financial health, see Federal Reserve Bank of New York, *Can Small Firms Weather the Economic Effects of COVID-19* (April 2020), p. 2, <https://tinyurl.com/y5s2q9cm> (PDF, 296 KB).

22. See Federal Deposit Insurance Commission, *Loans to Small Businesses and Farms, FDIC-Insured Institutions, 1995–2017*, <https://tinyurl.com/yy6ftv6w> (XLS, 115 KB). The number of loan originations to small businesses (firms with revenues of \$1 million or less) fell by almost three-quarters from 2007 to 2010: See Rebel A. Cole, *How Did Bank Lending to Small Business in the United States Fare After the Financial Crisis?* (Small Business Administration, Office of Advocacy, January 2018), pp. 34–35, <https://go.usa.gov/xGX9R>.

23. See Steven J. Davis and John C. Haltiwanger, *Dynamism Diminished: The Role of Housing Markets and Credit Conditions*, Working Paper 25466 (National Bureau of Economic Research, January 2019), www.nber.org/papers/w25466.

24. See Jorge Guzman and Scott Stern, "The State of American Entrepreneurship: New Estimates of the Quantity and Quality of Entrepreneurship for 32 U.S. States, 1988–2014," *American Economic Journal: Economic Policy*, vol. 12, no. 4 (November 2020) pp. 212–243, <https://tinyurl.com/y89aym6w>.

25. See Robert J. Shiller, *Irrational Exuberance*, 3rd ed. (Princeton University Press, 2015), and see U.S. home price and related data for Figure 3.1 as updated by the author (accessed December 10, 2020), <https://tinyurl.com/yyta4pul> (XLS, 281 KB).

26. See Arnobio Morelix, *3 Facts You Probably Didn't Know About Venture Capital and Entrepreneurship* (Kauffman Foundation, May 13, 2016), <https://tinyurl.com/y3etnlpj>; and PwC/CB Insights, *MoneyTreeReport Q3 2020*, <https://tinyurl.com/y5bk4my6>.

In addition to exacerbating the challenges that financial constraints pose to new firms, economic downturns can inhibit their formation and growth, causing current revenues and expectations for future earnings to fall. In the wake of the coronavirus pandemic, a leading indicator of new business activity declined by more than 25 percent before rebounding strongly. The full extent of the impact will depend on the course of the pandemic and the measures implemented to contain it, particularly whether additional waves of outbreaks will require the reimposition of mitigation measures and the length of time it takes to distribute treatments and an effective vaccine. As the economic downturn continues to cause high unemployment, many people may turn to entrepreneurial activities to earn income. As long as the pandemic does not produce a full-blown crisis in the financial sector, such as the one that precipitated the 2007–2009 recession, entrepreneurship could maintain its prepandemic levels—if not exceed them—as the economy recovers.

In the long term, changes in the prices of labor and materials or other factors that affect their supply can also affect the incentives to create new firms and innovative products. For example, an increase in workers' wages can prompt the invention of equipment to replace labor at a lower cost. Some observers have suggested that shortages of certain kinds of labor in some advanced economies may promote the development and use of new technologies, such as artificial intelligence.²⁷ Higher energy costs can likewise induce the development of more fuel-efficient products.

Demographics and Immigration

Demographic shifts can affect rates of entrepreneurship. For example, people in certain age groups are more likely to be entrepreneurs. Although a popular notion is that the typical entrepreneur is very young, middle-aged people are more likely to start a business and be successful with it. More specifically, an increase in the share of the population between the ages of 35 and 44 is associated with higher rates of starting businesses. People of that age, along with those between 45 and 54 years old, are most likely to start a firm with at least one employee and are also most likely to start one that grows quickly. Founders of rapidly growing start-ups (with employees) between 2007 and 2014 were 45 years old, on average, and founders of high-tech start-ups were of a similar

age.²⁸ Throughout the 1980s and 1990s, the share of the labor force in the 35-to-54 age group rose before declining to roughly the level it was in 1990 (see Figure 6, upper panel).²⁹

A lack of employees can also stymie entrepreneurship because the overall growth rate of the labor force affects the formation of new firms. A decrease in that growth rate, from 2.7 percent during the 1970s to 1.0 percent in the 2000s, has been linked to a decline in new businesses (see Figure 6, middle panel).³⁰ That rate fell further in the aftermath of the 2007–2009 recession to an average of 0.6 percent growth in the 2010s. Several factors contributed to the changes in labor supply, including slower population growth; slower growth in women joining the labor force beginning in the 1990s; and a decline in the overall labor force participation rate after 2000 (see Figure 6, lower panel).³¹ The imminent shortage of qualified job candidates with STEM (Science, Technology, Engineering, and Math) skills is a particular concern.³²

Immigration of people with high skill levels contributes to the pool of qualified workers for new firms, whose success often depends on access to highly skilled employees. In particular, the foreign-born population accounts

27. See Hal Varian, "Automation Versus Procreation (Aka Bots Versus Tots)" (VoxEU, March 30, 2020) <https://tinyurl.com/y45yydc4>.

28. See Ian Hathaway and Robert E. Litan, *What's Driving the Decline in the Firm Formation Rate? A Partial Explanation* (Economic Studies at Brookings, November, 2014), <https://tinyurl.com/y2xjqxvq>; and Pierre Azoulay and others, "Age and High-Growth Entrepreneurship," *American Economic Review: Insights*, vol. 2, no. 1 (March 2020), pp. 65–82, <https://tinyurl.com/y4t5lu3s>.

29. See Mitra Toossi, "A Century of Change: The U.S. Labor Force, 1950–2050," *Monthly Labor Review* (May 2002), pp. 15–28, <https://go.usa.gov/x7BF2>.

30. See Fatih Karahan, Benjamin Pugsley, and Ayşegül Şahin, *Demographic Origins of the Startup Deficit*, Working Paper 25874 (National Bureau of Economic Research, May 2019), www.nber.org/papers/w25874; and Hugo Hopenhayn, Julian Neira, and Rish Singhania, *From Population Growth to Firm Demographics: Implications for Concentration, Entrepreneurship and the Labor Share*, Working Paper 25382 (National Bureau of Economic Research, December 2018), www.nber.org/papers/w25382.

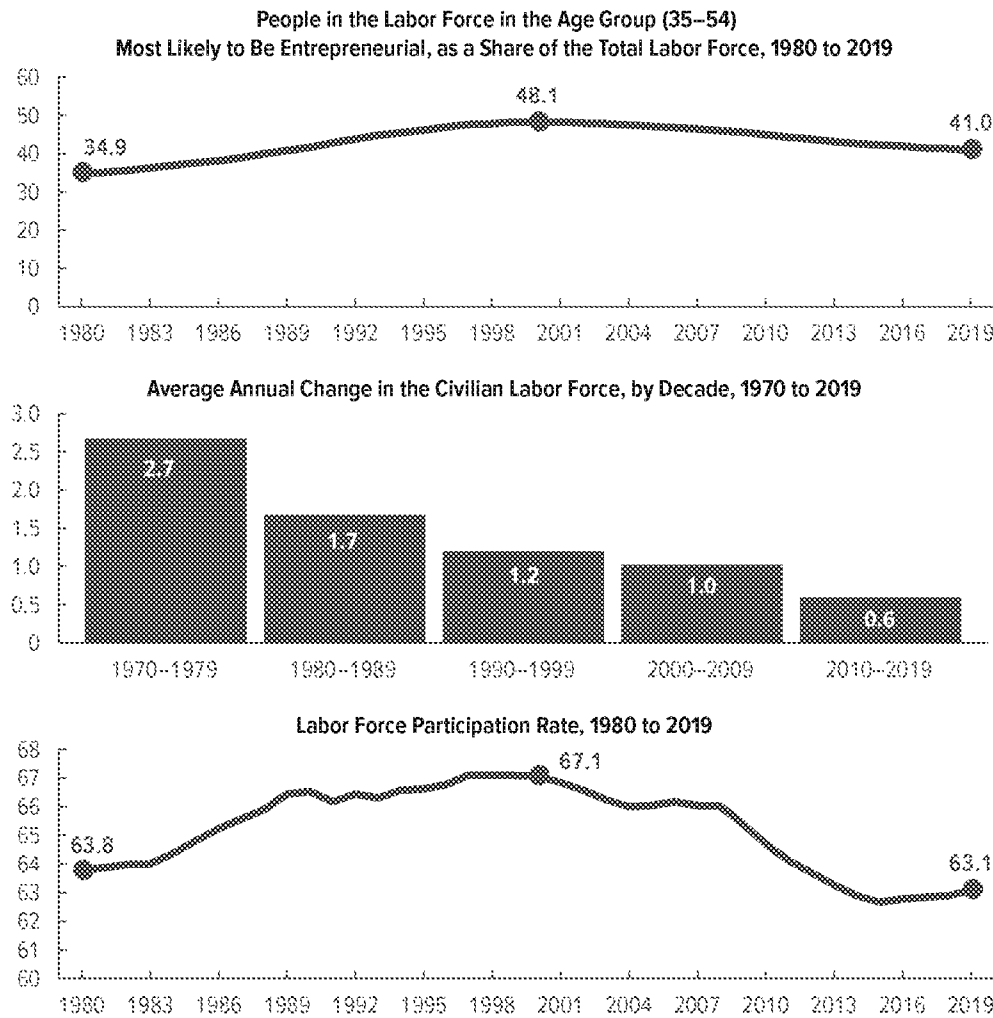
31. The labor force participation rate is the percentage of people in the civilian noninstitutionalized population who have jobs or are actively seeking work.

32. See National Academies of Sciences, Engineering, and Medicine, *Building America's Skilled Technical Workforce* (The National Academies Press, 2017), Chapter 2: Labor Market Patterns and Trends, pp. 21–38, <https://doi.org/10.17226/23472>.

Figure 6.

Demographic Trends Affecting Entrepreneurship

Percent



Demographic trends have also contributed to the decline in entrepreneurship over the past 20 years, including a falling share of the workforce among the age group most likely to be entrepreneurial, slower growth of the working-age population, and a lower rate of labor force participation.

Data source: Congressional Budget Office, using data from the Census Bureau and the Bureau of Labor Statistics. See www.cbo.gov/publication/56906#data.

The labor force participation rate is the percentage of people in the civilian noninstitutionalized population who have jobs or are actively seeking work.

for much higher shares of workers in science and engineering fields and those with doctorates than their share of the overall population. (And many foreign-born workers in the high-tech sector eventually start their own firms.)³³ New firms that fared better in the lottery for H-1B visas for highly skilled, foreign-born workers have

proven more successful in acquiring venture funding and in being acquired by a larger firm than new firms that were less lucky in their pursuit of such workers through that lottery.³⁴ Immigration is also linked to entrepre-

33. Workers with H-1B visas are not allowed to start a business unless they obtain a green card. For a more detailed discussion of the importance of foreign-born residents in the United States to innovation and productivity and of proposals to increase the immigration of skilled workers to the United States, see

Congressional Budget Office, *Federal Policies and Innovation* (November 2014), www.cbo.gov/publication/49487.
 34. See Stephen G. Dimmock, Jiekun Huang, and Scott J. Weisbenner, *Give Me Your Tired, Your Poor, Your High-Skilled Labor: H-1B Lottery Outcomes and Entrepreneurial Success*, Working Paper 26392 (National Bureau of Economic Research, October 2019), www.nber.org/papers/w26392.

neurship because immigrants have been more likely than native-born Americans to create new businesses, having founded an estimated one-quarter of all start-ups in recent years.³⁵ Foreign-born business owners have been especially successful in the high-tech sector.³⁶

Changes in immigration do not appear to have negatively affected entrepreneurship over the past half-century. The presence of foreign-born residents in the United States has increased fairly steadily, from roughly 5 percent of the total population in 1970 to 14 percent in 2018.³⁷ CBO projects a substantial drop in net immigration to the United States over the next several years because of the coronavirus pandemic, which could hamper both overall economic growth and the pace of business formation.³⁸

The Regulatory Environment

The regulatory environment also affects the creation and growth of new businesses. Federal regulations affect firms directly, and antitrust policies can influence the competition they face. Intellectual property protections, especially in the form of patents, can also have an impact on the viability of new firms. The regulatory environment includes state and local regulations pertaining to non-compete clauses, the growing use of which has received greater scrutiny in recent years.

Federal Regulation. An increasingly burdensome regulatory regime is often cited as a factor in the decline in new businesses over the past four decades. Because they are typically small, newer firms may face a competitive disadvantage from regulation when compliance requirements are the same for businesses of all sizes. Larger firms—which tend to be older—can also better absorb regulatory costs, and they are more likely than new firms to have the wherewithal to lobby for favorable regulatory treatment, which may affect the rate at which the latter are created.³⁹ In addition, new companies in particular may need to spend considerable time learning how to comply with existing regulations—a barrier to entry in the form of a onetime expense that incumbents have already incurred.⁴⁰

The federal government does try to limit the potential competitive disadvantage imposed on small companies by regulations. The Regulatory Flexibility Act of 1980 (Public Law 96-354), along with subsequent legislation and executive orders modifying it, requires federal agencies to assess the effects of regulations on small businesses. If a regulation is likely to affect a substantial number of small firms, an agency must evaluate the burden of the regulation and identify less-costly alternatives. The Office of Management and Budget is also required annually to analyze the effect of federal regulations on small businesses. In addition, small firms are exempt from certain environmental requirements and provisions for employee health insurance.⁴¹

Empirical assessments do not always agree about the impact of federal regulation on new and small businesses in the United States.⁴² Some individual regulations

35. See Sari Pekkala Kerr and William Kerr, “Immigrant Entrepreneurship in America: Evidence From the Survey of Business Owners 2007 and 2012,” *Research Policy*, vol. 49, no. 3 (April 2020), article 103918, <https://tinyurl.com/y2dqnan7>.

36. See Robert W. Fairlie, *Estimating the Contribution of Immigrant Business Owners to the U.S. Economy* (Small Business Administration, Office of Advocacy, November 2008), <https://tinyurl.com/y6kra9yv> (PDF, 247 KB); and William R. Kerr, “High-Skilled Immigration, Innovation, and Entrepreneurship: Empirical Approaches and Evidence,” in Carsten Fink and Ernest Miguelez, eds., *The International Mobility of Talent and Innovation: New Evidence and Policy Implications* (Cambridge University Press, 2017), pp. 193–221, <https://doi.org/10.1017/9781316795774.007>.

37. See Congressional Budget Office, *A Description of the Immigrant Population—2013 Update* (May 2013), www.cbo.gov/publication/44134, and *The Foreign-Born Population and Its Effects on the U.S. Economy and the Federal Budget—An Overview* (January 2020), www.cbo.gov/publication/55967.

38. See Congressional Budget Office, *The 2020 Long-Term Budget Outlook* (September 2020), p. 43, www.cbo.gov/publication/56516.

39. See Germán Gutiérrez and Thomas Philippon, *The Failure of Free Entry*, Working Paper No. 26001 (National Bureau of Economic Research, June 2019), www.nber.org/papers/w26001.

40. The federal tax code is criticized along much the same lines; see Steven J. Davis, “Regulatory Complexity and Policy Uncertainty: Headwinds of Our Own Making” (paper prepared for the 2017 Hoover Institution Conference on ‘Restoring Prosperity,’ February 9–10, 2017), pp. 14–15, <https://tinyurl.com/y4f7zec2>.

41. See Lloyd Dixon and others, “The Impact of Regulation and Litigation on Small Business and Entrepreneurship,” WR-317-ICJ (RAND Corporation, February 2006) www.rand.org/pubs/working_papers/WR317.html.

42. See Office of Management and Budget, *2017 Draft Report to Congress on the Benefits and Costs of Federal Regulations and Agency Compliance With the Unfunded Mandates Reform Act* (March 5, 2018), pp. 37–40, <https://go.usa.gov/xGX9t>.

discourage the entry of new firms, and some research finds that, overall, increasing regulation is harmful to the formation and growth of new businesses. One study found such a relationship for firms with between 10 and 500 employees, although it did not report results for the smallest firms (of fewer than 10 employees), which account for about 90 percent of newly created ones.⁴³ Other studies that found a negative relationship between regulation and entrepreneurship are limited in their applicability because they compare countries at very different stages of development than the United States.⁴⁴ Two recent studies found different results—one showing a negative relationship and the other finding no relationship—even though they examined the same data over the same time periods using similar methods.⁴⁵

Antitrust Policies. Federal antitrust laws, as implemented by the Department of Justice and the Federal Trade Commission, can affect competition and market power in an industry. Large firms with market power could exercise that power in a way that results in diminished competition and less favorable conditions for entrepreneurship. Frequently cited indicators of market power include widespread increases in measures of concentration, such as the share of sales accounted for by the largest firms in an industry. Another indication is an increase in the dispersion in profitability across firms in the same market.⁴⁶ Although both the market shares claimed by the largest firms in an industry, and

their profitability, appear to have grown over the past two decades, evidence of the impact of market power on competition and entrepreneurship is still the subject of debate.

A rise in concentration and profitability may reflect superior performance by a relatively small group of firms, such as those that can use new technologies to achieve greater economies of scale or realize network effects (conditions under which demand for a good or service increases in relation to the number of people using it).⁴⁷ Greater market power may also be associated with behavior that makes it harder for entrepreneurs to start new firms and for their businesses to grow. For example, large established firms may acquire smaller, emergent competitors to quash the threat they pose—either by selling the competing product themselves or by terminating its development or sale.⁴⁸ Concern about such anticompetitive behavior is especially pronounced in the high-tech sector. Internet “platform” firms, for example, may use information about their customers’ online activity to identify potential competitors and lucrative product markets. They may then either acquire their future rivals or enter those markets directly with their own products. In either case, entrepreneurship may be less viable than before (though it is also possible that the acquisition improves the product of the firm being acquired).⁴⁹

43. See Germán Gutiérrez and Thomas Philippon, *The Failure of Free Entry*, Working Paper 26001 (National Bureau of Economic Research, June 2019), www.nber.org/papers/w26001.

44. See Simeon Djankov and others, “The Regulation of Entry,” *Quarterly Journal of Economics*, vol. 117, no. 1 (February 2002), pp. 1–35, <https://academic.oup.com/qje/article/117/1/1/1851750>; and Leora Klepper, Luc Laevan, and Raghuram Rajan, “Entry Regulation as a Barrier to Entrepreneurship,” *Journal of Financial Economics*, vol. 82, no. 3 (December 2006), pp. 591–629, <https://tinyurl.com/y68sabcd>.

45. See Nathan Goldschlag and Alex Tabarrok, “Is Regulation to Blame for the Decline in American Entrepreneurship?” *Economic Policy*, vol. 33, no. 93 (January 2018), pp. 5–44, <https://tinyurl.com/y9ls7sy5>; and James Baily and Diana Thomas, “Regulating Away Competition: The Effect of Regulation on Entrepreneurship and Employment,” *Journal of Regulatory Economics*, vol. 52, no. 3 (December 2017), pp. 237–254, <https://tinyurl.com/yycgfax>.

46. See, for example, Jan De Loecker, Jan Eeckhout, and Gabriel Unger, “The Rise of Market Power and the Macroeconomic Implications,” *Quarterly Journal of Economics*, vol. 135,

no. 2 (May 2020), pp. 561–644, <https://doi.org/10.1093/qje/qjz041>.

47. See David Autor and others, “The Fall of the Labor Share and the Rise of Superstar Firms,” *Quarterly Journal of Economics*, vol. 135, no. 2 (May 2020), pp. 645–709, <https://doi.org/10.1093/qje/qjaa004>.

48. For evidence of the commercialization of new products being preempted through a competitor’s acquisition of the developing firm, see the research on the pharmaceutical industry by Colleen Cunningham, Florian Ederer, and Song Ma, “Killer Acquisitions,” *Journal of Political Economy* (forthcoming), <https://tinyurl.com/yb2m2rpv>.

49. See “American Tech Giants Are Making Life Tough for Startups,” *The Economist* (June 2, 2018), <https://tinyurl.com/y93nwkrrh>; and James Pethokoukis, “Incumbents vs. Startups: The Case That Big Tech Is Squashing Small Tech” (blog entry, American Enterprise Institute, June 4, 2018), <https://tinyurl.com/y6fz3c5d>. For regulators’ consideration of competition issues in digital technology markets, see Federal Trade Commission, “FTC Hearing 3: Oct. 17 Session 3 Nascent Competition: Economic Incentives and Business Strategies of Tech Firms” (accessed December 16, 2020), <https://tinyurl.com/y7ye4rbo>.

The Patent System. Patents provide a legal framework for protecting investments in intellectual property. They give an incentive for people to engage in innovative activity by granting them the exclusive use of the patented product (or process) for a period of time in exchange for disclosing the discovery. Patents can help new firms in particular to overcome financing constraints: A patent can make a new firm more attractive to investors because it provides greater assurance of the firm's success. In some industries, such as the pharmaceutical industry, where new products can be replicated at low marginal costs once the blueprint for them is known, the patent provides incentive for innovation.

Patents held by others can act as barriers to entrepreneurs. Patents have proliferated in high-tech industries since the early 1980s.⁵⁰ Large patent portfolios held by incumbent firms in those industries may deter start-ups because cash-constrained new businesses may have difficulty bearing the costs of patent-related litigation.⁵¹

Noncompete Clauses. Trends in state and local regulations may compound or, alternatively, offset developments at the federal level. For instance, state and local regulatory policies that allow restrictions on labor mobility, such as noncompete clauses, can affect entrepreneurship. Unfortunately, few data are available with which to examine whether the incidence of noncompete clauses has increased over time.

Under a noncompete clause in an employment contract, a worker must wait a certain amount of time after leaving an employer before joining another firm in a related industry (or geographic area)—or before starting a business that could compete with the former employer. Firms pursue such clauses because they constrain employees' ability to leave, thus affording a greater opportunity to recoup training costs and to invest in other activities that allow workers to learn (such as R&D). Noncompete clauses can also restrict the creation

of new businesses and their growth in markets that rely on skilled employees.⁵²

Health Insurance. Federal and state regulations pertaining to health insurance markets may also affect entrepreneurship. When the regulatory environment makes health insurance available at a lower cost or better quality through employers than is available to individuals, employees with health insurance coverage provided through their employer are likely to view starting or joining a firm as less attractive.

Evidence suggests that insurance coverage can affect self-employment, at least in some circumstances in which alternative sources of coverage are not as readily available and families expect to need it.⁵³ Health insurance considerations may or may not influence the self-employment decisions of individuals more broadly among all ages and in all circumstances, or the decisions of individuals who are specifically considering founding (or leaving a current employer to join) a new growth-oriented firm.⁵⁴

Certain changes to health care policy over the past four decades have supported entrepreneurship, in CBO's assessment. The Consolidated Omnibus Budget Reconciliation Act of 1985 (COBRA) provided for the continuation of group coverage upon separation from employment (although at greater expense to the individual than when the coverage was subsidized by the employer). From 1986 to 2003, the federal income

50. See Congressional Budget Office, *Federal Policies and Innovation* (November 2014), www.cbo.gov/publication/49487.

51. See Ian Appel, Joan Farre-Mensa, and Elena Simintzi, "Patent Trolls and Startup Employment," *Journal of Financial Economics*, vol. 133, no. 3 (September 2019), pp. 708–725, <https://doi.org/10.1016/j.jfineco.2019.01.003>; and Bronwyn H. Hall, Georg von Graevenitz, and Christian Helmers, "Technology Entry in the Presence of Patent Thickets," *Oxford Economic Papers* (September 2020), <https://doi.org/10.1093/oeq/gpaa034>.

52. See Department of the Treasury, *Non-compete Contracts: Economic Effects and Policy Implications*, (March 2016), <https://go.usa.gov/xAjFY> (PDF, 500 KB); and Evan Starr, *The Use, Abuse, and Enforceability of Non-Compete and No-Poach Agreements: A Brief Review of the Theory, Evidence, and Recent Reform Efforts* (Economic Innovation Group, February 2019) <https://eig.org/noncompetesbrief>.

53. For example, see Robert W. Fairlie, Kanika Kapur, and Susan Gates, "Is Employer-Based Health Insurance a Barrier to Entrepreneurship?" *Journal of Health Economics*, vol. 30, no. 1 (January 2011), pp. 146–162, <https://doi.org/10.1016/j.jhealeco.2010.09.003>. However, see also Douglas Holtz-Eakin, John R. Penrod, and Harvey S. Rosen, "Health Insurance and the Supply of Entrepreneurs," *Journal of Public Economics*, vol. 62, no. 1-2 (October 1996), pp. 209–235, [https://doi.org/10.1016/0047-2727\(96\)01579-4](https://doi.org/10.1016/0047-2727(96)01579-4).

54. For example, see Bradley T. Heim and Ithai Z. Lurie, "Did Reform of the Non-Group Health Insurance Market Affect the Decision to Be Self-Employed? Evidence From State Reforms in the 1990s," *Health Economics*, vol. 23, no. 7 (July 2014), pp. 841–860, <https://doi.org/10.1002/hec.2960>.

tax deduction provided to the self-employed for health insurance rose from 25 percent to 100 percent, making coverage less expensive for the entrepreneur.⁵⁵

The passage of the ACA in 2010 made health insurance coverage cheaper and more accessible for some entrepreneurs but more expensive for others.⁵⁶ Some potential entrepreneurs may have been more likely to start a business because they could no longer be denied coverage on the basis of preexisting health conditions (either their own or their family members') or be charged higher premiums because of their health. The ACA also subsidized health insurance for some entrepreneurs, making it less expensive, and caused premiums to rise for others—in some cases by substantial amounts.⁵⁷ Recently, the Trump Administration promulgated rules expanding association health plans allowing entrepreneurs to join with other small employers to purchase insurance at a lower cost in the large group market. Statutory changes in 2016 and recent changes to rules governing health reimbursement arrangements also now allow entrepreneurs and their employees to purchase health insurance in the nongroup market on a tax-preferred basis.⁵⁸ (Sometimes called health reimbursement accounts, health reimbursement arrangements are employer-funded group health plans from which employees receive tax-free reimbursements for qualified medical expenses up to a fixed dollar amount per year.) The effects of those rules on entrepreneurship are not yet known.

Federal Policies to Support Entrepreneurship

Federal policies can address a number of the factors that influence entrepreneurship. Policymakers could increase

access to financing for new firms or provide more financial support for those small firms that are likely to be innovative. Policymakers could also support entrepreneurship by facilitating the immigration of highly skilled workers and entrepreneurs to the United States. Finally, policymakers could modify regulations that affect the conditions under which firms are started and grow.

Different approaches would have their own advantages and disadvantages (see Table 1). For policies that would increase federal outlays, lawmakers might want to consider whether the costs of such policies would exceed the benefits from any improvement in the economy's performance. Those assessments are beyond the scope of this report.

Financing and Financial Support for New Firms

Policymakers could increase access to financing for new firms by creating programs that explicitly target them or through existing programs that target small firms. They could also provide other forms of financial support to new or small firms, either directly or indirectly. In either case, a key consideration is whether to try to channel assistance to those new firms that are likely to be innovative and grow or to support small firms more broadly.

Increase the Availability of Financing. Financing constraints can limit entrepreneurship. The federal government could create a program to provide financing specifically for innovative new firms. Policymakers could also provide greater access to financing for small firms more generally by expanding programs run by the SBA.

Provide a Credit Program Specifically for New Firms. New firms can face different—and sometimes more challenging—obstacles than existing businesses face. For example, it is often more difficult for promising new firms without a well-established track record of performance to obtain bank financing than older small firms of the same creditworthiness.⁵⁹ That is because new firms are more likely to fail than older firms. In fact, only about half of all new businesses survive past their first five years (see Figure 7). After that initial period, the chance of survival continues to decline, but more

55. However, health insurance premiums paid by the self-employed are not sheltered from payroll taxes. By comparison, most premiums paid by employers and employees for group health insurance are excluded from both federal income and payroll taxes.

56. For more discussion of the impact of the Affordable Care Act, see Congressional Budget Office *Private Health Insurance Premiums and Federal Policy* (February 2016), www.cbo.gov/publication/51130.

57. See Bradley T. Heim and others, "The Impact of the ACA on Premiums: Evidence From the Self-Employed," *Journal of Health Politics, Policy, and Law*, vol. 40, no. 5 (October 2015), pp. 1061–1085, <http://doi.org/10.1215/03616878-3161248>.

58. See Congressional Budget Office, *How CBO and JCT Analyzed Coverage Effects of New Rules for Association Health Plans and Short-Term Plans* (January 2019), www.cbo.gov/publication/54915.

59. Federal Reserve Bank of New York, *2016 Small Business Credit Survey: Report on Startup Firms* (August 2016), p. 19, <https://tinyurl.com/y4k394hv>.

Table 1.

Advantages and Disadvantages of Policies to Support Entrepreneurship

	Policy Approaches	Advantages	Disadvantages
Financial	1. Provide a credit program specifically for new firms	Would provide more targeted assistance to new firms	Would entail greater costs to the federal government and could run the risk of prominent failures
	2. Increase funding for SBA's 7(a) or SBIC credit assistance programs	Could increase new firms' access to financing	Most small firms are older firms
	3. Increase set-asides for SBIR and STTR programs	Some studies show these programs help firms start and grow	Could result in less-efficient allocation of federal funds
	4. Provide financial support for markets with innovative products	Could be particularly useful for products with social benefits	Could be difficult to identify promising technologies that would remain underdeveloped without government support
	5. Increase tax preferences such as those for capital gains, R&D investment, or depreciation allowances for capital spending	Would increase the after-tax return of starting a business	Would reduce tax revenues and accrue to many other firms in addition to new firms
Demographic	1. Allow for more immigration of skilled workers	Could increase the supply of qualified workers available to new firms—especially in the high-tech sector	Could depress the employment and wages of similarly skilled native-born workers
	2. Allow for more immigration of entrepreneurs	Could increase the number of people who start a firm in the United States	Could be difficult to identify foreign entrepreneurs and verify their activity in the United States
Regulatory	1. Reduce the regulatory burden for small or new firms	Could lower costs of regulatory compliance for new firms	Could limit the effectiveness of the regulations in question
	2. Expand the scope of antitrust enforcement	Could make it easier for start-ups in some markets to compete with larger, established firms	Could preclude some innovation- and productivity-enhancing mergers of new firms and incumbent firms
	3. Restrict the use of noncompete contracts	Could make it easier for workers to leave their employer and establish or join a new firm	Could weaken incentives for firms to invest in worker training or other types of intangible capital

Data source: Congressional Budget Office.

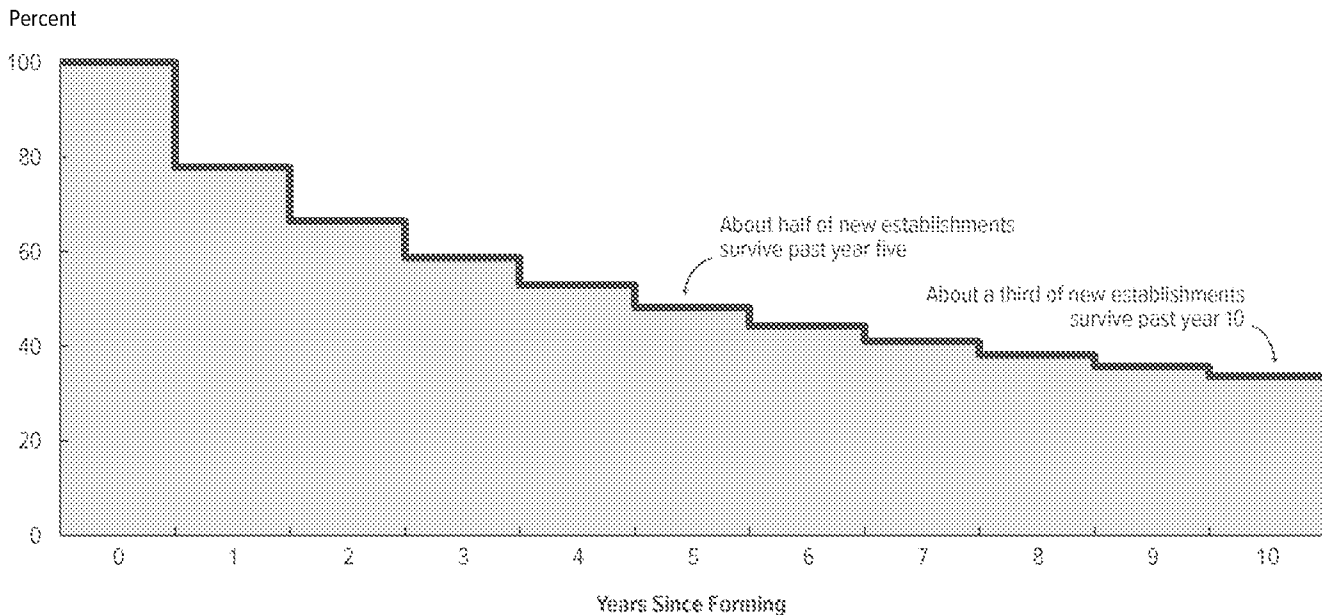
R&D = research and development; SBA = Small Business Administration; SBIC = Small Business Investment Corporation; SBIR = Small Business Innovation Research; STTR = Small Business Technology Transfer.

slowly.⁶⁰ (If it declined at the same rate as the first five years, only one-quarter of new businesses would survive after the second five years; but, on average, that percentage is instead about one-third.) Policymakers could help entrepreneurs overcome that impediment by establishing a loan guarantee program specifically targeting new businesses.

Choosing which new businesses to provide credit to can be particularly challenging, especially when choosing among those that are innovative and attempting to commercialize a new technology. The success of a new business often depends not only on an underlying technology but also on a broad array of other factors—including product design, the timing of product introduction, market placement, advertising, cost control, logistics, and product support. In the private sector, venture capital companies (which specialize in investing in new firms) anticipate that only one or two start-ups of every 10 they invest in will become high-growth firms. Three or four

60. The relatively low survival rate of new firms may reflect a process of experimentation in which new firms enter a market with an expectation about their profitability and learn over time whether they are in fact a viable business. See, for example, Boyan Jovanovic, "Selection and the Evolution of Industry," *Econometrica*, vol. 50, no. 3 (May 1982), pp. 649–670, <https://tinyurl.com/y5uncnpz>.

Figure 7.

Average Survival Rate of Establishments Formed Between 2000 and 2009

Data source: Congressional Budget Office, using data from the Bureau of Labor Statistics. See www.cbo.gov/publication/56906#data.

An establishment is a single place of business, and any one firm can own multiple establishments, although that is unlikely during the early years of a firm's life.

will likely fail.⁶¹ The federal government is generally in a worse position than private investors to determine which businesses will succeed. The government is at a disadvantage because it is likely to incorporate other goals into its credit decisions. Some firms that could have obtained financing on their own might thus receive federally subsidized credit, also preventing firms that warrant federal support from receiving a portion of the limited resources.

In the past, when federal programs have sought to promote technologies by providing credit to firms, the programs have frequently addressed some type of failure in the private market. A market failure occurs, for example, when those who buy something do not pay for the costs it imposes on society, like the costs of pollution or traffic congestion. Other federal programs have boosted technologies that served a governmental mission, such as defense.

61. For more discussion of the difficulty that even venture capital firms have in identifying which new companies will be successful, see William R. Kerr, Ramana Nanda, and Matthew Rhodes-Kropf, "Entrepreneurship as Experimentation," *Journal of Economic Perspectives*, vol. 28, no. 3 (Summer 2014), pp. 25–48, <https://tinyurl.com/y3bv8zr6>.

Regardless of its overall purpose, a program designed to minimize costs to the federal government would either be limited in its ability to take risks on promising new firms or would have to charge borrowers very high fees to cover those risks. Ultimately, the amount of assistance provided to new businesses through such a program would depend on how much risk—and thereby cost—policymakers were willing to accept.⁶² In fact, the SBA recently discontinued one such program. In 2012, the agency established a five-year initiative to promote early-stage Small Business Investment Corporations (SBICs; the SBIC program is described in more detail in the next section). Early-stage SBICs were required to invest at least 50 percent of their funds in early-stage small businesses, which are those that have never been cash-flow positive. By the end of fiscal year 2018, early-stage SBICs had invested only \$267.5 million in 82 small businesses because the SBICs had trouble attracting qualified investment funds. The SBA had stopped accepting new applications for the program in 2017, and it ceased

62. As discussed below, because most SBA lending programs are designed to minimize costs to the federal government, they are limited in their ability to take risks on promising new firms or would have to charge borrowers very high fees to cover that risk.

efforts to improve the program in 2018, citing excessive costs and a lack of support for the initiative.⁶³

Targeting firms by age instead of size raises other questions, such as what age to use to qualify a business as new. SBA programs typically have different thresholds, varying by industry, for the sizes of firms that qualify as small businesses.⁶⁴ Moreover, the potential for a firm to manipulate its age could be an issue. Depending on how age was measured, a very small business could reestablish itself as a new business by closing temporarily and then reopening. Databases used to research firms may not work well to administer a federal program.

Increase Funding for SBA Credit Programs. The federal government typically supports new firms on the basis of their size rather than their age—mainly through the SBA, which helps small businesses access financing, primarily through two programs. Under the Section 7(a) program, the SBA guarantees loans originated by banks and other financial institutions. In 2019, that program supported \$23.2 billion in loans. Under the SBIC program, investment companies borrow at reduced cost using an SBA guarantee and then use the proceeds to make debt and equity investments in small businesses.⁶⁵ Those investments amounted to \$5.9 billion in 2019.

To increase federally supported access to financing for new firms, policymakers could raise the loan limits on the SBA's lending programs or provide funding to subsidize access to credit. The fees that the SBA receives from its small-business and SBIC borrowers are intended to offset losses from loans that are not fully repaid. The SBA has subsidized loans in certain years, as it did after the 2007–2009 recession and is doing again in 2020 in response to the coronavirus pandemic. The rates at which

7(a) program loans have been written off over the past decade have typically been less than 2 percent of unpaid loan balances each year.⁶⁶ If the federal government explicitly shouldered some of the cost of the loan guarantees, those loans could be offered at lower rates and hence made more accessible to new businesses. And with additional funding, SBICs could guarantee more loans or support investments in potentially more innovative but riskier firms. Or the amount of SBIC investments could be increased by raising the amount that firms can borrow with an SBA guarantee.

Assessments of the SBA's ongoing efforts to support small firms have been hampered by limited data.⁶⁷ Firms that received 7(a) loans in the 1999–2001 time frame were found to have fared as well as, or better than, firms that did not receive such loans. Between 70 percent and 85 percent of the time, borrowers (regardless of their size) survived for at least four years after receiving their loans, whereas firms that did not receive loans had a four-year survival rate of roughly 70 percent.⁶⁸ Survival rates over the same time horizon for firms that received SBIC investments were lower—between 50 percent and 80 percent (depending on the method used for the estimate).

One problem with expanding extant small-business programs to increase support for new firms is that doing so will inadvertently benefit many older businesses that happen to be small. Although almost all new firms are small, so too are many older firms. In 2018, for example, virtually all firms less than five years old had no more than 100 employees, but only about 30 percent of businesses with fewer than 100 employees were less than five years old.⁶⁹

63. See Robert Jay Dilger, *SBA Small Business Investment Company Program*, Report for Congress R41456, version 87 (Congressional Research Service, August 31, 2020), <https://go.usa.gov/xACcn>.

64. To establish eligibility for its programs, the SBA relies on a variety of measures—such as employment, annual receipts, and assets—to determine a firm's size. The choice of measure may vary by industry as well as by the program or provision in question; see Robert Jay Dilger, *Small Business Size Standards: A Historical Analysis of Contemporary Issues*, Report for Congress R40860, version 91 (Congressional Research Service, August 28, 2020), <https://go.usa.gov/xACcH>.

65. See Robert Jay Dilger and Sean Lowry, *Small Business Administration: A Primer on Programs and Funding*, Report for Congress RL33243, version 116 (Congressional Research Service, October 6, 2020), <https://go.usa.gov/xACY3>.

66. See Small Business Administration, “Small Business Administration Loan Program Performance” (accessed April 26, 2020), <https://tinyurl.com/y6xqxrqi>.

67. See Robert Jay Dilger, *SBA Assistance to Small Business Startups: Client Experiences and Program Impact*, Report for Congress R43083, version 29 (Congressional Research Service, November 24, 2020), <https://go.usa.gov/xACcz>; and Government Accountability Office, *Priority Open Recommendations: Small Business Administration*, GAO-19-371SP (April 4, 2019), <https://go.usa.gov/xGsj6>.

68. Urban Institute, *A Performance Analysis of SBA'S Loan and Investment Programs* (January 2008), <https://tinyurl.com/yy5xmgsx> (PDF, 220 KB).

69. CBO analyzed data from the Census Bureau's Business Dynamics Statistics program. For the data, see www.census.gov/programs-surveys/bds.html.

Increase Other Forms of Federal Financial Support.

The federal government provides financial support for innovative small firms directly through research programs at federal agencies and indirectly through the tax code. Boosting that financial support could take the form of either increased funding for those programs or more favorable tax provisions for small businesses. Policymakers could target such support directly to new firms by substituting age requirements for size requirements. That approach would be more effective at providing support to new firms but would also entail more risk that recipients could fail despite the support.

Increase Set-Asides for Small-Business Research and Technology Programs. Policymakers could increase support for innovative new firms by modifying the Small Business Innovation Research (SBIR) program or the Small Business Technology Transfer (STTR) program. Under the SBIR program, every federal department with an R&D budget of \$100 million or more must allocate a fixed share (currently 3.2 percent) of that budget to pay for work done by small firms. The STTR program funds research proposals that are developed and executed cooperatively between small businesses and scientists in nonprofit research organizations. The program receives a set-aside (currently 0.45 percent) from the R&D budgets of federal departments that spend more than \$1 billion per year on joint R&D efforts between private firms and nonprofit scientists. Increasing the set-aside for the STTR program would increase the funding available to small high-tech businesses, and raising the share allocated for the SBIR program could increase the participation of such businesses in federal R&D efforts.

Some assessments of the SBIR program have found evidence that it helps firms start and grow. Start-up rates have risen in localities where at least one business has received an SBIR grant, and the program appears to have increased the share of R&D undertaken by smaller firms relative to larger ones.⁷⁰ The program has also been associated with several measures of future success

70. See Haifeng Qian and Kingsley E. Haynes, “Beyond Innovation: The Small Business Innovation Research Program as Entrepreneurship Policy,” *Journal of Technology Transfer*, vol. 39 (December 2014), pp. 524–543, <https://tinyurl.com/yclu62qn>; and Matthew R. Keller and Fred Block, “Explaining the Transformation in the U.S. Innovation System: The Impact of a Small Government Program,” *Socio-Economic Review*, vol. 11, no. 4 (September 2013), pp. 629–656, <https://tinyurl.com/y3qyvbhf>.

for businesses, such as a higher likelihood of increased revenue and venture capital funding.⁷¹

Other observers have been more skeptical, arguing that the SBIR program is susceptible to lobbying as well as waste, fraud, and other abuse.⁷² A potential disadvantage of modifying either the SBIR program or the STTR program is that doing so could result in a less-efficient allocation of federal funds. And administering a program that is selective can require substantial resources. For example, the 11 federal agencies participating in the SBIR program in fiscal year 2017 reviewed just over 19,000 initial-stage proposals, making funding awards to roughly one in six firms.⁷³

Provide Financial Support for Markets That Draw on Innovative Technologies. Another way to support leading-edge new companies is to subsidize demand for their products.⁷⁴ That approach could valorize products that embody technologies whose development would provide social benefits that the marketplace may not otherwise value, such as improving national security or reducing greenhouse gas emissions. One challenge for such an approach is the potential difficulty in determining which technologies will prove to be commercially successful and

71. See Sabrina T. Howell, “Financing Innovation: Evidence From R&D Grants,” *American Economic Review*, vol. 107, no. 4 (April 2017), pp. 1136–1164, <https://doi.org/10.1257/aer.20150808>.

72. See Marcy E. Gallo, *Small Business Research Programs: SBIR and STTR*, Report for Congress R43695, version 5 (Congressional Research Service, May 5, 2020), <https://go.usa.gov/xACYx>; and Josh Lerner, “Government Incentives for Entrepreneurship,” in Austan Goolsbee and Benjamin Jones, eds., *Innovation and Public Policy* (National Bureau of Economic Research, forthcoming), <https://tinyurl.com/ya4nyuf5>.

73. See Small Business Administration, *Small Business Innovation Research and Small Business Technology Transfer Annual Report: Fiscal Year 2017*, pp. 6–7, 12, <https://go.usa.gov/xGNzw> (PDF, 5.88 MB). In the initial stage of the SBIR program (referred to as Phase I), selected firms establish the technical merit, feasibility, and commercial potential of their proposed project. Awards at subsequent stages (Phases II and III), which are also competitively based, allow for continued R&D and commercialization, respectively.

74. For more discussion of this kind of approach, which is credited with fostering the development of integrated circuits in the 1960s, see Congressional Budget Office, *Federal Policies and Innovation* (November 2014), p. 16, www.cbo.gov/publication/49487.

which of those would remain underdeveloped without government support.

Modify the Federal Tax System. Small businesses currently benefit from several tax preferences. Those preferences could be made more generous and directed toward new firms instead of small ones. Major tax provisions that provide financial support to small firms include an expensing allowance for investment in qualifying equipment, which largely comprises machinery, equipment, and off-the-shelf software. Although businesses of all sizes can claim the allowance, it is capped at \$1 million and thus primarily benefits small firms (because the expensing limit represents a greater share of their investment than it does for larger firms). The \$1 million expensing allowance reflects a \$0.5 million increase brought about by the 2017 tax act (P.L. 115-97). However, because the 2017 tax act also provided 100 percent bonus depreciation (which also allows for expensing of equipment) for all firms through 2022, that increase will have little or no incremental value to small (and new) firms until 2023.⁷⁵

Other tax provisions that support small businesses include the option to use cash rather than accrual accounting; an exclusion for gains from the sale of qualified small business stock; and a tax credit for the cost of providing employees with health insurance. Those provisions accounted for nearly all of the roughly \$23 billion in federal tax expenditures for fiscal year 2019 associated with measures that favor small firms. The amount of forgone federal revenues from those major tax provisions varies widely. In fiscal year 2019, forgone revenues amounted to \$15.7 billion from the expensing allowance; \$6.0 billion from the option to use cash rather than accrual accounting; \$1.3 billion from the exclusion of gains from the sale of qualified small business stock; and less than \$50 million from the tax credit for the cost of providing employees with health insurance.⁷⁶

75. Additionally, the 2017 tax act replaced the previous corporate tax rates (which were increasing in corporate income) with a uniform rate of 21 percent. The lower, uniform rate could have hurt new (and small) firms relative to older (and large) ones; see Gary Guenther, *The 2017 Tax Law (P.L. 115-97) and Investment in Innovation*, Report for Congress IF10757, version 3 (Congressional Research Service, April 9, 2018), <https://go.usa.gov/xACYg>.

76. See Joint Committee on Taxation, *Estimates of Federal Tax Expenditures for Fiscal Years 2019-2023*, (December 18, 2019), <https://www.jct.gov/publications/2019/jcx-55-19/>.

The support provided by the tax system through those tax expenditures could be expanded to larger firms, which could also increase the financial support for growing entrepreneurial ventures. For instance, the \$1 million cap on depreciation allowances could be increased, which would subsidize a larger amount of investment by new firms.⁷⁷ Policymakers could also raise the gross asset ceiling used to treat a business's stock as a qualified small business stock (currently \$50 million) and expand the roster of eligible industries, making such tax-preferred investment available to more companies. The portion of capital gains that may be excluded from gross income could also be increased. Some observers have already argued that the criteria that businesses must meet to qualify for the provision are overly restrictive.⁷⁸

Finally, policymakers could alter tax laws that govern how businesses treat their expenses for conducting R&D. The 2017 tax act repealed the option for firms to expense research expenditures, requiring instead (starting in 2022) that they be capitalized and amortized over five years. The option to expense R&D may have been especially beneficial to small firms because it allowed them to avoid a relatively complex tax filing.⁷⁹ Another way to reduce the after-tax cost of R&D for new firms would be to allow them to receive a refund for the R&D tax credit. Because of start-up costs, young businesses are more likely than older ones to post losses and may not have sufficient tax liability to apply the credit fully (if at all).⁸⁰ A consideration weighing against such an approach is that experience with refundability in other contexts has shown that such allowances can be abused and difficult to administer.

77. The \$1 million cap represents an increase from the previous level of \$510,000 and was one result of the 2017 tax act (P.L. 115-97). For additional effects of the 2017 tax act on small firms, see Gary Guenther, *P.L. 115-97, the 2017 Tax Revision, and Small Business Taxation*, Report for Congress IF10723, version 7 (Congressional Research Service, February 9, 2018), <https://go.usa.gov/xACYD>.

78. See Alan D. Viard, "The Misdirected Debate and the Small Business Stock Exclusion," *Tax Notes*, vol. 134, no. 6 (February 2012), <https://tinyurl.com/y4vrw5zj>.

79. See Bronwyn H. Hall, "Tax Policy for Innovation," in Austan Goolsbee and Benjamin Jones, eds., *Innovation and Public Policy* (National Bureau of Economic Research, forthcoming), <https://tinyurl.com/yapzfwb9>.

80. See Gary Guenther, *The 2017 Tax Law (P.L. 115-97) and Investment in Innovation*, Report for Congress IF10757, version 3 (Congressional Research Service, April 9, 2018), <https://go.usa.gov/xACYZ>.

A result of modifying the federal tax system to benefit entrepreneurs would be a greater after-tax return when starting a business and the incentives that it would create to encourage start-ups. There are several potential disadvantages to that approach. Generally, greater federal financial support for new firms would either increase deficits or require reductions in spending for other government activities. More specifically, providing a tax preference for capital gains from the sale of new firms' stock could provide a windfall to investors who would have taken an equity stake in those businesses anyway. In addition, making the R&D tax credit more easily claimed by new firms could provide them with an incentive to mitigate revenue losses by falsely claiming the credit for non-R&D related activities. Moreover, because new firms may not be profitable right away, additional tax preferences might not be of immediate use to them.

Immigration of Highly Skilled Workers and Entrepreneurs

Another way to increase entrepreneurial activity is to increase immigration to the United States of highly skilled workers or of people who are particularly likely to start a business.

Facilitate Immigration of Highly Skilled Workers. The federal government could increase the supply of qualified workers available to new firms—especially in high-tech industries—by facilitating the immigration of highly skilled workers to the United States.⁸¹ One approach for expanding the pool of foreign-born, high-tech workers is to increase employment-based immigration under the H-1B visa program. An H-1B visa admits highly skilled, foreign-born workers into the United States to work for up to three years (with the possibility of renewal). The H-1B program has a cap of 65,000 visas.⁸² Applications submitted by businesses for new H-1B workers have routinely exceeded limits in recent years—during the

first week or even on the first day that they are accepted, in some cases.⁸³

One potential unintended consequence of increasing the cap on visas in the H-1B program is that it would admit foreign-born workers to work in more industries than just those that are high-tech. Other possible unintended consequences are the effect on the employment and wages of native-born workers and the potential for companies to abuse the program by hiring foreign-born workers to replace native-born workers at lower wages. Increasing the employment of young, highly skilled immigrants decreases the employment of older native-born workers, either absolutely or as a share of total employment. In the past, CBO has estimated that increasing the size of the H-1B program would decrease wages slightly over several years for workers in the top fifth of the skill distribution.⁸⁴

Another way to increase the number of highly skilled, foreign-born workers in the United States is to grant permanent residency (by issuing what are commonly called green cards) to more of those workers who qualify for it. Because of country-specific caps on green-card issuance, some foreign-born workers (from countries that supply a large share of work-based visa holders, who generally come with high skills) must wait a long time to become permanent residents. Such a delay could encourage those foreign-born workers to seek citizenship elsewhere and discourage nonresident, foreign-born workers from immigrating to the United States.⁸⁵

81. See Sari Pekkala Kerr and William R. Kerr, "Immigration Policy Levers for U.S. Innovation and Startups," in Austan Goolsbee and Benjamin Jones, eds., *Innovation and Public Policy* (National Bureau of Economic Research, forthcoming), <https://tinyurl.com/ybsn4mn9>.

82. Visa applicants for employment at universities and nonprofit research facilities are exempt from the cap, as are the first 20,000 applicants with a master's degree or doctorate from a university in the United States.

83. See Jill H. Wilson, *Temporary Professional, Managerial, and Skilled Foreign Workers: Policy and Trends*, Report for Congress R43735, version 7 (Congressional Research Service, August 9, 2016), <https://go.usa.gov/xACYK>. For a description of the H-1B visa program and its requirements, see Citizenship and Immigration Service, "H-1B Specialty Occupations, DOD Cooperative Research and Development Project Workers, and Fashion Models" (accessed December 10, 2019), <https://go.usa.gov/xGnW7>.

84. Economic studies have generally found that increases in the number of skilled workers raise the productivity of less-skilled workers. See Congressional Budget Office, *The Economic Impact of S. 744, the Border Security, Economic Opportunity, and Immigration Modernization Act* (June 2013), p. 20, www.cbo.gov/publication/44346.

85. See William A. Kandel, *Permanent Employment-Based Immigration and the Per-country Ceiling*, Report for Congress R45447, version 3 (Congressional Research Service, December 21, 2018), <https://go.usa.gov/xACY9>.

A third approach to facilitating immigration is to provide employment-based visas to foreign-born students in STEM fields, allowing them to stay in the United States after graduation. Such visas could be temporary or could confer permanent-resident status. Many students in STEM-related fields already remain in this country after graduation, but providing visas especially for such students would make it much easier for them to do so.⁸⁶ Linking education in the United States to immigration prospects would affect incentives for foreign-born people to study here and for U.S. colleges and universities to provide them opportunities to do so.

Facilitating the immigration of highly skilled, foreign-born workers would expand the pool of such employees for large incumbent firms as well as for new firms. The additional inflow of workers could, however, depress the employment and wages of similarly skilled, native-born workers—although such effects appear to be small.⁸⁷ Another concern is that some countries that have linked immigration to education have found that certain educational institutions increased their enrollment of foreign-born students substantially by lowering educational standards.⁸⁸

Facilitate Immigration of Entrepreneurs. Another approach that lawmakers could take to encourage foreign-born entrepreneurs to come to the United States would be to expand visa programs exclusively for immigrant business owners or investors. A program that reflects that approach (and simultaneously illustrates the difficulty of implementing such a narrowly targeted

policy) is the International Entrepreneur Rule (IER).⁸⁹ Under the IER, the Department of Homeland Security can extend for up to two and a half years (renewable once) the stay of foreign-born business owners who can establish that they have started a firm within the past five years that has the potential to grow rapidly, create jobs, and provide other benefits to the United States. According to federal regulations, evidence of such future business prospects is significant capital investment by U.S. citizens or the receipt of grants from a state, local, or federal government entity.⁹⁰ Although the IER was established in January 2017, it was suspended before it was due to take effect and is currently being considered for termination. Among the reasons cited for terminating the IER is the difficulty of implementing it.⁹¹ Identifying foreign-born individuals who can start successful businesses in the United States (and monitoring the performance of those firms to confirm that outcome) can entail a significant commitment of staff and time. The more demanding the eligibility criteria of a given program, the more effort is required to administer it.

Another program, the Immigrant Investor Program (also known as EB-5), targets immigrant investors who create a new commercial enterprise or invest in one of the federally designated regional centers that pools funds to make investments to promote economic growth. Under the program, entrepreneurs can apply for a visa for permanent residence. Each year, a maximum of 10,000 people (and their families) who invest at least \$1.8 million in a new business and create 10 jobs in the United States, or who invest \$900,000 in a new business in an underdeveloped or high-unemployment area, receive EB-5 visas. One source of concern is that, because the program defines “new” as any firm established after 1990, the investment that takes place through EB-5 visas may not always support new start-ups. In recent years, participation in the program has shifted away from directly creating new businesses to investing in the

86. Such an employment-based visa could be a replacement for, or an extension of, the Optional Practical Training (OPT) currently available to foreign-born students. The OPT is a temporary employment directly related to a student’s major area of study. Individuals can apply to receive up to 12 months of OPT employment authorization before or after completing their academic studies; foreign-born students with STEM degrees can generally apply for a 24-month extension of their post-completion OPT employment.

87. See Congressional Budget Office, *The Economic Impact of S. 744, the Border Security, Economic Opportunity, and Immigration Modernization Act* (June 2013), www.cbo.gov/publication/44346.

88. See Leslyanne Hawthorne, *Competing for Skills: Migration Policies and Trends in New Zealand and Australia* (New Zealand Department of Labour, 2011), pp. 108–114, <https://tinyurl.com/y676d93s> (PDF, 1.7 MB).

89. See Sari Pekkala Kerr and William R. Kerr, “Immigration Policy Levers for U.S. Innovation and Startups,” in Austan Goolsbee and Benjamin Jones, eds., *Innovation and Public Policy* (National Bureau of Economic Research, forthcoming), <https://tinyurl.com/ybsn4mn9>.

90. See International Entrepreneur Rule, 82 Fed. Reg. 5238 (January 17, 2017).

91. See Removal of International Entrepreneur Parole Program, 83 Fed. Reg. 24415 (2018).

regional centers. The program has also been subject to concerns about fraud.⁹²

An advantage of facilitating the immigration of foreign-born entrepreneurs to the United States is that it could increase the number of individuals who start a firm. A disadvantage is that the challenges of administering such programs may be substantial.

The Regulatory Environment's Effects on Competition

Policymakers could make the regulatory environment more favorable for new firms by reducing the burden of regulations that directly affect them, strengthening anti-trust policy, or limiting the use of noncompete clauses.

Reduce the Regulatory Burden on New Firms. The Congress could further lighten the burden of federal regulations on entrepreneurship by expanding the requirements for federal agencies to limit the impact of regulations on small firms. Alternatively, because not all small firms are new, a more focused approach could establish new eligibility criteria for existing regulatory exceptions so that only new firms, and not all small ones, would benefit from them. To help formulate such a policy, lawmakers could require that either the Office of Management and Budget or another federal agency explore how new firms are disproportionately burdened by federal regulations relative to incumbent firms generally and to older small firms in particular.

One advantage of reducing the regulatory burden on entrepreneurship is that it could lower the costs of regulatory compliance for new businesses, making it easier for them to start and grow. A disadvantage is that doing so could limit the effectiveness of the regulations in question.

Change Antitrust Enforcement. The Congress could require antitrust regulators to examine the effects of increasing concentration on the ability of start-ups to

compete and grow. For example, the Congress could revise the Clayton Act to require the Department of Justice and the Federal Trade Commission (the agencies responsible for applying antitrust laws) to investigate the likely effects of mergers and acquisitions on innovation, which could result in fewer mergers and acquisitions being allowed. But innovation, by its nature, is very difficult to predict, so the result of such a review would likely be inconclusive. A more immediate course of action would be to increase funding for those agencies' efforts to enhance their current monitoring of competition in U.S. technology markets, including competition among firms with online platforms.⁹³

Restrict the Use of Noncompete Contracts. Another way the federal government could address regulations that presumably constrain entrepreneurship is to move to limit the anticompetitive use of noncompete contracts. Although regulating noncompete contracts has traditionally been the responsibility of state governments, the Congress could preempt state labor laws to restrict the use of such contracts. That approach could increase the pool of potential employees for both new and incumbent firms.

One advantage of federal intervention in noncompete contracts is that although several states already limit the use of such contracts, employees are often unaware of that fact.⁹⁴ A federal measure could ensure that restrictions on noncompete contracts were uniform among states and were clearly communicated to both current employees and potential hires. While limiting or eliminating the use of noncompete contracts could make it easier for workers to leave their employers and establish or join new firms, it could also weaken incentives for employers to invest in worker training or other types of human capital.

92. U.S. Securities and Exchange Commission, "Investor Alert: Investment Scams Exploit Immigrant Investor Program," (accessed December 10, 2020), <https://go.usa.gov/xA37H>.

93. See Federal Trade Commission, "FTC's Bureau of Competition Launches Task Force to Monitor Technology Markets" (accessed May 2, 2019), <https://go.usa.gov/xGsZW>.

94. See Matt Marx and Lee Fleming, "Non-compete Agreements: Barriers to Entry ... and Exit?," in Josh Lerner and Scott Stern, eds., *Innovation Policy and the Economy*, Volume 12 (National Bureau of Economic Research, 2012), pp. 39–64, www.nber.org/chapters/c12452.



Appendix: Alternative Measures of Entrepreneurship

Several measures of entrepreneurship other than the one used in this report are used in different contexts. The measure of new firms with at least one employee, used here, is more likely than the others to capture full-fledged business operations with growth prospects. Although alternative measures include those of new entrepreneurs, nonemployer firms and self-employment, and sole proprietorships identified by tax filings, once they began hiring workers, those entities would be included in the measure of employer firms used here.¹

The Rate of New Entrepreneurs

One alternative measure of entrepreneurship is the rate at which entrepreneurs start their own businesses—with or without employees. An entrepreneur may not want to hire employees right away when commercializing a new product or service. With the attendant responsibilities of establishing a payroll, reporting tax information, and fulfilling other types of employer requirements, starting a new business is inherently risky, and the likelihood of success is uncertain.

The Kauffman Foundation’s “Rate of New Entrepreneurs,” which is based on data from the Current Population Survey, only includes individuals who report a weekly commitment of 15 hours or more to their new business.² By doing so, it attempts to distinguish between those people who have opted for self-employment as their primary activity and those pursuing a business interest “on the side.”

Over the 1996–2016 period, the Rate of New Entrepreneurs fluctuated but remained within a relatively narrow band around 0.3 percent. The divergence between the trend in that measure and the rate at which employer firms were created was most pronounced during the 2007–2009 recession. The Rate of New Entrepreneurs actually rose from 0.30 percent in 2007 to 0.34 percent in 2010, while the rate of new employer firms dropped sharply over that period. The upswing in entrepreneurship during that time primarily represented unemployed people who started their own businesses.

Nonemployer Firms and Self-Employment

Independent tradesmen and contractors who conduct essentially routine types of business may form firms without employees or otherwise be self-employed. As a result, measures of those activities may also overstate innovative and highly productive entrepreneurship.

The Census Bureau records statistics on nonemployer firms, and the Bureau of Labor Statistics compiles data on the incidence of self-employment. Neither source shows the declines in activity that the statistics on new employer firms show. The share of nonemployers in the total number of business establishments remained roughly constant, at around 75 percent, over the 2005–2014 period, and the incidence of self-employment (either incorporated or unincorporated) was fairly stable from 1990 through 2009.³

1. See Robert W. Fairlie, Javier Miranda, and Nikolas Zolas, “Measuring Job Creation, Growth, and Survival Among the Universe of Start-ups in the United States Using a Combined Start-up Panel Data Set,” *ILR Review*, vol. 72, no. 5 (October 2019), pp. 1262–1277, <https://doi.org/10.1177/0019793919862764>.

2. The Census Bureau supplies its own data on entrepreneurship using business tax filings to inform its *Annual Survey of Entrepreneurs*. See Census Bureau, “Annual Survey of Entrepreneurs (ASE)” (accessed December 10, 2020), www.census.gov/programs-surveys/ase.html.

3. See Census Bureau, “Having a Boss vs. Working for Yourself” (accessed December 10, 2020), <https://go.usa.gov/xGs9y>; and Steven F. Hipple, “Self-employment in the United States,” *Monthly Labor Review* (Bureau of Labor Statistics, September 2010), pp. 17–32, <https://go.usa.gov/xGs9f>. For data on nonemployer firms, see the Census Bureau, “Nonemployer Statistics (NES)” (accessed December 10, 2020), <https://go.usa.gov/xA3Ac>. For data on self-employment, see Bureau of Labor Statistics, “Labor Force Statistics From the Current Population Survey” (accessed December 10, 2020), www.bls.gov/cps/lfcharacteristics.htm#self.

Sole Proprietorships

Entrepreneurs can form sole proprietorships to test an idea before establishing a more formal business structure. Sole proprietorships may also serve purposes that have little relation to an active or innovative business pursuit, such as receiving payments from an avocation or a passive investment. Only about half of sole proprietorships engaged in activities that qualified for business tax deductions between 2007 and 2010.⁴

4. See Richard Prisinzano and others, *Methodology to Identify Small Businesses*, Technical Paper 4: Update (Department of the Treasury, Office of Tax Analysis, November 2016), Tables 1a and 1b, <https://go.usa.gov/xGn5N> (PDF, 1.14 MB); and see

Sole proprietorships are tallied through tax filings to the Internal Revenue Service. The share of business tax returns filed by sole proprietorships rose slightly between 1980 and 2015, from 69 percent to 72 percent, while net income reported on tax returns filed by sole proprietorships fell by about one-third, from 16 percent to 10 percent.⁵

the tables published by the Treasury at <https://go.usa.gov/xAgbt> (XLS, 338 KB). Changes in the tax regime may have affected how new firms are identified using those measures.

5. See Internal Revenue Service, “SOI Tax Stats 1—Integrated Business Data” (accessed July 9, 2020), <https://go.usa.gov/xGs97>.



List of Tables and Figures

Table

- | | |
|---|----|
| 1. Advantages and Disadvantages of Policies to Support Entrepreneurship | 21 |
|---|----|

Figures

- | | |
|--|----|
| 1. Rates at Which Firms Entered and Exited the Market, 1982 to 2018 | 4 |
| 2. New Firms as a Share of Total Firms and Total Employment, 1982 to 2018 | 5 |
| 3. New Firms' Share of Employment, by Sector, 1982 to 2018 | 8 |
| 4. Number of Weekly High-Propensity EIN Applications in 2020 | 9 |
| 5. Trends in New Firms' Share of Employment, by Sector, and Labor Productivity Growth in the Economy, 1995 to 2017 | 11 |
| 6. Demographic Trends Affecting Entrepreneurship | 16 |
| 7. Average Survival Rate of Establishments Formed Between 2000 and 2009 | 22 |



About This Document

This report was prepared in response to a request from the Chairman of the Senate Budget Committee and the Chairman of the Senate Committee on Small Business and Entrepreneurship. In keeping with CBO's mandate to provide objective, impartial analysis, it contains no recommendations.

Nathan Musick wrote the report, with guidance from Joseph Kile and Chad Shirley. Pranav Bhandarkar contributed to the analysis. Useful comments were provided by Rebecca Heller, Junghoon Lee, John McClelland, Allison Percy, Joseph Rosenberg, Robert Shackleton, Natalie Tawil, and Jeff Werling of CBO. Ryan Decker of the Federal Reserve Board of Governors, William Kerr of Harvard Business School, Huiyu Li of the Federal Reserve Bank of San Francisco, and Benjamin Pugsley of the University of Notre Dame also provided helpful comments. The assistance of external reviewers implies no responsibility for the final product, which rests solely with CBO.

Mark Doms, Wendy Edelberg (formerly of CBO), and Jeffrey Kling reviewed the report. Scott Craver was the editor, and Casey Labrack and Jorge Salazar were the graphics editors. The report is available on CBO's website (www.cbo.gov/publication/56906).

CBO continually seeks feedback to make its work as useful as possible. Please send any comments to communications@cbo.gov.

A handwritten signature in black ink, appearing to read "Phillip Swagel".

Phillip L. Swagel
Director
December 2020

SAVINGS

[Advertiser Disclosure](#)

Survey: Fewer than 4 in 10 Americans could pay a surprise \$1,000 bill from savings

SHARE:



Image by Getty Images; illustration by Bankrate

5 MINUTE READ
January 11, 2021

Written by **Jeff Ostrowski**



The Bankrate promise

[More info](#)

At Bankrate we strive to help you make smarter financial decisions. While we adhere to strict editorial integrity, this post may contain references to products from our partners. Here's an explanation [for how we make money.](#)

As the pandemic enters its most intensive phase yet and job growth stalls, just 39 percent of Americans surveyed say they could comfortably cover an unexpected expense of \$1,000.

Underscoring the shaky finances of many Americans, Bankrate's January Financial Security Index finds that fewer than 4 in 10 U.S. adults could absorb the cost of a four-figure car repair or emergency room visit by tapping into [savings](#).

Fully 18 percent of respondents said they would put the expense on a [credit card](#) and pay it off over time, incurring interest charges. Another 18 percent said they could handle a surprise expense without borrowing, but would have to make room in their budgets by scrimping on other items.

An additional 12 percent said they would borrow from family or friends, while 8 percent said they would take personal loans.

"The precarious state of Americans' emergency savings has been further set back by the pandemic, with nearly as many needing to borrow to cover a \$1,000 unplanned expense as those that can pay for it from savings," says Greg McBride, CFA, Bankrate chief financial analyst.

Key findings:

- [During booms and busts, the share of Americans struggling with financial instability has held steady.](#)
- [Nearly 4 in 10 Americans would need to borrow to pay an unexpected bill.](#)

- 44 percent expect their financial situations to improve this year.

During booms and busts, the share of Americans struggling with financial instability has held steady

The results of Bankrate's surveys have remained consistent for years. Since 2014, the percentage of U.S. adults who would tap cash reserves to cover a \$1,000 emergency has hovered between 37 percent to 41 percent.

The results echo findings from the Federal Reserve and the Pew Charitable Trusts. Both have reported that many Americans haven't managed to stash away rainy-day savings, and therefore must rely on credit cards or borrowing from friends and family for an unexpected expense.

The higher your household income, the more likely you would be to use savings to pay for unanticipated costs. For households earning \$75,000 or more annually, fully 58 percent have built enough of a nest egg to absorb a \$1,000 hit, the Bankrate survey found. For households making less than \$30,000, just 21 percent have a rainy-day fund to cover \$1,000.

Age also matters. Only a third of millennials could turn to emergency funds to pay \$1,000. By contrast, 46 percent of Gen Xers and 45 percent of baby boomers said they could cover a \$1,000 emergency.

Fortunes improve for some, decline for others

Affluent Americans have done well during the coronavirus recession. Many white-collar workers switched to working from home and continued to collect paychecks. With the stock market and home values at record levels, those who own assets have thrived.

However, service-sector workers haven't fared so well. Many have lost their jobs as customers stay home. Those who kept their positions have risked their health at workplaces that put them in close contact with other people.

Economists have begun to refer to that disconnect as the K-shaped economy: Those at the top enjoy improving fortunes, while those at the bottom see their financial situations decline.

"Things are getting better for some while getting worse for others," McBride says. "While things were certainly dire in 2020 for the millions of households that lost a job, had a health issue or suffered an income disruption due to the pandemic, for millions of other Americans, 2020 was a year of significantly boosting savings and paying down debt. Stimulus checks and money not being spent on vacations, ballgames and concerts helped a lot of households better secure their financial foundation."

On the food line in Florida: 'We're surviving'

Long lines of cars at food banks throughout the country illustrate the stark reality of life on the bottom of the K. At a recent food giveaway at a church in West Palm Beach, Florida, school secretary Pamela Bryant was among the two dozen or so motorists waiting in a queue that stretched for two blocks.

Bryant said she lost income to the COVID-19 pandemic and burned through her modest savings. "The little that I did have, I had to use," she said. "It's just a tough situation."

A few vehicles away, Casimira Rodriguez said she lost all of her income as a self-employed house cleaner when the pandemic struck in March. Her husband is still working, but the couple has exhausted their savings.

“We’re surviving,” Rodriguez said.

She hopes to go back to work when the spread of the coronavirus slows, but she’s unsure when that will happen.

Nearly 4 in 10 Americans would need to borrow to pay for an unexpected bill

For adults who would need to borrow to cover a \$1,000 emergency, the most common option is putting the expense on a credit card now and dealing with the financial consequences later. A credit card debt was the preferred payment method for 18 percent of Americans.

However, going into debt to deal with a rainy day is expensive. Even as interest rates have plunged on other types of debt, the average interest rate on a credit card remains north of 16 percent, according to Bankrate's national survey of lenders.

If you don't pay off that surprise expense quickly, credit card finance charges can add hundreds of dollars to the cost of that mechanic's bill or hospital visit.

As for the 8 percent who said they'd need a personal loan, many face unsavory options such as high-interest payday loans, says Signe-Mary McKernan, vice president at the nonprofit Urban Institute. She says consumers should look for better options, such as personal loans from credit unions or from employers that offer emergency loans as a worker benefit.

44 percent expect their financial situations to improve this year

Americans are optimistic that 2021 will be better for their finances, with 44 percent forecasting their finances will improve. That includes 12 percent who say their fortunes will get significantly better and 32 percent saying they will get somewhat better.

Just 14 percent expect their finances to get worse in 2020.

Optimism about an improved financial situation in 2021 declines with age. Younger millennials, ages 24 to 30, are the most optimistic, with 53 percent expecting an improved financial situation this year. Just 28 percent of Americans age 66 and older think things will improve in 2021.

Half of the highest-earning households expect an improved financial situation in 2021, while just 37 percent of the lowest-earning households feel that way.

Political differences are evident, too. While 56 percent of Democrats expect an improved financial situation in 2021, just 33 percent of Republicans feel that way.

Preparing for tough times

The U.S. economy endured a sharp downturn in 2020. Unemployment spiked into the double digits.

The coronavirus recession served as a stark reminder of the wisdom of stashing away three to six months of living expenses in a no-risk account. While that basic bit of advice hasn't changed for decades, many Americans still struggle to save, whether it's because their incomes are barely enough to cover their cost of living or they lack financial discipline.

The results of Bankrate's latest survey serve as a reminder to get into the habit of saving, McBride says.

"Establishing the habit via direct deposit from your paycheck or automatic transfer from checking into savings is critically important, especially if you're starting from a position of little or no savings," McBride says. "If you wait until the end of the month and try to save what is left over, too often there is nothing left over. Becoming a good saver isn't just a switch that you flip one day."

While you'll ultimately want to amass a few thousand dollars in savings, McKernan says workers who are living paycheck to paycheck should start with a modest goal.

"When you hear three months of living expenses, that can be daunting," she says. "Even small amounts of savings help."

The Urban Institute's research shows that setting aside as little as \$250 can help consumers avoid eviction, disconnected utilities and other financial calamities, McKernan says.

Methodology

This study was conducted for Bankrate via telephone by SSRS on its Omnibus survey platform. The SSRS Omnibus is a national, weekly, dual-frame bilingual telephone survey. Interviews were conducted from Dec. 8-13, 2020, among a sample of 1,003 respondents in English (970) and Spanish (33). Telephone interviews were conducted by

landline (293) and cell phone (710, including 469 without a landline phone). The margin of error for total respondents is +/-3.58 percent at the 95 percent confidence level. All SSRS Omnibus data are weighted to represent the target population.

Written by

Jeff Ostrowski

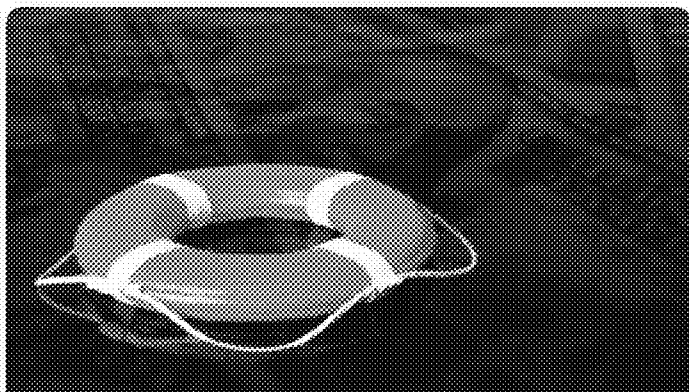
Senior mortgage reporter



[Read more from Jeff](#)

Jeff Ostrowski covers mortgages and the housing market. Before joining Bankrate in 2020, he wrote about real estate and the economy for the Palm Beach Post and the South Florida Business Journal.

Related Articles



Survey: Nearly 4 in 10 Americans would borrow money to cover a \$1K emergency



Survey: Less than half of Americans have savings to cover a \$1,000 surprise expense

CLIENT ALERT

Happy Anniversary, DTSA: The Defend Trade Secrets Act at Five

25 May 2021

Keep up with the latest legal and industry insights, news, and events from MoFo



SIGN UP



On May 11, 2016, the Defend Trade Secrets Act (“DTSA”) was signed into law with sweeping bipartisan support, passing unanimously in the Senate, and by a vote of 410-2 in the House.^[1] In the current political climate, passing any significant piece of legislation by such a lopsided vote seems almost unthinkable. The Act’s popularity can be attributed to an objective that seemingly everybody could agree upon—protecting the intellectual property of United States companies—and the crucial decision to create a federal statute that largely mirrored existing state law.

Three key goals of the statute were: (1) to create a federal civil remedy, bringing the rights of trade secret owners “into alignment with those long enjoyed by owners of other forms of intellectual property,” (2) to promote uniformity by providing for a “single, national standard for trade secret misappropriation with clear rules and predictability for everyone involved,” (3) and to better address the concerns of “a globalized and national economy” where trade secrets can readily be spirited across state lines and “beyond the reach of American law.”^[2] With the DTSA’s fifth anniversary this month, we take stock of how much progress has been made towards accomplishing those goals.

Access to Federal Courts

ing, or intended for use in interstate or foreign commerce. [6] This means nearly all trade secret plaintiffs now have a right to file suit in federal court. The more interesting question is the extent to which parties have taken advantage of that right.

A recent Lex Machina study found that, after holding steady for several years, trade secret filings in federal court increased by 30% between 2015 and 2017—jumping from 1,075 in 2015 to 1,396 in 2017. Federal trade secret filings have remained at that level since, with 1,397 cases filed in 2018, and 1,401 cases filed in 2019. [4]

The Lex Machina report does not address state court trade secret filings, which are more challenging to track due to variability in electronic filing. However, our review of data obtained through *Courthouse News* shows that state court trade secret filings stayed roughly the same during this time period, or declined slightly. [5] This confirms that the increase in federal trade secret cases can be attributed to the DTSA, not just to an increase in trade secret cases generally.

Uniformity

While the DTSA was intended to promote uniformity, the Act explicitly does not preempt state law. [6] The statute also adopted definitions of key terms such as “misappropriation” from the Uniform Trade Secrets Act (“UTSA”), on which state statutes are modelled, “to make clear that this Act is not intended to alter the balance of current trade secret law or alter specific court decisions.” [7] These choices greatly facilitated the Act’s passing, but also significantly undermined the goal of uniformity.

Rather than providing consistent outcomes across jurisdictions, so far the DTSA has taken the shape of the law of the forum state. In a typical example, the Ninth Circuit has stated that the California UTSA is “analogous” to the DTSA, [8] and that federal and state claims can be analyzed together because the elements of each are “substantially similar.” [9] Other courts agree, [10] and a number have analyzed federal and state trade secret claims exclusively under state law. [11] In *In re Patriot National*, for example, the court mirrored its rulings under the state statutes for the DTSA claims. [12]

The result is that state law differences are being imported into the DTSA, creating the same “patchwork” of law the statute was intended to rectify. Perhaps the best example of this is the courts’ differing treatment of “inevitable disclosure,” one of the key areas where trade secret law differs among states. The doctrine essentially allows a plaintiff to prove misappropriation by showing a defendant’s new job is so similar to their prior one that they “inevitably” will make use of trade secrets. The DTSA steered clear of inevitable disclosure by prohibiting injunctions that prevent “a person from entering into an

knows. It also precludes injunctions that would otherwise conflict with state law prohibiting restraints of trade.^[13] In California, where the doctrine has long been rejected, federal courts find that there is no claim based on inevitable disclosure under the DTSA.^[14] But a series of cases in the Northern District of Illinois, which sits in the Seventh Circuit—author of the country’s leading opinion upholding inevitable disclosure^[15]—have reached the opposite conclusion.^[16] Other federal courts have also followed state law on this issue^[17] and, where it is undecided under state law, have declined to determine its applicability under the DTSA.^[18]

Federal trade secret law also lacks uniformity on the important issue of when, and to what extent, a plaintiff must identify its trade secrets—an issue which the DTSA does not expressly address. A growing number of courts are requiring plaintiffs to identify their trade secrets with reasonable particularity before taking discovery.^[19] This has long been the case in California, where many district courts have taken their cue from California’s state statute requiring a plaintiff to identify its trade secrets with “reasonable particularity.”^[20] But some courts reject pre-discovery identification altogether.^[21]

The DTSA is still young, however. Uniformity may develop over time as more cases make their way to appellate courts. If the Supreme Court gets a chance to weigh in, it may set the statute back on course to providing more of the consistency Congress hoped to achieve.

Protection for a Global Economy

So far the DTSA has achieved more success with its third goal. Led by the Northern District of Illinois’s decision in *Motorola Sols., Inc. v. Hytera Commun. Corp. Ltd.*,^[22] several district courts around the country have now held that the DTSA has extraterritorial effect.^[23] These holdings are based on 18 U.S.C. § 1837, which applies where (1) the offender is a citizen or permanent resident of the U.S. or organized under U.S. law, or (2) “an act in furtherance of the offense” was committed in the U.S.^[24]

Courts have found acts “in furtherance” where meetings in the U.S. purportedly led to later misappropriation, even though those contacts did not themselves constitute elements of misappropriation.^[25] Directing communications related to trade secrets to the U.S. and accessing U.S.-based servers may also be sufficient.^[26] One court has applied the DTSA where the purported act in furtherance was performed by a third party, rather than the defendant.^[27] A defendant’s activities in the U.S. are unlikely to satisfy the “in furtherance” requirement, however, when there is not a clear nexus with a DTSA violation.^[28] While the case law on extraterritoriality may shift as these issues work their

The DTSA's *ex parte* civil seizure provision was also seen as a tool for fighting international misappropriation, and was "expected to be used in instances in which a defendant is seeking to flee the country."^[20] This provision goes beyond traditional trade secrets law, providing for seizure by a Federal law enforcement officer, who may be aided by a private technical expert,^[30] after a true *ex parte* hearing.

The seizure provision contains many requirements and is to be used only in extraordinary circumstances.^[31] That has been true to date: our review of electronic databases reveals only ten reported orders granting *ex parte* seizures, out of 21 total applications. As a general matter, courts have followed the DTSA's guidance and refrained from seizure orders if a less intrusive TRO or injunction would be sufficient.^[32] Still, it remains a powerful weapon under the right circumstances. Courts that grant seizure orders often reference concrete allegations establishing that the defendant would not follow a court order, such as past dishonesty or a defendant's technical proficiency and ability to conceal evidence.^[33]

At five the DTSA is still finding its legs, but it is already a frequently used and effective tool for trade secret enforcement, particularly against misappropriation with an international scope.

[1] 162 Cong. Rec. S-1631, H-2046.

[2] S. Rept. 114-220; H. Rept. 114-529 (2016).

[3] 18 U.S.C. § 1836(b)(1).

[4] *Lex Machina Trade Secret Litigation Report* (April 2020) at 3. According to our own analysis of more recent Lex Machina data, federal trade secret filings have remained on a similar pace, with 1,369 in 2020 and 320 in the first quarter of 2021.

[5] *Courthouse News* data shows that 1,161 trade secret cases were filed in state court in 2015, compared to 1,188 in 2016, 1,195 in 2017, 1,252 in 2018, and 1,103 in 2019. *Courthouse News* currently tracks data from over 2,875 state courts.

[6] 18 U.S.C. § 1838.

[7] S. Rept. 114-220 (2016).

[8] *ExamWorks, LLC v. Baldini*, 835 F. App'x 251, 252 (9th Cir. 2020).

[9] *InteliClear, LLC v. ETC Glob. Holdings, Inc.*, 978 F.3d 653, 657 (9th Cir. 2020).

Green, Inc. v. Intel Corp. Photographic Res. Corp., Inc., 2018 WL 5776551, at *3 (M.D. Fla. Nov. 2, 2018); *Medidata Sols., Inc. v. Veeva Sys. Inc.*, 2018 WL 6173349, at *3 (S.D.N.Y. Nov. 26, 2018); *ActivEngage, Inc. v. Smith*, 2019 WL 5722049, at *3 (M.D. Fla. Nov. 5, 2019).

[11] *Kuryakyn Holdings, LLC v. Ciro, LLC*, 242 F. Supp. 3d 789, 797–98 (W.D. Wis. 2017). See also *In re Patriot Nat'l Inc.*, 592 B.R. 560, 577 (Bankr. D. Del. 2018).

[12] 592 B.R. at 577.

[13] 18 USC § 1836(b)(3)(A)(i)(I)-(II).

[14] See, e.g., *EL T Sight, Inc. v. Eyelight, Inc.*, 2020 WL 7862134, at *16 (C.D. Cal. Aug. 28, 2020) (collecting cases).

[15] *PepsiCo, Inc. v. Redmond*, 54 F.3d 1262 (7th Cir. 1995).

[16] See, e.g., *Inventus Power, Inc. v. Shenzhen Ace Battery Co.*, 2020 WL 3960451, at *11 (N.D. Ill. July 13, 2020) (and cases cited therein).

[17] See, e.g., *Sunbelt Rentals, Inc. v. Love*, 2021 WL 82370, at *68 (D.N.J. Jan. 11, 2021), *appeal pending*, No. 21-1233 (3rd Cir. filed Feb. 8, 2021).

[18] See, e.g., *AWP, Inc. v. Henry*, 2020 WL 6876299, at **4-5 (N.D. Ga Oct. 28, 2020).

[19] See *A&P Tech., Inc. v. Lariviere*, 2017 WL 6606961, at *9 (S.D. Ohio Dec. 27, 2017) (collecting cases).

[20] *Alta Devices, Inc. v. LG Elecs., Inc.*, 2019 WL 176261, at *1 (N.D. Cal. Jan. 10, 2019) (applying Cal. Code Civ. Proc. § 2019.210).

[21] See *A&P*, 2017 WL 660691, at *9 (noting “divergent rulings from various federal courts”).

[22] 436 F. Supp. 3d 1150, 1160 (N.D. Ill. 2020).

[23] See, e.g., *Herrmann Intl., Inc. v. Herrmann Intl. Europe*, 2021 WL 861712, at *16 (W.D.N.C. Mar. 8, 2021); *Medcenter Holdings Inc. v. WebMD Health Corp.*, 2021 WL 1178129, at *6 (S.D.N.Y. Mar. 29, 2021); *Syntel Sterling Best Shores Mauritius Ltd., v. Trizetto Group, Inc.*, 2021 WL 1553926, at *14 (S.D.N.Y. Apr. 20, 2021).

[24] This provision was originally part of the criminal Economic Espionage Act, which the DTSA amended to add a civil cause of action.

Aug. 27, 2020).

[27] *vPersonalize Inc. v. Magnetize Consultants Ltd.*, 437 F. Supp. 3d 860, 878 (W.D. Wash. 2020) (DTSA “does not require the *defendant* to have committed such act”).

[28] *ProV Intl. Inc. v. Lucca*, 2019 WL 5578880, at *3 (M.D. Fla. Oct. 29, 2019) (attendance at trade show and submission of resignation in U.S. did not satisfy statute).

[29] S. Rept. 114-220.

[30] 18 U.S.C. § 1836(b)(2)(E).

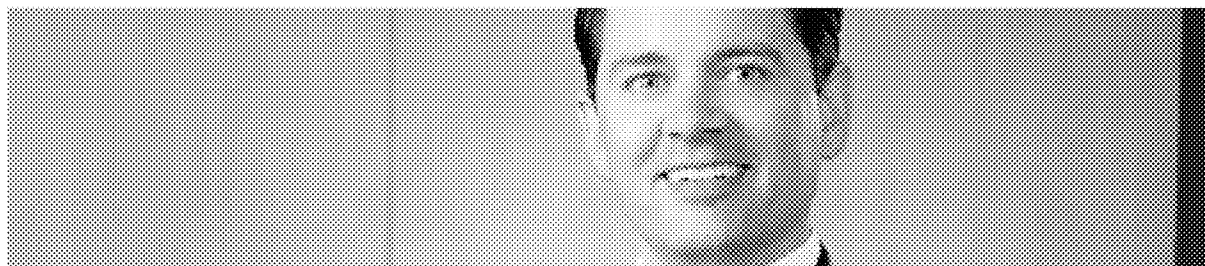
[31] 18 U.S.C. § 1836(b)(2).

[32] 18 U.S.C. § 1836(b)(2)(A)(ii)(I). *See, e.g., OOO Brunswick Rail Mgt. v. Sultanov*, 2017 WL 67119, at *2 (N.D. Cal. Jan. 6, 2017) (issuing preservation order and TRO rather than DTSA seizure).

[33] *See Mission Capital Advisors LLC v. Romaka*, 2016 WL 11517104, at *2 (S.D.N.Y. July 29, 2016) (defendant falsely represented he removed trade secrets from computer); *Solar Connect, LLC v. Endicott*, 2018 WL 2386066, at *2 (D. Utah Apr. 6, 2018) (defendants lied and hid information; also had “high level of computer and technical proficiency”); *Axis Steel Detailing, Inc. v. Prilex Detailing LLC*, 2017 WL 8947964, at *2 (D. Utah June 29, 2017) (similar).



Kenneth A. Kuwayti
Partner



.....

PRACTICES

Trade Secrets

Mobility Restrictions, Bargaining, and Wages: Evidence from the National Longitudinal Survey of Youth 1997

Donna S. Rothstein*
Bureau of Labor Statistics

Evan Starr
University of Maryland

November 2021

Abstract

We examine the use of noncompete agreements (NCAs) and their relationship with wage bargaining and wage outcomes using new data from the National Longitudinal Survey of Youth 1997. NCAs cover 18% of the workers in our sample and adoption patterns are broadly consistent with prior research. The NCA-wage correlation is positive and highly sensitive to controls for demographics and job characteristics, suggesting selection into NCAs causes positive bias in the estimates. While it is not obvious what the baseline level of the NCA-wage differential is, some heterogeneous effects are more stable: the NCA-wage differential is lower for workers that do not bargain over wages, have less education, have lower ability, or live in a state that enforces NCAs. Notably, wage bargaining—which is only marginally more likely with NCAs in our most saturated model—does not explain the heterogeneous effects across subgroups. We discuss these findings in light of competing theories of the social value of NCAs, and describe future directions for research on NCAs as more waves of data are collected.

Keywords: Noncompete Agreements, Bargaining, Wages
JEL Classifications: J3, J41, J42, K31

* The views expressed are those of the authors and do not reflect the policies of the BLS or the views of other BLS staff members.

1. Introduction

Amid a decades-long decline in job mobility and wage stagnation, the last few years has witnessed renewed policy and research interest in the use of noncompete agreements (“NCAs”), which are employment provisions that prohibit departing workers from joining or starting competing businesses, often within time and geographic limits (US Treasury 2016, White House 2020). Since the 2014 discovery of NCAs in low-wage jobs (Greenhouse 2014, Jamieson 2014, Starr et al. 2021, Johnson and Lipsitz 2020), more than 69 new state or federal NCA policies have been proposed, including bans on NCAs for all or a subset of the workforce.¹ These proposals join a centuries-long debate over the value of NCAs, which juxtaposes the potential for NCAs to constrain the upward mobility of workers and their potential to restore incentives for firms to invest in the development and sharing of valuable information (Rubin and Shedd 1981).

A growing stream of academic research has aided this debate by seeking to understand how NCAs and the policies that regulate them influence economic activity. The vast majority of this research examines NCA policies alone, however, without any information on the actual use of NCAs (Bishara and Starr 2016).² This omission is critical, given that the limited data we do have on NCAs suggests that they are frequently found in states where they are per se unenforceable (Sanga 2018, Starr et al. 2021, Colvin and Shierholz 2020), that workers perceive

¹ For example, in 2021 the Uniform Law Commission promulgated the Uniform Restrictive Employment Agreement Act for adoption by state legislatures. The proposed act bans NCAs and related restrictive agreements for low-wage workers and mandates notice and other requirements for other workers (<https://www.uniformlaws.org/committees/community-home?communitykey=f870a839-27cd-4150-ad3f-51d8214f1cd2&tab=groupdetails>). For other state and federal policies, see generally <https://faircompetitionlaw.com/changing-landscape-of-trade-secrets-laws-and-noncompete-laws/>. In addition, state Attorneys General have investigated more than a dozen NCA cases (Madigan and Flanigan 2018), the Federal Trade Commission has considered making a rule related to regulating NCAs, and several federal agencies have written reports on the topic (US Treasury 2016, White House 2016).

² See Starr (2019) for a review of this literature. Indeed, only a handful of studies possess data on the use of NCAs (Posner and Krueger 2018), and most examine a specific occupational context, such as executives (Schwab and Thomas 2006, Bishara et al. 2015), engineers (Marx 2011), physicians (Lavetti et al. 2020), and hair stylists (Johnson and Lipsitz 2020).

their NCAs to be enforceable when they are not (Prescott and Starr 2021), and that NCAs can limit employee mobility regardless of the law (Starr et al. 2020). More broadly, existing data on NCAs have four limitations: (1) they are not publicly available; (2) they come from either selected occupations or non-random sampling schemes; (3) they are cross-sectional; (4) they are not repeated cross-sections of the same population or sampling frame. As a result, researchers have not been able to study the evolution of NCA use and how NCAs affect a variety of economic dynamics, like the historical decline in business dynamism and wage stagnation.

To address these concerns, in 2017 the Bureau of Labor Statistics (BLS) added a question on NCAs to the National Longitudinal Survey of Youth 1997 (NLSY97)—a panel dataset consisting of individuals born between 1980 and 1984. The first NLSY97 wave with NCA data was published in December 2019, and data collection efforts are ongoing. These data address the gaps highlighted above by providing a publicly available, longitudinal dataset which will allow researchers to develop new evidence on this important labor market friction.

In this manuscript we introduce the first wave of these data.³ We begin with a brief discussion of the theoretical tensions related to NCAs, focusing on bargaining and hold-up. Then we describe the NLSY97 and the new NCA question. In our empirical work, we examine the use of NCAs and its correlates, drawing parallels to prior work where possible. We then focus on how NCAs relate to wages, in light of the competing predictions made by extant theories. *Our estimates here should not be interpreted causally*—indeed, one of our key findings is that the sensitivity of the NCA-wage relationship to controls suggests significant selection into NCA use. In our analysis we leverage a unique question on wage bargaining to understand how NCAs relate to wage bargaining, and its role in explaining (a) heterogeneity in the overall NCA-wage

³ Concurrent work by Boesch et al. (2021) also examine the incidence of NCAs using the NLSY97 data.

relationship and (b) for heterogeneous effects observed by gender, education, ability, and NCA enforceability. We conclude with a discussion of research directions as future waves of data become available.

2. Guiding Theory and Institutional Background

Since the first legal case dating back to 1414, NCAs have been a topic of significant theoretical debate (Blake 1960). The essence of the debate is to understand whether and under what circumstances it is worth preventing workers from deploying their full set of human capital in a competing firm (typically within some time and geographic boundaries). Courts have generally been concerned that noncompetes, like other restraints of trade, can impose significant hardship on workers, since workers who wish to leave the firm without violating their NCA will either have to change industries (Marx 2011), leave the geographic area (Marx et al. 2015, Balasubramanian et al. 2020), or sit out of the labor market. Moreover, since NCAs increase the costs of moving to a competitor, they shield the firm from labor market competition, potentially curtailing wage growth for workers (Starr 2019, Lipsitz and Starr 2021, Johnson et al. 2021).

However, theories rooted in efficient contracting posit that NCAs will only be observed when they are mutually beneficial to firms and workers. The theories tend to have two components: First, workers have the “freedom to contract,” such that they would only agree to an NCA if it made them better off (Friedman 1991, Callahan 1985). Second, firms would never pay a worker a compensating differential for an NCA unless they too were benefiting from it. And the reason firms might benefit from NCAs is that they resolve an investment hold-up problem (Rubin and Shedd 1981). If a firm were to share valuable information with a worker, without an NCA the worker could hold-up the firm by threatening to misappropriate that information at a competitor. As a result, the firm may be unwilling to develop such information in the first place,

or unwilling to share it with the worker, both of which may reduce the worker's productivity. Accordingly, under this view NCAs can *only* be productive for both workers and firms, because they give firms stronger incentives to invest in worker training and in the development of valuable information (Barnett and Sichelman 2020).

Despite a burgeoning literature on NCAs, which of these theories is most accurate is still an open question. These competing theories make different predictions both about where NCAs should be observed, and (among other things) how NCAs relate to wages. Regarding the use of NCAs, the hold-up theory suggests that NCAs will be used mostly in jobs that have access to valuable information (e.g., trade secrets, client lists), and only in places where they can be enforced (since court enforcement underlies firm confidence that NCAs will resolve the hold-up problem).⁴ In contrast, theories that firms are using NCAs as value extraction tools (Balan 2021) posit that they will be used much more broadly—potentially even with low-wage workers who have no access to valuable information, and in places where they cannot be enforced.

With regards to wages, three possibilities arise: (1) workers may receive a compensating differential (whether they had to negotiate for it or if it was included in the offer) for signing an NCA, but then suffer lower wage growth as the NCA shields them from competitors; (2) wage growth may rise if NCAs indeed spur productivity-enhancing investments, and wages are tied to productivity; (3) workers may not receive any compensating differential (because e.g., they just sign the NCA when asked, as in Arnow-Richman 2006), and experience lower wage growth.

Prior research finds some evidence in favor of each of these arguments. NCAs are adopted widely, even though they tend to be more common in states that enforce them and for

⁴ In practice, unenforceable noncompetes may resolve hold-up problems to some degree if, (a) workers are unaware of the law (Prescott and Starr 2021), (b) workers cannot access legal counsel or otherwise face costs of breaking an unenforceable contract (Starr et al. 2020). Classic efficient contracting theories do not consider these possibilities.

workers in technical jobs (Starr et al. 2021, Balasubramanian et al. 2021). Regarding wage outcomes (summarized in Starr 2021), prior research on NCA enforceability finds negative effects on wage levels and wage growth (Balasubramanian et al. 2020, Lipsitz and Starr 2021, Johnson et al. 2020, Starr 2019), while studies of NCA use find positive wage effects and positive wage growth (Starr et al. 2021, Lavetti et al. 2020, Kini et al. 2020, Shi 2020). The discrepancy in wage results could arise either from the specific occupations studied, differences between the actual effects of NCA enforceability and NCAs themselves, the time period studied, selection into NCA use, the cross-sectional nature of the studies of NCA use, and lack of data on key variables (i.e., wage bargaining, job tasks, ability, etc.).⁵ In this regard, new data collected via the National Longitudinal Survey of Youth 1997 offer an important opportunity to push this literature forward, especially as more waves of data are collected over time.

3. Data

Background on the NLSY97 and NCA Question Design

The National Longitudinal Survey of Youth 1997 (NLSY97) is a nationally representative sample of 8,984 men and women born in the years 1980 to 1984. Sample members were first interviewed in 1997 when they were ages 12-17; the latest data available when we began this paper are from the 2017-18 interview, when they were ages 32 to 38. A particular strength of the NLSY97 is the collection of respondents' employment histories from their teenage years until the present. The employment module of the NLSY97 contains a core set of questions that are asked in each survey round about each job held since the date of the last interview, but certain additional modules of interest to research and public policy rotate in and out.

⁵ Among these explanations, Balasubramanian et al. (2021) use data on NCAs and three other restrictions and show that selection effects likely underlie the positive average NCA-wage differential, while the true effect is negative.

Recent added questions include those on NCAs, job tasks, and wage bargaining. The NCA questions first appeared in the 2017-18 survey and are also in the 2019-20 survey (data released in November 2021). In the 2017-18 survey the NCA questions were asked of all jobs that were not military or self-employed. In the 2019-20 survey, the NCA questions were restricted to newly reported jobs since the date of the last interview.

In the 2017-18 survey, for each job held since the date of the last interview, the respondent is asked about a series of job characteristics. The NCA question is as follows:

Some employers try to restrict what their employees can do after they leave their job. In this job, did you agree that if you [leave/left] your employer, you [will/would] not start or join a competing business? This is often called a non-compete agreement.

Because prior research has documented uncertainty in who signs NCAs (Starr et al. 2021), a follow-up question asks, “*How confident are you in your answer?*” The wording of the two NLSY97 questions on NCA agreements were based on those asked in prior surveys on the same topic (Balasubramanian et al. 2021).

Sample Construction

To construct our sample, we take the full NLSY97 sample (N=8,984) and keep those who responded to the 2017-18 (round 18) interview (N = 6,734). We then restrict the sample to those who reported a job in the interview (N=5,970). We drop the self-employed, government, and military workers and those who are working for their family without pay (N=4,481). We also drop those whose geographic region is missing (do not reside in the US at the 2017-18 interview date) (N = 4,443). We then restrict our sample to those working at their main job at least 30 hours per week (N = 3,589). We use the CPI-U to inflation-adjust hourly wages to 2017 dollars, and we drop those who earn less than \$2 an hour and those who make above \$250 an hour or those missing wage information (N = 3,490). Finally, we drop those whose NCA variable is

missing (N = 3,426), those with missing wage bargaining questions (whether they bargained over pay when they were first offered their job) (N = 3,092), and a few observations with an occupation code of 9990. Our final sample consists of 3,090 individuals. We use the NLSY97 weights for the 2017-18 interview, which account for the oversamples of black and Hispanic individuals in the NLSY97 data and the complex survey design.

4. The Incidence of Noncompetes

We begin by examining the incidence of NCAs. Table 1 provides summary statistics on NCA incidence from the NLSY97 in columns 1 and 2, and, for comparison purposes, data from the 2014 Noncompete Survey Project (Starr et al. 2021) in column 3 and data from the 2019 Cornell National Social Survey (CNSS) in column 4, collected by Schwab and Starr (2019).⁶ Overall, 18.1 percent of the NLSY97 sample is bound by an NCA, identical to the overall multiple imputation estimates reported in Starr et al. (2021), but slightly larger than the lower-bound estimates for this age group. The estimates are also nearly identical to the CNSS estimates. With regards to uncertainty regarding whether they have an NCA, 90.4 percent are very confident in their answer, whereas 9 percent and 0.7 are somewhat and not confident.⁷

We briefly describe some of the NLSY97 NCA incidence results from Table 1. In the NLSY97 men are about 5 percentage points more likely than women to report signing an NCA at their job (20 vs. 15 percent), while non-black, non-Hispanic workers are 4 percentage points

⁶ Data from the 2014 Noncompete Survey Project is described in greater detail in Prescott, Bishara, and Starr (2016) and covers in total 11,505 respondents. It derives from an online survey that the authors created and deployed via Qualtrics in 2014. Note that data from the 2014 Noncompete Survey Project include both imputed and lower bound estimates, which differ in how they treat individuals who are unaware whether they have signed a noncompete; here we emphasize the lower-bound estimates. Data from the CNSS derives from a random digit dial survey of 1,000 respondents. The noncompete question from the CNSS data is very similar to the one in the NLSY97. Note that, relative to the NLSY97, which is cohort-specific, these surveys cover all age categories. Accordingly in the Noncompete Survey Project Data we limit to the same age range as the NLSY97 and in the CNSS we limit to 25-50, in order to keep a large enough sample to say anything meaningful.

⁷ We also examined confidence levels by gender, education, wages, and NCA status. Across all these cuts, at least 81 percent of workers are very confident in their NCA answer and at most 1.6 percent are not confident.

more likely to be bound by an NCA than either black or Hispanic workers. Figure 1 shows that NCA incidence rises with education, with 15 percent of those without a bachelor's degree signing one, compared to 24 percent with at least a bachelor's degree.

In terms of worker and firm characteristics, Table 1 shows that NCAs rise with tenure, and that NCAs are 12 percentage points more common for those working in the for-profit sector than the nonprofit sector (19.6 vs 7.4 percent). Interestingly, unionized workers are only somewhat less likely to sign NCAs (16.6 vs. 18.6 percent). With regards to wages, Figure 2 shows that the incidence of NCAs is 9-11 percent for those in the bottom two wage deciles and rises with wages such that those with wages in the top decile (at least \$45 dollars per hour) have a 32 percent chance of having an NCA. Overall, NCAs are still found at the low-end of the wage distribution, with 14.4 percent of workers earning less than median hourly wages signing one.

Figure 3 and 4 show the distribution of NCAs by two-digit occupation and industry codes (conditional on having at least 20 observations in the occupation or industry). Consistent with hold-up theories, occupations where NCAs are found most frequently are in more technical areas such as engineering (38 percent), computer science (36 percent), sales (28 percent), and management (24 percent), whereas occupations such as food preparation (7 percent) and social services (4 percent) have very low reported NCA incidences.⁸ Similarly, Figure 4 shows that workers in industries such as professional services and information have high rates of NCAs (33 percent and 30 percent, respectively) in contrast to workers in social services, food services (10 percent), or agriculture (6 percent).

We also consider whether NCAs are deployed even in states that would not sanction them. Only three states—California, North Dakota, and Oklahoma—will void all noncompetes

⁸ Interestingly, legal jobs have the lowest use of NCAs (4 percent), which likely arises because they are the only occupation in which NCAs are unenforceable in all 50 states (Buffkin 1999, Starr et al. 2017).

agreed to in the employment context, and these policies have been in place since the 1800s. (Arnow-Richman 2020).⁹ Table 1 shows that 15 percent of workers who live in these states are bound by NCAs, compared to 18.5 percent elsewhere.

Overall, while there are some discrepancies between the magnitude or direction of the NLSY97 results relative to both the 2014 NSP and the 2019 CNSS (which themselves do not always agree), the general patterns and magnitudes are roughly in line.

In Panel B of Table 1 we examine variables unique to the NLSY97. First, although investing in worker training is an oft-referenced rationale for using NCAs (Rubin and Shedd 1981), workers whose employers have provided at least some training in the past are only marginally more likely to have noncompetes (19.8 percent to 17.7 percent). Second, the NLSY97 includes a unique measure of ability—the Armed Forces Qualification Test (AFQT) (math/verbal percentile score).¹⁰ Figure 5 breaks down AFQT scores by decile, showing that the incidence of NCAs is 11 percent for those with the lowest AFQT scores, but rises roughly monotonically such that those with the highest AFQT score have a 23 percent likelihood of agreeing to an NCA.

Lastly, job tasks (similar to those in the 2008 Princeton Data Improvement Initiative Survey; Autor and Handel 2013) show considerable variation with NCA use: Individuals in jobs that require more physical and repetitive tasks are about 7 percentage points less likely to report signing an NCA, whereas individuals in jobs with more problem solving, reading long documents, and supervising are much more likely to sign one.

⁹ Other states have NCA bans for some sets of workers, though most of these started in 2017 or later (Beck 2021).

¹⁰ The Armed Forces Qualification Test (AFQT) covers four sections of the Armed Services Vocational Aptitude Battery (ASVAB) and measures math and verbal aptitude. This test was given to NLSY97 respondents in 1997–1998.

Since many of the characteristics described above are likely to be correlated with each other, in Table 2 we incorporate these variables into a linear probability model to assess which characteristics are correlated with NCA use, conditional on the other variables. We cluster the standard errors by state. Several patterns emerge: First, across all models having a bachelor's degree is associated with a greater chance of signing an NCA, even though AFQT scores are uncorrelated with NCA use.¹¹ Non-profit jobs are also far less likely to have NCAs relative to for-profit jobs (9.1 percentage points in the most saturated model). While the use of NCAs appears to be lower in states that do not enforce NCAs per se, this difference becomes statistically insignificant with more controls. We also see that, even conditional on occupation and industry, several job tasks are still correlated with NCA use, including face-to-face contact with others (+4.4 percentage points), reading longer documents (+4.5 percentage points), solving problems daily (+6.3 percentage points), or frequent physical tasks (-3.3 percentage points).

5. NCAs, Bargaining, and Wages

Empirical Approach

In this section we leverage the fact that the NLSY97 has data on both wages and wage bargaining to examine how NCAs relate to wage bargaining and wage outcomes. We begin with a discussion of the ideal empirical designs to estimate the effect of NCAs, what our approach is, and why, ultimately, our results should be thought of as correlational and not causal in nature.

The ideal empirical design to estimate the causal effect of NCAs on bargaining and wages is to randomly ask some workers to sign NCAs and not others. Then one could consider who turns down the offer outright, who negotiates over the NCA or the terms of the offer, and

¹¹ Theoretically, one may worry that workers sort into NCAs on the basis of unobserved ability; and since unobserved ability also drives wages, such sorting will cause upward bias in the NCA-wage effect. The results in Table 2 suggest that workers are not sorting in NCAs by ability, conditional on demographic characteristics.

wage outcomes. If NCAs were randomly assigned, then no other firm or worker characteristics (observed or unobserved) would differ between who received an NCA and who didn't—at least before the NCA was deployed—allowing us to isolate the effect of NCAs. To our knowledge, such an experiment has yet to be run in the real world.

An alternative approach to estimating the causal effect of NCAs is to find an instrument—something that would randomly cause some firms to use NCAs but would not be correlated with wages or bargaining through any other pathway. The most natural instrument, it seems, would be the enforceability of NCAs, which might exogenously increase the firm's willingness to use them. However, the fact that firms still use NCAs relatively frequently in states that do not enforce them poses some challenges for this approach. The exclusion restriction is also likely to be violated if the instrument is just cross-sectional state NCA enforceability, since other state characteristics might be correlated with the policy and outcomes of interest. Variation over time in state NCA enforceability, combined with variation over time in NCA use, is likely to be a more plausible identification strategy. Another approach that future data collection makes possible could leverage Bartik-style instruments that interact industry shares with national growth rates (Goldsmith-Pinkham et al. 2020).

To date no research has been able to use these research designs, mostly because of the cross-sectional nature of data on NCAs. Instead prior work documents conditional correlations. With just one cross-section of data, we face the same challenges, even though the NLSY97 contains some rich measures of job attributes, and so also estimate conditional correlations.

We estimate models of the form $y_i = b_0 + b_1 NCA_i + A X_i + e_i$ using ordinary least squares, where X_i is a vector of covariates. In order for b_1 to estimate the true causal effect of an NCA, we

need a conditional independence assumption to hold—that $cov(NCA, e) = 0$, conditional on X_i .¹² This assumption is highly unlikely to hold. Based on where we see NCAs being deployed, our estimates of the NCA-wage differential will likely be seriously biased upwards. For example, since NCAs are more common in technical jobs or for workers with more education, a worker bound by an NCA is highly likely to earn more than a worker not bound by NCA—but this difference is perhaps mostly or entirely due to differences in their human capital, the type of job they are in, and the tasks they are asked to perform. We can control for some of these variables at a broad level, which should mitigate these concerns, but because we cannot hold constant all of the variables that determine both NCA use and wages, the positive bias will likely persist.

Nevertheless, inclusion of different covariates can be informative of the extent of selection into NCAs and thus the extent to which the NCA-wage differential is biased upward. Accordingly, we estimate two sets of models, one with “basic” controls, which are exogenous demographic characteristics. These are education, gender, race, AFQT score at or above 50th percentile, and whether the state enforces NCAs. We also estimate models which seek to compare workers who are in the same type of job and doing the same set of tasks. To do this we add “advanced” controls to the basic controls. These are the for-profit status of the firm, job tasks (as shown in Table 1), and 2-digit occupation and industry fixed effects. We note that some of the advanced controls may be “bad controls” (Cinelli et al. 2020) in that they may be endogenous to agreeing to an NCAs (i.e., the tasks a worker does may depend on whether they agree to an NCA). Due caution is required when interpreting the NCA coefficient with these controls.

Wage Bargaining and Wage Outcomes

¹² Practically, this condition means that anything else that affects wages must not also be related to NCAs. This condition will be violated if there are omitted variables that affect NCAs and wages, if there is reverse causality, etc.

We focus first on bargaining as an outcome of NCA use, and later as a mediator and moderator of the NCA-wage relationship. Bargaining is relevant because NCAs give firms power only *after* an NCA is signed. As a result, NCAs put some pressure on the initial negotiations for workers to receive compensation for their postemployment concessions. Before we turn to the results, it is worth considering why bargaining may or may not arise in response to NCAs.

Different models of the labor market differ in how they consider bargaining. For example, wage posting models (Burdett and Mortensen 1998) assume employers simply post a take it or leave it offer, precluding the possibility of bargaining. In these models, as long as the NCA is sufficiently observable and perceived as costly to the worker, a compensating differential may be “built into” the posted wages, rendering bargaining unnecessary. Other “wage bargaining” models assume that workers bargain for some proportion of the surplus from the job (Morentensen and Pissarides 1994), but these models are agnostic to the precise mechanics of how the bargaining occurs. Such a process may look as follows in the case of NCAs: The firm may initially offer an NCA paired with a wage offer that is at or slightly above the wages offered by firms that do not use NCAs. In this situation, the worker may either accept the contract as presented, turn it down, or ask for a larger compensating differential. In the third case, we might observe a positive relationship between bargaining and NCAs.

To set a baseline, prior research suggests that only approximately one-third of workers bargain over their wages at all (Hall and Krueger 2012) and the only evidence on negotiation over NCAs suggests that only 10 percent of NCA signers report negotiating over the terms of their NCA or for other benefits in exchange for signing (Starr et al. 2021). In the NLSY97, 36 percent of workers report that their wage was bargained over, while the rest indicate that it was a

take it or leave it offer. Figure 6 shows that the likelihood of wage bargaining rises effectively monotonically across the wage distribution, with 15 percent of the lowest earners bargaining over their wages, compared to 61 percent of the highest.

In light of this discussion, we begin by assessing whether NCAs are associated with a greater chance of wage bargaining. Table 3 Panel A shows that while NCAs are associated with a 9.5 percentage point increase in the likelihood of wage bargaining, controlling for basic controls and advanced controls, reduces the differential to 2.1 percentage points and becomes statistically insignificant. Thus the positive relationship between NCAs and wage bargaining seems largely driven by certain individual or job-specific characteristics.

Columns (4)-(6) of Table 3 examine the baseline wage results. Unconditionally, those bound by NCAs earn about 25 percent more ($\exp(0.22)-1$). However, as in the case of bargaining, the inclusion of basic controls reduces this coefficient by half to 12.7 percent ($\exp(0.12)-1$), and the inclusion of advanced controls reduces it to just 5 percent ($\exp(0.049)-1$). Given the precipitous drop in the coefficient on NCAs as controls are added, the correlation between NCAs and wages is highly susceptible to unobserved variables. That is, there are many other variables that we cannot observe (i.e., access to valuable trade secrets, clients, etc.) which might drive both NCA use and wage outcomes. Such omitted variables will positively bias the NCA-wage correlation, even with the relatively granular controls we do observe in the NLSY97.

There are two unanswered questions that follow with regards to NCAs, wages and bargaining. First, how much of the NCA-wage differential can be explained by baseline differences in bargaining behavior? Second, do workers with NCAs who bargain actually end up with higher wages, perhaps because they asked for a *greater* compensating differential? Put differently, does wage bargaining mediate and/or moderate the NCA-wage relationship?

Columns (1)-(3) of Panel B of Table 3 address the mediation question. Column (1) shows that, in the unconditional specification, controlling for bargaining causes the NCA coefficient to fall by 13 percent (from 0.221 to 0.192). However, the NCA-wage differential explained by bargaining falls to 7.5 percent and 4.1 percent when we include controls (columns (2)-(3)); and so too does the extent to which bargaining itself positively relates to wages. Thus, bargaining only modestly mediates the NCA-wage relationship.

Columns (4)-(6) of Panel B considers moderation, allowing for bargaining to be differentially related to wages for those with NCAs. Column (4) shows that, unconditionally, the NCA-wage differential for workers *who do not bargain* over wages is 16.8 percent ($\exp(0.155)-1$)—a 29.9 percent decrease from the baseline—while the NCA-wage differential is 9.5 percent ($\exp(0.091)-1$) higher among those who do bargain. Moreover, while the controls reduce the NCA-wage differential for those who do not bargain—reducing it by 63.3 percent in the most saturated model relative to the main effect in Panel A (0.018 vs. 0.049)—the NCA-wage differential for those who bargain remains 7 percent higher.

Taken together, this suite of results suggests that NCAs are positively correlated with wages, but that there is significant selection into NCA use. As a result, our analysis does not show that NCAs cause higher wages; in fact, it may be that NCAs reduce wages but that we cannot account for all the variables that confound the NCA-wage relationship. Our results also show that wage bargaining can explain a significant amount of the NCA-wage relationship; not because workers with NCAs are necessarily more likely to bargain over wages, but because those with NCAs who do bargain drive much of the positive baseline relationship.

6. Heterogeneous Wage Effects

In this section we examine several potential heterogeneous effects discussed in the prior literature as well as novel heterogeneous effects made possible by the rich data in the NLSY97. The prior literature has emphasized the potential for historically disadvantaged populations to be especially harmed by NCAs. For example, Lipsitz and Starr (2021) find that women particularly benefit when NCAs are banned, while Johnson et al. (2020) find that both women and black workers are better off when NCA enforceability is weakened. Lastly, Starr (2019) finds evidence that those with less education are *more* likely to be harmed when NCAs are more likely to be enforced. Several rationales for these findings have been proposed, including that disadvantaged populations may be more likely to voluntarily abide by an NCA, that firms may selectively target such groups for enforcement, and that such workers are less likely to bargain over the NCA.

However, all of these studies examine *state NCA policies* and none of the studies of NCA *use* have examined similar predictions. Accordingly, in Table 4 we present analyses examining how, in the cross-section, NCAs differentially relate to wages for various subgroups. As before, we estimate models that include the same basic and advanced controls, clustering the standard errors by state.

The results largely accord with what we observed in the case of wage bargaining: The main effects are highly sensitive to the inclusion of controls, while in most cases the heterogeneous effects of NCAs are more stable (and often line up with the prior literature). For example, Panel A shows that relative to the NCA-wage differential for those with less than a bachelor's degree, the NCA-wage differential for those with a bachelor's degree is practically no different, while for those with more than a bachelor's degree it is 19 percent ($\exp(0.0175)-1$) to 25 percent ($\exp(0.227)-1$) higher.

The heterogeneous effects of NCAs are more sensitive when it comes to race and gender. Panel B shows that the NCA-wage differential for minority (black or Hispanic) workers is lower than the NCA-wage differential for non-black-non-Hispanic workers, but the estimates are noisy and fall close to zero in the most saturated model. Similarly, Panel C shows that, at baseline, men bound by NCAs earn between 7 percent and 16 percent more than men without NCAs. The same differential for women, however, ranges from 5 to 10 percent lower than that of men, with the difference being statistically significant only in the model with basic controls.

Given the literature's focus on disadvantaged workers, in Panel D we consider whether higher ability workers, as measured by having an AFQT score at or above the 50th percentile, have higher NCA-wage differentials than lower ability workers. Indeed, relative to the NCA-wage differences for those with low-AFQT scores, the same difference for those with high AFQT scores is approximately 10 percent higher.

Finally, we consider heterogeneous NCA-wage effects based on the extent of enforceability of the NCA. Under the efficient contracting theories, it is the actual law (i.e., whether a contract will be held up in court) that determines whether the firm can ultimately be protected from a hold-up problem. Accordingly, under this theory, workers should be better off where NCAs are more enforceable—either because of being more likely to bargain or because access to valuable information makes them more productive. In contrast, if NCAs are value extraction tools, then NCAs might more effectively extract value when the firm can legitimately threaten the worker with a lawsuit for violating a noncompete agreement. Panel E shows that relative to states that enforce NCAs, the wage differential associated with NCAs when they are not enforceable is 4 to 7 percent higher.

Taken together, because the base rates are so sensitive to controls, it is not obvious whether the baseline positive NCA-wage relationship is driven by selection or treatment. However, the more consistent heterogeneous effects for these groups suggest that, whatever the baseline effect is, the wage-differentials associated with NCAs is lower for those with less education, women, those with low AFQT scores, and in states more likely to enforce NCAs.

In columns (3) and (4) of Table 4 we consider the plausible theory that the observed NCA-wage differentials are driven by group differences in bargaining. If, for example, women are less likely to bargain over wages, or when they do bargain ask for smaller compensating differentials, then these baseline bargaining differences may explain why NCAs are more harmful to women than men. Accordingly, we rerun our heterogeneous effects models controlling for whether the individual bargained for their wages, and we allow for different subgroups (as defined for each panel) to have differential effects from bargaining. In each case, we observe that subgroup bargaining patterns do not explain the NCA differentials, since the estimated NCA-wage differentials budge little when including these controls.

Discussion

This study is motivated by the recent and historical debates over the value of NCAs and the relative paucity of data on NCAs themselves, amidst a growing literature studying state NCA policies. Using new data collected on NCAs as part of the NLSY97, we examine who signs NCAs, how NCAs are related to wages and wage differentials between subgroups, and the role of bargaining in explaining these differentials. Our results both support the prior literature on NCAs and extend it in new, important ways.

At a broad level, we find that 18.1 percent of workers aged 32-38 in 2017 were bound by NCAs, very similar to prior estimates (Starr et al. 2021, Schwab and Starr 2019). We also

document similar patterns to the prior literature—that the use of NCAs is more common for workers with more education, that NCAs are more common in technical occupations and industries, but that NCAs are still used for a wide swath of workers at the low-end of the wage distribution (Johnson and Lipsitz 2020) or even workers in states that would never sanction such an agreement. We extend these findings by showing that NCAs are also more common for workers with high ability, and that even within a job-type, variation in job tasks (such as problem solving) are strongly associated with NCA use. Interestingly, our results suggest little selection into NCA use by ability after conditioning on broad demographics.

Examining wage outcomes, we find that NCAs are positively associated with wages, but that this association is highly sensitive to demographic and job characteristics, as in prior work (Starr et al. 2021, Balasubramanian et al. 2021). As a result, we recommend interpreting the main correlations with due caution. Heterogeneous effects in the NCA-wage differential are more stable however. For example, the wage differential associated with NCAs is lower for those with less education (relative to more education), lower for those with lower ability (relative to higher ability) and in some models lower for women (relative to men). Interestingly, while we also find that the NCA-wage differential is lower for those who do not bargain over wages (relative to the NCA-wage differential for those who do bargain), bargaining differentials across groups do not explain the aforementioned NCA-wage differentials across groups. Finally, our results suggest that the enforceability of NCAs reduces NCA-wage differentials, as in Starr et al. 2021.

Taken together, our results are consistent with elements of the efficient contracting perspective, for example that NCAs are more common in high skilled jobs, and that NCAs are associated with higher wages on average. But our results also challenge that narrative because (a) the use of NCAs is widespread, (b) the NCA-wage effect is highly sensitive to demographic and

job controls, and (c) the fact that the positive wage associations with NCAs dissipate where NCAs are more enforceable suggest that the baseline positive wage estimates may be highly selected. Our results also suggest that since bargaining power differentials are unlikely to underlie the NCA-wage differentials for the subgroups we study, alternative theories may be considered, such as differential access to legal services or acquiescence to legal threats.

Limitations and Future Directions

While these analyses advance our understanding of NCA adoption, wage setting, and wage bargaining, they have several important shortcomings. Notably, the data are from a single cross-section, making it difficult to extract anything but correlational relationships, and precluding a study of longitudinal earnings or job mobility dynamics. However, as more data are collected, there are several clear opportunities to exploit the richness of the NLSY97. In this section we lay out a broader research agenda that these data will allow researchers to fill.

First, one of the major challenges in this literature is finding exogenous variation in the use of NCAs, given the existence of only cross-sectional data. However, as more states change their policies on NCAs, and as more data on NCAs are collected in the NLSY97, it seems natural that such policy variation could be used to instrument for NCA use. For example, between 2017 and 2021 several states banned noncompete agreements for low-wage workers (and several more are considering it). These policy changes will likely exogenously reduce the use of NCAs among the low-wage population, especially those policies that impose penalties for using NCAs deemed illegal. As long as these policies leave unaffected the status quo enforcement for those above the wage threshold (which some seem to do), then the exclusion restriction may plausibly hold (i.e., that these policies affect various outcomes only through their effect on NCA use). With this and

perhaps Bartik-style approaches, we can hopefully begin to tease out the selection and treatment effects of NCAs, and compare it with the better identified work on NCA policies.

Second, as longitudinal data are collected, the scope of variables one can analyze grows significantly, including analyses of within-individual wage profiles, job mobility choices, entrepreneurial behavior, and variation in non-wage benefits. NLSY97 data on moves and their timing also allow one to explore the relationship between NCAs and migration. Data on spouse and partner labor supply could be used to study the role job restrictions like NCAs play in dual labor market decisions. The NLSY also has unique data on several other dimensions that could be used to examine unique heterogeneous effects (such as job tasks, bargaining, AFQT, etc.), which would not be possible with other data. Moreover, those interested in understanding the causal drivers of NCA use will be better positioned to use time-variant identification strategies. For example, one can examine how changes in minimum wages over time affect NCA use, or how changes in subsidies or tax incentives for investment might drive firms into using NCAs.

Third, as longitudinal data on the use of NCAs becomes available, one can calculate estimates of the growth of NCAs and relate them to various outcomes relevant to multiple disciplines. For example, one important question is what are the downstream effects of NCAs? How does the rise of NCAs affect prices (Lipsitz and Tremblay 2021, Hausman and Lavetti 2021), product quality, R&D expenditures and innovation, and consumer welfare more broadly? Another set of questions relates to the patterns of wage stagnation and economic dynamism and the role that NCAs played in those dynamics. Note that care should be given to these estimates because estimates of the growth of NCAs will track the use of NCAs among a given age cohort, and so may just reflect how NCA adoption changes as a cohort ages. Thus it may be helpful to

benchmark the results to other nationally representative cross-sections (Balausbramanian et al. 2021, Starr et al. 2021) to separate out the trends from cohort-specific effects.

As the NLSY97 data continues to accumulate, so too will the opportunities to learn more about how these contractual restrictions on employee mobility affect many important economic dynamics.

References

- Arnow-Richman, Rachel S, “Cubewrap Contracts and Worker Mobility: The Dilution of Employee Bargaining Power Via Standard Form Noncompetes,” *Michigan State Law Review*, 2006, (07–01).
- Arnow-Richman, Rachel, “The New Enforcement Regime: Revisiting the Law of Employee Mobility (and the Scholarship of Charles Sullivan) with 2020 Vision,” *50 Seton Hall L. Rev.* 1223 (2020).
- Autor, David H. and Michael J. Handel, 2013, “Putting Tasks to the Test: Human Capital, Job Tasks, and Wages, *Journal of Labor Economics*, Vol. 31, No. 2, pp. S59-S96.
- Balan, David. “Labor Non-Compete Agreements: Tool for Economic Efficiency, or Means to Extract Value from Workers?” *The Antitrust Bulletin* 2021.
- Balasubramanian, Natarajan, Jin Woo Chang, Mariko Sakakibara, Jagadeesh Sivadasan, and Evan Starr, “Locked in? The enforceability of covenants not to compete and the careers of high-tech workers,” *Journal of Human Resources*, 2020, pp. 1218–9931R1.
- Balasubramanian, Natarajan, Evan Starr, and Shotaro Yamaguchi, “Bundling Employment Restrictions and Value Capture from Employees.” Working paper 2021.
- Barnett, Jonathan M., and Ted Sichelman. “The case for noncompetes.” *University of Chicago Law Review* 87 (2020): 953.
- Beck, Russell “The Changing Landscape of Trade Secrets Laws and Noncompete Laws Around the Country” (2021) Available at <https://faircompetitionlaw.com/changing-landscape-of-trade-secrets-laws-and-noncompete-laws/>
- Bishara, N.D., Martin, K.J. and Thomas, R.S., 2015. An empirical analysis of noncompetition clauses and other restrictive postemployment covenants. *Vand. L. Rev.*, 68, p.1.
- Bishara, Norman and Evan Starr, “The Incomplete Noncompete Picture,” *Lewis and Clark Law Review*, 2016.
- Blake, Harlan M, “Employee Agreements Not to Compete,” *Harvard Law Review*, 1960, pp. 625–691.
- Boesch, Tyler, Katherine Lim, and Ryan Nunn, 2021. “Non-compete contracts sideline low-wage workers.” <https://www.minneapolisfed.org/article/2021/non-compete-contracts-sideline-low-wage-workers>
- Buffkin, R.C., 1999. Non-Competition Clauses in Law Firm Partnership Agreements: How Far Can Partnership Agreements Control Future Conduct of Lawyers. *J. Legal Prof.*, 23, p.325.
- Burdett, Kenneth, and Dale T. Mortensen. "Wage differentials, employer size, and unemployment." *International Economic Review* (1998): 257-273.
- Callahan, M.B., 1985. Post-employment restraint agreements: A reassessment. *The University of Chicago Law Review*, 52(3), pp.703-728.
- Cinelli, Carlos, Andrew Forney, and Judea Pearl. "A crash course in good and bad controls." Available at SSRN 3689437 (2020).
- Colvin, A., and H. Shierholz. “Noncompete agreements: Ubiquitous, harmful to wages and to competition, and part of a growing trend of employers requiring workers to sign away their rights.” *Economic Policy Institute*, 2019.
- Friedman, David, “Non-Competition Agreements: Some alternative Explanations,” Working Paper, 1991.

- Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. "Bartik instruments: What, when, why, and how." *American Economic Review* 110, no. 8 (2020): 2586-2624.
- Greenhouse S (2014) Noncompete clauses increasingly pop up in array of jobs. *New York Times*
- Hall, Robert E., and Alan B. Krueger. 2012. "Evidence on the Incidence of Wage Posting, Wage Bargaining, and On-the-Job Search." *American Economic Journal: Macroeconomics*, 4 (4): 56-67.
- Hausman, Naomi, and Kurt Lavetti. "Physician practice organization and negotiated prices: evidence from state law changes." *American Economic Journal: Applied Economics* 13, no. 2 (2021): 258-96.
- Jamieson, Dave. "Jimmy John's makes low-wage workers sign 'oppressive' noncompete agreements." *Huffington Post*, October 13, 2014.
- Johnson, Matthew S., Kurt Lavetti, and Michael Lipsitz. "The labor market effects of legal restrictions on worker mobility." 2020. Available at SSRN 3455381.
- Johnson, Matthew S., and Michael Lipsitz. "Why are low-wage workers signing noncompete agreements?" *Journal of Human Resources* (2020): 0619-10274R2.
- Kini, Omesh, Ryan Williams, and David Yin. "CEO noncompete agreements, job risk, and compensation." *Review of Financial Studies* (2020): hhaa103.
- Krueger, A.B. and Posner, E.A., 2018. A proposal for protecting low-income workers from monopsony and collusion. *The Hamilton project policy proposal*, 5.
- Lavetti, Kurt, Carol Simon, and William D. White. "The impacts of restricting mobility of skilled service workers evidence from physicians." *Journal of Human Resources* 55, no. 3 (2020): 1025–1067.
- Lipsitz, Michael, and Evan Penniman Starr. "Low-wage workers and the enforceability of non-compete agreements." *Management Science*, forthcoming.
- Lipsitz, Michael and Mark Tremblay. "Noncompete Agreements and the Welfare of Consumers." *Working paper*.
- Madigan, Lisa and Jane Flanagan. "Protecting Competition on Behalf of the People." *Inequality and the Labor Market: The Case for Greater Competition* (2021): 107.
- Marx, Matt. "The firm strikes back: Non-compete agreements and the mobility of technical professionals." *American Sociological Review* 76, no. 5 (2011): 695–712.
- Marx, M., Singh, J. and Fleming, L., 2015. Regional disadvantage? Employee non-compete agreements and brain drain. *Research Policy*, 44(2), pp.394-404.
- Mortensen, Dale T., and Christopher A. Pissarides. "Job creation and job destruction in the theory of unemployment." *The review of economic studies* 61, no. 3 (1994): 397-415.
- Posner, E.A., 2020. The Antitrust Challenge to Covenants Not to Compete in Employment Contracts. *Antitrust Law Journal*, 83(1), pp.165-200.
- Prescott, J.J., and Evan Starr. "Subjective Beliefs about Contract Enforceability." *Working paper*.
- Prescott, J.J., Norman Bishara, and Evan Starr, "Understanding Noncompetition Agreements: The 2014 Noncompete Survey Project," *Michigan State Law Review*, 2016, pp. 369–464.

- Rubin, Paul H and Peter Shedd, "Human Capital and Covenants Not to Compete," *Journal of Legal Studies*, 1981, 10, 93.
- Sanga, Sarath. "Incomplete contracts: An empirical approach." *Journal of Law, Economics, and Organization* 34, no. 4 (2018): 650–679.
- Schwab, Stewart and Evan Starr, 2019. Noncompete Data Collection as part of Cornell National Social Survey.
- Schwab, S.J. and Thomas, R.S., 2006. An empirical analysis of CEO employment contracts: What do top executives bargain for. *Wash. & Lee L. Rev.*, 63, p.231.
- Shi, Liyan. *The Macro Impact of Noncompete Contracts*. Working paper (2020).
- Starr, Evan. "Are noncompetes holding down wages?" In *Inequality and the Labor Market: The Case for Greater Competition*, ed. Ben Harris and Sharon Block. Washington, DC: Brookings Institution Press, 2021.
- Starr, Evan, Natarajan Balasubramanian, and Mariko Sakakibara. "Screening spinouts? How noncompete enforceability affects the creation, growth, and survival of new firms." *Management Science* 64, no. 2 (2018): 552-572.
- Starr, E., Frake, J. and Agarwal, R., 2019. Mobility constraint externalities. *Organization Science*, 30(5), pp.961-980.
- Starr, Evan, "Consider this: Training, wages, and the enforceability of covenants not to compete," *ILR Review*, 2019, 72 (4), 783–817.
- Starr, Evan, J.J. Prescott, and Norman Bishara, "The Behavioural Effects of (Unenforceable) Contracts," *Journal of Law, Economics, and Organization*, 2020.
- Starr, E.P., Prescott, J.J. and Bishara, N.D., 2021. Noncompete agreements in the US labor force. *The Journal of Law and Economics*, 64(1), pp.53-84.
- U.S. Treasury, "Non-compete Contracts: Economic Effects and Policy Implications," 2016.
- The White House, "Non-Compete Agreements: Analysis of the Usage, Potential Issues, and State Responses," 2016.

Figure 1: Incidence of NCAs, By Highest Educational Degree, NLSY97

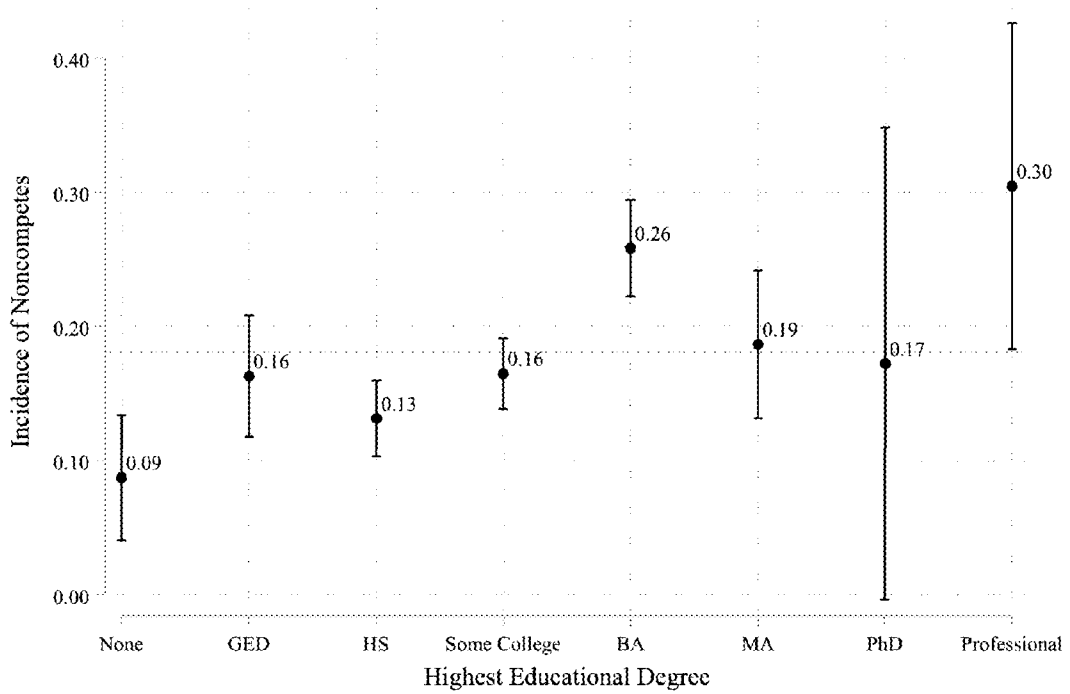


Figure 2: Incidence of NCAs, By Hourly Wage Decile, NLSY97

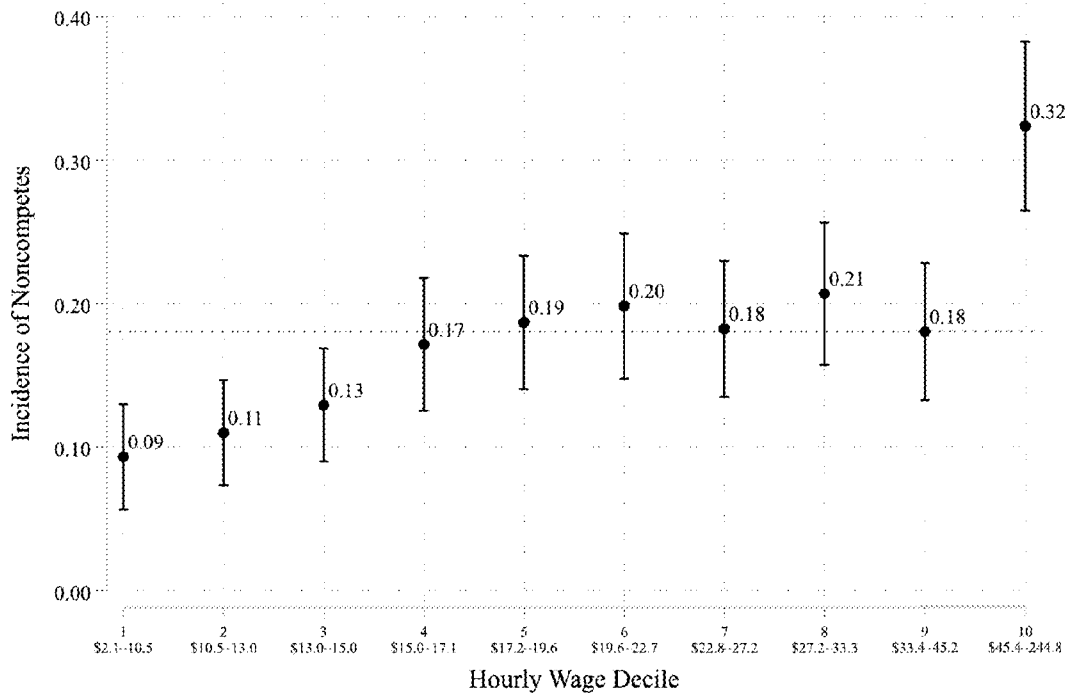
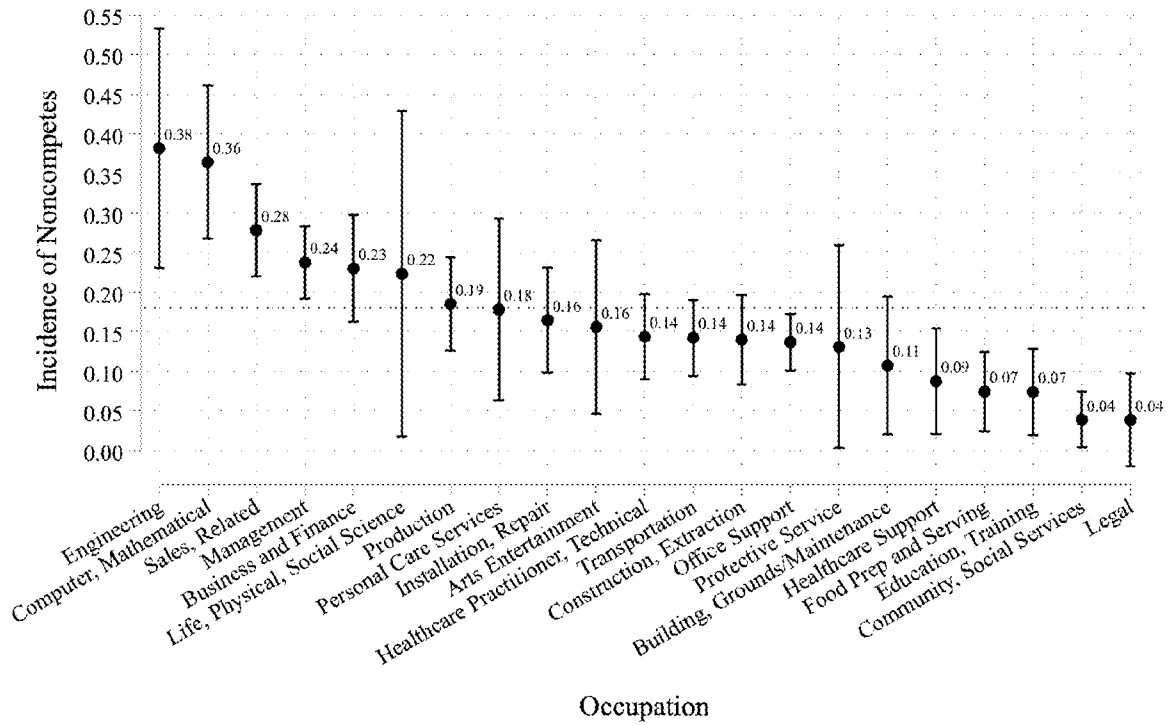
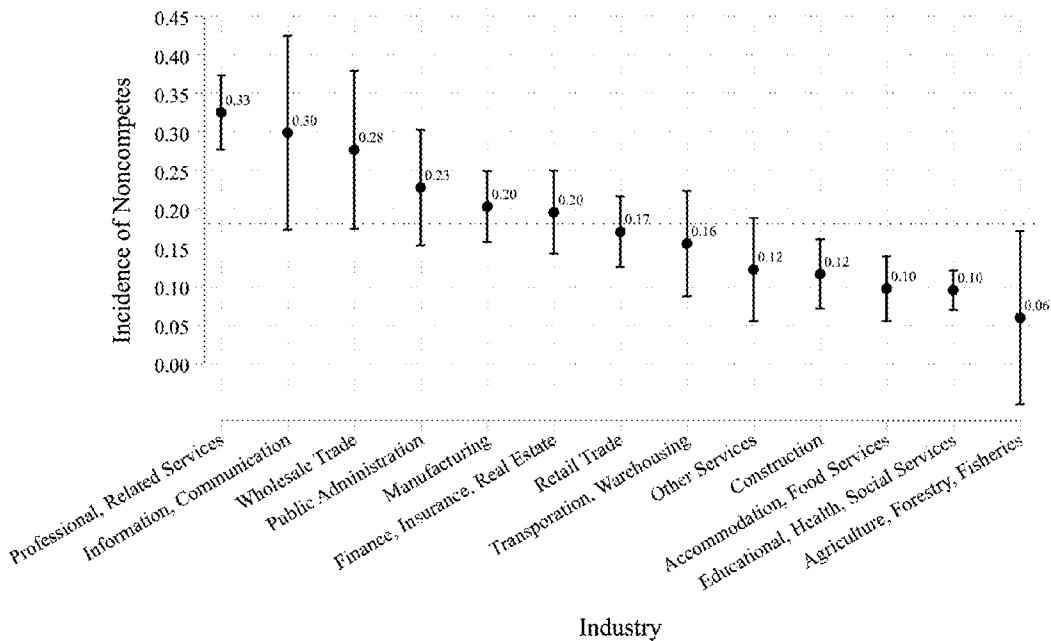


Figure 3: Incidence of NCAs in NLSY97, By Occupation



Only occupations with at least 20 observations are included.

Figure 4: Incidence of NCAs in NLSY97, By Industry



Only industries with at least 20 observations are included.

Figure 5: Incidence of NCAs, By AFQT Percentile Score Decile, NLSY97

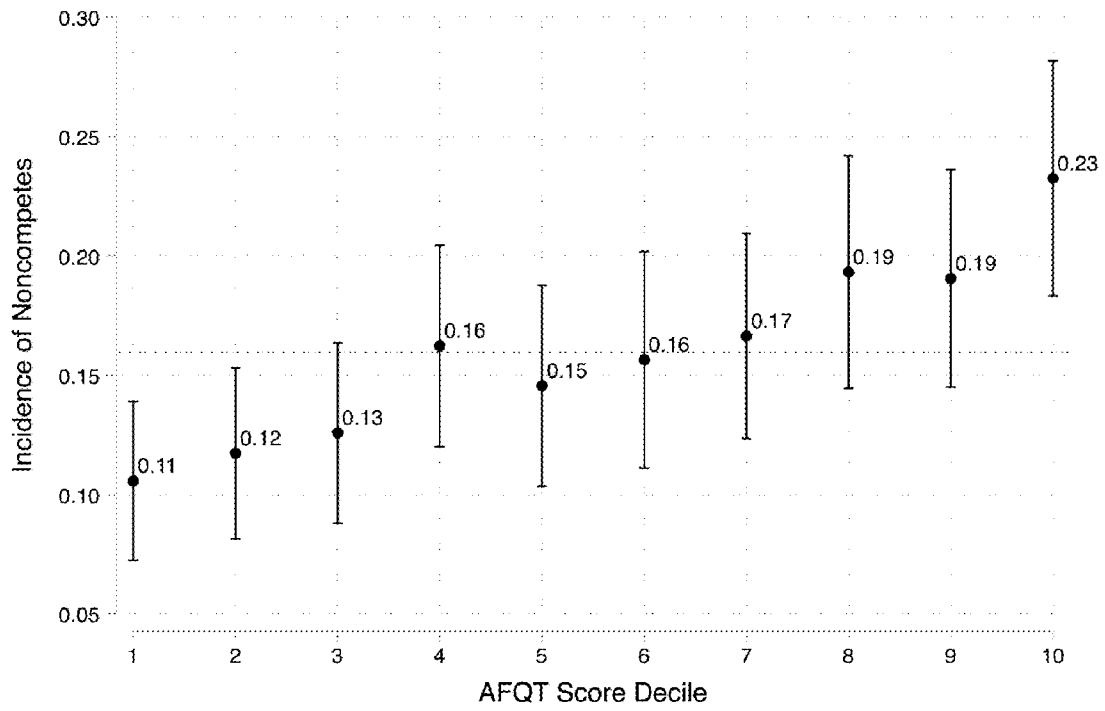


Figure 6: Likelihood of Wage Bargaining by Hourly Wage Decile

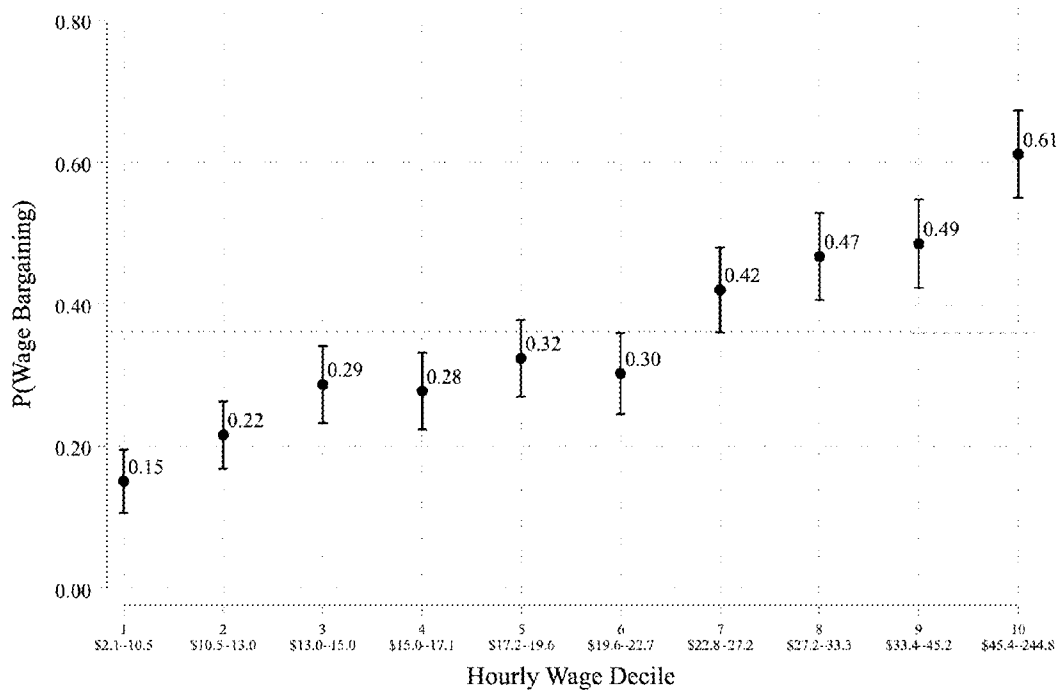


Table 1. The Incidence of NCAs

Data set	(1)	(2)	(3)	(4)	(5)	(6)	
	NLSY97		2014 NSP	2019 CNSS	NLSY97		
Characteristics	NCA Incidence (%)	N	Lower Bound NCA Incidence (%)	NCA Incidence (%)	Characteristics	NCA Incidence (%)	N
<i>Panel A. Worker and firm characteristics, and comparison to other surveys</i>				<i>Panel B. Variables specific to the NLSY97</i>			
Overall	18.07	3090	16.09	19.23	1(Ever employer-provided training)	19.76	523
1(Male)	20.08	1665	16.56	18.95	1(Never employer-provided training)	17.68	2567
1(Female)	15.37	1425	15.50	19.86	1(AFQT percentile score < 50)	15.08	1346
1(Non-black, Non-Hispanic)	19.21	1625	15.01	10	1(AFQT percentile score ≥ 50)	20.91	1177
1(Black, Non-Hispanic)	15.69	777	16.26	11.9	<i>Job Tasks</i>		
1(Hispanic)	14.84	665	---	14.29	1(Repetitive tasks > half workday)	14.86	1511
1(Less than a bachelor's degree)	14.77	2125	11.06	13.85	1(Repetitive tasks < half workday)	21.39	1434
1(Bachelor's degree or higher)	24.27	953	25.15	22.22	1(Physical tasks > half workday)	14.18	1490
1(State enforces NCAs)	18.45	2683	15.62	19.94	1(Physical tasks < half workday)	22.2	1463
1(State does not enforce NCAs)	15.19	407	19.02	11.11	1(Supervise or manage > half wd)	20.53	1037
1(Hourly wage < 20)	14.36	1687	10.81	---	1(Supervise or manage < half wd)	17.12	1913
1(Hourly wage ≥ 20)	21.74	1403	20.34	---	1(Problem solve every day)	23.99	1255
1(Tenure < 3 years)	16.76	1431	10.96	---	1(Problem solve < every day)	13.6	1697
1(Tenure ≥ 3 years)	19.51	1619	18.75	---	1(Read long documents)	23.8	635
1(Private sector)	19.64	2653	17.05	---	1(Not read long documents)	16.51	2315
1(Non-profit sector)	7.41	325	4.78	---	1(A lot of face-to-face contact with non-coworkers)	17.85	1467
1(Union)	16.57	254	20.33	---	1(Not a lot of face-to-face contact with non-coworkers)	18.73	1487
1(No union)	18.59	2499	15.7	---			
1(Employer size < 20)	17.22	747	12.95	---			
1(Employer size ≥ 20 and <	17.84	725	18.89	---			
1(Employer size ≥ 100)	19.58	1168	16.5	---			

2014 NSP stands for 2014 Noncompete Survey Project, described in Prescott et al. (2016) and Starr et al. (2021). Data are limited to workers age 32-38 in 2014; N=1649. Incidence estimates from the NSP are lower bound estimates. 2019 Cornell National Social Survey (CNSS), collected by Stewart Schwab and Evan Starr in 2019 via random digit dial survey. Limited to age 25-50 in 2019; N=338

Table 2. Multivariate Model of Noncomplete Agreement (NCA) Incidence

Model: OLS	(1)	(2)	(3)	(4)
	Dependent Variable: 1(NCA)			
1(At least a Bachelor's Degree)	0.093*** (0.017)	0.085*** (0.021)	0.063*** (0.021)	0.064*** (0.022)
1(Hispanic)	-0.015 (0.019)	-0.015 (0.019)	-0.018 (0.018)	-0.018 (0.018)
1(Black-non-Hispanic)	-0.011 (0.017)	-0.001 (0.016)	0.002 (0.016)	0.001 (0.017)
1(Mixed Race)	-0.054 (0.057)	-0.045 (0.053)	-0.027 (0.052)	-0.057 (0.057)
1(Female)	-0.057*** (0.014)	-0.033** (0.016)	-0.035** (0.017)	-0.013 (0.017)
1(AFQT Percentile Score-25% to less than 50%)	0.011 (0.018)	0.006 (0.020)	-0.005 (0.020)	-0.004 (0.021)
1(AFQT Percentile Score-50% to less than 75%)	0.024 (0.018)	0.013 (0.017)	0.000 (0.018)	-0.002 (0.019)
1(AFQT Percentile Score-75% or higher)	0.019 (0.025)	0.006 (0.025)	-0.012 (0.025)	-0.018 (0.024)
1(State does not enforce NCAs)	-.031* (0.018)	-.035* (0.020)	-.031 (0.020)	-.026 (0.018)
1(Non-profit)	---	-0.144*** (0.019)	-0.156*** (0.021)	-0.091*** (0.025)
1(2nd Quartile Hourly Wage)	---	0.051** (0.020)	0.042** (0.021)	0.043** (0.020)
1(3rd Quartile Hourly Wage)	---	0.040** (0.018)	0.018 (0.017)	0.011 (0.018)
1(4th Quartile Hourly Wage)	---	0.073*** (0.018)	0.032* (0.018)	0.016 (0.021)
1(Employment 21-100)	---	0.003 (0.022)	0.001 (0.023)	0.014 (0.022)
1(Employment >100)	---	0.002 (0.019)	0.001 (0.020)	0.003 (0.021)
1(Employer ever trained worker)	---	-0.003 (0.021)	-0.013 (0.022)	-0.017 (0.022)
1(Unionized)	---	-0.005 (0.032)	0.012 (0.032)	0.026 (0.034)
1(Tenure \geq 3 Years)	---	0.007 (0.012)	0.010 (0.012)	0.010 (0.011)

1(Frequency with which contact with others is 'A Lot')	---	---	0.023 (0.019)	0.044** (0.020)
1(Longest document read at work is at least 11 pages)	---	---	0.024 (0.019)	0.045** (0.019)
1(Use math to solve problems at least once a day)	---	---	-0.009 (0.024)	-0.015 (0.022)
1(Solve problems at least once a day)	---	---	0.073*** (0.017)	0.063*** (0.017)
1(Supervise/manage others more than half the time)	---	---	0.014 (0.019)	0.014 (0.019)
1(More than half of tasks are physical)	---	---	-0.045*** (0.017)	-0.033* (0.019)
1(Short and repetitive tasks more than half the time)	---	---	-0.012 (0.016)	-0.010 (0.017)
Occupation and Industry FE	No	No	No	Yes
Observations	3,090	3,090	3,090	3,090
R-squared	0.022	0.040	0.054	0.099

Standard errors clustered by state of residence in parentheses. *** p<0.01, ** p<0.05, * p<0.10.

Regressions are weighted with round 18 survey weights. If the variable of interest is missing for some values, an indicator is included (but not reported) which equals 1 if the variable is missing. Results are available from the authors.

Table 3. NCAs, Bargaining, and Wages

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: Baseline Bargaining and Wages</i>						
Dependent variable	1(Bargain over wages)			Log(Hourly Wage)		
NCA	0.095*** (0.027)	0.069** (0.027)	0.021 (0.026)	0.221*** (0.024)	0.120*** (0.020)	0.049** (0.018)
Observations	3,090	3,090	3,090	3,090	3,090	3,090
R-squared	0.006	0.040	0.129	0.022	0.346	0.526
Controls	None	Basic	Advanced	None	Basic	Advanced
<i>Panel B: Wages as a function of bargaining</i>						
Dependent Variable:	Log(Hourly Wage)					
NCA	0.192*** (0.018)	0.111*** (0.015)	0.047*** (0.017)	0.155*** (0.028)	0.074** (0.030)	0.018 (0.022)
1(Bargain over wages)	0.287*** (0.024)	0.175*** (0.017)	0.101*** (0.019)	0.271*** (0.026)	0.171*** (0.020)	0.087*** (0.018)
NCA*1(Bargain over wages)	---	---	---	0.091* (.052)	0.079* (.047)	0.070* (.036)
% of NCA-Wage Differential Explained by Bargaining	13.1%	7.5%	4.1%	---	---	---
Observations	3,090	3,090	3,090	3,090	3,090	3,090
R-squared	0.081	0.379	0.532	0.082	0.370	0.532
Controls	None	Basic	Advanced	None	Basic	Advanced

Notes: Basic controls include three education categories, indicators for race and ethnicity, AFQT score at 50th percentile or more, gender, and an indicator for whether the State of residence does not enforce NCAs. Advanced controls add an indicator for for-profit or non-profit status, occupation and industry fixed effects (2 digit SOC and NAICS), and indicators for job tasks including indicators for repetitive work, frequency of contact with others, the length of the longest document read on the job, solving problems, using math to solve problems, supervising others, and the extent of physical tasks. If the variable of interest is missing for some values, an indicator is included (but not reported) which equals 1 if the variable is missing. Results are available from the authors. Standard errors clustered by state of residence in parentheses. *** p<0.01, ** p<0.05, * p<0.10. Regressions are weighted with round 18 survey weights. The “% of NCA-Wage Differential Explained by Bargaining” row takes the NCA coefficients from Columns (1)-(6) of Panel B and divides them by the corresponding NCA coefficient in columns (4)-(6) in Panel A.

Table 4. Heterogeneous Wage Effects

	(1)	(2)	(3)	(4)
<i>Panel A: Education</i>				
NCA	0.088*** (0.022)	0.024 (0.019)	0.076*** (0.023)	0.022 (0.019)
1(Bachelor's Degree)	0.446*** (0.029)	0.278*** (0.028)	0.414*** (0.036)	0.252*** (0.031)
1(>Bachelor's Degree)	0.703*** (0.041)	0.474*** (0.036)	0.706*** (0.053)	0.482*** (0.053)
NCA*1(Bachelor's Degree)	0.008 (0.051)	0.011 (0.048)	0.010 (0.050)	0.012 (0.049)
NCA*1(>Bachelor's Degree)	0.227** (0.085)	0.175** (0.078)	0.224* (0.085)	0.176** (0.078)
R-squared	0.349	0.527	0.372	0.534
<i>Panel B: Race and Ethnicity</i>				
NCA	0.134*** (0.027)	0.053** (0.025)	0.116*** (0.026)	0.049* (0.024)
1(Black or Hispanic)	-0.091*** (0.020)	-0.063*** (0.021)	-0.078*** (0.025)	-0.052** (0.025)
NCA*1(Black or Hispanic)	-0.058 (0.040)	-0.017 (0.037)	-0.043 (0.042)	-0.012 (0.038)
R-squared	0.344	0.525	0.368	0.532
Observations	3,090	3,090	3,090	3,090
Controls	Basic	Advanced	Basic	Advanced
Bargaining Indicator	No	No	Yes	Yes
Bargaining*Group Indicator(s)	No	No	Yes	Yes

Notes: The dependent variable is log hourly wage. Basic controls include three education categories, indicators for race and ethnicity, AFQT score at 50th percentile or more, gender, and an indicator for whether the State of residence does not enforce NCAs. Advanced controls add an indicator for for-profit or non-profit status, occupation and industry fixed effects (2 digit SOC and NAICS), and indicators for job tasks including indicators for repetitive work, frequency of contact with others, the length of the longest document read on the job, solving problems, using math to solve problems, supervising others, and the extent of physical tasks. If the variable of interest is missing for some values, an indicator is included (but not reported) which equals 1 if the variable is missing. Results are available from the authors. Standard errors clustered by state in parentheses. *** p<0.01, ** p<0.05, * p<0.10. Regressions are weighted with round 18 survey weights.

Table 4. Heterogeneous Wage Effects (Continued)

	(1)	(2)	(3)	(4)
<i>Panel C: Gender</i>				
NCA	0.158*** (0.026)	0.067*** (0.025)	0.144*** (0.028)	0.066*** (0.025)
1(Female)	-0.159*** (0.025)	-0.116*** (0.028)	-0.150*** (0.033)	-0.117*** (0.030)
NCA*1(Female)	-0.102** (0.047)	-0.049 (0.044)	-0.099** (0.049)	-0.051 (0.046)
R-squared	0.347	0.526	0.371	0.532
<i>Panel D: AFQT</i>				
NCA	0.082*** (0.020)	-0.007 (0.019)	0.076*** (0.020)	-0.007 (0.019)
1(AFQT Score 50%+)	0.139*** (0.023)	0.036 (0.024)	0.122*** (0.024)	0.029 (0.029)
NCA*1(AFQT Score 50%+)	0.092** (0.040)	0.113*** (0.037)	0.079* (0.041)	0.108*** (0.038)
R-squared	0.359	0.527	0.380	0.533
<i>Panel E: State NCA Enforceability</i>				
NCA	0.117*** (0.019)	0.042** (0.020)	0.106*** (0.016)	0.040** (0.018)
1(State not Enforce NCAs)	0.134** (0.066)	0.131 (0.081)	0.130** (0.055)	0.130* (0.070)
NCA*1(State not Enforce NCAs)	0.040 (0.024)	0.065** (0.031)	0.043** (0.020)	0.065** (0.030)
R-squared	0.358	0.526	0.380	0.532
Observations	3,090	3,090	3,090	3,090
Controls	Basic	Advanced	Basic	Advanced
Bargaining Indicator	No	No	Yes	Yes
Bargaining*Group Indicator(s)	No	No	Yes	Yes

Notes: The dependent variable is log hourly wage. Basic controls include three education categories, indicators for race and ethnicity, AFQT score at 50th percentile or more, gender, and an indicator for whether the State of residence does not enforce NCAs. Advanced controls add an indicator for for-profit or non-profit status, occupation and industry fixed effects (2 digit SOC and NAICS), and indicators for job tasks including indicators for repetitive work, frequency of contact with others, the length of the longest document read on the job, solving problems, using math to solve problems, supervising others, and the extent of physical tasks. If the variable of interest is missing for some values, an indicator is included (but not reported) which equals 1 if the variable is missing. Results are available from the authors. Standard errors clustered by state in parentheses. *** p<0.01, ** p<0.05, * p<0.10. Regressions are weighted with round 18 survey weights.



Working Papers
RESEARCH DEPARTMENT

WP 21-26

July 2021

<https://doi.org/10.21799/frbp.wp.2021.26>

Do Non-Compete Covenants Influence State Startup Activity? Evidence from the Michigan Experiment

Gerald A. Carlino

Emeritus Economist, Federal Reserve Bank of Philadelphia Research Department

ISSN: 1962-5361

Disclaimer: This Philadelphia Fed working paper represents preliminary research that is being circulated for discussion purposes. The views expressed in these papers are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Any errors or omissions are the responsibility of the authors. Philadelphia Fed working papers are free to download at: <https://philadelphiafed.org/research-and-data/publications/working-papers>.

“DO NON-COMPETE COVENANTS INFLUENCE STATE STARTUP ACTIVITY?
EVIDENCE FROM THE MICHIGAN EXPERIMENT”

Gerald A. Carlino*

Federal Reserve Bank of Philadelphia

July 2021

Abstract

This paper examines how the enforceability of employee non-compete agreements affects the entry of new establishments and jobs created by these new firms. We use a panel of startup activity for the U.S. states for the period 1977 to 2013. We exploit Michigan’s inadvertent policy reversal in 1985 that transformed the state from a non-enforcing to an enforcing state as a quasi-natural experiment to estimate the causal effect of enforcement on startup activity. In a difference-in-difference framework, we find little support for the widely held view that enforcement of non-compete agreements negatively affects the entry rate of new firms or the rate of jobs created by new firms. We find that increased enforcement had no effect on the entry rate of startups, but a positive effect on jobs created by these startups in Michigan relative to a counterfactual of states that did not enforce such covenants pre- and post-treatment. Specifically, we find that a doubling of enforcement led to an increase of about 8 percent in the startup job creation rate in Michigan. We also find evidence that enforcing non-competes positively affected the number of high-tech establishments and the level of high-tech employment in Michigan. Extending our analysis to consider the effect of increased enforcement on patent activity, we find that enforcement had differential effects across technological classifications. Importantly, increased enforcement had a positive and significant effect on the number of Mechanical patents in Michigan, the most important patenting classification in that state.

Keywords: Startup activity, Non-compete agreements, Regional economic growth.

JEL Codes: O30, O38, R11

*I thank Thorsten Drautzburg, Fernando Ferreira, Lester Lusher, Enrico Moretti, and participants of the East-West Center/Korea Development Institute’s conference on “Job Creation Strategy in a Knowledge Economy,” May 12-14, 2021, for comments and helpful suggestions. This paper has benefited from outstanding research assistance by Adam Scavette. The views expressed here are those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Philadelphia Fed working papers are free to download at <https://philadelphiafed.org/research-and-data/publications/working-papers>.

1. INTRODUCTION

Business startups play an important role in job creation. For example, on average, startups created almost 4 million jobs over the past four decades in the U.S. economy. As Haltiwanger, Jarmin, and Miranda (2013) show, while many of these startups will fail within a few years, a small percentage of fast growers will ultimately contribute disproportionately to job creation in the U.S. Recent examples of fast-growing startups include Google, Amazon, and Microsoft. One channel for growth of startup activity and entrepreneurship is through employees' leaving their current employers to form new establishments. Concerned about competition with former employees, many employers require their employees to sign non-compete covenants. In contract law, a post-employment non-compete covenant is a clause whereby one party (typically an employee) agrees not to start or join a similar business that would be in competition with another party (usually the employer). Typically, non-competes restrict an employee's job mobility for a limited time and within a narrowly defined geographic region.

An important consideration is that non-competes may hinder knowledge flows among workers and firms in states that enforce such agreements. Marx, et al. (2015) find that employee non-compete agreements are responsible for the migration of knowledge workers from states enforcing these contracts to non-enforcing states. Further, workers covered by such agreements may feel constrained about sharing information with outsiders, further limiting an important source of knowledge spillovers.

Despite the fact that non-competes represent a restraint on trade and may limit knowledge spillovers, such agreements are common for many types of workers in the U.S. (Stone, 2002, and

Starr, et al., 2020).¹ There are no federal laws governing the enforceability of non-competes; enforcement is left to the states, and states differ in the manner and the extent of non-compete enforcement. The courts in many U.S. states tend to enforce employee non-compete agreements because they recognize them as a way to safeguard the legitimate business interest of firms. An important issue is whether and to what extent does judicial enforcement of non-compete clauses impede entrepreneurial activity and employment growth.

The impact of enforceability on entry and employment of new firms is theoretically ambiguous. The literature has identified two channels in which the enforcement of non-compete agreements could affect startup activity. Starr, et al. (2015) identify a negative channel, referred to as a “screening effect,” in which greater enforcement lowers the expected returns to spinoff activity by raising the probability of losing a lawsuit for violating the terms of a non-compete agreement.² Kang and Fleming (2020) point out that startups could avoid states with stronger non-compete laws, since the entrepreneurs typically lack the resources necessary to educate and train employees and may prefer to hire experienced employees from nearby competitors.

On the other hand, to the extent that non-compete clauses help companies protect their investments, this protection may stimulate startup activity and employment growth (Starr, et al., 2015, refer to this channel as an “investment-protection effect”).³ Kang and Fleming (2020) point out that startups tend to be small, having few assets other than their ideas and intellectual

¹ Starr et al. (2020) report that 18 percent of all U.S. workers are covered by non-compete agreements and that 37 percent say they have been covered by such an agreement during their career.

² These costs would include any payments an employee (or a third party) makes to his parent firm to be released from a non-compete agreement.

³ The higher expected profits associated with the investment channel will be reduced if firms have to pay a wage premium to entice potential workers to move to enforcing states.

property, and may be attracted to states with relatively stronger enforcement that may deter employees from departing.⁴

The overall effect of non-compete covenants on startup activity is an open question given the competing forces of the screening effect and the investment-protection effect.⁵

The purpose of this paper is to provide evidence on the effect of judicial enforcement of non-compete covenants on the rate of entry of startups and the job creation rate of new firms. In the main analysis, we use a panel of startup activity in U.S. states for the period 1977 to 2013 and exploit the Michigan Antitrust Reform Act (MARA) of 1985 (which inadvertently “legalized” non-compete agreements) as a quasi-natural experiment to estimate the causal effect of enforcement on startup activity. To evaluate whether the observed changes in startup activity is being driven by a response to changed enforcement policy, we need to identify a comparison state or states that trace the counterfactual path of startup trends for Michigan. The quality of our analysis is obviously tied to how well we estimate the comparison group. There are a number of strategies for constructing a comparison group, and we start by identifying three alternative control groups of states (states sharing a land border with Michigan; states sharing a land border or a water boundary with Michigan; and the 10 “non-compete” states identified by Marx, et al., 2009). In a difference-in-difference (hereafter, DID) analysis, we find that enforcement had a positive and, in some cases, significant effect on the startup job creation rate but little or no

⁴ In general, Meccheri (2009) shows that non-compete covenants can be justified on efficiency ground as they attempt to solve a “hold-up” problem. *Ex ante*, both the employee and the employer benefit from worker training and the sharing of trade secrets. But *ex post*, the employee has an incentive to “hold up” his employer for additional compensation by threatening to divulge confidential information. Forward-looking employers would be unwilling to invest (or would under-invest) in education and training and be less willing to share trade secrets with employees unless they had some form of legal recourse provided by non-compete agreements.

⁵ Also, strict enforcement may limit agglomeration economies, by limiting knowledge spillovers, and the benefits associated with labor market matching and pooling. An analysis of the effect of enforcement on agglomeration economies is beyond the scope of this paper.

effect on the entry rate of new firms. Specifically, depending on the control group, a doubling of enforcement led to a 6 percent to 8 percent increase in the startup job creation rate in Michigan and to essentially no change in the startup entry rate.

A crucial assumption underlying the DID strategy is that the outcome in treatment and control groups would follow the same time trend or a parallel trend in the absence of the treatment. A standard DID analysis would result in biased estimates if the treatment and control groups did not meet the parallel-trends assumption. This leads us to consider a fourth alternative control group, identified using a data-driven search routine: the Synthetic Control Method (SCM). The basic idea underlying the SCM is that often a combination of states produces a better control group than any single state or arbitrary group of states, such as states bordering Michigan. When using the SCM method, we find that changes in both the startup entry rate and the job creation rate, while positive, following MARA, are not significantly different from pre-MARA findings. Taken together, these findings offer little support for the view that enforcement of non-compete agreements negatively affects the entry rate of new firms or the rate of jobs created by new firms.

We extend our analysis to consider the effect of increased enforcement on high-tech establishments and to employment by these establishments.⁶ A common view since the work of Saxenian (1994) is that employee non-compete covenants may serve to suppress high-tech development in states such as Massachusetts, where historically, non-competes tend to be enforced by the courts, but did not hold back the tech boom in states such as California, where such covenants are much less likely to be enforced. As with startup activity, we found some evidence that enforcement of non-competes in Michigan after 1985 positively affected both high-

⁶ We thank Enrico Moretti for suggesting the extension of the analysis to high-tech activity.

tech establishments and employment at these establishments, but only for one of the four control groups considered.

Finally, patent data can be used to study entrepreneurial activity and are available for total patent activity and for six technology classifications. We find that enforcement had a significant positive effect on the total number of patents issued to Michigan inventors, and in the Mechanical category and in the other classifications. We find the number of patents issued in the drug classifications is negatively and significantly affected by enforcement. The patent findings are important in that they demonstrate the importance of considering subcategories, something we are not able to do with the publicly available startup data.

Previous studies have found mixed evidence regarding the importance of non-compete clauses on worker and inventor mobility. Most state courts enforce some form of non-compete clauses, with California being an important exception. Thus, worker mobility, or “job hopping,” could be unusually high in California because of the unenforceability of non-compete clauses under California law (Gibson, 1999). Fallick, et al. (2006) find much greater mobility of college-educated males employed in the computer industry in Silicon Valley compared with the inter-firm mobility of similarly educated workers in the computer industry in other areas outside of California. Part of this turnover could be induced as firms and workers seek better matches. It is important to note that Fallick, et al. (2006) find that employee turnover in other industries is no higher in California than in other locations, suggesting that a lack of enforcement of non-compete clauses is not the primary reason for the job-hopping observed in California. Still, a number of other studies offer evidence that tends to suggest the enforcement of non-competes limits worker mobility. Balasubramanian, et al. (2019) find that increased enforcement is positively associated with longer job tenure in high-tech industries, without increased wages.

Garmaise (2011) finds that stronger state enforcement tends to reduce mobility of U.S. executives and lowers their compensation. Marx, et al. (2009) find that the 1985 policy reversal that transformed Michigan from a non-enforcing state to an enforcing state resulted in an 8 percent decrease in within-state mobility of inventors, and Marx, et al. (2015) find that Michigan's policy reversal not only restricted within-state mobility but also led to increased inter-state mobility of inventors (a "brain drain"). Bozkaya and Kerr (2014) more broadly show how rigid employment law can hinder the development of innovative sectors that rely on rapid labor turnover. Samila and Sorenson (2011) find that local supply of venture capital in states that limit the scope of non-compete agreements positively influences innovative activity, firm entry, and job creation. Conti (2014) finds that firms are more likely to undertake riskier, potentially path-breaking, R&D projects in states that tend to enforce non-competes than in states with less-restrictive enforcement policies. Starr, et al. (2015) find that greater enforceability is associated with fewer within-industry spinoffs compared with cross-industry spinoffs, providing evidence for the screening channel.

In a study more closely related to ours, Kang and Fleming (2020) use Florida's 1996 legislative change that eased restrictions on their enforcement to study the effects of increased enforcement both on large firms and on small firms. The researchers find that following the enforcement change, large firms were more likely to establish businesses in Florida, while small firms were not. It's important to note that we look at startup activity as opposed to focusing on small firms, as do Kang and Fleming (2020).

Our study also differs from past research in that we focus on how non-competes affect startup activity rather than how they affect worker mobility or investment activity. While knowledge of how non-competes limit mobility and investment is useful, this research does not inform us

about the effects of non-competes on firm entry or about the employment created by new firms in states enforcing such agreements.

2. EMPIRICAL METHODOLOGY AND DATA

Michigan had a long history of prohibiting the enforcement of non-compete agreements. Section 1 of Act No. 329 of the Public Acts of 1905 prohibited the enforcement of non-compete covenants. The act states:

“All agreements and contracts by which any person, copartnership or corporation promises or agrees not to engage in any avocation, employment, pursuit, trade, profession or business, whether reasonable or unreasonable, partial or general, limited or unlimited, are hereby declared to be against public policy and illegal and void.”

This act governed the enforcement of non-compete clauses until March 1985, when Michigan’s legislature inadvertently eliminated the statute when it passed MARA. While the main purpose of MARA was to consolidate Michigan’s antitrust statutes, in doing so, the legislature unintentionally repealed numerous statutes, including Public Act No. 329. According to Marx, et al., (2009): “More than 20 pages of legislative analysis of MARA by both House and Senate subcommittees does not mention non-competes as a motivation for the bill.” A number of researchers conclude that the repeal of Public Act 329 was unintentional given that antitrust reform was the main motivation for MARA. Marx, et al. (2009) persuasively argue that changes in Michigan’s enforcement policy can be viewed as an exogenous event allowing one to test for causal influence of non-competes on startup activity.

In December 1987, the Michigan Legislature reversed course and passed MARA Section 4a adopting a “reasonableness standard” in that non-compete agreements can be enforceable to the extent that they are “reasonable as to its duration, geographic area, and the type of employment

or line of business.” Moreover, if the non-compete clauses in the agreement are “found to be unreasonable in any respect, a court may limit the agreement to render it reasonable.”⁷

2.1 Empirical Methodology

In the research described below, Michigan’s seeming unintended reversal of its non-compete enforcement policy is used as a quasi-natural experiment in a DID analysis. If the judicial enforcement of non-compete agreements initiated by the passage of MARA (the treatment) had a measurable effect on startup activity (the outcome) in Michigan, we expect to observe differences between startup activity pre- and post-treatment compared with a control group of other states. The DID estimation can be expressed in regression form as:

$$Y_{st} = \beta_0 + \beta_1 T_s + \beta_2 P_t + \beta_3 P_t * T_s + \mu_{st},$$

where Y_{st} represents the outcome of interest for state s in time period t . T_s is a dummy variable equal to one for Michigan observations, the treatment state, and zero otherwise. P_t represent a dummy variable equal to unity beginning in 1985 and zero otherwise. The interaction term $P_t * T_s$ is essentially an indicator variable equal to unity for Michigan observations post-treatment. $\hat{\beta}_3$ is an estimate of the average differential change in Y from the pre-treatment period 1977 to 1984 to the post-treatment period 1985 to 2013 for Michigan relative to the control group. Under the assumption that the treatment is randomly assigned, $E(\mu_{s,t} | Y_{s,t}) = 0$, the OLS estimator of β_3 is unbiased.

⁷ Reasonable covenants also may protect trade secrets, confidential information, and employers’ customers or customer lists.

One concern is that other state characteristics, such as state income growth and state population growth, may be important for determining the outcome of the experiment. Including these additional covariates helps to ensure that there is no omitted variable bias:

$$Y_{st} = \beta_0 + \beta_1 T'_s + \beta_2 P_t + \beta_3 P_t * T'_s + \gamma_j \sum_{j=1}^J X_{jst} + \mu_{st}$$

For the additional covariates, denoted j , we included state level values for: the nine one-digit SIC industry share of total state employment; the percentage of a state's population with a college degree; the percentage of a state's population aged 15 to 64 years old; the state's unemployment rate; the state's labor force participation rate; real per capita state income growth; and state population growth. We include year fixed effects to control for common aggregate sources of variation in startup activity. The variables are in logs, with the exception of real per capita income growth and population growth, which are in levels. Bertrand, et al. (2004) demonstrate the importance of using cluster-robust standard errors in a DID framework. We adopt this approach and cluster the standard errors at the level of treatment, which is the state.

As already noted, the sign on β_3 is uncertain. To the extent that the enforcement of non-compete agreements impedes entrepreneurial activity, there could have been a decrease in entry and job creation in Michigan compared with the control group following the passage of MARA. Alternatively, the opposite (a positive) effect on startup activity is anticipated if enforcement of non-compete clauses leads companies to invest more in Michigan. How these opposing forces net-out is an empirical issue central to the analysis considered in this paper.

2.2 Data

In this section, we will discuss the data and sources for startup activity. We will discuss the data for high-tech activity and patents in subsequent sections. We use annual state-level data from the U.S. Census Bureau's Business Dynamics Statistics (BDS) for the period 1977–2013 on the entry of new establishments and the number of private sector jobs created by these establishments. The BDS consists of longitudinal data covering all private non-farm U.S. establishments and firms. For the 50 U.S. states, this gives a panel consisting of 1,850 observations. The BDS provides annual measures of business dynamics (such as the number of startups, firm closures, and job creation and destruction) for states, aggregated by establishment and firm characteristics. We are limited to looking at aggregate state-level startup activity because a state-industry-level breakdown of the data is not publicly available. The outcome variables used in our analysis are defined as:

$$\text{Establishment Entry Rate}_{s,t}^0 = \frac{\text{New Establishments}_{s,t}^0}{1/2(\text{No. of Estabs}_{s,t} + \text{No. of Estabs}_{s,t-1})} \quad (1)$$

$$\text{Job Creation Rate}_{s,t}^0 = \frac{\text{Job Created}_{s,t}^0}{1/2(\text{Employment}_{s,t} + \text{Employment}_{s,t-1})}, \quad (2)$$

where the Establishment Entry Rate $_{s,t}^0$ refers to the number of startups in state s in time t by age zero establishments relative to the total number of establishments in state s . Similarly, the Job Creation Rate $_{s,t}^0$ refers to the number of jobs created by startups relative to total employment in the state. Following Haltiwanger, et al. (2013), we define rates relative to a denominator that averages employment of the number of firms in the current and previous year. We supplement these data with additional covariates predictive of startup activity, such as

economic and demographic characteristics of states. The share of a state's employment by the one-digit industry, state unemployment rates, and state labor-force participation rates are obtained from the Bureau of Labor Statistics. Data for state population, state population share aged 15-64, and the share of state population with a college degree are obtained from the Census Bureau. State-level GDP is obtained from the Bureau of Economic Analysis Regional Economic Accounts.

3. IDENTIFYING THE CONTROL GROUPS

To evaluate whether startup activity is responsive to changed enforcement policy, we need to identify a comparison state or states that trace the counterfactual path of the outcome variables of interest. There are a number of strategies for constructing a comparison group, all of which have merit, but also concerns.

3.1 States Sharing a Land Border with Michigan

One control group used in the analysis consists of U.S. states that border Michigan, as these states may have similar economic, demographic, and social characteristics (Indiana, Ohio, and Wisconsin).⁸ It is likely that states sharing a border with Michigan share many unobservable characteristics, helping to ensure the parallel trend requirement for this control. We refer to this group of states as Border States.

⁸ Michigan borders the Canadian province of Ontario, but Ontario is excluded from the control group to maintain consistency of the included data.

3.2 States Sharing a Land or Water Boundary with Michigan

Two states that share a water boundary with Michigan are added to give us five states in this control group (Illinois, Indiana, Ohio, Minnesota, and Wisconsin). For brevity, we refer to this control group as Expanded Border States.

3.3 Non-Enforcing States

Based on Table 1 in Stuart and Sorenson (2003), Marx, et al. (2009) and Marx, et al. (2015) identify 10 states with statutes that claimed to limit the enforcement of non-compete agreements both pre- and post-MARA. These 10 states (Alaska, California, Connecticut, Minnesota, Montana, Nevada, North Dakota, Oklahoma, Washington, and West Virginia) constitute the control group used by Marx, et al. (2015) and Marx, et al. (2009) to study the migration of knowledge workers from Michigan to other states. Given Michigan's switch from a non-enforcing state pre-MARA to an enforcing state post-MARA, other states that did not enforce such agreements both pre- and post-MARA constitute an appropriate comparison group. We use these 10 states, referred to as Non-Enforcement States, as the main control groups in the analysis to follow. The dynamic consistency in non-compete legislation is a main reason we prefer using the Non-Enforcement States to chart the counterfactual path for the variables of interest.

3.4 Parallel Trends

Another condition for a good control group is that the group should display similar or parallel trends during the pre-treatment period compared with the treatment state (Michigan in our case). Figure 1a illustrates the trends in the job creation rate for the states sharing only a land border with Michigan, while Figure 1b shows the trends for the entry rate. As the figures show, each state in this control group closely tracks movements in Michigan both for the job creation rate

and for the entry of startups during the pre-MARA period. The levels for these variables for Michigan and the states in this control group are similar, too, strengthening the assumption that these untreated states provide an appropriate “counterfactual Michigan.”

Figure 1c shows the trends in the job creation rate for the states in the potential control group called Expanded Border States, while Figure 1d shows the trends in the job creation rate for the 10 states that make up the potential control group called Non-Enforcement States. As the Figure 1c and Figure 1d reveal, the trends in job creation rate in the pre-treatment period for either of these potential control groups do not as closely parallel those for Michigan during this period as do the trends for the Border States. Still, the Non-Enforcement States is the preferred comparison group given that these states did not enforce non-competes pre- and post-MARA. For consistency, we will conduct a DID analysis using these three alternative control groups, both for completeness and for comparison to other studies, such as Marx, et al. (2015).^{9, 10}

Before proceeding to the formal analysis, it is important to consider what happened to startup activity in Michigan just after the 1985 legislation. There is little evidence of a sharp break in Michigan’s trend either in the job creation rate (Figure 1a) or in the entry rate (Figure 1b) immediately after 1985. This lack of an immediate effect on startup activity is most likely due to a provision in the legislation that non-competes in effect at the time of repeal remained unenforceable. Because of this, the number of employees in Michigan who were actually subject to enforcement was relatively small for a significant period of time following the passage of MARA. In December 1987, the reasonableness standard was made retroactive to the 1985

⁹ In Section 5, we use the SCM as an alternative approach for selecting a control group.

¹⁰ To conserve space, figures for the entry rate are not presented for either the Expanded Border States or the Non-Enforcement States. The trends for the entry rate are similar to those shown for the job creation rate. These additional figures are available upon request.

passage of MARA. Given these legislative provisions, it seems reasonable to expect that startup activity post-MARA would initially continue to closely track that of the other control groups, but eventually start to diverge from counterfactual Michigan. However, Figure 1a provides little evidence that startup activity in Michigan diverged relative to the dynamic path of the other control groups. This visual inspection of the data suggests that MARA had little or no effect on startup activity in Michigan, relative to the control groups.

We turn to the DID analysis to see if it reveals a similar lack of an enforcement effect on startup activity.

FINDINGS

Our analysis is at the state level, since non-compete legislation is determined at the state level, and as such, non-competes' occurrence and enforcement will vary across states. As indicated, a number of different groupings of the states are used to construct the various comparison groups. The alternative control groups are used in a DID framework during the period 1977 to 2013 to estimate the causal effects of enforcement on startup activity. The null hypothesis we test is:

H₀ : The Passage of MARA did not affect startup activity in Michigan relative to the control group

H_A : The passage of MARA did have an effect on startup activity in Michigan relative to the control group

Table 1a presents summary statistics for Michigan, and Table 1b presents these statistics for states other than Michigan. The panel on the left side of the tables shows the summary statistics for the pre-MARA period, while the panel on the right shows these statistics for the post-MARA period. The tables show that startup activity has been declining over time. In Michigan, the mean

job creation rate fell from just over 3 percent pre-MARA to about 2.5 percent post-MARA. In the nation, the average job creation rate fell from 4.4 percent pre-MARA to just under 3 percent post-MARA. As the tables show, the entry rate also fell post-MARA relative to the pre-MARA period in both Michigan and the nation.

Table 2 summarizes the findings of the DID analysis. The first two columns of Table 2 present the findings for job creation and the startup entry rate, respectively, when the control group consists of the states sharing a land border with Michigan. The next two columns in Table 2 give the findings relative to a counterfactual based on the Expanded Border States, while the final two columns show the results for the Non-Enforcement States.

The results of interest are given by the interaction of the Michigan dummy variable and a dummy variable for the post-MARA period (shown in the third row of Table 2). Recall that pre-MARA refers to the period 1977 – 1984 and the post-MARA period is 1985 – 2013. Beginning with the job creation rate for Michigan relative to its land neighbors, the coefficient on the interaction of Michigan and the post Michigan indicator is positive for the job creation rate, but it's not statistically significant. Using the Expanded Border States as the control group, we find a positive and statistically significant effect for the job creation rate, suggesting that a doubling in enforcement results in a 6.2 percent increase in job creation in Michigan, relative to counterfactual Michigan. Similarly, we find that a doubling in enforcement leads to a 7.8 percent increase in the job creation rate by startups relative to a counterfactual based on the Non-Enforcement States.

Turning to the startup entry rate, the estimated coefficient on the interaction between the Michigan dummy variable and the dummy variable for the post-MARA period is not statistically significant for any of the alternative control groups considered.

Taken together, the findings summarized in Table 2 suggest that increased enforcement of non-competes had no effect on the entry rate of startups, but had a positive effect on jobs created by these startups in Michigan relative to a counterfactual of Non-Enforcement States. Importantly, we find little support for the widely held view that enforcement of non-compete agreements negatively affects the entry rate of new firms or the rate of jobs created by new firms.

SYNTHETIC CONTROLS

An important requirement of the DID approach is that, in the absence of treatment, the outcomes for the treated and control groups follow parallel trends through time (i.e., the effects of the unobserved variables are fixed over time). Thus, the parallel-trends requirement implies that without treatment, the outcomes of interest for the treated and control groups are expected to evolve at the same rate. However, it is likely that many of the unobserved variables may have time-varying effects on the outcomes of interest. This could be one reason that startup activity in Michigan post-MARA fails to grow faster than that in the states sharing a land border with Michigan, for example. The SCM developed by Abadie, et al. (2010) is an alternative method that accounts for the effects of confounders changing over time. The SCM is a data-driven search routine designed to construct a comparison group based on pre-treatment economic and demographic trends. In our application, the SCM is a technique for constructing a counterfactual or “Synthetic Michigan” based on a linear combination of algorithmically derived weights assigned to the most representative or most similar states (using all 49 other states) that did not

receive the treatment. For our purposes, the SCM matches Michigan to potential candidate states having comparable pre-treatment-period predictor variables. The weighting assigned to each predictor variable is determined by regression analysis that minimizes the root mean square prediction error (RMSPE) in the pre-treated sample period. Importantly, the SCM is used to forecast a counterfactual post-MARA path for Synthetic Michigan.

Our objective is the construction of a Synthetic Michigan prior to 1985 based on a composite of all U.S. states.¹¹ For each outcome variable, we search for a synthetic match using the covariates given in first column of Table 1. Table 3 gives the states and their relative weights used to construct the synthetic control groups. Four states — Illinois, Ohio, Pennsylvania, and West Virginia — received positive weights in the construction of a Synthetic Michigan when the jobcreation rate is the outcome of interest. Two states — Ohio and Pennsylvania — contribute to the synthetic control group when the startup entry rate is the outcome of interest. As Table 3 shows, for example, Ohio receives the highest weight for both outcome variables, ranging from 0.58 to 0.87. Figure 2 shows a graph of outcome variables for Synthetic Michigan (the broken line) juxtaposed with the graph of actual outcomes in Michigan. The figure reveals that Michigan and its synthetic track one another tightly during the pre-treatment period both for the job creation rate (Figure 2a) and for the startup entry rate (Figure 2b).

Figure 2 shows there was not much change in either the job creation rate (Figure 2a) or the startup entry rate (Figure 2b) immediately following the passage of MARA, relative to Synthetic Michigan. As already mentioned, this lack of an immediate effect on startup activity is most likely due to a provision in the legislation that all statutes that were repealed by MARA would

¹¹ We use STATA Synth's procedure developed by Abadie, et al. (2010) to conduct the synthetic control analysis.

remain in force. Given these legislative provisions, we expect that startup activity post-MARA would initially continue to closely track Synthetic Michigan. But we also expect that, over time, startup activity in Michigan would diverge from the path of its synthetic. This is in fact what we observe. Both the job rate (Figure 2a) and the entry rate (Figure 2b) in Michigan increase relative to the synthetic control starting in about the mid-1990s.

The Synthetic Control Method uses a simple DID estimator:

$$DID_{MI} = (\bar{Y}_{Post}^{MI} - \bar{Y}_{Pre}^{MI}) - (\bar{Y}_{Post}^C - \bar{Y}_{Pre}^C) = \Delta\bar{Y}_{treatment} - \Delta\bar{Y}_{control},$$

where \bar{Y}_{Pre}^{MI} represents the sample average of Y in Michigan, MI , before treatment, Pre . Similarly, \bar{Y}_{Post}^{MI} is the sample average for MI after treatment, $Post$. Correspondingly, \bar{Y}_{Pre}^C and \bar{Y}_{Post}^C represent averages of Y pre- and post-treatment for the control group, C .

Table 4 presents the estimates of the DID analysis both for the startup entry rate and for the startup job creation rate using Synthetic Michigan as the control group. For each of the outcomes considered, the first column presents the average difference between Michigan and Synthetic Michigan during the pre-treatment period. The second column gives the mean difference during the post-treatment period, while the third column presents the DID estimates of the effect of MARA on startup activity.

Starting with the job creation rate, the first column of Table 4 shows a pre-MARA difference between Michigan and its counterfactual of 0.067 percent. The gap between Michigan and Synthetic Michigan widens to 0.125 percent during the post-MARA period, producing a DID estimate of 0.0575 percent. This result is similar to what we found for the regression-based DID analysis reported in Table 2. Specifically, a doubling of enforcement leads approximately to a 6

percent increase jobs created by startups. The last row of Table 4 shows the findings for the entry rate. Once again, we find that increased enforcement had essentially no effect on the entry rate of startups in Michigan.

To formally test the significance of the DID estimates, we follow Abadie, et al. (2010) and use a placebo test (or a falsification test) where the treatment is applied iteratively to each of the states. The fourth column of Table 4 shows the rank of the estimates and the *p*-values of the post-treatment change for Michigan relative to the distribution of all other U.S. states taken from the placebo tests. The final column in Table 4 presents the pre-treatment RMSPE. Based on Michigan's placebo ranking shown in the fourth column of Table 4, the DID-estimated coefficient for the job creation rate is not significantly different from zero. The last row of Table 4 shows that the DID estimate for the startup entry rate is small and not significantly different from zero.

To summarize, our findings based on the SCM also offer little support for the widely held view that enforcement of non-compete agreements negatively affects startup activity. If anything, increased enforcement appears to have had positive effects on the job creation rate of startups in Michigan, although placebo tests are not statistically significant.

6. HIGH-TECH EMPLOYMENT AND ESTABLISHMENTS

High-technology businesses are often considered pillars of growth for regions and the nation. High-tech activity is associated with high-value-added production, highly skilled workers, and relatively high wages. Many cities and states view high-tech activity as a source of growth and have offered generous subsidies to attract high-tech firms to their locations. A prominent

example is the 238 city leaders who offered subsidies, some quite substantial, for Amazon's second headquarters.

Many occupations in high-tech firms tend to be disproportionately covered by non-compete agreements. For example, Starr, et al. (2020) reported that more than 40 percent of electrical and electronics engineers were covered by non-competes. Because of both the increased coverage of non-competes for workers in the high-tech industry and policy considerations, it seems reasonable to consider the effect of non-compete agreements on high-tech activity.

Unfortunately, official statistics identifying high-tech activity do not exist and must be estimated. According to Goldschlag and Miranda (2020), “[a]n interagency workshop held by U.S. statistical agencies in 2004 identified a set of important factors that contribute to the concept of High Tech. These include disproportionately high employment of STEM workers, disproportionately high employment of R&D workers and capital, the production of High Tech products, and the use of High Tech production methods, including the use of High Tech capital goods and services.” Implementing this definition in practice requires identifying economic activity based on the use of high-tech workers, or the production of high-tech goods or services.

Most attempts to estimate high-tech industries are based on input rather than on output.¹² For the nation, Hecker (1999) identified a list of “High-Tech” three-digit SIC industries using employment in both R&D and technology-oriented occupations as reported in the Occupational Employment Statistics (OES) surveys (now referred to as STEM occupations). Based on OES surveys, Hecker identified 29 three-digit SIC industries in which the number of R&D workers and technology-oriented occupations accounted for a proportion of employment that was at least

¹² See Goldschlag and Miranda (2020) for a review of literature identifying high-tech activity based on the industry's use of input versus industrial output.

twice the average for all industries surveyed. We apply Hecker’s 29 industries to state-level employment and establishment data found in County Business Patterns for the period 1977–2013.¹³ Figure 3a shows the series for high-tech employment in Michigan and the 10 Non-Enforcement States, while Figure 3b shows the time series for establishments. As the figures show, other than the level, each state in this control group more or less tracks movements in Michigan for both the jobs in high-tech and the number of high-tech establishments during the pre-MARA period, indicating these untreated states provide an appropriate “counterfactual Michigan.”¹⁴

High-tech is an important source of employment in Michigan, averaging almost 400,000 jobs during the period 1977-2013, second only to California among the 10 Non-Enforcement States. In terms of the share of total employment accounted for by high-tech jobs, this share averaged 11.8 percent in Michigan, compared with an average rate of 7.5 percent for the nation. The share of employment accounted for by high-tech is almost 60 percent greater in Michigan than the average state.

Table 5 summarizes our findings of the regression-based DID analysis. The first two columns of Table 5 present the findings for high-tech employment and high-tech establishment, respectively, when the control group consists of the states sharing a land border with Michigan. The next two columns in Table 5 give the findings relative to a counterfactual based on the Expanded Border States, while the final two columns show the results for the Non-Enforcement States.

¹³ See Table 1 in Hecker (1999) for a list of the 29 three-digit SIC code industries designated high-tech activity used in this study.

¹⁴ The 29 SIC codes used in this study were converted to the 1997 edition of NAICS by the Office of Technology Policy and the Census Bureau. These NAICS codes are reported in Table Appendix B in Goldschlag and Miranda (2016). We found no appreciable difference in the findings when the NAICS definitions are used instead of the SIC code definitions. The NAICS findings are available upon request.

The results of interest are given by the interaction of a Michigan dummy variable and a dummy variable for the post-MARA period (shown in the third row of Table 5). Beginning with the creation of high-tech jobs in Michigan relative to the states sharing a land border with Michigan, the coefficient on the interaction of Michigan and the post-Michigan indicator is negative, but insignificant for high-tech employment and high-tech establishments. Using the Expanded Border States as the control, we find a positive effect for both the job creation rate and for the startup entry rate, but neither coefficient is statistically significant. We find positive and statistically significant effects of enforcement on both high-tech jobs and establishments when counterfactual Michigan is based on the Non-Enforcement States.

Taken together, the findings summarized in Table 5 suggest that increased enforcement of non-competes had no effect on either high-tech jobs or high-tech establishments in Michigan relative to a counterfactual of states sharing borders with Michigan. Only when we use the alternative definitions of counterfactual Michigan based on the Non-Enforcement States do we find any evidence of a positive and statistically significant effect of increased enforcement on high-tech activity. Specifically, we find that a 10 percent increase in enforcement led to a 5 percent increase in high-tech jobs and to a 4 percent increase in the number of high-tech firms.

Table 6 presents the states and relative weights used to construct “Synthetic Michigan” for high-tech activity. Figures 3c and 3d show a graph of high-tech variables for Synthetic Michigan (the broken line) juxtaposed with the graph of actual outcomes in Michigan. The figures reveal that Michigan and its synthetic track one another tightly during the pre-treatment period, both for employment (Figure 3c) and for establishments (Figure 3d).

Table 7 presents the estimates of the DID analysis for the high-tech employment and high-tech establishments using Synthetic Michigan as the control group. The second column of the table gives the mean difference during the post-treatment period, while the third column presents the DID estimates of the effect of MARA on high-tech activity. The rank and p -values associated with the placebo tests are shown in the fourth column of Table 7. Based on Michigan's placebo ranking shown in the fourth column of Table 7, the DID-estimated coefficient is not significantly different from zero.

7. PATENTS AND PATENT CITATIONS

So far, we have considered the effects of enforcement on aggregate startup activity. However, enforcement may have differential effects on startups across industries, and these effects may get muted in the aggregation of industry-level startup to total startup activity. Employers are much more likely to be concerned with mobility of their employees to other firms in the same industry than they are with employees changing industries altogether. In interviews with 52 randomly sampled patent holders in a single industry, Marx (2011) found that 25 percent of those who signed non-compete agreements changed industries when changing jobs. In comparison, individuals not covered by such covenants were significantly less likely to change industries when changing jobs. Marx's finding indicates that we need industrial-level data to more fully address this question of whether and to what extent non-compete covenants influence startup activity. Barnett and Sichelman (2016) point out that "no state enforces non-competes that

purport to proscribe employment at non-competing firms.” Unfortunately, the Census Bureau does not make industry-level data on startup activity publicly available.¹⁵

Fortunately, patent data can be used to study entrepreneurial activity and are available by technology classifications. In this section, we consider the effects of enforcement of patent activity in Michigan for total patents per 10,000 people (referred to as patents per capita) and for six patent technology classifications identified by Hall, et al. (2001). The null hypothesis we test is:

H₀ : The Passage of MARA did not affect patents per capita or patent citations per capita in Michigan relative to the control group

H_A : The passage of MARA did have an effect on patents per capita or patent citations per capita in Michigan relative to the control group.

7.1 Patent Data

We use data on patent applications obtained from the NBER Patent Data Project. The data span the years 1976–2006.¹⁶ Hall, et al. (2001) aggregated various patent classes into six main technological categories: Chemical (excluding Drugs); Computers and Communications; Drugs and Medical; Electrical and Electronics; Mechanical; and Others. Figure 4 shows the path of total patents per capita and for the six technological classifications in Michigan. For total patents, we see that there were less than three patents per 10,000 people in Michigan between 1977 through the late 1980s. Patents per capita started rising in the mid-1990s to reach a peak of 4.4

¹⁵ Starr, et al. (2020) report the results for a 2014 survey they conducted finding that incidence of non-compete varies across industries: The percentage of workers covered by such covenants ranged from 9 percent in agriculture and hunting, to over 30 percent in information; mining and extraction; and professional and scientific.

¹⁶ U.S. Patent and Trademark Office. Overview of the U.S. Patent Classification System (USPC). Washington, D.C. (2012), <http://www.uspto.gov/patents/resources/classification/overview.pdf>.

patents per capita in 2003 before declining to a value around 3.6 in 2006. Figure 4 shows that Mechanical patents were a major contributor to the run-up in total patents per capita observed since the late 1990s. There was less than 1 Mechanical patent per 10,000 people in Michigan during the period 1977–1988. Mechanical patents per capita started rising more rapidly in the late 1990s and reach a peak of 1.7 patents per capita during the period 2001–2004 before declining a bit after that. Mechanical patents accounted for 30 percent of total patents in Michigan in the mid-1970s to the early 1980s. That share steadily rose to account for 42 percent of all Michigan’s patents in 2006. During our sample period, Mechanical patents on average accounted for 35 percent of total Michigan patents, compared with 22 percent for the average state. The Chemical and the Others categories accounted for 19 percent and 22 percent of Michigan patents, respectively, during our sample period, whereas the Drugs category accounted for only 5 percent.

One concern about using patents as an innovation indicator is that the value of patents is highly skewed. Most patents are not worth much, while a few are valuable (e.g., Harhoff, et al., 1999). If a patent has value, we would expect it to be renewed before the patent expires. Serrano (2010) calculates that 78 percent of U.S. patents granted during 1983–2001 were not renewed, indicating that most patents are of low value. Fortunately, researchers can adjust for patent quality in their innovation metrics by weighting patents by the number of citations they receive. In the analysis, we exclude self-citations from these counts (i.e., a Microsoft patent that cites another Microsoft patent). In the analysis that follows, we look at both patents per capita and citations-weighted patents per capita. The regressions we use in the patent analysis are similar to the regressions used for startup activity as summarized in Table 2, except patents (citation-weighted patents) replace the startup job (entry) rate as the dependent variable.

Table 8 summarizes our findings by technology category when Non-Enforcement States is the control group.¹⁷ The first two columns of Table 8 present the findings for total private patents per capita and for the citation-weighted version. Subsequent columns in Table 8 show the results for the six major technological categories we investigate. Starting with total private patents per capita, the coefficient of interest (the estimated coefficient on interaction of Michigan and the post-MARA indicator) is positive and significant for total patents. Regarding the subcategories, the coefficient on patents per capita is negative in three cases and negative and significant in one case (Drugs). We also find that the estimated coefficient on patents per capita is positive for three categories and positive and significant for two classifications (Mechanical and Other).

Considering citations, we find that citation-weighted patents per capita are negative but insignificant for total patents per capita. For the subcategories, the coefficient on citations per capita is negative in four cases and significantly negative for Computer patents. The coefficient is positive in two cases and significantly positive for Mechanical patents.

Table 9 presents the estimates of the DID analysis for the citation-weighted patents using the SCM to form the control group. For each of the technological categories considered, the first column presents the average difference between Michigan and Synthetic Michigan during the pre-treatment period. The second column gives the mean difference during the post-treatment period, while the third column presents the DID estimates of the effect of MARA for the various technological categories considered. The rank and *p*-values associated with the placebo tests are shown in the fourth column of Table 9.

¹⁷ To conserve space, we present only tables with Non-Enforcement States as the comparison group.

The most statistically significant results are found for Mechanical patents. Table 9 shows a pre-MARA difference between Michigan and its counterfactual of 0.2176 percent. The gap between Michigan and Synthetic Michigan widens to 0.3747 percent during the post-MARA period, producing a DID estimate of 0.1570 percent. This estimate is quite similar to the 0.1899 estimate reported in Table 8 for Mechanical patents. Based on Michigan's placebo ranking for Mechanical patents shown in the fourth column of Table 9, the DID-estimated coefficient is marginally significantly different from zero. The *p*-values associated with Michigan's relative rank in the distribution of placebo states shows, the DID estimate is not significantly different from zero for the other five technology categories, as well as for the total category. Still, the finding of a positive and significant effect of increased enforcement on Mechanical patents is important given that Mechanical patents in 2006 accounted for more than 40 percent of all patent activity in Michigan.

8. CONCLUSION

In this study, we considered how state-level enforcement of non-compete agreements affects the entry of new establishments and the jobs created by these new firms at the state level. Exploiting Michigan's inadvertent reversal of its enforcement policy as a quasi-natural experiment, we find no evidence that increased enforcement negatively affected the aggregate startup entry rate or the overall job creation rate of new firms. If anything, increased enforcement appears to have had positive effects on the job creation rate of startups in Michigan. In a standard DID analysis, we find that a doubling in enforcement led to a 6 percent to 8 percent increase in the startup job creation rate in Michigan and to essentially no change the startup entry rate.

We also find evidence that enforcing non-competes positively affected the number of high-tech establishments and the level of high-tech employment in Michigan.

Extending our analysis to consider the effect of increased enforcement on patent activity, we find that enforcement had differential effects across technological classification. Importantly, increased enforcement has a positive and significant effect on Mechanical patents in Michigan. The Mechanical category is an important technological classification in Michigan: In 2006, it accounted for over 40 percent of the state's total patent applications.

While our findings suggest that increased enforcement may promote job creation, there is room for future research. Our findings for patents suggest that enforcement may have differential effects on startup activity across industries. Yet, only researchers with projects approved by the Census Bureau can access the industry-level data. This is one area for future research.

An interesting question is how does enforceability affect productivity of cities and ultimately of the nation? Cities may be less productive and less innovative if enforcement limits employees' outside options, even if enforcement does not limit startup activity. For example, knowledge spillovers may be limited if employees feel constrained by non-compete agreements from exchanging information with outsiders.

Patent data could be used to shed light on this issue. Often, patent citations are used to measure the extent of localized knowledge spillovers (see Jaffe, et al., 1993, and Buzard, et al., 2017). Using patent-citation data, Singh and Marx (2013) provide tantalizing evidence that knowledge diffusion is subdued in regions where non-competes are enforceable. Still, more work needs to be done.

A related issue is whether the ability of firms and workers to form better matches is constrained in local labor markets that enforce non-competes. Reduced “job-hopping” resulting from non-competes is a concern if reduced churning lowers labor productivity through less-efficient matching among firms and workers. Davis and Haltiwanger (2015) report that job seekers have fewer opportunities to meet prospective employers if startup activity is less fluid. Similarly, non-competes also may limit labor market pooling by tying workers to their current employers and by giving rise to a “brain-drain” from enforcing to non-enforcing states (for a review of the evidence, see Marx, et al., 2015). Policymakers need to not only balance the interest of firms and workers, but also consider the broader issues associated with the effects of non-competes on the productivity and the growth of cities.

REFERENCES

- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. (2010). "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*, 105 (490), pp. 493-505.
- Balasubramanian, Natarajan, Chang, Jin Woo, Sakakibara, Mariko, Sivadasan, Jagadeesh, & Starr, Evan. (2019). "Locked in? The Enforceability of Covenants Not to Compete and the Careers of High-Tech Workers," SSRN Working Paper No. 2905782.
- Barnett, Jonathan, and Sichelman, Ted M. (2016). "Revisiting Labor Mobility in Innovation Markets," *University of Southern California Legal Studies Research Paper Series*.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. (2004). "How Much Should We Trust Differences-in-Differences Estimates?," *Quarterly Journal of Economics*, 119 (1), pp. 249-275.
- Bozkaya, Ant, and Kerr, William. (2014). "Labor Regulations and European Venture Capital." *Journal of Economics & Management Strategy*, 23 (4), pp. 776-810.
- Buzard, Kristy, Gerald A. Carlino, Robert H. Hunt, Jake K. Carr, and Tony E. Smith. (2017). "The Agglomeration of American R & D Lab," *Journal of Urban Economics*, 101, pp. 14-26.
- Conti, Raffaele. (2014). "Do Non-Competition Agreements Lead Firms to Pursue Path-Breaking Inventions?" *Strategic Management*, Vol. 35 (8), pp. 1230-1248.
- Davis, Steven J., and John Haltiwanger. (2015). "Labor Market Fluidity and Economic Performance," *Reevaluating Labor Market Dynamics*, 2014 Jackson Hole Symposium Volume: Federal Reserve Bank of Kansas City, pp. 17-108.
- Fallick, Bruce, Fleischman, Charles, and Rebitzer, James. (2006). "Job-Hopping in Silicon Valley: Some Evidence Concerning the Microfoundations of a High-Technology Cluster." *Review of Economics and Statistics*, 88 (3), pp. 472-81.
- Garmaise, Mark J. (2011). "Ties that Truly Bind: Noncompetition Agreements, Executive Compensation, and Firm Investment," *Journal of Law, Economics, and Organization*, Vol. 27 (2), pp. 376-425.
- Gilson, Ronald J. (1999). "The Legal Infrastructure of High-Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not to Compete." *New York University Law Review*, 74, pp. 575-629.
- Goldschlag, Nathan, and Miranda, Javier. (2020). "Business Dynamics Statistics of High Tech Industries," *Journal of Economics & Management Strategy*, Vol. 29 (1), pp. 3-30.
- Goldschlag, Nathan, and Miranda, Javier. (2016). "Business Dynamics Statistics of High Tech Industries," *Center for Economic Studies*, No. 16-55, December 2016.

- Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg. (2001). "The NBER Patent Citations Data File: Lessons, Insights, and Methodological Tools," NBER Working Paper No. 8498.
- Haltiwanger, John, Ron S. Jarmin, and Javier Miranda. (2013). "Who Creates Jobs? Small Versus Large Versus Young," *Review of Economics and Statistics*, 95 (2), pp. 347-361.
- Harhoff, Dietmar, Francis Narin, F.M. Scherer, and Katrin Vopel. (1999). "Citation Frequency and the Value of Patented Inventions," *Review of Economics and Statistics*, 81, pp. 511-515.
- Hecker, Daniel. (1999). "High-Technology Employment: A Broader View," *Monthly Labor Review*, June, pp. 18-28.
- Hecker, Daniel. (2005). "High-Technology Employment: A NAICS-based Update," *Monthly Labor Review*, July, pp. 57-72.
- Jaffe, Adam, M. Trajtenberg, and R. Henderson. (1993) "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, 108, pp. 577-598.
- Kang, Hyo, and Fleming, Lee. (2020) "Non-Competes, Business Dynamism, and Concentration: Evidence from a Florida Case Study," *Journal of Economics & Management Strategy*, Vol. 29 (3), pp. 663-685.
- Marx, Matt, Jasjit Singh, and Lee Fleming. (2015). "Regional Disadvantage? Employee Non-Compete Agreements and Brain Drain," *Research Policy*, Vol. 44, pp. 394-404.
- Marx, Matt. (2011). "The Firm Strikes Back: Non-compete Agreements and the Mobility of Technical Professionals," *American Sociological Review*, Vol. 76 (5), pp. 695-712.
- Marx, Matt, Deborah Strumsky, and Lee Fleming. (2009). "Mobility, Skills, and the Michigan Non-Compete Experiment," *Management Science*, Vol. 55 (6), pp. 875-889.
- Meccheri, Nicola. (2009). "A Note on Non-Competes, Bargaining, and Training by Firms," *Economics Letters*, Vol. 102, pp. 198-200.
- Samila, Sampsa, and Olav Sorenson. (2011). "Non-Compete Covenants: Incentives to Innovate or Impediments to Growth," *Management Science*, 57 (3), pp. 425-438.
- Saxenian, AnnaLee. (1994). "Regional Advantage: Culture and Competition in Silicon Valley and Route 128," Harvard University Press, Cambridge, MA.
- Serrano, Carlos J. (2010). "The Dynamics of the Transfer and Renewal of Patents," *Rand Journal of Economics*, 41 (1), pp. 686-708.
- Singh, Jasjit, and Matt Marx. (2013). "Geographic Constraints on Knowledge Spillovers: Political Borders vs. Spatial Proximity," *Management Science*, 59 (9), pp. 2056-2078.
- Starr, Evan, Prescott, J.J., and Bishara, Norman. (2020). "Non-Compete Agreements in the U.S. Labor Force," *Journal of Law and Economics*, *Forthcoming*.

Starr, Evan. (2019). "Consider This: Training, Wages, and the Enforceability of Covenants Not to Compete." *ILR Review*, 72 (4), pp. 783-817.

Starr, Evan, Natarajan Balasubramanian, and Mariko Sakakibara. (2015). "Screening Spinouts? How Non-Compete Enforceability Affects the Creation, Growth, and Survival of New Firms," U.S. Census Bureau CES Working Paper No. 14-27.

Stone, Katherine V.W. (2002). "Knowledge at Work: Disputes Over the Ownership of Human Capital in the Changing Workplace," *Connecticut Law Review*, 34, pp. 721-763.

Stuart, Toby, and Olav Sorenson. (2003). "Liquidity Events and the Geographic Distribution of Entrepreneurial Activity," *Administrative Science Quarterly*, Vol. 48, pp. 175-201.

TABLE 1a: Summary Statistics for Michigan										
	Michigan, 1977–1984					Michigan, 1985–2013				
	OBS.	MEAN	SD	MIN	MAX	OBS.	MEAN	SD	MIN	MAX
Job Creation Rate	8	0.0313	0.0071	0.0256	0.0478	29	0.0258	0.0042	0.0176	0.0332
Est. Entry Rate	8	0.1169	0.0163	0.1013	0.1486	29	0.0895	0.0150	0.0628	0.1215
Share Agri.	8	0.0024	0.0003	0.0020	0.0027	29	0.0053	0.0010	0.0030	0.0066
Share Mining/Extraction	8	0.0440	0.0048	0.0375	0.0497	29	0.0457	0.0124	0.0258	0.0720
Share Light Mfg.	8	0.0742	0.0023	0.0723	0.0793	29	0.0568	0.0094	0.0439	0.0702
Share Heavy Mfg.	8	0.2754	0.0304	0.2359	0.3136	29	0.1915	0.0248	0.1418	0.2436
Share Trans./Communications	8	0.0516	0.0012	0.0493	0.0527	29	0.0254	0.0214	0.0048	0.0511
Share Trade	8	0.2629	0.0078	0.2533	0.2745	29	0.3551	0.1020	0.2146	0.4641
Share Depository Inst.	8	0.0577	0.0044	0.0526	0.0636	29	0.0523	0.0099	0.0362	0.0671
Share Services	8	0.0802	0.0058	0.0750	0.0907	29	0.0846	0.0317	0.0468	0.1450
Share Health/Legal/Ed. Services	8	0.1516	0.0182	0.1291	0.1728	29	0.1834	0.0290	0.1338	0.2286
Share Pop Aged 15 - 64	8	0.6428	0.0271	0.6101	0.6625	29	0.6612	0.0034	0.6590	0.6691
Percent College Grad	8	0.1247	0.0252	0.0943	0.1430	29	0.1942	0.0362	0.1430	0.2520
Unemployment Rate	8	0.1109	0.0312	0.0696	0.1537	29	0.0742	0.0024	0.0366	0.1378
Labor Force Part. Rate	8	0.6363	0.0046	0.6281	0.6433	29	0.6505	0.0230	0.6001	0.6871
Real Per Capita GSP Growth	8	0.0153	0.0350	-0.0343	0.0654	29	0.0138	0.0226	-0.0465	0.0489
Pop Growth	8	-0.0009	0.0059	-0.0102	0.0051	29	0.0031	0.0040	-0.0054	0.0095
Patents Per 10,000 People	8	2.2039	0.2911	1.8488	2.5161	29	3.3178	0.6826	2.2763	5.1695
Citations per 10,000 People	8	23.6368	2.4554	19.6388	28.5580	21	25.106	11.7822	0.1335	35.9402

TABLE 1b: Summary Statistics for All Other, Excluding Michigan										
	States Other than Michigan, 1977–1984					States Other than Michigan, 1985–2013				
	OBS.	MEAN	SD	MIN	MAX	OBS.	MEAN	SD	MIN	MAX
Job Creation Rate	392	0.0445	0.0160	0.0218	0.1145	1421	0.0295	0.0085	0.0128	0.0989
Est. Entry Rate	392	0.1332	0.0285	0.0885	0.2423	1421	0.0981	0.0293	0.0508	0.1802
Share Agri.	392	0.0044	0.0019	0.0018	0.0104	1421	0.0067	0.0020	0.0027	0.0177
Share Mining/Extraction	392	0.0866	0.0489	0.0331	0.3248	1421	0.0705	0.0332	0.0230	0.2558
Share Light Mfg.	392	0.1122	0.0589	0.0194	0.3326	1421	0.0712	0.0368	0.0118	0.2621
Share Heavy Mfg.	392	0.1350	0.0682	0.0057	0.3184	1421	0.1003	0.0457	0.0054	0.2497
Share Trans./Communications	392	0.0639	0.01538	0.0368	0.1458	1421	0.0341	0.0286	0.0034	0.1378
Share Trade	392	0.2867	0.0345	0.2264	0.4042	1421	0.3688	0.1108	0.1621	0.5703
Share Depository Inst.	392	0.0677	0.0136	0.0439	0.1221	1421	0.0612	0.0189	0.0256	0.1683
Share Services	392	0.0981	0.0509	0.0529	0.4254	1421	0.0956	0.0556	0.0369	0.4279
Share Health/Legal/Ed. Services	392	0.1455	0.0297	0.0728	0.2266	1421	0.1917	0.0399	0.0660	0.3154
Share Pop Aged 15 - 64	392	0.6403	0.2850	0.5825	0.7018	1421	0.6594	0.0175	0.6012	0.7034
Percent College Grad	392	0.1406	0.0379	0.0662	0.2300	1421	0.2187	0.0533	0.1040	0.3900
Unemployment Rate	392	0.0709	0.0221	0.0276	0.1779	1421	0.0574	0.0190	0.0230	0.1353
Labor Force Part. Rate	392	0.6451	0.0381	0.5097	0.7332	1421	0.6687	0.0401	0.5136	0.7537
Real Per Capita GSP Growth	392	0.0200	0.0299	-0.1383	0.1926	1421	0.0185	0.0238	-0.1022	0.1344
Pop Growth	392	0.0122	0.0134	-0.0098	0.0845	1421	0.0099	0.0097	-0.0598	0.0739
Patents Per 10,000 People	392	1.4950	2.7115	0.1173	21.0235	1029	2.2927	3.7465	0.0930	30.4258
Citations Per 10,000 People	392	14.6505	25.3367	0.6961	197.7697	1029	18.0755	30.3086	0	260.4539

Table 2: Estimated Effects of the Michigan Antitrust Reform Act on Startup Entry Rate and the Job Creation Rate, 1977–2013[†]

	States Sharing a Land Border with Michigan ^{††}		States Sharing a Land or Water Border with Michigan ^{†††}		Non-Enforcement States ^{††††}	
	Job Rate	Entry Rate	Job Rate	Entry Rate	Job Rate	Entry Rate
Michigan	0.1691 (0.0405)***	0.0842 (0.0175)***	0.0933 (0.0301)***	0.0432 (0.0156)***	0.1118 (0.0593)*	0.0673 (0.0410)*
Post-MARA	-0.3815 (0.2973)	-0.5477 (0.1071)***	-0.5682 (0.1785)***	-0.9266 (0.1040)***	-0.9465 (0.1043)***	-0.5570 (0.0825)***
Mich.*Post-MARA	0.0172 (0.0331)	-0.0074 (0.0163)	0.0623 (0.0281)**	0.0005 (0.0139)	0.0780 (0.0405)**	0.0054 (0.0265)
Percent College Grad	-0.0075 (0.0156)	0.0013 (0.0068)	0.0143 (0.0060)**	0.0212 (0.0035)***	0.0029 (0.0047)	-0.0018 (0.0037)
Unemployment Rate	0.0016 (0.0094)	0.0076 (0.0045)*	-0.0007 (0.0078)	0.0139 (0.0038)	0.0095 (0.0066)	0.0177 (0.0048)***
Labor Force Part. Rate	-0.2000 (0.0062)***	-0.0104 (0.0029)***	-0.0105 (0.0037)***	-0.0041 (0.0021)*	-0.0065 (0.0027)**	0.0026 (0.0018)
Real Per Capita Income (levels)	0.2581 (0.6967)	-0.0605 (0.2857)	-0.3172 (0.5026)	-0.0124 (0.2729)	0.1530 (0.2240)	-0.0778 (0.1627)
Share Pop age 15 to 64	-0.0944 (0.0205)***	-0.0122 (0.0076)	-0.0564 (0.0146)***	0.0111 (0.3893)	-0.0527 (0.0069)***	-0.0196 (0.0049)***
Pop Growth (levels)	2.5810 (2.8308)	3.5853 (1.5780)**	3.5763 (2.2605)	5.1755 (1.1003)***	5.4376 (0.7615)***	5.4120 (0.5667)***
Constant	2.60 (4.0200)	-3.83 (1.8791)**	3.85 (1.8259)**	-3.71 (0.7242)	4.35 (0.8405)***	0.44 (0.6187)
No. of Obs.	148	148	222	222	406	406
R^2	0.9712	0.9929	0.9523	0.9793	0.9006	0.9092

[†]Robust standard errors in parentheses.

*, **, *** Represent statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

^{††} States Sharing a Land Border with Michigan are IN, OH, and WI.

^{†††} States sharing a land border or a water border with Michigan are IN, OH, WI, IL, and MN.

^{††††} Limited Non-Enforcement States are AK, CA, CT, MN, MT, NV, ND, OK, WA, and WV.

All regressions include one-digit SIC industry share of total employment and year fixed-effects.

Table 3: States Receiving Positive Weights for the Synthetic Control Group	
Startup Activity	
Job Rate	Entry Rate
Illinois = 0.264	
Ohio = 0.576	Ohio = 0.865
Pennsylvania = 0.126	Pennsylvania = 0.135
W. Virginia = 0.034	

Table 4: Estimated Effects of Change in Michigan's Non-Compete Enforcement on the Startup Entry Rate and the Startup Job Creation Rate Relative to Synthetic Michigan					
	Average Difference Relative to Control Group Pre-Treatment Period	Average Difference Relative to Control Group Post-Treatment Period	Change Post-Treatment	Rank	Pre-Treatment Period RMSPE
High-Tech Jobs	0.0671	0.1245	0.0575	14/50 $p = 0.28$	0.0925
High-Tech Firms	0.0873	0.0872	0.0001	12/50 $p = 0.48$	0.0877

Table 5: Estimated Effects of the Michigan Antitrust Reform Act on High-Tech Employment and the High-Tech Firms, 1977–2013 [†]						
	States Sharing a Land Border with Michigan ^{††}		States Sharing a Land or Water Border with Michigan ^{†††}		Non-Enforcement States ^{††††}	
	JOBS	FIRMS	JOBS	FIRMS	JOBS	FIRMS
Michigan	0.0063 (0.0601)	-0.1365 (0.0661)**	-0.1582 (0.0851)*	-0.2732 (0.0963)***	-0.2565 (0.3764)	0.1437 (0.3190)
Post-MARA	-2.2494 (0.4585)***	-1.2834 (0.5292)**	-2.4802 (0.3988)***	-2.1529 (0.4060)***	1.9780 (0.5079)***	-2.5906 (0.4867)***
Mich. *Post-MARA	-0.0714 (0.0516)	-0.0127 (0.0536)	0.0651 (0.0911)	0.0692 (0.1009)	0.4952 (0.2181)**	0.4032 (0.1257)**
Percent College Grad	0.0836 (0.0255)***	0.1069 (0.0280)***	0.0171 (0.0143)	0.0391 (0.0153)**	-0.0384 (0.0252)	-0.0124 (0.0246)
Unemployment Rate	-0.0185 (0.0144)	-0.0230 (0.0158)	0.0181 (0.0186)	0.0178 (0.0210)	0.1491 (0.0350)***	0.1316 (0.0320)***
Labor Force Part. Rate	0.3778 (0.0098)***	0.0480 (0.0107)***	-0.0622 (0.0111)***	-0.0654 (0.0119)***	0.0170 (0.0158)	-0.0070 (0.0146)
Real Per Capita Income Growth (Levels)	0.2657 (0.9062)	0.4382 (1.0681)	0.3389 (1.1932)	0.7927 (1.3388)	-1.0653 (1.2300)	-1.2964 (1.1968)
Share Pop age 15 to 64	0.0238 (0.0345)	0.0365 (0.0387)	0.2113 (0.0350)***	0.2001 (0.0361)***	0.0144 (0.0328)	-0.0244 (0.0294)
Pop Growth (levels)	-5.2326 (7.6608)	-7.1501 (4.8486)	-14.7027 (5.9348)**	-22.2128 (6.5415)***	24.4059 (4.7466)***	23.0728 (4.5025)***
Constant	14.30 (5.6513)**	23.89 (6.6384)***	10.21 (4.3542)**	8.48 (4.6874)*	-16.37 (3.9803)***	-8.16 (3.7040)**
No. of Obs.	148	148	222	222	406	406
R^2	0.9755	0.9867	0.9129	0.9529	0.8260	0.7755

[†]Robust standard errors in parentheses.

*, **, *** Represent Statistical Significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

^{††}Shares a land border with Michigan: IN, OH, WI.

^{†††}Shares a land or a water border: IL, IN, OH, MN, WI.

^{††††}Non-Enforcement States are AK, CA, CT, MN, MI, NV, ND, OK, WA, and WV.

All regressions include one-digit SIC industry share of total employment and year fixed-effects.

Table 6: States Receiving Positive Weights for the Synthetic Control Group High-Tech Activity	
Job Rate	Entry Rate
Indiana = 0.156	Indiana = 0.119
Ohio = 0.828	Ohio = 0.805
W. Virginia = 0.016	W. Virginia = 0.076

Table 7: Estimated Effects of Change in Michigan's Non-Compete Enforcement on High-Tech Entry Rate and Job Creation Rate Relative to Synthetic Michigan					
	Average Difference Relative to Control Group Pre- Treatment Period	Average Difference Relative to Control Group Post- Treatment Period	Change Post- Treatment	Rank	Pre- Treatment Period RMSPE
High-Tech Jobs	0.1797	0.1514	-0.0283	22/50 $p = 0.44$	0.1892
High-Tech Firms	0.0871	0.1395	0.0524	21/50 $p = 0.42$	0.0875

Table 8: Estimated Effects of the Michigan Antitrust Reform Act on Private Patent and Patent Citations, 1976–2006^{†, ††}

	Total Private Patents Per Capita		Private Chemical Patents Per Capita		Private Computer Patents Per Capita		Private Drugs Patents Per Capita	
	Patents	Citations	Patents	Citations	Patents	Citations	Patents	Citations
Michigan	-0.0460 (0.0386)	-0.7063 (0.3882)*	-0.0171 (0.0105)*	-0.1792 (0.0868)**	-0.0104 (0.0117)	0.0799 (0.1348)	-0.0195 (0.0072)***	-0.4118 (0.1060)***
Post-MARA	0.0232 (0.0796)	-0.7564 (0.8233)	-0.0183 (0.0226)	-0.0653 (0.2001)	-0.0536 (0.0258)	-0.5262 (0.2673)**	0.0024 (0.0171)	-0.1691 (0.1818)
Mich. *Post-MARA	0.0537 (0.0249)**	-0.2094 (0.2425)	-0.0041 (0.0077)	-0.0696 (0.0651)	-0.0056 (0.0064)	-0.2635 (0.0957)**	-0.0076 (0.0039)**	-0.1364 (0.0708)
Share of High-Tech Jobs	0.0003 (0.0091)	0.0893 (0.0827)	-0.0100 (0.0017)***	-0.0588 (0.0161)***	0.0150 (0.0039)***	0.0948 (0.0383)**	0.0072 (0.0013)***	0.1455 (0.0224)***
Percent College Grad	0.0173 (0.0041)***	-0.0641 (0.0354)*	-0.0039 (0.0012)***	-0.0136 (0.0088)	0.0085 (0.0014)***	0.0126 (0.0128)	0.0027 (0.0007)***	0.0017 (0.0090)
Unemployment Rate	-0.0092 (0.0044)**	-0.1014 (0.0443)**	-0.0037 (0.0014)**	-0.0351 (0.0122)**	0.0003 (0.0011)	-0.0087 (0.0136)	0.0004 (0.0006)	-0.053 (0.0104)
Labor Force Part. Rate	-0.0033 (0.0022)	0.0683 (0.0211)***	-0.0014 (0.0006)**	0.0089 (0.0045)**	-0.0034 (0.0007)***	-0.0131 (0.0075)*	0.0014 (0.0006)**	0.0381 (0.0073)***
Real Per Capita Income (levels)	0.1762 (0.1570)	2.9400 (1.7308)*	0.0400 (0.0365)	0.5564 (0.3448)	0.0462 (0.0456)	0.6155 (0.5535)	0.0137 (0.0414)	0.9758 (0.6446)
Share Pop age 15 to 64	0.0152 (0.0055)**	0.3647 (0.0484)***	0.0042 (0.0014)***	0.0719 (0.0128)***	0.0008 (0.0016)	0.0804 (0.0174)***	-0.0007 (0.0009)	0.0129 (0.0114)
Pop Growth (levels)	-0.5240 (0.3913)	-10.8540 (5.2508)**	-0.1769 (0.1204)	-2.6602 (1.1868)**	-0.1370 (0.1246)	-6.2048 (2.2102)**	0.1275 (0.0876)	1.096 (1.2278)
Constant	-1.64 (0.6560)**	-49.03 (21.4671)	-0.1065 (0.1792)	-7.07 (1.5427)***	0.3351 (0.2057)**	-9.77 (2.3063)	-0.2898 (0.1279)**	-7.09 (1.6116)***
No. of Obs.	330	330	330	330	330	330	330	330
R^2	0.8400	0.8336	0.8414	0.8184	0.6619	0.6622	0.5927	0.6369

Table 8: Continued						
	Private Electrical Patents Per Capita		Private Mechanical Patents Per Capita		Private Other Patents Per Capita	
	Patents	Citations	Patents	Citations	Patents	Citations
Michigan	0.0036 (0.0099)	-0.0434 (0.0939)	0.0218 (0.0076)***	0.1231 (0.0767)	-0.0234 (0.0085)***	-0.2748 (0.0983)**
Post-MARA	0.0127 (0.0171)	-0.1617 (0.1669)	0.0230 (0.0145)	0.0315 (0.1354)	0.0570 (0.0053)	0.1344 (0.1829)
Mich.*Post-MARA	0.0066 (0.0063)	-0.01780 (0.0578)	0.0430 (0.0094)***	0.1899 (0.0668)***	0.0214 (0.0053)***	0.0881 (0.0609)
Share of High-Tech Jobs	0.0012 (0.0028)	0.0007 (0.0206)	-0.0106 (0.0012)***	-0.802 (0.0128)***	-0.0026 (0.0015)**	-0.0127 (0.0125)
Percent College Grad	0.0023 (0.0008)**	-0.0141 (0.0077)*	0.0003 (0.0006)	-0.0252 (0.0056)***	-0.0005 (0.0008)	-0.0256 (0.0070)***
Unemployment Rate	-0.0019 (0.0010)**	-0.0231 (0.0098)**	-0.0023 (0.0009)**	-0.0137 (0.0079)*	-0.0019 (0.0008)**	-0.0155 (0.0089)*
Labor Force Part. Rate	-0.0016 (0.0004)***	-0.0020 (0.0041)	0.0002 (0.0003)	0.0147 (0.0030)***	0.0014 (0.0004)***	0.0217 (0.0039)***
Real Per Capita Income (levels)	0.0459 (0.0301)	0.6840 (0.3020)**	0.0252 (0.0299)	0.6778 (0.2647)	0.0053 (0.0301)	0.3205 (0.3351)
Share Pop age 15 to 64	0.0051 (0.0011)	0.0991 (0.0102)***	0.0055 (0.0010)***	0.0765 (0.0079)***	0.0003 (0.0009)	0.0239 (0.0087)**
Pop Growth (levels)	-0.1884 (0.0867)**	-3.2678 (1.1549)**	-0.1067 (0.671)	-0.9804 (0.7786)	-0.0424 (0.0764)	1.1632 (1.2422)
Constant	-0.5049 (0.1296)*	-10.33 (1.2826)***	-0.6568 (0.1095)***	-10.06 (0.9419)***	-0.1166 (0.1108)	-4.71 (1.2897)***
No. of Obs.	330	330	330	330	330	330
R^2	0.7948	0.8186	0.8665	0.8594	0.7831	0.7539

†Robust standard errors in parentheses.

*, **, *** Represent Statistical Significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

††Control Group is Non-Enforcement States (AK, CA, CT, MN, MT, NV, ND, OK, WA, and WV).

All regressions include year effects and one-digit industry controls.

Table 9: Estimated Effects of Change in Michigan's Non-Compete Enforcement on Total Patent Citations, and by Technology Classifications, Relative to Synthetic Michigan					
	Average Difference Relative to Control Group Pre-Treatment Period	Average Difference Relative to Control Group Post-Treatment Period	Change Post-Treatment	Rank (p-value)	Pre-Treatment Period RMSPE
Total Patent Citations	0.2290	0.1705	0.0586	28/50 (0.5490)	0.2582
Chemical Citations	-0.1005	-0.1019	-0.0014	24/50 (0.4706)	0.1293
Computer Citations	0.0615	-0.0470	-0.1085	25/50 (0.4902)	0.0903
Drugs Citations	-0.0279	-0.1058	-0.0778	45/50 (0.8824)	0.0493
Electrical Citations	0.0343	-0.0177	-0.0520	38/50 (0.7451)	0.0454
Mechanical Citations	0.2176	0.3747	0.1570	4/50 (0.0784)	0.2210
Other Citations	-0.0077	0.0487	0.0564	7/50 (0.1373)	0.0429

Figure 1a: Percent of Jobs Created by Startups
Michigan and the Border States, 1977 - 2013

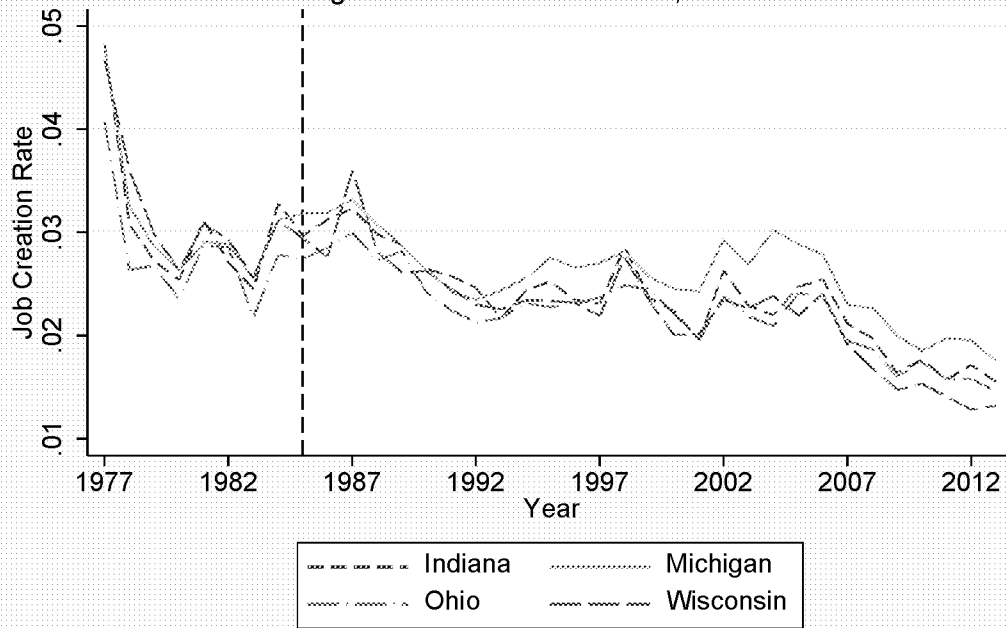


Figure 1b: Startup Entry Rate
Michigan and the Border States, 1977 - 2013

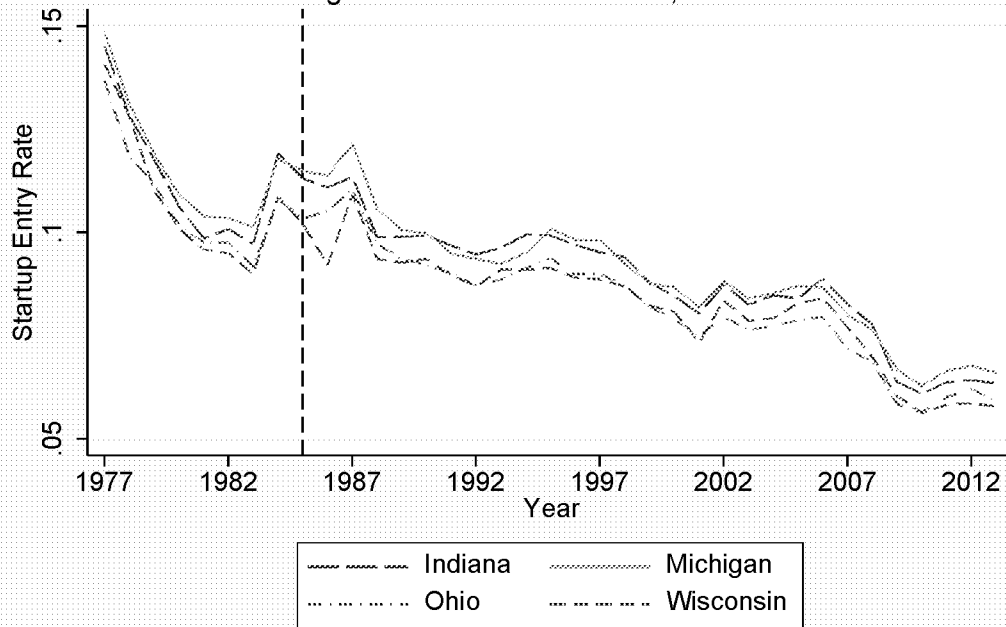


Figure 1c: Percent of Jobs Created by Startups
Michigan and the Expanded Border States, 1977 - 2013

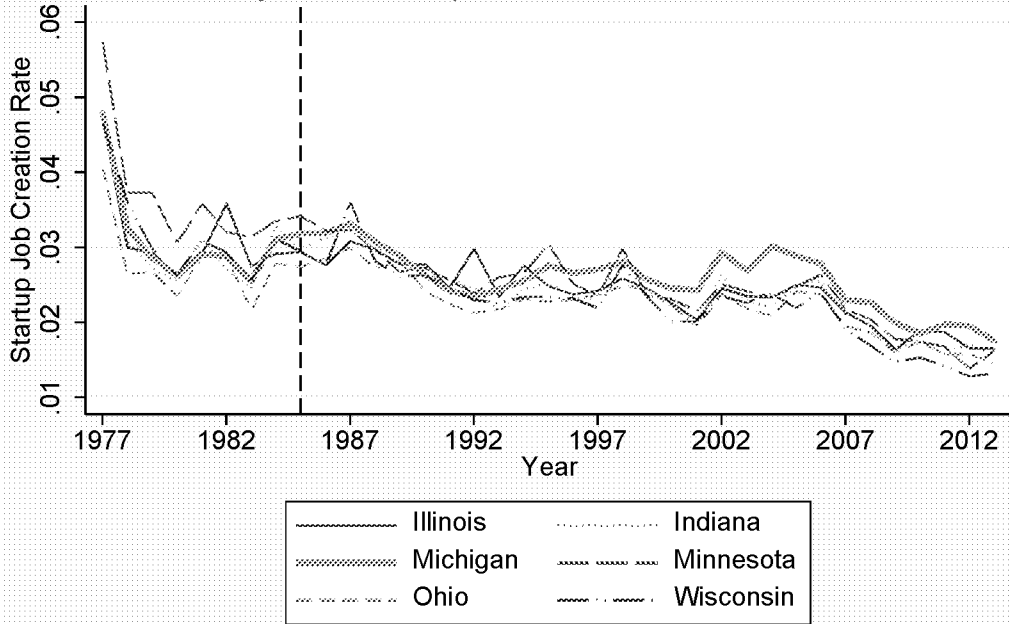


Figure 1d: Percent of Jobs Created by Startups
Michigan and the Non-Enforcement States, 1977 - 2013

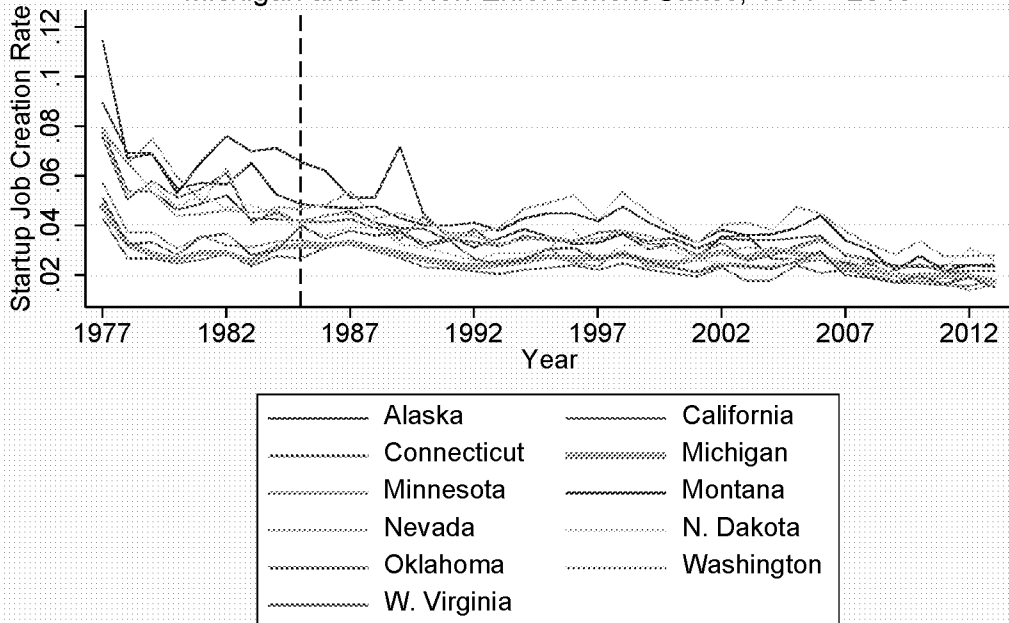


Figure 2a: Percent of Jobs Created by Startups
Michigan and the Synthetic Control Group, 1977 - 2013

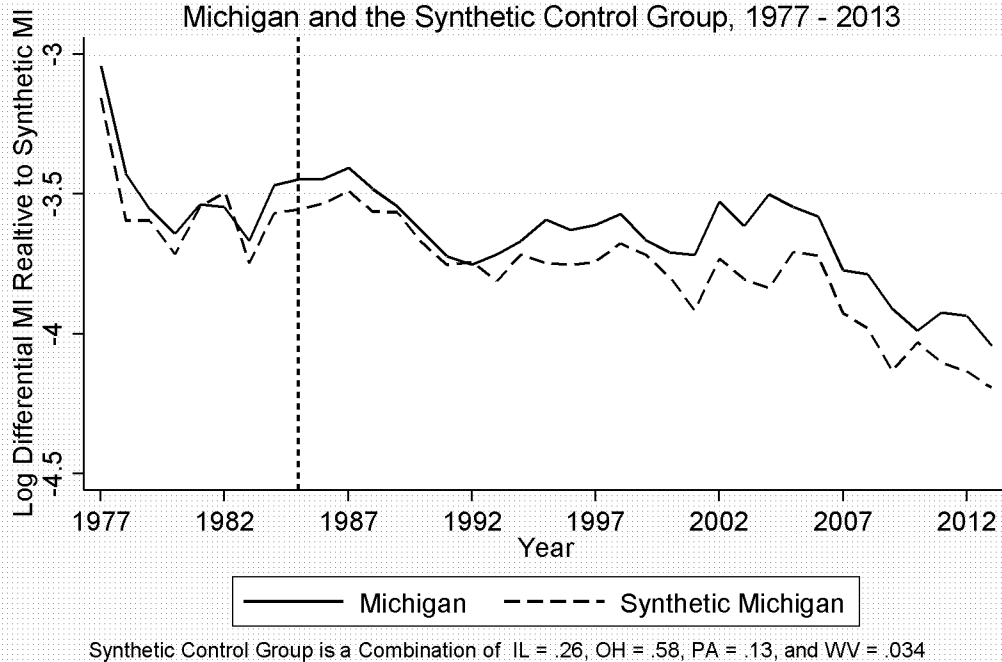


Figure 2b: Startup Entry as a Percent of All Establishments
Michigan and the Synthetic Control Group, 1977 - 2013

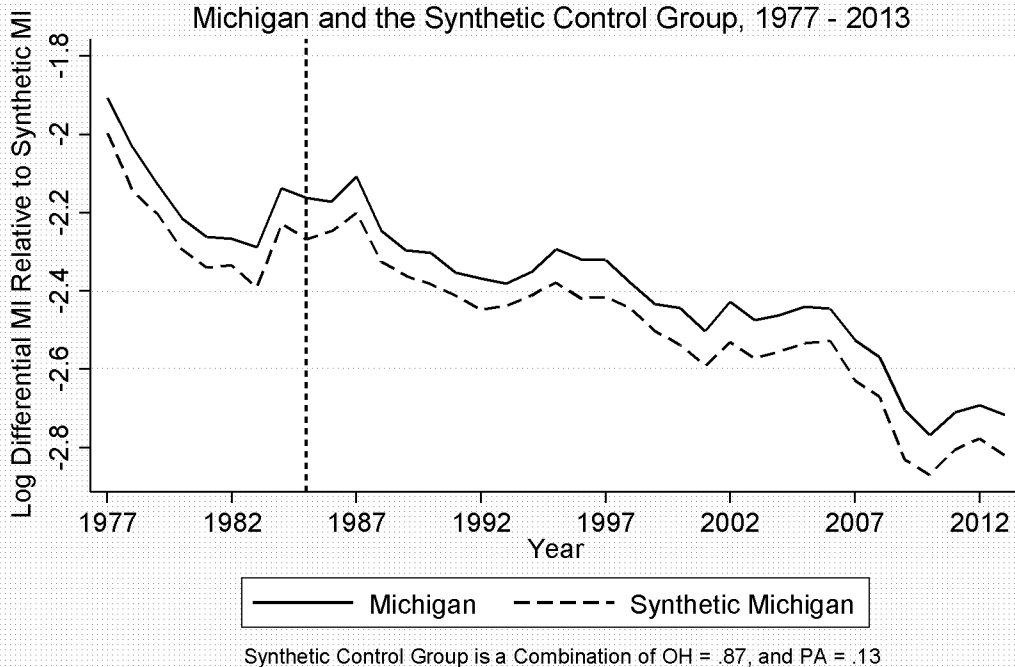


Figure 3a: High-Tech Employment

Michigan and the Non-Enforcement States, 1977 - 2013

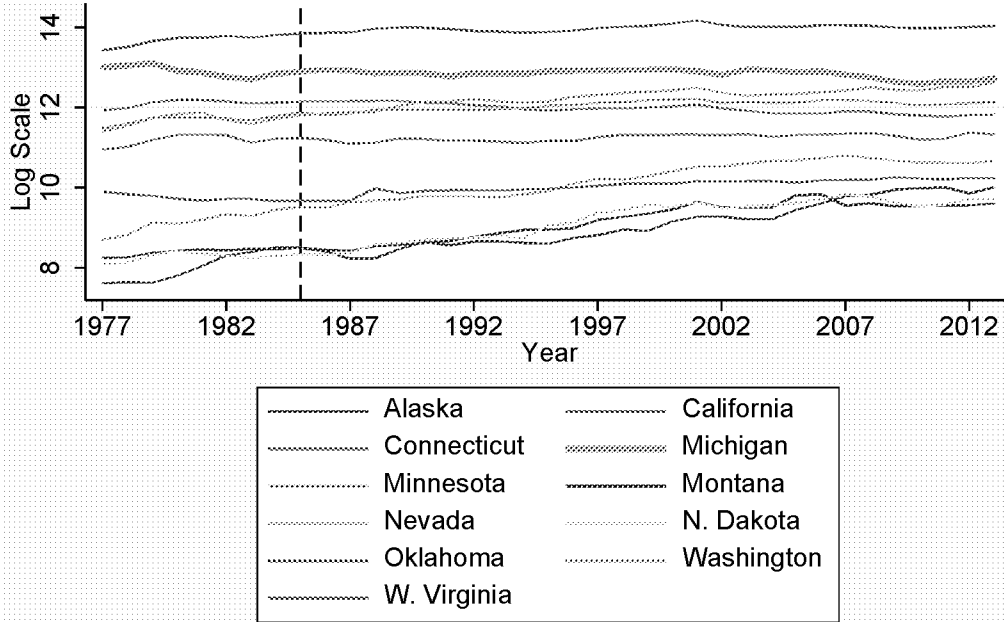


Figure 3b: High-Tech Establishments

Michigan and the Non-Enforcement States, 1977 - 2013

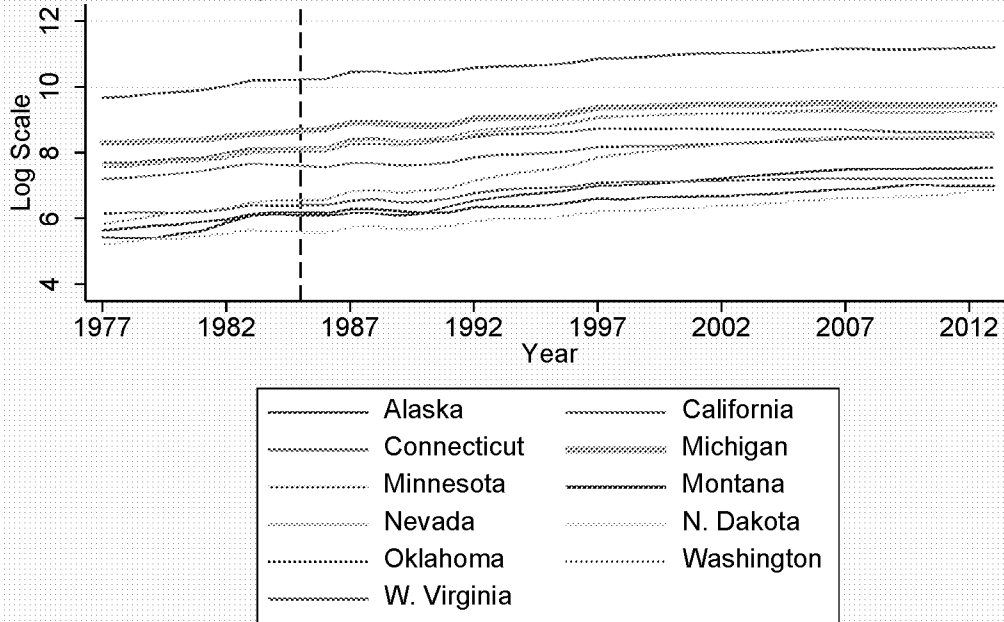


Figure 3c: High-Tech Employment
Michigan and Synthetic Michigan, 1977 - 2013*

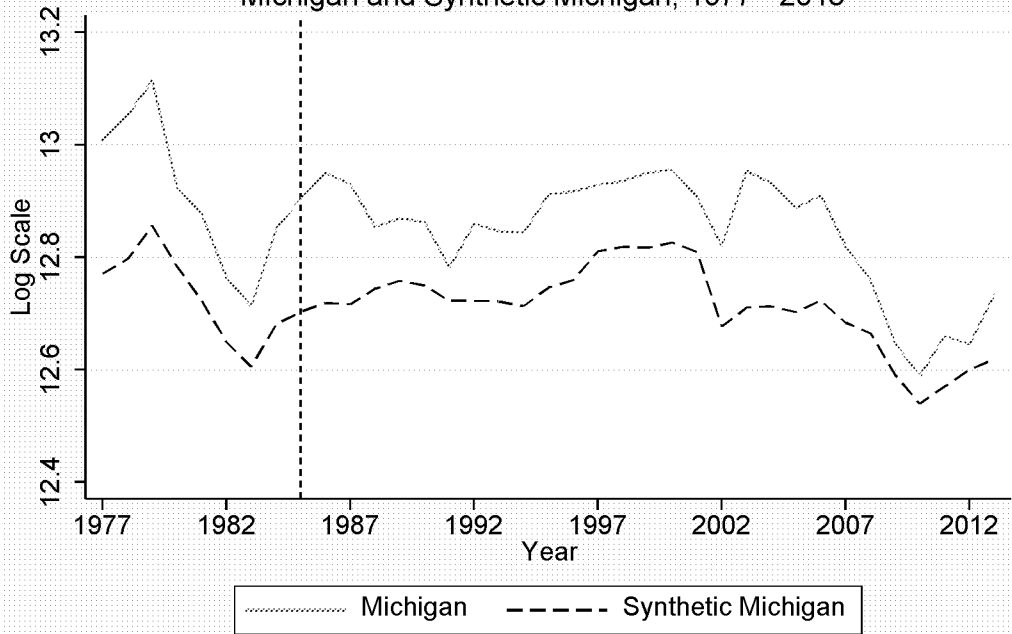


Figure 3d: High-Tech Establishments
Michigan and Synthetic Michigan, 1977 - 2013*

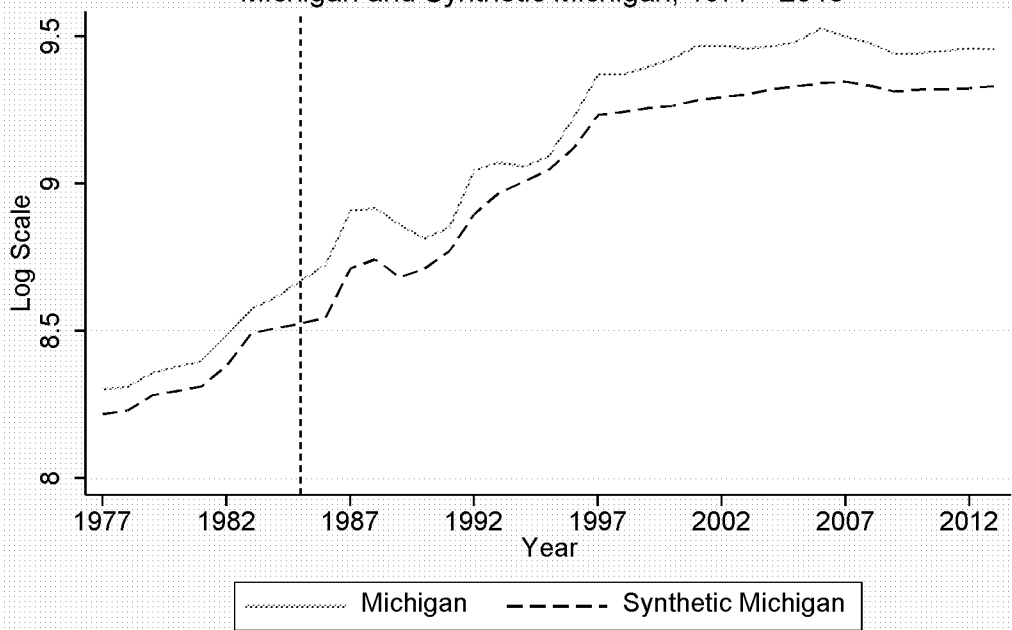
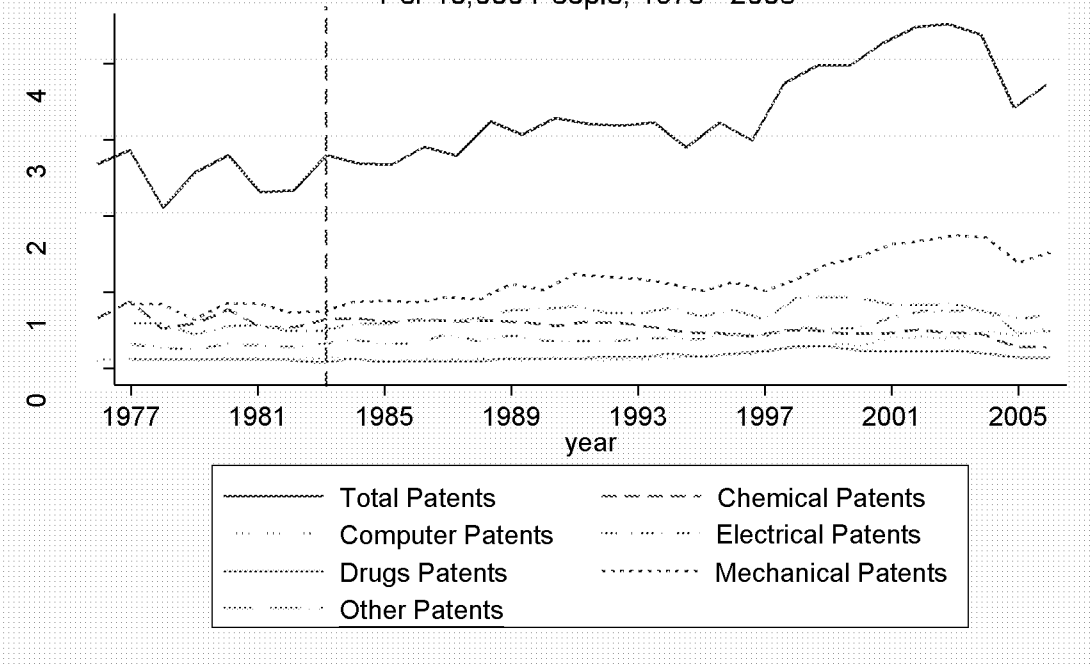


Figure 4: Total Patents and Patents by Technology Class
Per 10,000 People, 1976 - 2006



Noncompete Agreements and the Welfare of Consumers*

Michael Lipsitz
Federal Trade Commission

Mark Tremblay
Miami University

December 1, 2021

Abstract

Employee spinoffs may harm incumbent firm owners for two reasons: first, they increase competition in relevant product markets, potentially decreasing rents associated with market power. Second, the threat of an employee spinoff may prevent a firm owner from making costly, productivity-enhancing investments in their workers. Noncompete agreements (NCAs) solve both problems. From the perspective of a consumer, NCAs may increase prices by decreasing competition, but the investments made by firm owners have the potential to mitigate competitive harms. We develop a model which formally demonstrates this tradeoff to assess the impacts of NCA policy on consumers, and discuss when a ban on NCAs is most likely to be beneficial for consumers. We show that the competitive environment, the nature of investment pass-through, and the benefits of investment play large roles. Counterintuitively, increased benefits of costly investments have the potential to *harm* consumers, such that industries where NCAs are most important to firms may also be those where harm is greater. Finally, we draw two analogies between NCAs and antitrust (merger analysis and pay-for-delay agreements) and show how insights in those areas may inform NCA policy.

Keywords: Noncompete agreements, entrepreneurship, employee spinoffs, antitrust, consumer welfare

JEL Classifications: J41, J53, K21, K31, L26, L41

*Emails: mlipsitz@ftc.gov; tremblmj@miamioh.edu. We thank Dan Hosken, David Schmidt, Evan Starr, and participants at the LERA Annual Meeting for helpful comments. The views expressed in this article are those of the authors and do not necessarily reflect those of the Federal Trade Commission or any individual Commissioner.

1 Introduction

Employee spinoffs make up a significant and important portion of new firm births (Muendler et al., 2012; Franco, 2005). Typically studied as engines of innovation, spinoffs may also foster competition in industries that are otherwise highly concentrated: as noted by Fulghieri and Sevilir (2011), “...in many industries, the majority of new firms are created by workers of established firms, and such firms end up competing in similar industries as their parent firm.” By using the human capital they accrue at their parent firms, employee entrepreneurs are uniquely able to found spinoffs which may have success where entrepreneurs with less industry experience may fail.

Fulghieri and Sevilir (2011) go on to ask “...why established firms do not prevent the creation of such new firms that lead to greater competition...”. While they provide one answer to this question, another answer is that established firms do, in fact, prevent creation of employee spinoffs: they do so using employee noncompete agreements (NCAs). An NCA may stipulate that a worker is prohibited from founding a competing firm, typically for a set duration of time. NCAs are historically explained as a means by which to guarantee returns on investments: without the assurance that a worker will not leave to join or found a competing firm, an owner may not be willing to make a costly investment in a worker’s human capital, or to impart trade secrets which aid in a firm’s creation of value. In other words, NCAs solve an investment holdup problem on the part of the owner (Williamson, 1975; Klein et al., 1978; Rubin and Shedd, 1981; Hart and Moore, 1990).

The tradeoff between value creation and competition is the focus of this paper. Courts and policymakers who wish to determine optimal levels of NCA enforceability may be concerned with the overall impact of NCAs on the economy: for example, President Biden recently issued an executive order that in part seeks to “...curtail the unfair use of non-compete clauses”.¹ While the impacts of NCA enforceability on workers (Balasubramanian

¹See <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/07/09/executive-order-on-promoting-competition-in-the-american-economy>.

et al., 2020; Starr, 2019; Johnson et al., 2021; Lipsitz and Starr, 2021) and firms (Samila and Sorenson, 2011; Starr et al., 2018; Jeffers, 2019; Marx and Fleming, 2012) have been examined extensively, the impact of NCAs on consumers has largely been ignored. We provide a unified framework to examine the welfare consequences of NCA enforceability on consumers, and provide analogies to assist policymakers in weighing the resulting tradeoff. While the prior literature has often leaned on behavioral reasons (Starr et al., 2020; Lipsitz and Starr, 2021) or liquidity constraints (Rauch and Watson, 2015; Wickelgren, 2018) as reasons that NCAs are used or have the potential to harm economic actors, we demonstrate that harms may arise even absent these explanations.

As motivation, we begin our analysis by documenting the extent to which NCAs and market outcomes are related. Using data from the US Census Bureau's County Business Patterns, we show that stricter NCA enforcement (making it easier for employers to legally enforce NCA contracts to prevent spin-offs) can result in increases in concentration.

To explain the mechanisms underlying this finding, we introduce a model of a firm operating in an industry in which the primary barrier to entry is industry-specific knowledge. The owner of this incumbent firm employs a worker who accumulates the knowledge necessary to compete in the industry, and who may ultimately decide to found their own competing firm (an employee spinoff). The owner and worker may agree to use an NCA, which legally prohibits the worker from spinning off, barring a negotiated buy-out payment (a payment from the worker to the firm which nullifies the NCA). NCAs have two effects on the market: on one hand, NCAs may increase the likelihood that the market remains a monopoly, preventing price competition. On the other hand, under an NCA, the owner does not face a holdup problem with respect to investment in their worker, and therefore is willing to make productivity-enhancing investments (for example, investments which reduce marginal costs) which may pass through to lower prices or higher quality for consumers.

We consider several comparative statics to analyze when a ban on NCAs may ultimately benefit consumers, and when such a ban may be harmful. We show that, in most cases, when

price setting in duopoly markets closely mirrors perfectly competitive price setting, and when investments made by firms do not pass through to consumer prices, NCA bans are more attractive. Then, we show that while enhanced benefits from investment may pass through to lower consumer prices, they also increase the use of NCAs. Counterintuitively, increased investment benefits may therefore exacerbate consumer harms from NCAs, especially when investment pass-through to consumers is low. Industries with low investment pass-through and high investment benefits are likely to find NCAs attractive, which suggests that solving a significant hold-up problem is not sufficient to render NCAs nonproblematic from a consumer perspective.

The tradeoff between productivity enhancement and market competition mirrors tradeoffs present in antitrust policy. The Horizontal Merger Guidelines published by the U.S. Department of Justice and the Federal Trade Commission note that “the Agencies consider whether cognizable efficiencies likely would be sufficient to reverse the merger’s potential to harm customers in the relevant market, e.g., by preventing price increases in that market.”² Cognizable efficiencies refer to enhancements in productivity (for example, cost reductions or increased potential for new innovations) that a merged firm will achieve but non-merged firms will not. To clarify this analogy, we introduce a basic model of pass-through (drawing on Weyl and Fabinger (2013)). Based on this model, we show that the tradeoffs inherent in merger analysis are also present in NCA analysis. We note that firm wide NCA use may prevent several spinoffs from occurring (especially if the firm has a large number of workers), which increases the potential for harms associated with NCAs.

We also consider an analogy to pay for delay agreements, which are agreements between a patent holder and a generic producer under which the generic producer agrees not to enter a market in exchange for monetary compensation. We show that arguments supporting the procompetitive nature of pay for delay agreements do not apply to NCAs, but arguments supporting the anticompetitive nature of pay for delay agreements do. The literature and

²The Horizontal Merger Guidelines are available online at <https://www.justice.gov/atr/file/810276/download>.

court decisions regarding pay for delay agreements are particularly important in assessing NCA policy when workers are sophisticated in their contract negotiations, with an ability to assess costs and benefits. In this case, high compensation for NCAs to sophisticated agents (such as the literature indicates CEOs may receive — see, e.g., Kini et al. (2020)) may indicate payment made to preserve anticompetitive market power.³

This paper is closest in spirit to Kräkel and Sliwka (2009), Rauch and Watson (2015), and Rauch (2016), all of which examine how NCAs may limit worker entrepreneurship. We contribute to this literature an analysis of the *competitive* effects that arise due to NCAs, and how those effects may be evaluated against the gains from investment. Hausman and Lavetti (2021) also consider the competitive effects of NCAs, demonstrating that NCAs cause establishment-level HHIs to decrease, and firm-level HHIs to increase in the context of physicians. Finally, our model may be viewed as similar to that of Bernheim and Whinston (1998), which assesses the reasons for and effects of exclusive contracting: in particular, they show that exclusive contracting may act as a way for a manufacturer and retailer to collude against other retailers. In our model, an NCA may be viewed as a way for a firm owner and worker to collude against consumers in the market by committing to share monopoly rents.

2 Noncompetes and Market Concentration

To motivate our analysis, we investigate the connection between market concentration and NCA enforceability. In particular, we present the finding that strict NCA enforceability (i.e., a high likelihood that a given NCA will be upheld by the court) causes high industrial concentration.⁴ We do not claim that our model is the *only* model that can explain this

³Note that large payments may instead be compensation for forgoing the option of working for a competitor, rather than forgoing the option of spinning off. The mechanisms explained in this paper apply, to some extent, in this situation, as well: insofar as the CEO may allow the outside firm to compete more effectively, there is a tradeoff between the incumbent investing in its CEO and increasing effective levels of competition in the market.

⁴By “enforceability of NCAs”, we mean the legal framework under which NCAs are assessed. At the lower extreme (lax enforceability), if an NCA is contended in front of a court, the court would rule that it will not be enforced: in other words, the worker is free to violate the NCA as it was written, and may therefore

finding, nor do we suggest that the magnitudes presented here are predictive of changes that would occur if NCA policy were changed. Instead, we present this analysis to demonstrate that the issues with which we are concerned (namely, the effect of NCAs on market power) are of a magnitude important enough to merit discussion.

To demonstrate this, we combine data on NCA enforceability with county-level industrial data from the US Census Bureau's County Business Patterns (CBP) dataset. To measure the enforceability of NCAs, we use data collected by Johnson et al. (2021), which extends data initially collected by Hausman and Lavetti (2021). The dataset is a state by year panel of a measure, normalized to a zero to one scale, which reflects the relative enforceability of NCAs in each state (initially constructed in Bishara (2010)). The exogeneity to product market and labor market outcomes (a prerequisite for the forthcoming analysis) has been extensively vetted in Johnson et al. (2021) and Hausman and Lavetti (2021): indeed, part of the analysis in Hausman and Lavetti (2021) is comparable to the analysis presented here (see Figure 2 of that paper). Each of the aforementioned two papers contains a more detailed description of the database.

The CBP dataset measures the number of establishments with different employment levels in each county, in each year, for each NAICS code. We use data from 1998 to 2018 to construct an employment-based HHI⁵ for 3-digit NAICS industries (indexed by i) in each county (indexed by c) in each year (indexed by y).⁶ An establishment, e , consists of a building with a full address that is designated a 3-digit NAICS code; thus, we let the county

compete by joining or founding a competing firm. On the other end of the spectrum (strict enforceability), NCAs are much more likely to be upheld by the court (in which case the worker may not compete, or may face penalties for doing so). The enforceability of NCAs varies at the state level. Examples of states which do not enforce NCAs (with limited exceptions) are California (see California Business and Professions Code Section 16600) and North Dakota (See North Dakota Century Code Section 9-08-06.). At the other side of the spectrum, many NCAs are explicitly admissible under Florida law (See Florida Statute 542.335.).

⁵Employment-based HHIs, rather than revenue or unit sales based HHIs, are used for data availability reasons. Widespread data on revenue and sales are not available at the level of detail required for this analysis. Employment serves as an imperfect proxy.

⁶We note that the HHIs we calculate here are likely different from what an antitrust agency would calculate when investigating a merger. First, that HHI would likely not be based on employment, which is used here for data availability reasons. Second, an antitrust agency would define markets according to the *Horizontal Merger Guidelines*, as opposed to using a blanket method which is employed in this context for consistency across industries and markets.

containing establishment e be defined by $c(e)$ and the industry containing e be defined by $i(e)$. For each establishment-year, we impute employment (as discussed in Appendix A) in that establishment-year, $n(e, y)$, and calculate the percentage share of employment in that establishment-year, $s(e, y)$, according to $s(e, y) = 100 * \frac{n(e, y)}{\sum_{j \in E(i(e), c(e))} n(j, y)}$, where $E(i(e), c(e))$ represents the set of all establishments in industry $i(e)$ and county $c(e)$. The *HHI* in industry i , county c , and year y is therefore calculated as:

$$HHI_{icy} = \sum_{e \in E(i, c)} s(e, y)^2.$$

We estimate the impact of NCA law changes on HHI using a distributed lag model (Schmidheiny and Siegloch, 2019), which examines the dynamic effects of NCA policy changes over time.⁷ The resulting coefficients, shown with corresponding 95% confidence intervals in Figure 1, may be interpreted identically to an event study model. The coefficients in years $t - k$ represent the effect on HHI in year y of a law which will be changed k years in the future. We therefore expect, if NCA law changes are conditionally exogenous to past HHI, that the “pre-trend” in years $t - k$ will be essentially flat, and roughly equal to zero. Indeed, this is what we see in each of those years.

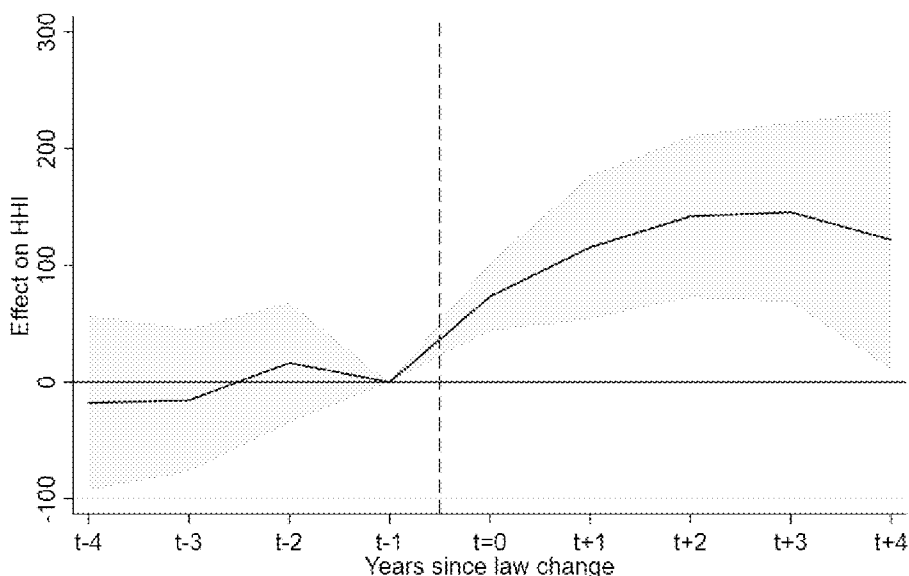
Following law changes (in years $t + k$), the coefficients may be interpreted as the effect of a law change which occurred k years in the past, beginning with the year of the law change, $t = 0$. Here, we see that concentration rises substantially following NCA law changes, and appears to remain at a high level several years into the future.⁸

The effect is consistent and relatively large: at its peak, an NCA score increase of 1 (i.e., from the minimum observed enforceability level to the maximum) is associated with an HHI increase of just under 150 points relative to baseline. Compare this, for example, with the

⁷See Appendix A for details of the model, as well as details on sample selection and data imputation.

⁸Note that an identical analysis based on 4-digit NAICS level HHIs yields similar results. Using finer NAICS classifications may more accurately identify markets where concentration may affect outcomes; however, this comes at the cost of losing data due to privacy-related data omissions introduced by the Census Bureau.

Figure 1: Distributed Lag Estimates of the Effect of NCA Enforceability Changes on HHI



This figure plots distributed lag model estimates of the dynamic effects of an NCA law change on employment HHI at the county-by-3 Digit NAICS level. The coefficient representing one year prior to law changes is normalized to zero. The underlying regressions are weighted by employment in the county-industry-year cell, and also include Census division by year by 3-Digit NAICS fixed effects. Standard errors are clustered by state.

U.S. Department of Justice & FTC 2010 Horizontal Merger Guidelines, which state that, in moderately and highly concentrated markets, HHI increases of 150 points “potentially raise significant competitive concerns and often warrant scrutiny.”⁹

Several limitations of this analysis must be noted. First, the CBP data only contains employment information for establishments, rather than firms. It is therefore possible that firm-level HHIs (which are more likely to be reflective of competitive conditions in an industry) would react differently to changes in NCA enforceability (as, e.g., is found in Hausman and Lavetti (2021)). Running a comparable analysis to the one presented above using the US Census Bureau’s Statistics of US Businesses (which presents data for firms, but only at the state level) returns directionally identical results. We present the analysis based on the CBP to isolate the local nature of markets which are likely to be affected by state-level law changes.

⁹The 2010 Horizontal Merger Guidelines are available at <https://www.justice.gov/atr/horizontal-merger-guidelines-08192010>.

Second, the HHI measure we construct is based on employment, rather than unit sales or revenue. While it is likely that shares based on employment are correlated with shares based on some measure of sales, the analysis would benefit from direct data on sales, which are not easily available for broad swaths of the American economy.

Finally, our calculation of HHI uses imputed employment, which may be relatively imprecise. Employment is not reported directly in the CBP to protect privacy, and we therefore use the midpoint of the reported cells as a proxy for actual employment.

3 Model

Motivated by the impact of NCA law changes on HHI, we develop a model in which the owner of an incumbent monopoly firm hires a worker that has the potential to found a competing firm using industry-specific knowledge naturally acquired on the job. The owner may make a costly, productivity-enhancing investment in the worker's human capital (e.g., providing skills training for the worker or imparting a valuable trade secret), but faces a hold-up problem (Williamson, 1975; Klein et al., 1978; Hart and Moore, 1990): if the worker is the claimant on rents associated with the investment (because she can threaten to spin off a competing firm or go to work for a competitor), the owner will not have sufficient incentive to invest. This situation is often used as a justification for the use of an NCA, which the firm may decide to use. By ensuring that a worker cannot spin off a competing firm or work for a competitor, an NCA ensures that owners receive returns on their investments, encouraging the owner to make that investment in the first place (Rubin and Shedd, 1981).

To model this explicitly, suppose that the worker and incumbent firm may agree to use an NCA when they begin the employment relationship. We denote the employment contract by $\{w, NCA\}$, where w is the wage that the owner pays the worker and $NCA \in \{0, 1\}$ indicates the existence (1) or nonexistence (0) of an NCA. If a contract includes an NCA, then the worker may not form a spinoff without permission from the incumbent.¹⁰ We allow

¹⁰We focus attention on employee spinoffs as a leading example, and to simplify the language used in the

contracts to be freely renegotiable, however: permission to found a spinoff may be granted if the worker makes a buy-out payment to the incumbent that is mutually agreed upon. We denote such a buy-out payment by BO . Similarly, without an NCA, the incumbent firm may make a *buy-in* payment to the worker to induce that worker *not* to found a spinoff: we denote such a side payment by BI .¹¹

After the initial contract decision is made, both players observe the spinoff marginal cost draw, c_s , which is drawn from the distribution $F(\cdot)$, and then the worker makes the spinoff decision. One input into this decision may be the incumbent firm's marginal cost, c_i , which is exogenously given and observed by all players at the start of the game. If the worker pursues a spinoff, then the incumbent firm (which continues to face a marginal cost of c_i) competes with the spinoff, which operates at marginal cost c_s . Instead, if the worker remains employed by the incumbent, then the single firm operates at marginal cost c_i . We implicitly assume that the incumbent cannot use the cost draw of the spinoff: if it could, spinoffs would not occur in equilibrium, which is not true in reality. In practice, this restriction can be explained by nontransferable aspects of costs, such as differences in business philosophies or organizational or managerial differences, which would be costly to change in the incumbent firm.

Prior to the realization of c_s , but after the initial contract, $\{w, NCA\}$, has been agreed upon, the owner of the incumbent firm must decide whether or not to make a costly investment. The owner may choose to pay a one-time, exogenously determined cost, κ , to make a binary, relationship-specific investment which increases profit for the owner-worker pair if they operate as one firm. The specificity of the investment means that if the worker spins off, all value from the investment will be lost, but if the worker remains employed by the incumbent, profit will be greater than without the investment.

paper. However, the worker may comparably bring their knowledge to a new or existing firm they do not own (and which is therefore not a spinoff), and the mechanisms presented in this paper are largely unchanged.

¹¹In reality, the renegotiability of an NCA or a payment made for a worker not to found a spinoff may be limited for a variety of reasons, such as liquidity constraints, reputational concerns, behavioral factors, or the law. We abstract away from these to focus attention on a model which explains the effects of NCAs without relying on, for example, behavioral explanations or liquidity constraints.

Altogether, the formal timing of the game is as follows. First, the incumbent owner makes a take-it-or-leave-it contract offer, $\{w, NCA\}$, to the worker, where $w \in \mathbb{R}$ and $NCA \in \{0, 1\}$. The worker may accept or reject the contract. If the worker rejects, each agent receives their outside option (normalized to zero for each agent) and the game ends. If the worker accepts the contract, then the owner pays the worker w and makes its investment decision, $I \in \{0, 1\}$. After the owner's investment decision, the worker's spinoff marginal cost, c_s , is observed by both parties. When $NCA = 0$, the worker may spin off, or may make a take-it-or-leave-it offer¹² of a buy-in payment to the owner, BI , which, if accepted, will cause the worker to remain employed by the incumbent. With $NCA = 1$, the worker may decide to remain employed by the owner, or make a take-it-or-leave-it offer, BO , to the owner which, if accepted, allows the worker to buy out of the NCA and spin off. If the worker remains employed by the incumbent, then the owner receives gross profit $\Pi_{i,M}(I)$ from the final goods market,¹³ where M represents that the market is a monopoly.¹⁴ If the worker spins off, then the worker receives gross profit $\Pi_s(c_s)$ and the owner receives gross profit $\Pi_{i,D}(c_s)$ from the final goods market.

We assume that agents are risk-neutral, seek to maximize payoffs, and that information is perfect. To allow for a variety of market structures we make the following assumptions:

Assumption 1. *We assume that*

a. $\Pi_{i,M}(1) > \Pi_{i,M}(0) \geq \Pi_{i,D}(c_s)$ for all c_i, c_s

b. $\frac{d\Pi_s(c_s)}{dc_s} < 0$

¹²The assumption that buy-in and buy-out payment offers are take-it-or-leave-it, made by the worker, is unimportant: the alternative assumptions that the owner makes the offer, or that bargaining occurs, would not change the predictions of our model except for the ex-post allocation of cash between the worker and owner.

¹³This formulation of the incumbent firm's monopoly profit tacitly assumes that profit does not rely on c_s (the spinoff's marginal cost). Many of our results do not rely on this assumption; results are qualitatively equivalent if we instead assume that $\Pi_{i,M}(\cdot)$ depends on c_s . We use the current formulation for tractability and notational brevity. Similarly, since c_i is exogenously given at the outset of the model, and not a major focus of exploration in this paper, we omit it when writing the gross profit functions for brevity.

¹⁴To highlight the competitive issues involved in our model, we use the subscripts M and D to denote "monopoly" and "duopoly" markets. However, M may alternatively be understood to signify a market with $N \in \mathbb{N}$ firms, and D to signify a market with $N + 1$ firms. As long as Assumption 1 is satisfied, our model is robust across a variety of oligopoly settings where other competing firms may exist.

$$c. \frac{d\Pi_{i,D}(c_s)}{dc_s} > 0 \text{ but } \frac{d(\Pi_{i,D}(c_s) + \Pi_s(c_s))}{dc_s} < 0$$

d. For all I , there exists a $\bar{c}_s(I)$ such that $\Pi_s(c_s) + \Pi_{i,D}(c_s) > \Pi_{i,M}(I)$ for all $c_s < \bar{c}_s(I)$ and $\Pi_s(c_s) + \Pi_{i,D}(c_s) < \Pi_{i,M}(I)$ for all $c_s > \bar{c}_s(I)$.¹⁵

Qualitatively, Assumption 1a requires that investment by the incumbent owner improves gross profits when the incumbent firm is a monopoly, as well as requiring that monopoly profits exceed duopoly profits for the incumbent owner. Assumption 1b simply ensures that spinoff profits are decreasing in marginal cost. Assumption 1c states that the incumbent owner's profits are increasing in *spinoff* marginal costs in a duopoly, but the sum of gross profit in the market is decreasing in spinoff costs. Finally, Assumption 1d requires a threshold value of c_s , above which monopoly outperforms duopoly from the perspective of the pair, and below which duopoly outperforms monopoly.

4 Equilibrium

Our solution concept is Subgame Perfect Nash Equilibrium. Solving the model backwards, the last stage of the game is the spinoff decision of the worker. Given frictionless renegotiation, this decision is straightforward: whenever the sum of profits in duopoly is greater than the sum of profits in monopoly, the worker finds a spinoff. Formally, the worker spins off whenever

$$\Pi_s(c_s) + \Pi_{i,D}(c_s) > \Pi_{i,M}(I).$$

Buy-in or buy-out payments are made if necessary to ensure that the owner receives as much as she would have had the “default” action (i.e., the market remaining a monopoly when $NCA = 1$ or the market becoming a duopoly when $NCA = 0$) occurred. So, when $NCA = 1$, the incumbent would receive $\Pi_{i,M}(I)$ if the worker does not spin off. If the worker opts to spin off, the payment BO is set so that $\Pi_{i,D}(c_s) + BO = \Pi_{i,M}(I)$. Note

¹⁵The direction of the inequalities stem from Assumption 1c.

that, by Assumption 1a, $\Pi_{i,M}(I) \geq \Pi_{i,D}(c_s)$, so $BO \geq 0$. Similarly, when $NCA = 0$, the incumbent would receive $\Pi_{i,D}(c_s)$ if the worker spins off, and so BI is set such that $\Pi_{i,M}(I) + BI = \Pi_{i,D}(c_s)$. Again, since $\Pi_{i,M}(I) \geq \Pi_{i,D}(c_s)$, $BI \leq 0$ (i.e., BI is a payment made from the incumbent to the worker).

Beyond the spinoff decision, the next important feature of equilibrium is the hold-up problem (the relationship between NCA use and the incumbent owner's investment decision). Solving the investment decision subgame generates the following result:

Lemma 1. *If $NCA = 1$, then $I = 1$ if and only if $\kappa \leq \Pi_{i,M}(1) - \Pi_{i,M}(0)$. Instead, if $NCA = 0$, then $I = 0$ for all $\kappa > 0$.*

All proofs may be found in Appendix B.

The investment decision is intuitive: with an NCA, regardless of whether the worker spins off, the owner controls the rents related to the investment and will be compensated accordingly. Therefore, the only thing that matters in the owner's investment decision is whether the investment contributes more to monopoly profits than it costs. Without an NCA, since the worker effectively controls the rents related to investment (due to her ability to credibly threaten to spin off), the owner is unable to recoup any of the rents associated with investment and is therefore unwilling to invest at any cost; in other words, the holdup problem prevents investment without an NCA.

Turning to the contracting stage, both the owner and the worker know that an NCA, with $\kappa \leq \Pi_{i,M}(1) - \Pi_{i,M}(0)$, results in investment by the owner. Given that renegotiation is frictionless, an NCA is used whenever the expected value for the pair is greater with an NCA than without. Naturally, if $\kappa > \Pi_{i,M}(1) - \Pi_{i,M}(0)$ so that investment does not occur with or without an NCA (by Lemma 1), then use of an NCA makes no difference. However, if $\kappa \leq \Pi_{i,M}(1) - \Pi_{i,M}(0)$, then the pair may or may not agree to an NCA in equilibrium:

Proposition 1. *If $\kappa > \Pi_{i,M}(1) - \Pi_{i,M}(0)$, then frictionless renegotiations cause contracts to be "inconsequential": the owner and worker receive identical expected utility with $NCA = 0$ or $NCA = 1$.*

Instead, if $\kappa \leq \Pi_{i,M}(1) - \Pi_{i,M}(0)$, then there exists a $\tilde{\kappa} \leq \Pi_{i,M}(1) - \Pi_{i,M}(0)$ so that $NCA = 1$ in equilibrium if $\kappa \leq \tilde{\kappa}$ and $NCA = 0$ otherwise.

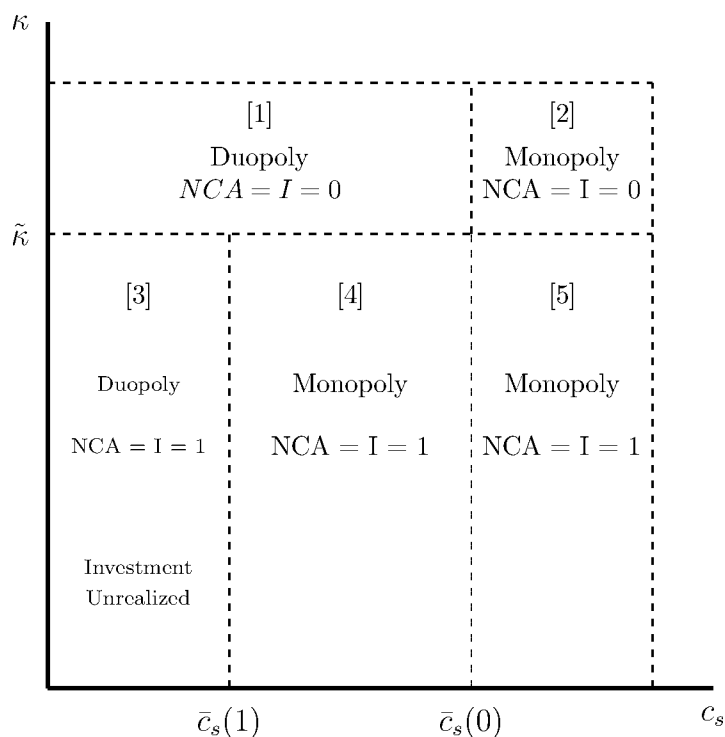
To understand why there exist values of κ close to but potentially less than $\Pi_{i,M}(1) - \Pi_{i,M}(0)$ where the owner and worker choose not to use an NCA, and then invest, consider a case where $\kappa = \Pi_{i,M}(1) - \Pi_{i,M}(0) - \epsilon$ for some small $\epsilon > 0$. In this case, an NCA followed by investment generates only a minor gain when the realization of c_s is such that a monopoly occurs; however, the cost κ generates no benefit when the realization of c_s is such that a duopoly occurs.¹⁶ Thus, this loss in surplus from an investment that results in a duopoly restricts the pair's willingness to adopt an NCA in the initial contracting stage where the pair evaluates contracts based on the joint expected profit.

To complete the game and fully characterize the equilibrium outcomes across all realizations of c_s , we now discuss critical values that we use extensively: $\bar{c}_s(I)$ (for $I = 0$ and 1), as defined in Assumption 1d. The critical values $\bar{c}_s(I)$ represent the values of c_s such that $\Pi_{i,D}(c_s) + \Pi_s(c_s) = \Pi_{i,M}(I)$. Therefore, $\Pi_{i,D}(c_s) + \Pi_s(c_s) > \Pi_{i,M}(I)$ if and only if $c_s < \bar{c}_s(I)$. Furthermore, Assumption 1a implies that $\bar{c}_s(1) < \bar{c}_s(0)$. Intuitively, $\bar{c}_s(I)$ represents the cutoff at which duopoly joint profits equal monopoly profits (so the pair is indifferent between the worker founding a spinoff or not).

Given these thresholds, we describe the realized market outcomes (that depend on the realization of c_s) in Figure 2. In cells [1] and [2], we see that $\kappa > \bar{\kappa}$ so that NCAs are not used and investment does not occur; in addition, market structure depends on the total profit comparison so that a duopoly occurs whenever $c_s < \bar{c}_s(0)$ (i.e., whenever $\Pi_{i,D}(c_s) + \Pi_s(c_s) > \Pi_{i,M}(0)$). In cells [3]-[5], NCAs are used and the owner invests in the worker; however, the benefits of investment are only realized when a monopoly occurs, which happens whenever $\Pi_{i,D}(c_s) + \Pi_s(c_s) < \Pi_{i,M}(1)$ (i.e., whenever $c_s > \bar{c}_s(1)$).

¹⁶Of course, the difference in monopoly profit from investment, $\Pi_{i,M}(1) - \Pi_{i,M}(0)$, also impacts the c_s required for a duopoly to occur under $I = 1$ versus $I = 0$. To see a richer discussion, consider Equation (5) and the paragraphs that follow in Appendix B.

Figure 2: Equilibrium Realizations



5 Policy and Heterogeneity

To illustrate the extent to which our model can inform policy, we consider an outright ban on NCAs.¹⁷ Our model does not consider the full range of economic benefits and harms that would result from such a ban; however, it does shed light on one particular tradeoff for consumers when NCAs are banned: consumers benefit from an increased likelihood that a duopoly will emerge, but they are harmed by decreased investment on the part of the owner, which may lower marginal costs and pass through to lower prices.

In this section, we first consider comparative statics with respect to characteristics of firms and markets which may drive a ban on NCAs to be more or less attractive when measured by consumer welfare. In Section 5.1, we look at the nature of competition in duopolies and how investments pass through to consumers. In Section 5.2, we look at the

¹⁷This policy is comparable to the “Ban on Non-Compete Agreements Amendment Act of 2020”, recently passed in Washington, DC. Note that most recently passed bans on NCAs (e.g., in Virginia, Maryland, Illinois, Massachusetts, and others) apply only to workers making below a certain income threshold, or who satisfy other restrictions.

size of the benefit from investment. Then, we derive a generalized expression indicating when NCAs should optimally be banned, based solely on a consumer welfare criterion.

5.1 Competition and Pass-Through

The tradeoff related to NCA policy (as it pertains to *consumer* welfare) is governed directly by market structure and the benefits from firm investment. The extent to which the harm and benefit pass through to consumers is unclear without further details of the market.

On the market power side of the tradeoff, the primary characteristic of the market which impacts the extent to which NCA policy impacts consumers is the nature of duopoly competition between the incumbent and the spinoff. Consider, for example, a spectrum where, at one end, the duopoly competes *à la* homogenous Bertrand, and at the other end, *à la* homogeneous Cournot. This distinction is important because differing levels of duopoly competition change the market power harms of NCAs. Broadly speaking, when the difference between a monopoly and duopoly is large (because the firms Bertrand compete), a given NCA will cause a larger gross harm than if that difference is small (Cournot competition), since price differences are amplified when moving from a monopoly incumbent to a more competitive market. However, note that the nature of competition also affects the cutoff values which determine when spinoffs will occur, as well as the investment decision and the decision of whether or not to use an NCA in the first place.

To consider these features in our context, we borrow the conduct parameter and pass-through approach from Weyl and Fabinger (2013). This allows us to abstract away from the exact nature of competitive interaction and investment benefits by assuming that there exists a market power parameter (known as the conduct parameter), θ , and an investment pass-through parameter, ρ , that captures the extent to which investment benefits pass onto consumers (e.g., if investment reduces marginal costs by ΔC , then $\rho \cdot \Delta C$ of those savings are passed onto consumers in the form of lower prices).¹⁸ The conduct parameter $\theta \in [0, 1]$

¹⁸On the cost savings side of the tradeoff, several market characteristics may impact pass-through to

captures the amount of market power so that $\theta = 0$ corresponds to homogenous Bertrand competition, $\theta = 1$ corresponds to a monopoly or perfect collusion, and $\theta = 1/N$ corresponds to homogenous Cournot competition.¹⁹

Given this approach, we immediately see that (a) the duopoly price is increasing in market power ($\frac{dp_D}{d\theta} > 0$) while the monopoly price is unaffected by duopoly conduct ($\frac{dp_M(I)}{d\theta} = 0$) for $I \in \{0, 1\}$, and (b) prices are decreasing in consumer investment pass-through only when investment occurs ($\frac{dp_M(I=1)}{d\rho} < 0$ and $\frac{dp_M(0)}{d\rho} = \frac{dp_D}{d\rho} = 0$). In addition, we now rewrite the profit functions so that the incumbent receives $\Pi_{i,D}(c_s, \theta)$ under a duopoly, the spinoff receives $\Pi_s(c_s, \theta)$ under a duopoly, and the incumbent receives $\Pi_M(I, \rho)$ under a monopoly with $\Pi_M(0, \rho) = \Pi_M(I, 1)$ capturing the cases where investment generates no profit either because investment does not occur ($\Pi_M(0, \rho)$) or because all benefits from investment pass-through to consumers ($\Pi_M(I, 1)$). Under this characterization, we have that (a) $\frac{d\Pi_{i,D}(c_s, \theta)}{d\theta}, \frac{d\Pi_s(c_s, \theta)}{d\theta} > 0$ and $\frac{d\Pi_M(I, \rho)}{d\theta} = 0$ and (b) $\frac{d\Pi_M(I, \rho)}{d\rho} \leq 0$ and $\frac{d\Pi_{i,D}(c_s, \theta)}{d\rho}, \frac{d\Pi_s(c_s, \theta)}{d\rho} = 0$. Altogether we see that, conditional on market structure and c_s , θ only impacts duopoly markets and ρ only impacts the monopoly market; however, these parameters will play important roles in determining consumer surplus, NCA tradeoffs, and market outcome thresholds.

Since the model introduced in Section 3 was agnostic about the nature of competition and investment pass-through, we examine the effect of θ and ρ on outcomes. We first focus on *within-box* differences, where boxes [1]-[5] are those depicted in Figure 2. In other words, we first take as given the NCA use and investment decisions, and focus on how consumer welfare changes when competition and pass-through change, considering those changes across NCA policies.

Regardless of NCA policy, [1] will always contain a duopoly without investment, [2] will always contain a monopoly without investment, and [3] will always contain a duopoly with (unrealized) investment, and all three are therefore uninteresting for the purposes of this section. In [4], both θ and ρ play a role: when NCAs are allowed, the market is a monopoly

consumers (see, e.g., Gron and Swenson (2000); Bonnet et al. (2013); Miller et al. (2017)).

¹⁹Weyl and Fabinger (2013) provide a rich set of micro-foundations that support such a parameterization.

with (realized) investment, whereas when NCAs are banned, the market is a duopoly without investment. In other words, the differential impact on price of banning NCAs is the sum of the differential impacts from θ and ρ . Finally, in [5], the market will always be a monopoly (whether or not NCAs are allowed), and therefore only ρ plays a role: when ρ is positive (i.e., when investment pass-through to consumers exists), an NCA ban will harm consumers by reducing investment without increasing market competition.²⁰

Moving now to changes in the boundaries between the boxes, consider first the thresholds $\bar{c}_s(0)$ and $\bar{c}_s(1)$. Recall that these thresholds are the critical values of the spinoff marginal cost which govern the spinoff decision of the worker, depending on the investment decision of the firm. These thresholds are determined by the joint profitability of the incumbent and spinoff firms, versus the incumbent firm remaining a monopolist: $\bar{c}_s(I)$ is given by $\Pi_s(\bar{c}_s(I), \theta) + \Pi_{i,D}(\bar{c}_s(I), \theta) = \Pi_M(I, \rho)$. Therefore, as the duopoly structure gains market power, $\Pi_s(\bar{c}_s(I), \theta) + \Pi_{i,D}(\bar{c}_s(I), \theta)$ increases, meaning that the value of c_s which makes the two sides equal is now larger. Put simply, greater market power in the duopoly allows for more spinoffs so that $\frac{d\bar{c}_s(I)}{d\theta} > 0$ for $I = 0, 1$.

The value of ρ does not play a role in determination of $\bar{c}_s(0)$, since investment does not play a role. However, ρ does play a role when considering $\bar{c}_s(1)$, which is defined by $\Pi_s(\bar{c}_s(1), \theta) + \Pi_{i,D}(\bar{c}_s(1), \theta) = \Pi_M(1, \rho)$. In this case, as the investment pass-through to consumers increases, $\Pi_M(1, \rho)$ decreases, meaning that the value of c_s which makes the two sides equal is now smaller. In other words, for large ρ , investment benefits pass-through in large part to consumers meaning that producers benefit less from investment. Therefore, the difference between $\Pi_M(1, \rho)$ and $\Pi_M(0, \rho)$ is smaller so that a large ρ results in a lower $\bar{c}_s(1)$ that approaches $\bar{c}_s(0)$.

To summarize, increased competition within the duopoly (through a reduction in market

²⁰Given that a ban on NCAs harms consumers for equilibria in [5], it is worth noting that we implicitly assume that $c_i \geq \bar{c}_s(1)$. In this case, if the spinoff also has no chance of being more efficient than the incumbent (so that every c_s draw is greater than c_i and $\bar{c}_s(1)$), then only the outcomes in [2] and [5] occur so that a ban on NCAs is always detrimental to consumers. Altogether, this implies that NCAs that foreclose inefficient spinoffs are always beneficial to consumers as they prevent the investment holdup and allow investment benefits to pass-through to consumers.

power captured by θ) lowers both $\bar{c}_s(0)$ and $\bar{c}_s(1)$. More competitive duopoly structures therefore decrease the attractiveness of an NCA ban from a consumer welfare perspective by increasing the mass that will be on [5], which is where the unambiguous benefits of NCAs accrue. Factors that drive larger investment pass-through to consumers will decrease $\bar{c}_s(1)$, but do not change $\bar{c}_s(0)$: therefore, when investment pass-through to consumers is larger, an NCA ban is less attractive as the mass on [5] increases.

Finally, we consider $\tilde{\kappa}$. The value $\tilde{\kappa}$ represents the critical value of κ , the cost of investment, at which the owner/worker pair is indifferent between using an NCA and not. This determination is made by considering all the effects of investing: a monopoly with reduced costs and a diminished probability of a spinoff (see Equation (5)). Both investment pass-through to consumers and market power play a role in this determination. First, they impact the thresholds $\bar{c}_s(0)$ and $\bar{c}_s(1)$, as described above, which factor into $\tilde{\kappa}$. They also impact the magnitudes of profit, which determine whether the firm wishes to use an NCA or not. Ultimately, the impact of increased duopoly competition and increased investment pass-through is ambiguous. However, from a policy perspective, it is less important to pin down the exact nature of the impact on $\tilde{\kappa}$: while $\tilde{\kappa}$ influences whether or not a firm uses an NCA, it does not impact whether it is harmful or beneficial. In other words, consider two states of the world with identical values of $\tilde{\kappa}$: one in which $\kappa < \tilde{\kappa}$ (and therefore the firm uses an NCA) and one in which $\kappa > \tilde{\kappa}$ (and therefore the firm does not). If the firm in the latter state of the world accidentally used an NCA, it would be equally beneficial or harmful to consumers as the firm's NCA in the state of the world where the profit-maximizing decision was not to use an NCA.

Corollary 1. *In summary, we formally have the following results:*

1. *Greater duopoly competition (conditional on the spinoff occurring) implies that monopoly outcomes — with and without investment — are more likely to occur in equilibrium:*

$$\frac{d\bar{c}_s(I)}{d\theta} > 0 \text{ for } I = 0, 1.$$

2. Greater investment pass-through to consumers does not impact the spinoff decision without investment (i.e., $\frac{d\bar{c}_s(0)}{d\rho} = 0$). However, greater investment pass-through to consumers causes spinoff decisions with investment to approach spinoff decisions without investment (i.e., $\frac{d\bar{c}_s(1)}{d\rho} > 0$ and $\bar{c}_s(1) = \bar{c}_s(0)$ when $\rho = 1$). Graphically, as ρ increases, [3] expands while [4] contracts in Figure 2, and when $\rho = 1$, [4] disappears so that [3] and [5] mirror [1] and [2] in Figure 2.
3. Greater competition between the incumbent and the spinoff or greater investment pass-through to consumers have ambiguous effects on NCA usage: ρ and θ have ambiguous effects on $\tilde{\kappa}$.

Taken all together, under what circumstances should policymakers expect that a ban on NCAs would be most likely to benefit consumers? When duopoly markets generate relatively low prices (competition is fierce), consumers are helped more by a ban on NCAs because the duopoly is so beneficial; however, firms are more likely to retain their monopoly structure since the resulting duopoly would be less profitable for them. On the other hand, when investment pass-through is high, consumers are hurt more by a ban on NCAs because highly beneficial investments will never occur; nevertheless, consumers are helped more by a ban on NCAs because spinoffs are relatively more likely to occur.

While the circumstances of any given industry, occupation, or slice of the labor market are important to consider, it is likely that the first-order impacts (i.e., that Bertrand-style duopolies cause an NCA ban to benefit consumers more than in a Cournot-style duopoly, and that high investment pass-through causes an NCA ban to hurt consumers more than low investment pass-through) dominate the second-order impacts. If this is indeed the case, when considering an NCA ban, policymakers should weight more highly slices of the economy where increased competition is likely to have a large impact on prices and where investment pass-through to consumers is relatively low.

5.2 The Benefit of Investment

The benefits or harms of banning NCAs are also affected by the relative payoff of investment. In our baseline model, we operationalize investment by assuming that the monopoly firm's profit (not necessarily net of the cost of investment itself) increases after it invests in the worker. In Section 5.1, we further assume that some percentage of the investment benefit, ρ , passes through to the consumer. We now introduce a parameter, $\iota > 0$, which represents the magnitude of the surplus benefit from investment so that $\Pi_M(I, \rho)$ from 5.1 becomes $\Pi_M(I, \rho, \iota)$, where ι captures the surplus generated in the market from investment (either through a cost reduction or an increase in demand). We accordingly assume that $\frac{d\Pi_M(I, \rho, \iota)}{d\iota} > 0$, and we now have that $\Pi_M(0, \rho, \iota)$, $\Pi_M(I, 1, \iota)$, and $\Pi_M(I, \rho, 0)$ all capture the case where the incumbent earns no benefit from investment.

Ignoring for the time being the impacts of ι on $\tilde{\kappa}$ and $\bar{c}_s(I)$, the first order impact of an increase in ι on consumers occurs either through a price reduction (from investment reducing costs) or through an increase in price and consumer surplus due to an increase in demand. Thus, while the price effect is ambiguous (depending on the type of investment), consumers clearly benefit from investments that generate greater surplus. At the same time, the prices $p_M(0)$ and p_D are unaffected.

The size of ι is also going to impact the size of each of the boxes in Figure 2. First consider the values of $\bar{c}_s(0)$ and $\bar{c}_s(1)$ which generate the vertical lines that divide [3] from [4], and [1] and [4] from [2] and [5]. Extending the expression to include ι , $\bar{c}_s(I)$ is defined by $\Pi_{i,D}(\bar{c}_s(I)) + \Pi_s(\bar{c}_s(I)) = \Pi_M(I, \rho, \iota)$. Given that no investment occurs when $I = 0$, we clearly have that $\bar{c}_s(0)$ is unaffected by ι . However, the value of $\bar{c}_s(1)$ is affected by changes in ι . Since joint profits of the duopoly are decreasing in c_s , it must be the case that $\frac{d\bar{c}_s(1)}{d\iota} < 0$. Again, this is intuitive: as the value of investment increases, spinoffs are less likely to occur, as the monopoly is more profitable and it is more difficult for the pair to find an acceptable buyout payment which makes both parties better off. Hence, under an NCA (and only under an NCA), greater ι implies that a monopoly occurs under a larger mass of

c_s draws (increasing the size of [4] and decreasing the size of [3]).

Corollary 2. *In summary, we formally have the following results:*

1. *A larger surplus from investment increases consumers surplus (either through a price reduction or through an increase in demand): $\frac{dCS}{d\iota} > 0$.*
2. *Greater surplus from investment does not impact the spinoff decision without investment (i.e., $\frac{d\tilde{c}_s(0)}{d\iota} = 0$). However, greater surplus from investment causes spinoff decisions with investment to shrink (i.e., $\frac{d\tilde{c}_s(1)}{d\iota} < 0$). Graphically, as ι increases, [4] expands while [3] contracts in Figure 2.*

Lastly, consider $\tilde{\kappa}$ which generates the horizontal line that divides [1] and [2] from [3]-[5]. In this case we obtain the following result:

Proposition 2. *Greater benefits from investment result in greater NCA use: (i.e., $\frac{d\tilde{\kappa}}{d\iota} \geq 0$).*

How do these relationships inform NCA policy? Based solely on the first order effect of ι on consumer welfare when NCAs are allowed, a ban becomes more harmful as ι increases (since the ban prevents investment which generates added consumer benefit with higher ι). While there are impacts on the thresholds to consider, when ι is low, a ban is (weakly) beneficial to consumers (with certainty when $\iota = 0$). As ι increases, a ban becomes less beneficial as $p_M(1)$ decreases or quality increases.

The effects on thresholds are more nuanced. Counterintuitively, and in contrast with historical arguments about NCAs and the hold-up problem, it is possible that larger benefits from investment may actually harm consumers. For example, when investment pass-through to consumers is low, the increase in monopoly prevalence due to greater investment benefit may hurt consumers more (by encouraging monopoly markets) than they help (due to low pass-through of investment). More formally, consider $\rho = 0$ with marginal cost reducing investment. In this case, $p_M(0) = p_M(1)$, since none of the benefit from investment passes through to consumers, and therefore $\frac{dp_M(0) - p_M(1)}{d\iota} = 0$. However, consumers are still harmed

by the increase in the probability that a monopoly will occur so that the expected price increases in the size of the investment.²¹ For small values of ρ , this relationship still holds, since $\frac{dp_M(0)-p_M(1)}{dt}$ will be small, while harm due to monopoly may still be quite large, though at some point the relationship may flip such that increased investment benefits consumers for sufficiently high levels of investment pass-through to consumers.

Formally, Corollary 3 summarizes this result. For simplicity in this Corollary, we assume that investment benefits pass through to prices (as opposed to product quality), such that monopoly price may be written as a function of ρ and ι , i.e., $p_M(\rho, \iota)$.

Corollary 3. *There exists $\tilde{\rho} > 0$ such that $\frac{d(1-\mathcal{P}^{NCA})p_M(\rho, \iota)+\mathcal{P}^{NCA}p_D}{dt} > 0$ if and only if $\rho < \tilde{\rho}$: increases in the benefits from investment cause consumers to pay higher expected prices when investment pass-through to consumers is sufficiently low.*

A similar result is that changes in ι affect $\tilde{\kappa}$ (see Proposition 2). Therefore, there is an analogous “extensive margin” result with the benefit from investment: when ι increases, use of NCAs increases due to an increase in $\tilde{\kappa}$. Whenever the additional harms caused by increased use of NCAs outweigh the benefits obtained by consumers (due to low pass-through, for example), the increase in ι will further harm consumers.

These counterintuitive results are especially important because industries with low investment pass-through to consumers and high benefit of investment are exactly those where protection of investments with NCAs is most important, since firms reap substantial benefit from being able to invest. In other words, while a high rate of NCA use is not proof positive, it is evidence that is consistent with consumer harm due to NCAs: if investment pass-through were high (which would be most beneficial to consumers), then firms would have less incentive to use NCAs, as they would reap less of their benefit.

²¹Formally, using notation from the proofs of Propositions 1 and 2, we have that $\frac{d(1-\mathcal{P}^{NCA})p_M(1)+\mathcal{P}^{NCA}p_D}{dt} > 0$ since $\frac{\mathcal{P}^{NCA}}{dt} < 0$.

5.3 Analysis of an NCA Ban

Graphically, an NCA ban means that the investment decisions in cells [3]-[5] in Figure 2 change to the respective investment decisions in cells [1] and [2]: in all of cells [3]-[5], $NCA = I = 0$ when NCAs are banned (Lemma 1). Additionally, the market in cell [4] will be a duopoly: since investment does not occur even for $\kappa < \tilde{\kappa}$ when NCAs are banned, the benefits of investment do not drive the market structure to remain a monopoly.

Consider an extension of the model: at time $t = 0$, the policymaker chooses policy (whether or not to ban NCAs). Then, after observing the policymakers decision, a representative firm and worker decide whether or not to use an NCA. Time $t = 1$ is the first period of the firm's existence and production, during which a cost-saving investment may be made. At time $t = 2$, the worker decides whether or not to spin off. Finally, let time $t = 3$ be a second period of production where the worker may have spun off to form a new firm.

While there may be further concerns for a policymaker, we consider a policymaker focused solely on price: what is the expected change in price due to allowing NCAs versus not? Simplifying, let \mathcal{P}^{NCA} be the probability of a spinoff with an NCA (i.e., $Pr(c_s < \bar{c}_s(1))$), and let \mathcal{P}^{Free} be the probability of a spinoff when the worker does not have an NCA (i.e., $Pr(c_s < \bar{c}_s(0))$).²² Let $\Delta\theta \cdot \frac{P}{-\epsilon}$ be the change in price due to market power²³ and $\rho\Delta C$ be the change in price due to increased investment (incentivized by NCAs)²⁴. Here, $\Delta\theta \cdot \frac{P}{-\epsilon}$ is interpreted as the difference in price between a monopoly and duopoly market structure, holding fixed any reductions in cost due to investment. Similarly, $\rho\Delta C$ is interpreted as the difference in price for a monopolist, with and without the investment having been made. Finally, let γ be the probability that the representative firm signs an NCA (supposing,

²²At the extreme, if spinoffs never occur (because $\bar{c}_s(1) > c_i$, for example), NCAs are clearly beneficial and a ban would harm consumers. This is because there is no market power benefit to banning NCAs, but a ban would eliminate beneficial investment. As the probability of a spinoff increases (e.g., because the distribution of c_s shifts downwards), the benefits of an NCA ban decrease, since the negative effects of NCAs with respect to market power increase.

²³Under the conduct parameter approach, equilibrium prices are given by $P = C + \theta \frac{P}{-\epsilon}$ so that a change in market power by $\Delta\theta$ (holding costs fixed) results in a price change of $\Delta\theta \cdot \frac{P}{-\epsilon}$.

²⁴We use ρ as pass-through akin to our explanation of investment benefits in Section 5.1

for example, that exogenous parameters may be randomly drawn prior to $t = 1$, driving differences in κ).

The expected price in the market when NCAs are allowed is therefore given by

$$P + \gamma \left[\mathcal{P}^{NCA} \left(-\Delta\theta \cdot \frac{P}{-\epsilon} \right) + (1 - \mathcal{P}^{NCA})(-\rho\Delta C) \right] + (1 - \gamma) \left[\mathcal{P}^{Free} \left(-\Delta\theta \cdot \frac{P}{-\epsilon} \right) \right],$$

where P represents the price that would prevail in the monopoly market with no cost savings.

The expected price in the market when NCAs are not allowed is given by

$$P + \mathcal{P}^{Free} \left(-\Delta\theta \cdot \frac{P}{-\epsilon} \right).$$

The expected price difference is therefore

$$E(\Delta P) = \gamma \left[(\mathcal{P}^{NCA} - \mathcal{P}^{Free}) \left(-\Delta\theta \cdot \frac{P}{-\epsilon} \right) + (1 - \mathcal{P}^{NCA})(-\rho\Delta C) \right],$$

and the policymaker optimally bans NCAs (based solely on the price criterion) whenever

$$(\mathcal{P}^{Free} - \mathcal{P}^{NCA})\Delta\theta \cdot \frac{P}{-\epsilon} > (1 - \mathcal{P}^{NCA})\rho\Delta C. \quad (1)$$

6 Analogies with Antitrust

In this section, we provide two analogies to antitrust theory and practice in order to guide policymakers. The link between NCAs and consumers has not been deeply explored in policy spaces. Therefore, we intend for this section to allow policymakers to leverage their familiarity with two more deeply explored practices to better understand the mechanisms by which NCAs affect consumers. The analogies are not meant to be perfect facsimilies of the NCA mechanisms described above: rather, they are meant to be demonstrative of important pieces of the puzzle, especially with respect to how those pieces have been treated by policy

in the past. We reiterate that NCAs may have many effects, including effects on workers, firms, and entrepreneurs, and the analogies presented here are only intended to consider the impacts on *consumers*, in line with the rest of the paper. We therefore do not claim that the analogies presented here contain a full view of the many aspects of the economy which may be considered by policymakers, all of which may be part of their decisions.

6.1 Merger Analysis

The first analogy is to merger analysis. In the past several years, one of the primary goals of merger analysis has been to quantify the tradeoff between harms and benefits to consumers, often in the form of price effects. In a horizontal merger (the merger of two firms selling substitutable goods or services), harms to consumers are typically due to increased market power in the market in which they participate. Benefits to consumers which have historically been used to justify otherwise anticompetitive mergers are often due to variable cost efficiencies, which are synergies from the merger that decrease variable costs and therefore may theoretically decrease equilibrium prices.

Merger analysis typically occurs on a case-by-case basis, whereas the conversation surrounding NCAs has been focused more on whether NCAs should be allowed for broader swathes of the workforce (e.g., if NCAs should be banned for all workers, for low-wage workers, allowed for all workers, etc.). To put the two analyses on equal footing, we consider a hypothetical choice facing a policymaker: should all mergers passing a given bright line test be banned? For example, one could imagine a ban on all mergers for retailers selling at least one overlapping product.

Consider a simple model of mergers, comparable to the simplified model of NCAs introduced in Section 5.3. Let $t = 0$ be a baseline period in which policy is made (i.e., whether mergers are banned). Let $t = 1$ be the first period of two representative companies' existence and production, let $t = 2$ be the point in time at which the companies decide whether or not to merge, and let $t = 3$ be another period of production during which the companies

may have merged. Abstracting away from case-by-case antitrust decisions, policymakers at $t = 0$ ask: is the expected value of harm to consumers that is generated by allowing the representative firms to merge greater than the expected value of the benefit?

More specifically, at $t = 3$, if the companies merged, consumers will be harmed by market power pass-through (we label the *increase* in price due to the increase in market power by $\Delta\theta \cdot \frac{P}{-\epsilon}$, where $\frac{P}{-\epsilon}$ denotes the price markup term, as in Section 5.3), and will benefit due to cost efficiency pass-through (we label the *decrease* in price due to cost efficiency pass-through as $\rho \cdot \Delta C$, as in Section 5.3), where these effects are measured relative to a counterfactual world in which the two companies did not merge. The companies may or may not merge, even when allowed, so we allow the merger to happen with probability \mathcal{P} (which may represent, for example, the probability that the merged firm will be more profitable than the unmerged firms).

Though there are other possible objectives, at $t = 0$, the policymaker may therefore be interested in $E(\Delta P)$, where ΔP is the difference in price between a world at $t = 3$ in which mergers are allowed versus a world where they are not. Since the merger itself generates both the benefit and the harm, the expected price difference is therefore given by:

$$E(\Delta P) = \mathcal{P} \cdot \left[\Delta\theta \cdot \frac{P}{-\epsilon} - \rho\Delta C \right],$$

The policymaker may therefore opt to ban mergers whenever $E(\Delta P) > 0$, i.e., when

$$\Delta\theta \cdot \frac{P}{-\epsilon} > \rho\Delta C. \tag{2}$$

While the values of the parameters may vary based on the context (mergers or NCAs), and while Inequalities 2 and 1 are not identical, there are four main components common across the two inequalities that would simultaneously cause a ban on mergers or on NCAs to help consumers by a greater amount: (i) a high $\Delta\theta$ (e.g., if the industry is highly concentrated, then a spinoff or merger will result in a large $\Delta\theta$), (ii) a large markup term

(associated with highly profitable industries and ones with inelastic demand), (iii) a small ρ (which may occur when demand is inelastic or when industry concentration is low), and (iv) a low ΔC (which is likely the case in mature industries with low concentration and limited opportunities for innovation).²⁵ Altogether, the main takeaway is that industries with already high concentration levels, inelastic demand, and few opportunities for investments to pass-through to consumers are the industries in which a ban on NCAs would have the greatest positive impact.

We raise two additional nuances related to this analysis, which help distinguish between merger analysis and NCA analysis. First, while a merger from duopoly to monopoly requires two specific companies to agree to terms on a merger (which we represent with probability \mathcal{P} above), the market power price effects due to a spinoff may potentially come from one of many workers at a given firm. In other words, \mathcal{P}^{Free} and \mathcal{P}^{NCA} qualitatively represent not the probabilities that a *given* worker will form a spinoff (without or with an NCA, respectively), but the probabilities that *any* worker at the firm will form a spinoff. For a monopolist whose market power is generated by industry-specific knowledge, many workers may be privy to the information necessary to build a startup, making the value of \mathcal{P}^{Free} possibly quite large.²⁶ Additionally, while Inequalities 2 and 1 indicate *whether* it is likely that mergers and NCA cause harm, the size of the expected harm (given that it exists) is outside the purview of those inequalities. By the argument above (that many workers in any given industry are prohibited from increasing competition due to an NCA, whereas, for example, a ban on mergers to monopoly would impact two specific firms in any given industry), it is possible that the harms associated with NCAs, in the case that they are indeed harmful, could be substantially larger, especially since an NCA may act as a merger between several entities (if multiple workers are capable of forming spinoffs), while a merger

²⁵From Weyl and Fabinger (2013) we have that $p = c + \theta \frac{p}{-\epsilon}$ so that $\rho = \frac{dp}{dc} = \frac{1}{1 - \frac{\theta}{-\epsilon}}$ which implies that $\frac{dp}{d\theta} > 0$ and $\frac{d\rho}{d(-\epsilon)} < 0$.

²⁶Here, we are concerned with the probability that greater than or equal to one employee spins off (i.e., the probability that the market is no longer a monopoly). A further consideration is whether additional firms may be added to the market beyond the second.

proposal typical involves only two firms. In this regard, the $\Delta\theta$ for NCAs (in Inequality 1) might be substantially larger than $\Delta\theta$ for mergers (in Inequality 2), especially for firms with large labor forces.

Second, for both merger analysis and NCA analysis, there are harms and benefits that may not be captured in this simple example. While much of merger analysis has focused on consumer prices, there is a literature showing harms to workers from increased levels of employer power in the labor market (Azar et al., 2020; Rinz, 2020; Prager and Schmitt, 2021). Similarly, the literature has found that NCAs may cause harms to workers (Lipsitz and Starr, 2021; Balasubramanian et al., 2020; Johnson et al., 2021) by impeding their mobility and therefore their bargaining power. If a policymaker considers the welfare of workers in their objective function, Inequalities 2 and 1 must include terms on the left hand side that account for harms to workers, as well.

6.2 Reverse Payment Patent Settlements (Pay for Delay)

Merger analysis considers a market whose firm count moves from some n to $n-1$. Temporally, a more apt antitrust analogy for NCAs is the practice of reverse payment patent settlements, colloquially known as “pay-for-delay” (hereafter abbreviated as P4D). In a P4D case, an alleged or potential patent infringer receives a payment from the patent holder in return for staying out of the market (often a pharmaceutical market) for a specified period of time. Consider the landmark Supreme Court case, *FTC v. Actavis, Inc.*, as an example. In that case, Solvay Pharmaceuticals held a patent on a low-testosterone treatment called Androgel. Actavis, a generic drug manufacturer, created a generic version of Androgel and sought to bring it to market. Solvay ultimately entered into a P4D agreement with Actavis, under the terms of which Actavis would delay bringing their drug to market, in return for receiving compensation from Solvay.

The intellectual history of P4D is worth reviewing, insofar as it relates to the analogy to NCAs. Generally speaking, settlements of disputes which might otherwise require a court

hearing may be economically efficient: settlements can reduce litigation costs and alleviate court clogging (see, e.g., Bebchuk (1984)). However, in the case of P4D, the literature went on to point out several anticompetitive effects which may exceed the efficiency typically associated with settlement. For example, Shapiro (2003) points out that P4D agreements often act to preserve monopoly profits by allowing a monopolist and a potential entrant to collude by splitting monopoly profits, rather than sharing (lesser) duopoly profits. This theory informed the Actavis decision: the court ruled that outside P4D payments are indicators that the agreement is based on avoiding competition (Edlin et al., 2015).

Some recent economic research calls into question the view that P4D agreements are necessarily anticompetitive on net. Consider Dickey and Rubinfeld (2012) who note that the possibility of a future P4D agreement may incentivize generic manufacturers to innovate more, possibly causing an increase in innovation. When P4D is not allowed, generic manufacturers may choose to take a gamble: they can innovate potentially infringing generic drugs, understanding that they may be able to bring the drug to market, or they may leave empty handed (if the courts determine that the generic clearly infringes the patent). P4D agreements improve the upside of this gamble. A patent holder may wish to avoid costly litigation or an uncertain outcome and offer a P4D agreement, which increases the payoff to a generic producer.

NCAAs have similarities to and differences from P4D agreements. An NCA is also an agreement which is intended to keep a potential competitor from entering the market. In contrast to mergers, the timing of an NCA bears a resemblance to P4D agreements: an incumbent monopolist is attempting to prevent a future duopoly. However, NCAs differ from P4D agreements in two important ways. First, NCAs are not intended to avoid costly litigations (in fact, NCAs likely cause more litigation than they prevent, since NCAs are sometimes disputed²⁷). Second, NCAs are not subject to the argument leveled by Dickey and Rubinfeld (2012). In P4D agreements, the risky and costly investment is made *prior*

²⁷See <https://faircompetitionlaw.com/2021/01/02/new-trade-secret-and-noncompete-case-growth-graph-updated-january-2-2021>

to receiving payment for the P4D agreement. With NCAs, especially those signed at the conception of an employment relationship or promotion, the payment (a pay increase in exchange for signing an NCA, assuming it is made) is made prior to the worker having an ability to engage in investment in themselves and a potential future spinoff, which is the parallel to generic drug development. Therefore, due to the timing of payment, there is no parallel argument to be made that NCAs can encourage competition by incentivizing workers to invest in potential future spinoffs.

In relation to NCAs, the piece of the P4D literature which remains is that P4D agreements represent an anticompetitive means for a patent holder to collude with a potential entrant, decreasing competition in the market and harming consumers. This argument, as modeled in the rest of this paper, also applies to NCAs. However, NCAs also differ from P4D agreements in one way which suggests that consumers may benefit: in particular, P4D agreements do not encourage patent holders to invest resources in potential entrants, whereas NCAs may.

The test based on *FTC v. Actavis* is often called the Actavis Inference (in which P4D agreements are likely to be anticompetitive when an outsize payment is made). The Actavis Inference does allow for some insight into a possible test regarding NCAs, though such a test may be difficult to implement. Using an analogy to the Actavis Inference, policymakers could institute policy which regards NCAs as presumptively illegal when a large payment is made for the NCA. In this case, “large” would be assessed relative to the potential benefits to productivity stemming from the NCA: if the productivity benefits of an NCA (due to increased training provided by the firm) are equal to X , then payments substantially larger than X would likely reflect the value of lost competition. Note that the productivity benefits must be assessed net of lost investment on the part of the worker, which is disincentivized by an NCA, though it is not explicitly modeled in this paper.

This test may be difficult to implement for two reasons, though. First, the productivity benefits of an NCA may be difficult to measure, which means that determining whether a payment is large is also difficult. Second, workers who sign NCAs are often uncompensated

for them (possibly due to informational asymmetries or possibly because they are asked to sign NCAs after their bargaining power has declined — i.e., after they have turned down alternative employment offers). This means that using a test of whether a payment to workers is large is likely to return a large number of false negatives, since NCAs which lead to consumer harm may not be accompanied by large payments to the worker.

These caveats suggest a tradeoff with respect to evaluation of NCAs. When payments for NCAs are small or nonexistent, workers may not be receiving compensation in line with NCA law (which in many jurisdictions requires that workers receive adequate “consideration”, or compensation, for their NCA). When payments for NCAs are large, they may indicate that the NCA is anticompetitive. Of course, there may be large payments made for NCAs which are solely or primarily due to possible benefits associated with NCA use (in other words, if there is no anticompetitive effect from an NCA, a large payment could represent the owner and worker sharing the rents from solving the investment hold-up problem). Additionally, a rule presuming that large payments indicate anticompetitive activity would create the perverse incentive for firms to underpay their workers for signing an NCA.

One final implication of the analogy to P4D bears noting. The literature on NCAs has pointed out several harms to workers in the form of decreased earnings. There are notable exceptions, though, one of which is CEOs (Kini et al., 2020), who have been shown to receive increased pay when they sign NCAs, or when NCAs are more enforceable. However, CEOs are workers that are highly likely to have advanced knowledge of contracting and labor practices, and may even consult attorneys before signing NCAs. They are therefore highly likely to receive compensation which is appropriately negotiated, and likely reflects a mutually agreeable arrangement.

With the framework we have developed in this paper, and the analogy to P4D, we may shed light not on the question of whether CEOs and firms benefit from using NCAs, but whether consumers may be harmed by such NCAs. Since payments to CEOs are likely to be more reflective of the true benefits to the worker and the firm (versus, for example, a

low-wage worker who is unable to hire counsel to consult on their labor contract), outsize payments are much more likely to be present and therefore to reveal an anticompetitive NCA. Of course, the size of the payment must still be weighed against the value of training (and whatever allocation of that value would likely be made to the CEO). But, evidence that CEOs are highly compensated for signing NCAs may in fact be evidence that those NCAs are most likely to be indicative of arrangements which, similar to P4D, are intended to collude over monopoly profits.

7 Conclusion

Employee spinoffs are a primary means by which concentrated industries become more competitive. NCAs may prevent employee spinoffs, simultaneously creating value for firms by encouraging investment (by solving holdup problems) and lessening future competition. While increased levels of investment may benefit consumers, decreased levels of competition do not. To assess this tradeoff, we posit a model that demonstrates the extent to which each of these two factors impact consumers.

Policymakers who are interested in the role that NCAs play in the economy may consider a variety of affected individuals, including low- and high-wage workers, entrepreneurs, and established firm owners. We add to this conversation by discussing the role that NCAs may have on *consumers*, specifically with respect to competition. In other words, the economic and policy debate over NCAs has largely focused on the apparent trade off between the solution of the hold-up problem (which favors NCA use) and harms to workers (which disfavors NCA use). As we show in this paper, while consumers may benefit from the solution to the hold up problem, the third element which should be considered in this conversation is the competitive harm which arises from NCA use. Holistically, NCAs may act as a way for firm owners and workers to collude by maintaining a monopoly market (using costly investment as a commitment device) to the detriment of consumers. Furthermore, as has been high-

lighted in the literature, workers may be unable to properly negotiate over compensation for an NCA. If this is the case, the sole beneficiaries of NCA use may be firm owners, while workers and consumers may suffer.

We show that consumers in industries with high levels of *potential* competition, and industries with low levels of investment pass-through are those that would likely benefit from an NCA ban. Increased benefits from investment increase NCA use, and may increase consumer welfare, but may counterintuitively decrease consumer welfare when investment pass-through is relatively low. We also compare NCA policy to merger policy and law governing pay-for-delay agreements. Both analogies reveal that NCAs may present puzzles akin to those studied in the antitrust literature: for example, the existence of outsize payments made to CEOs for signing NCAs may reveal that those NCAs are used for anticompetitive reasons.

References

- Azar, J., I. Marinescu, and M. Steinbaum (2020). Labor market concentration. *Journal of Human Resources*, 1218–9914R1.
- Balasubramanian, N., J. W. Chang, M. Sakakibara, J. Sivadasan, and E. Starr (2020). Locked in? the enforceability of covenants not to compete and the careers of high-tech workers. *Journal of Human Resources*, 1218–9931R1.
- Bebchuk, L. A. (1984). Litigation and settlement under imperfect information. *The RAND Journal of Economics*, 404–415.
- Bernheim, B. D. and M. D. Whinston (1998). Exclusive dealing. *Journal of political Economy* 106(1), 64–103.
- Bishara, N. D. (2010). Fifty ways to leave your employer: Relative enforcement of covenants not to compete, trends, and implications for employee mobility policy. *U. Pa. J. Bus. L.* 13, 751.
- Bonnet, C., P. Dubois, S. B. Villas Boas, and D. Klapper (2013). Empirical evidence on the role of nonlinear wholesale pricing and vertical restraints on cost pass-through. *Review of Economics and Statistics* 95(2), 500–515.
- Dickey, B. M. and D. L. Rubinfeld (2012). Would the per se illegal treatment of reverse payment settlements inhibit generic drug investment? *Journal of Competition Law and Economics* 8(3), 615–625.
- Edlin, A., S. Hemphill, H. Hovenkamp, and C. Shapiro (2015). The actavis inference: Theory and practice. *Rutgers UL Rev.* 67, 585.
- Farrell, J., C. Shapiro, et al. (2010). Antitrust evaluation of horizontal mergers: An economic alternative to market definition. *The BE Journal of Theoretical Economics* 10(1), 1–41.
- Franco, A. (2005). Employee entrepreneurship: recent research and future directions. *Handbook of entrepreneurship research*, 81–96.
- Fulghieri, P. and M. Sevilir (2011). Mergers, spinoffs, and employee incentives. *The Review of Financial Studies* 24(7), 2207–2241.
- Gron, A. and D. L. Swenson (2000). Cost pass-through in the us automobile market. *Review of Economics and Statistics* 82(2), 316–324.
- Hart, O. and J. Moore (1990). Property rights and the nature of the firm. *Journal of political economy* 98(6), 1119–1158.
- Hausman, N. and K. Lavetti (2021). Physician practice organization and negotiated prices: evidence from state law changes. *American Economic Journal: Applied Economics*.
- Jeffers, J. (2019). The impact of restricting labor mobility on corporate investment and entrepreneurship. Available at SSRN 3040393.
- Johnson, M. S., K. Lavetti, and M. Lipsitz (2021). The labor market effects of legal restrictions on worker mobility. Available at SSRN 3455381.
- Kini, O., R. Williams, and D. Yin (2020). Ceo non-compete agreements, job risk, and compensation. *The Review of Financial Studies*.
- Klein, B., R. G. Crawford, and A. A. Alchian (1978). Vertical integration, appropriable rents, and the competitive contracting process. *The journal of Law and Economics* 21(2), 297–326.
- Koszegi, B. (2014). Behavioral contract theory. *Journal of Economic Literature* 52(4), 1075–1118.
- Kräkel, M. and D. Sliwka (2009). Should you allow your employee to become your competitor? on noncompete agreements in employment contracts. *International Economic Review* 50(1), 117–141.
- Lipsitz, M. and E. Starr (2021). Low-wage workers and the enforceability of non-compete agreements. *Management Science*.
- Marx, M. and L. Fleming (2012). Non-compete agreements: Barriers to entry... and exit? *Innovation policy and the economy* 12(1), 39–64.
- Miller, N. H., M. Osborne, and G. Sheu (2017). Pass-through in a concentrated industry: empirical evidence and regulatory implications. *The RAND Journal of Economics* 48(1), 69–93.
- Muendler, M.-A., J. E. Rauch, and O. Tocoian (2012). Employee spinoffs and other entrants:

- Stylized facts from Brazil. *International Journal of Industrial Organization* 30(5), 447–458.
- Prager, E. and M. Schmitt (2021). Employer consolidation and wages: Evidence from hospitals. *American Economic Review* 111(2), 397–427.
- Rauch, J. E. (2016). Dynastic entrepreneurship, entry, and non-compete enforcement. *European Economic Review* 86, 188–201.
- Rauch, J. E. and J. Watson (2015). Client-based entrepreneurship. *The Journal of Law, Economics, & Organization* 31(1), 30–60.
- Rinz, K. (2020). Labor market concentration, earnings, and inequality. *Journal of Human Resources*, 0219–10025R1.
- Rubin, P. H. and P. Shedd (1981). Human capital and covenants not to compete. *The Journal of Legal Studies* 10(1), 93–110.
- Samila, S. and O. Sorenson (2011). Noncompete covenants: Incentives to innovate or impediments to growth. *Management Science* 57(3), 425–438.
- Schmidheiny, K. and S. Siegloch (2019). On event study designs and distributed-lag models: Equivalence, generalization and practical implications.
- Shapiro, C. (2003). Antitrust limits to patent settlements. *RAND Journal of Economics*, 391–411.
- Starr, E. (2019). Consider this: Training, wages, and the enforceability of covenants not to compete. *ILR Review* 72(4), 783–817.
- Starr, E., N. Balasubramanian, and M. Sakakibara (2018). Screening spinouts? how noncompete enforceability affects the creation, growth, and survival of new firms. *Management Science* 64(2), 552–572.
- Starr, E., J. Prescott, and N. Bishara (2020). The behavioral effects of (unenforceable) contracts. *The Journal of Law, Economics, and Organization* 36(3), 633–687.
- Starr, E. P., J. Prescott, and N. D. Bishara (2021). Noncompete agreements in the US labor force. *The Journal of Law and Economics* 64(1), 53–84.
- Weyl, E. G. and M. Fabinger (2013). Pass-through as an economic tool: Principles of incidence under imperfect competition. *Journal of Political Economy* 121(3), 528–583.
- Wickelgren, A. L. (2018). A novel justification for legal restrictions on non-compete clauses. *International Review of Law and Economics* 54, 49–57.
- Williamson, O. E. (1975). Markets and hierarchies: analysis and antitrust implications: a study in the economics of internal organization. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*.

A Empirical Appendix

In this Appendix, we describe the steps taken to perform the analyses reported in Section 2.

We begin with the CBP county-level data. We flag industries, using US Census Bureau concordances,²⁸ which had nonconstant NAICS codes over the relevant period, and omit these industries to avoid data abnormalities associated with changing NAICS codes.²⁹ We omit data prior to 1998 for this same reason: NAICS was first used in 1998, and prior to that year, SIC codes were used, which grouped industries somewhat differently.

Next, we impute employment per establishment in each size bin as the midpoint of the range given by the bin: e.g., 2.5 employees per establishment when the bin indicates between 1 and 4 employees, 7 employees per establishment when the flag indicates between 5 and 9 employees, and so on. For the bin indicating more than 5000 employees, we impute 6000 employees. We additionally impute total employment in the industry by county cell, instead of relying on reported employment or the employment flags provided in the data: this is done to avoid instances in which HHI is calculated to be greater than 10,000 (the theoretical maximum), often greatly so. Such an issue would arise if, for example, there was just one firm with 5 employees. The data would report one firm in the 5-9 employee size range, and we would impute a size of 7 employees for it. Its share would then be calculated as $100 \cdot 7 / 5 = 140$, and HHI would be calculated as 19,600. If, instead, the total employment in the county-industry cell were calculated as the sum of imputed employment (here, 7 employees total), the HHI would (correctly) be calculated as 10,000.

We further omit industry-county cells which had a calculated HHI of zero at some point during the time period: an HHI of zero represents a nonexistent industry. We would not want to conflate, for example, the advent of a new industry in a given county with an increase in concentration, which would occur if we observed a positive change after an HHI of zero. Furthermore, several of the industry-county-year observations with HHIs of zero appear to

²⁸Available at <https://www.census.gov/naics/?68967>.

²⁹Note that harmonizing NAICS codes across time and performing the analysis using those harmonized NAICS codes yields noisier, but qualitatively similar, results.

be misreported data (perhaps due to changing industrial classification). Finally, we omit industry-county observations which experience HHI changes greater than 6000 in a one-year period. The imputation procedure outlined above, combined with issues similar to but not identical to the zero-HHI issue, cause infrequent but enormous HHI changes year over year, which likely do not reflect true competitive conditions. Note that if we remove both of these sample restrictions, we find qualitatively similar results, but with some strange patterns attributable to artifacts of the cleaning and imputation procedures.

We run the distributed lag model outlined in Schmidheiny and Siegloch (2019) by calculating annual changes in HHIs in county by industry cells, and changes in NCA enforceability in state cells. We identify five leads and five lags of the corresponding event study coefficients, and report the coefficients which do not correspond to binned endpoints. We use Census division by year by industry fixed effects to alleviate concerns that treatment effects simply represent varying industrial trends, varying regional trends, or even varying industrial trends within specific regions. We weight the regressions by imputed employment in the industry-county-year cell and cluster standard errors by state, the level of the treatment.

B Appendix of Proofs

Proof of Lemma 1: Using backward induction, we first consider the decision to spin off given the values of NCA , I , and c_s . Due to free renegotiation, the pair will select a renegotiation transfer payment and opt for a spin-off whenever it is pairwise efficient: $\Pi_s(c_s) + \Pi_{i,D}(c_s) > \Pi_{i,M}(I)$. From Assumption 1d, we have that $\bar{c}_s(I)$ defines the cutoff value for c_s such that $\Pi_s(c_s) + \Pi_{i,D}(c_s) = \Pi_{i,M}(I)$. This implies that for any $c_s < \bar{c}_s(I)$ we have that $\Pi_s(c_s) + \Pi_{i,D}(c_s) > \Pi_{i,M}(I)$. Moving backwards, the owner will opt for $I = 1$ whenever investment nets a greater payoff than no investment.

Thus, if $NCA = 1$ and c_s is unrealized, then the owner invests whenever

$$P(c_s < \bar{c}_s(1))E[\Pi_{i,D}(c_s) + BO|c_s < \bar{c}_s(1)] + (1 - P(c_s < \bar{c}_s(1)))E[\Pi_{i,M}(1)|c_s > \bar{c}_s(1)] - \kappa \\ > P(c_s < \bar{c}_s(0))E[\Pi_{i,D}(c_s) + BO|c_s < \bar{c}_s(0)] + (1 - P(c_s < \bar{c}_s(0)))E[\Pi_{i,M}(0)|c_s > \bar{c}_s(0)].$$

Instead, if $NCA = 0$ and c_s is unrealized, then the owner invests whenever

$$P(c_s < \bar{c}_s(1))E[\Pi_{i,D}(c_s)|c_s < \bar{c}_s(1)] + (1 - P(c_s < \bar{c}_s(1)))E[\Pi_{i,M}(1) + BI|c_s > \bar{c}_s(1)] - \kappa \\ > P(c_s < \bar{c}_s(0))E[\Pi_{i,D}(c_s)|c_s < \bar{c}_s(0)] + (1 - P(c_s < \bar{c}_s(0)))E[\Pi_{i,M}(0) + BI|c_s > \bar{c}_s(0)],$$

Since the worker makes a take-it-or-leave-it offer to the owner if she wishes to spin off (if she has an NCA), or if she wishes to remain employed by the owner (if she does not have an NCA), the worker will capture all profit above and beyond what the owner would achieve with the other action. So, if the worker has an NCA and wishes to spin off, she will offer the owner exactly the difference in profit that the owner would receive if the worker remained employed (i.e., $BO = \Pi_{i,M}(I) - \Pi_{i,D}(c_s)$). And, if the worker does not have an NCA and wishes to remain employed, she will appropriate all of the profits above and beyond what the owner could receive in a duopoly (i.e., $BI = -\Pi_{i,M}(I) + \Pi_{i,D}(c_s)$). Taking this into account, the inequalities above reduce to

$$E[\Pi_{i,M}(1)] - \kappa > E[\Pi_{i,M}(0)] \text{ for } NCA = 1,$$

$$E[\Pi_{i,D}(c_s)] - \kappa > E[\Pi_{i,D}(c_s)] \text{ for } NCA = 0.$$

This implies that if $NCA = 1$, then $I = 1$ if and only if $\Pi_{i,M}(1) - \Pi_{i,M}(0) > \kappa$ (since $\Pi_{i,M}(I)$ is independent of the r.v. c_s), and if $NCA = 0$, then $I = 0$ for all $\kappa > 0$. \square

Proof of Proposition 1: Starting with the case where $\kappa > \Pi_{i,M}(1) - \Pi_{i,M}(0)$, by Lemma 1 we have that investment does not occur ($I = 0$). Given frictionless renegotiations, if

$NCA = 1$, the worker will propose a payment $BO = \Pi_{i,M}(0) - \Pi_{i,D}(c_s)$. If $NCA = 0$, the worker will propose a payment which guarantees the owner a profit of $\Pi_{i,D}(c_s)$. If the worker remains employed by the owner, profit of $\Pi_{i,M}(0)$ will accrue to the owner, meaning that $BI = \Pi_{i,D}(c_s) - \Pi_{i,M}(I)$. Moving back to the contracting stage prior to the draw of c_s , we have that the owner will decide between a contract with an NCA, $\{w^1, 1\}$, and one without, $\{w^0, 0\}$, by comparing:

$$\mathcal{P}^{NCA} \cdot E[\Pi_{i,D}(c_s) + BO | c_s < \bar{c}_s(1)] + (1 - \mathcal{P}^{NCA})\Pi_{i,M}(1) - w^1, \quad (3)$$

$$\mathcal{P}^{Free} \cdot E[\Pi_{i,D}(c_s) | c_s < \bar{c}_s(0)] + (1 - \mathcal{P}^{Free})E[\Pi_{i,M}(0) + BI | c_s > \bar{c}_s(0)] - w^0, \quad (4)$$

where $\mathcal{P}^{NCA} = Pr(c_s < \bar{c}_s(1))$ denotes the probability of a spinoff with an NCA and $\mathcal{P}^{Free} = Pr(c_s < \bar{c}_s(0))$ denotes the probability of a spinoff when the worker does not have an NCA. Note that $\mathcal{P}^{NCA} = P(c_s < \bar{c}_s(1)) < P(c_s < \bar{c}_s(0)) = \mathcal{P}^{Free}$ since $\bar{c}_s(1) < \bar{c}_s(0)$.³⁰

To determine wages, note that the worker's IR constraints are given by:

$$w^1 + \mathcal{P}^{NCA}\mathcal{P}^{NCA} \cdot E[\Pi_s(c_s) - BO | c_s < \bar{c}_s(1)] \geq 0, \quad (\hat{IR}_E^1)$$

$$w^0 + \mathcal{P}^{Free} \cdot E[\Pi_s(c_s) | c_s < \bar{c}_s(0)] + (1 - \mathcal{P}^{Free})E[-BI | c_s > \bar{c}_s(0)] \geq 0, \quad (\hat{IR}_E^0)$$

Solving for the w^1 and w^0 that make Inequalities \hat{IR}_E^1 and \hat{IR}_E^0 hold with equality, and plugging into Expressions 3 and 4 generates identical expressions for the owner's expected utility, regardless of contract. In fact, in either case, the owner captures all of the expected utility associated with the pairwise efficient market structure:

$$P(c_s > \bar{c}_s(I))\Pi_{i,M}(I) + P(c_s < \bar{c}_s(I))(\Pi_{i,D}(c_s) + \Pi_s(c_s)).$$

Furthermore, given that w^1 and w^0 are such that Inequalities \hat{IR}_E^1 and \hat{IR}_E^0 hold with equality, the worker is also indifferent between the two contracts.

³⁰Recall that $\bar{c}_s(I)$ solves $\Pi_{i,D}(c_s) + \Pi_s(c_s) = \Pi_{i,M}(I)$ so that $\bar{c}_s(1) < \bar{c}_s(0)$ since $\Pi_{i,M}(1) > \Pi_{i,M}(0)$.

For the more interesting case where $\kappa \leq \Pi_{i,M}(1) - \Pi_{i,M}(0)$, note that Lemma 1 implies that investment occurs if and only if $NCA = 1$. Moving back to the contracting stage prior to the draw of c_s , we have that the pair will agree to a contract that includes an NCA whenever the joint payoff with an NCA/investment is greater than the joint payoff without an NCA/investment:

$$\begin{aligned} & \mathcal{P}^{NCA} \cdot E[\Pi_{i,D}(c_s) + \Pi_s(c_s) | c_s < \bar{c}_s(1)] + (1 - \mathcal{P}^{NCA})\Pi_{i,M}(1) - \kappa \\ & > \mathcal{P}^{Free} \cdot E[\Pi_{i,D}(c_s) + \Pi_s(c_s) | c_s < \bar{c}_s(0)] + (1 - \mathcal{P}^{Free})\Pi_{i,M}(0). \end{aligned}$$

Rewriting and simplifying, the pair will use an NCA whenever

$$\begin{aligned} & \underbrace{\Pi_{i,M}(1) - \Pi_{i,M}(0)}_{\text{monopoly benefit from } I = 1} - \kappa \\ & > \underbrace{\mathcal{P}^{Free} E[\Pi_{i,D}(c_s) + \Pi_s(c_s) - \Pi_{i,M}(0) | c_s < \bar{c}_s(0)]}_{\text{duopoly benefits for } I = 0} \\ & \quad - \underbrace{\mathcal{P}^{NCA} E[\Pi_{i,D}(c_s) + \Pi_s(c_s) - \Pi_{i,M}(1) | c_s < \bar{c}_s(1)]}_{\text{duopoly benefits for } I = 1}, \end{aligned} \tag{5}$$

Given that we are considering the case where $\kappa \leq \Pi_{i,M}(1) - \Pi_{i,M}(0)$, the inequality given by Equation (5) holds whenever the right-hand side is less than or equal to zero. Given our general characterization, the right-hand side is ambiguous.³¹ Note that the right-hand side of Equation (5) captures the difference in the expected duopoly benefits across investment. This generates a clear interpretation for when Equation (5) fails: if the cost from investment κ and the benefits from the duopoly without investment (relative to those with investment) are sufficiently large, then investment does not occur.

Let $\bar{\kappa}$ be the κ so that Equation (5) holds with equality. As a technicality, we also assume that the inequality in Equation (5) holds when $\kappa = 0$. Otherwise an NCA never

³¹Assumption 1c and $\bar{c}_s(1) < \bar{c}_s(0)$ imply that $E[\Pi_{i,D}(c_s) + \Pi_s(c_s) | c_s < \bar{c}_s(1)] > E[\Pi_{i,D}(c_s) + \Pi_s(c_s) | c_s < \bar{c}_s(0)]$ but we also know that $\Pi_{i,M}(1) > \Pi_{i,M}(0)$ so the difference in $E[\Pi_{i,D}(c_s) + \Pi_s(c_s) | c_s < \bar{c}_s(1)] - \Pi_{i,M}(1)$ and $E[\Pi_{i,D}(c_s) + \Pi_s(c_s) | c_s < \bar{c}_s(0)] - \Pi_{i,M}(0)$ is ambiguous.

occurs in equilibrium and the investment exercise is non-existent. In addition, let $\tilde{\kappa} = \min\{\bar{\kappa}, \Pi_{i,M}(1) - \Pi_{i,M}(0)\}$. These results imply that $NCA = 1$ for all $\kappa \in [0, \tilde{\kappa}]$ and $NCA = 0$ for all $\kappa \in (\tilde{\kappa}, \Pi_{i,M}(1) - \Pi_{i,M}(0)]$.³² \square

Proof of Proposition 2: For notational simplicity, we denote the joint duopoly profit by $\Pi_D(c_s)$ so that $\Pi_D(c_s) := \Pi_{i,D}(c_s) + \Pi_s(c_s)$. Rewriting the definition of $\tilde{\kappa}$ based on Equation (5) gives the following:

$$\begin{aligned} \tilde{\kappa} &= \Pi_M(1) - \Pi_M(0) \\ &\quad + \mathcal{P}^{NCA} E[\Pi_D(c_s) - \Pi_M(1) | c_s < \bar{c}_s(1)] - \mathcal{P}^{Free} E[\Pi_D(c_s) - \Pi_M(0) | c_s < \bar{c}_s(0)], \end{aligned}$$

where, as defined in the proof of Proposition 1, $\mathcal{P}^{NCA} = Pr(c_s < \bar{c}_s(1))$ denotes the probability of a spinoff with an NCA and $\mathcal{P}^{Free} = Pr(c_s < \bar{c}_s(0))$ denotes the probability of a spinoff when the worker does not have an NCA. Simplifying implies that

$$\begin{aligned} \tilde{\kappa} &= (1 - \mathcal{P}^{NCA})\Pi_M(1) - (1 - \mathcal{P}^{Free})\Pi_M(0) \\ &\quad + \mathcal{P}^{NCA} E[\Pi_D(c_s) | c_s < \bar{c}_s(1)] - \mathcal{P}^{Free} E[\Pi_D(c_s) | c_s < \bar{c}_s(0)], \end{aligned}$$

implying

$$\begin{aligned} \tilde{\kappa} &= (1 - \mathcal{P}^{NCA})\Pi_M(1) - (1 - \mathcal{P}^{Free})\Pi_M(0) \\ &\quad + \int_{-\infty}^{\bar{c}_s(1)} (\Pi_D(c_s)) dF(c_s) - \int_{-\infty}^{\bar{c}_s(0)} (\Pi_D(c_s)) dF(c_s), \end{aligned}$$

which finally gives

$$\tilde{\kappa} = (1 - \mathcal{P}^{NCA})\Pi_M(1) - (1 - \mathcal{P}^{Free})\Pi_M(0) - \int_{\text{bar}c_s(1)}^{\bar{c}_s(0)} (\Pi_D(c_s)) dF(c_s).$$

³²Note that $[\tilde{\kappa}, \Pi_{i,M}(1) - \Pi_{i,M}(0)]$ is the empty set if the right-hand side of Equation (5) is negative.

Note that $\frac{d\mathcal{P}^{Free}}{dt} = \frac{dE[\Pi_D(c_s)|c_s < \bar{c}_s(0)]}{dt} = \frac{d\Pi_M(0)}{dt} = 0$. Therefore we have that

$$\frac{d\tilde{\kappa}}{dt} = \frac{d \left[(1 - \mathcal{P}^{NCA})\Pi_M(1) - \int_{\text{bar}c_s(1)}^{\bar{c}_s(0)} (\Pi_D(c_s))dF(c_s) \right]}{dt}$$

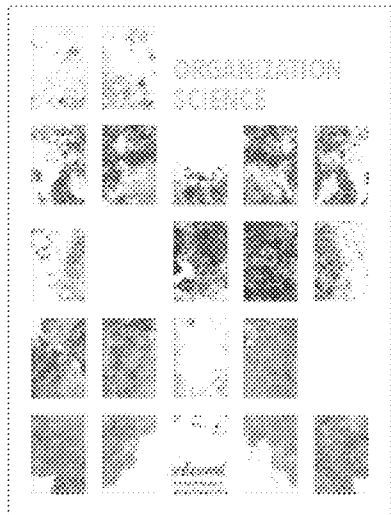
Differentiating using the Leibniz Rule, we arrive with

$$\frac{d\tilde{\kappa}}{dt} = (1 - \mathcal{P}^{NCA})\frac{d\Pi_M(1)}{dt} + \frac{d\mathcal{P}^{NCA}}{dt}(\Pi_M(1)) + (\Pi_D(\bar{c}_s(1)))\frac{d\bar{c}_s(1)}{dt}$$

With $\mathcal{P}^{NCA} = \int_{-\infty}^{\bar{c}_s(1)} dF(c_s)$, we have that $\frac{d\mathcal{P}^{NCA}}{dt} = \frac{d\bar{c}_s(1)}{dt}$. And, by definition of $\bar{c}_s(1)$, $\Pi_M(1) = \Pi_D(\bar{c}_s(1))$. Therefore, the latter two terms cancel, and we are left with:

$$\frac{d\tilde{\kappa}}{dt} = (1 - \mathcal{P}^{NCA})\frac{d\Pi_M(1)}{dt}$$

Finally, since $\mathcal{P}^{NCA} \in [0, 1]$ and $\frac{d\Pi_M(1)}{dt} > 0$ (by assumption), we have that $\frac{d\tilde{\kappa}}{dt} \geq 0$. □



Organization Science

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Employee Non-compete Agreements, Gender, and Entrepreneurship

Matt Marx

To cite this article:

Matt Marx (2021) Employee Non-compete Agreements, Gender, and Entrepreneurship. Organization Science

Published online in Articles in Advance 03 Sep 2021

. <https://doi.org/10.1287/orsc.2021.1506>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2021 The Author(s)

Please scroll down for article—it is on subsequent pages




With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Employee Non-compete Agreements, Gender, and Entrepreneurship

Matt Marx^a

^aCornell University, Ithaca, New York 14853

Contact: mmarx@cornell.edu,  <https://orcid.org/0000-0002-6173-4142> (MM)

Received: December 29, 2018

Revised: March 31, 2020; January 14, 2021;
June 11, 2021

Accepted: July 24, 2021

Published Online in Articles in Advance:
September 3, 2021

<https://doi.org/10.1287/orsc.2021.1506>

Copyright: © 2021 The Author(s)

Abstract. I contribute to the literature on institutions, gender, and entrepreneurship by showing that macrolevel institutional policies that do not explicitly target women nonetheless discourage them from leveraging prior professional experience—their own and that of others—in founding new ventures. Most ventures fail, but chances of success are greater if founders can bring to bear their professional expertise. However, employee non-compete agreements enjoin workers from leaving their employer to found a rival business in the same industry. Thus, non-competes add legal risk to business risk. To the extent that women exhibit greater risk aversion, the threat of litigation from their ex-employer may act as a sharper brake on startup activity than for men. Examining all workers who were employed exclusively within 25 states and the District of Columbia from 1990 to 2014, I find that women subject to tighter non-compete policies were less likely to leave their employers and start rival businesses. Non-competes increase the risk of entrepreneurship by making it harder to hire talent with relevant experience, shifting women away from higher potential ventures. A review of thousands of filed lawsuits suggests that firms do not target women in non-compete cases. Rather, it appears that non-competes disproportionately discourage women from leveraging their professional networks in hiring the sort of talent necessary for high-growth startups to succeed.



Open Access Statement: This work is licensed under a Creative Commons Attribution 4.0 International License. You are free to copy, distribute, transmit and adapt this work, but you must attribute this work as “*Organization Science*. Copyright © 2021 The Author(s). <https://doi.org/10.1287/orsc.2021.1506>, used under a Creative Commons Attribution License: <https://creativecommons.org/licenses/by/4.0/>.”

Funding: This research uses data from the Census Bureau’s Longitudinal Employer Household Dynamics Program, which was partially supported by the National Science Foundation [Grants SES-9978093, SES-0339191, and ITR-0427889], the National Institute on Aging [Grant AG018854], and grants from the Alfred P. Sloan Foundation. This work was moreover supported in part by the National Science Foundation Science of Science Policy [Grant 1735669].

Supplemental Material: The online appendix is available at <https://doi.org/10.1287/orsc.2021.1506>.

Keywords: non-compete agreements • entrepreneurship • gender

Despite many efforts to spur entrepreneurship among women, there remains a substantial gender gap. The Organisation for Economic Co-operation and Development’s (2017) “Missing Entrepreneurs” report notes that women in the European Union are less than half as likely as men to be self-employed. Similarly, the Global Entrepreneurship Monitor (2017) survey claims that only 11% of women are attempting to start a business compared with 16% of men. Guzman and Kacperczyk (2019) report that only 20% of new firms in California and Massachusetts are headed by women. Such disparities are magnified among high-growth, venture-backed activity, in which the share of startups led by women has yet to approach double digits (Brush et al. 2014). Focusing on high-tech entrepreneurship, Miric and Yin (2020) find that the share

of female founders lags the share of female workers in the industry. The gender gap in entrepreneurship has attracted the attention of scholars and policymakers alike (Coleman and Robb 2009, Kanze et al. 2018), who have sought to understand and address factors that depress entrepreneurial activity among women.

Much scholarship documents that the gender gap in entrepreneurship can be attributed in part to individuals—in particular, powerful individuals, such as investors—engaging in gender-based discrimination (Gupta and Bhawe 2007, Thébaud 2010, Brooks et al. 2014, Ewens and Townsend 2020). These preferences, whether based on personal taste or statistical expectation, not only explain gaps in founding and funding rates, but also illustrate the origins of structural differences, such as access to capital and advisor networks

(Belcourt et al. 1991). But, as Small and Pager (2020) observe, differential outcomes for groups are not only attributable to discrimination by individuals, whether intentional or implicit. Instead, divergent group outcomes may reflect the impact of what Small and Pager (2020) term *institutional discrimination*—that differential treatment/outcomes can arise via organizational practices, sometimes codified into law, even when there is no original intention to discriminate.

In this paper, I investigate whether lower rates of entrepreneurial entry among women might be explained in part by an institutional factor: state sanction of the widespread organizational practice whereby employers enjoin their employees from leaving to start rival businesses. Postemployment covenants not to compete (hereafter, “non-competes” or NCs) date back to the 15th century and are widely used in the modern economy. To my knowledge, no legal scholar has documented gender as a motivating factor for using non-compete agreements, nor does the usage of non-competes among men and women differ substantially. Nonetheless, I claim that non-competes may discourage entrepreneurship among women.

Most startups fail, making entrepreneurship an inherently risky activity that attracts founders who are risk loving (Ahmed 1985, Hvide and Panos 2014). Entrepreneurs perform better when they found firms that draw directly upon their professional expertise and experience (Klepper 2009), that is, by founding a firm in the same industry. But non-compete agreements are explicitly designed to prevent employees from leaving to launch a startup that competes with the ex-employer—that is, in the same industry in which the founders’ experience is directly applicable. Workers who nonetheless found a rival firm can be sued by their former employer, resulting in cessation of involvement with the startup and possibly damages (in addition to legal costs). Thus, non-compete agreements add legal and financial risk to the business risk inherent in entrepreneurship. Moreover, new ventures benefit from hiring workers with expertise in the industry, which is also more difficult given non-competes and heightens the risk of failure. An extensive literature documents higher levels of risk aversion among women in both experimental and observational studies (for reviews, see Eckel and Grossman 2008, Bertrand 2011, Charness and Gneezy 2012). Moreover, and independent of risk aversion, women are penalized more severely for failure (Egan et al. 2017, Sarsons 2020). Heightened aversion to the risk of a non-compete lawsuit may discourage women from leaving their employers to form businesses in the same industry.

I examine this hypothesis using employment histories for all workers whose careers were only in 25 U.S. states and/or the District of Columbia from 1990 to 2014, coupled with a state-by-state non-compete

enforceability index over the same period. I find that women in states with stricter non-compete enforceability are less likely than men to leave their jobs and start rival ventures. (No such effect is obtained for starting nonrivalrous ventures, for which non-competes would not bind.) This effect is not explained by firms targeting women in non-compete lawsuits; rather, two mechanisms underscore the potential role of risk aversion. First, although rivalrous startups hire more employees with industry experience, reflecting the importance of relevant human capital in startup success, this trend reverses and, thus, increases the risk of founding a rival startup when subject to stricter non-competes. Women are particularly less likely to draw on their prior professional networks to hire employees with relevant experience. Second, non-competes shift female founders away from higher potential, higher risk ventures toward those that grow less but survive longer. Thus, non-competes not only affect the rate of women’s participation in entrepreneurship, but also change the nature of startups founded.

The paper proceeds as follows. I review the literature on entrepreneurship and gender as well as the literature on non-compete agreements, noting the lack of overlap. I then sketch a theory, largely based on risk aversion, of why non-competes disproportionately discourage women from starting new ventures. I described the restricted-access Census data used to test this hypothesis and review the main results. I then delve into the mechanisms, ruling out bias in lawsuits while ruling in the difficulty of attracting experienced talent as a reason why non-competes redirect women away from higher potential, higher risk ventures. I conclude by reviewing the implications for organizational theory as well as policy.

Institutional Discrimination and the Gender Gap in Entrepreneurship

Although women outnumber men among college students (Goldin et al. 2006) and comprise nearly half of the U.S. workforce,¹ they are underrepresented among entrepreneurs. The Global Entrepreneurship Monitor (2017) reports that men are nearly 1.5 times more likely to start new businesses. The disparity in the United States appears even sharper: Guzman and Kacperczyk (2019) report that only one out of five newly registered businesses is woman owned. Sharper still is the disparity among ventures backed by professional investors, for which the share of female founders is in single digits. Given that women are near parity in education and workforce participation, what explains the gender gap in entrepreneurship?

Much scholarship on this point focuses on discrimination by individuals. Brooks et al. (2014) show that investors prefer the same startup pitch when delivered

by a man. Similarly, Ewens and Townsend (2020) find that male investors on AngelList express less interest in observably similar founders who are female even though those startups founded by women outperform those founded by men. Not only resource providers, but also potential cofounders or employees may discriminate because entrepreneurship is strongly male-typed (Kacperczyk and Younkin 2017). Such discrimination can contribute to weaker entrepreneurial networks for women (Belcourt et al. 1991), which, in turn, exacerbate the gender gap. Given the strong role of peer effects in encouraging entrepreneurs (Nanda and Sorensen 2010), coupled with the fact that peer effects in entrepreneurship are strongly homophilous by gender (Markussen and Roed 2017), discrimination by individuals can ripple forward and helps to explain the persistence of the gender gap among founders.

The literature regarding individual discrimination has helped to identify several reasons for disparity in entrepreneurial activity among men versus women. As Small and Pager (2020) note, economists have historically considered the sort of individual discrimination studied to date to compose the bulk, if not the entirety, of discriminative practices. Sociologists, however, have embraced a broader set of mechanisms at play. That is, differential group outcomes are not solely explained by discrimination conducted by individuals but also from *institutional* discrimination. Small and Pager (2020, p. 49) define institutional discrimination as “differential treatment . . . that is either perpetrated by organizations or codified into law. [I]t need not result from personal prejudice [or] from rational guesses on the basis of group characteristics.” Important to their formulation is that discrimination can occur because of organizational practices, possibly sanctioned by the state, even though these were not intended to discriminate against subgroups. In their treatise, which focuses primarily on racial discrimination, they note that the recent swell of sociological inquiry into institutional discrimination has not been matched by similar interest from economists.

Within entrepreneurship, the bulk of scholarship on the gender gap has explored either statistical or taste-based discrimination by individuals. There have been calls to examine institutional factors (Minniti and Arenius 2003), such as government policies, and a few scholars investigate the impact of institutional policies on entrepreneurship generally (Henrekson and Rosenberg 2001, Acs et al. 2008, Aidis et al. 2008). To my knowledge, however, only a very few articles have explored the intersection of public policy, entrepreneurship, and gender. Using survey data from the Global Entrepreneurship Monitor in two dozen industrialized countries, Elam and Terjesen (2010) present the puzzle that state-sponsored childcare appears to *reduce* levels of entrepreneurship among women. This

apparent anomaly is subsequently unpacked by Thébaud (2015), who decomposes the dependent variable (DV) into low- versus high-growth entrepreneurship. She finds that, although certain policies designed to address work/family conflict discourage entrepreneurship among women, the deleterious effect is exclusively among those with limited employment options. Castellana et al. (2020) find that simplifying governmental rules governing incorporation boosts entrepreneurship among women even more than men.

I extend this line of work by investigating whether a widespread organizational practice, sanctioned by the state in most locales, disproportionately discourages women from engaging in entrepreneurship. Specifically, I study employee non-compete agreements. Non-competes are employment contracts in which an employee commits not to found or join a rival business. Nearly one in five U.S. workers is subject to a non-compete (Starr et al. 2020b) and nearly half of high-tech workers (Marx 2011). Non-competes have been previously shown to act as a brake on entrepreneurship (Stuart and Sorenson 2003, Samila and Sorenson 2011, Starr et al. 2018), but I am unaware of any prior work regarding the potential for a differential effect by gender. In the following section, I describe why the combination of high failure rates among entrepreneurs coupled with higher risk aversion among women could suggest institutional discrimination against women via non-competes.

Entrepreneurial Failure, Prior Experience, and Non-competes

Most startups fail. Even among venture-backed startups, 75% have a liquidation value of zero (Hall and Woodward 2010). But the chances of success can be raised when startup founders and employees have relevant prior experience. Klepper (2009) and other scholars (e.g., Agarwal et al. 2004, Chatterji 2009) show that intraindustry “spinoffs” routinely outperform entrants whose founders did not have prior experience in the industry. Prior experience affords relevant resources and routines that address the fragility of early stage ventures. Indeed, Bhidé (2000) reports that 71% of the Inc. 500 found their entrepreneurial idea during their existing employment. For example, Fairchild Semiconductor was famously founded by the “traitorous eight” who left Shockley Semiconductor on November 8, 1957, and brought with them extensive knowledge of the industry. Thus, starting a firm in an industry in which one has professional experience and hiring employees with experience in that same industry reduce risk.

Of course, Fairchild’s gain was Shockley’s loss. Firms worry that their investments in employees’ human capital (e.g., via training) might be expropriated when those workers leave to found rival firms. Employers cannot ensure a return on their investment in

training by compelling workers to stay at the firm, but they can provide incentives to stay, such as increased compensation or workplace amenities. On the other hand, the employer might discourage departure by circumscribing professional opportunities elsewhere via restraints on the ex-employee. Although the U.S. Constitution bans indentured servitude, including “general” restraints on being employed elsewhere, the *Mitchel v. Reynolds* decision (1711) established that “particular” or limited restraints may be permissible. For example, although it would be unreasonable to restrain a software engineer from ever working at any software company for the rest of the engineer’s career, one might find it less unreasonable to restrain that same software engineer from working for a software company that produced a directly competitive product for one or two years.

At least in the United States, employers who want to impose restraints on the behavior of (ex-)employees must enter into an agreement with them. Such an agreement is typically referred to as a “postemployment restraint.” Postemployment restraints fall into two categories. First, some postemployment restraints govern the use of the ex-employer’s resources. These include non-disclosure agreements, which protect confidential information, and nonsolicitation agreements, which limit the ex-employee’s ability to solicit either the firm’s customers or its workers. These agreements place no stipulations on *where* the ex-employee may work—only on *how* workers may use their ex-employer’s resources.

A second, complementary category of postemployment restraints does not explicitly restrict the use of the former employer’s resources but rather limits the type of firms an ex-employee may found or join. The most well-known in this category is generally referred to as a “postemployment covenant not to compete” or “employee non-compete agreement.” A non-compete restricts employment possibilities by specifying a field in which the ex-employee is not allowed to work for a defined period of time—usually one to two years. Recent surveys estimate that 20%–50% of workers, depending on occupation, have signed a non-compete (Marx 2011, Starr et al. 2020b).

If an employer believes that an ex-employee has violated a non-compete by joining or founding a rival, the employer can sue the ex-employee for breaching the non-compete and possibly also the new employer for inducing the breach. Although, in some cases, damages may be sought, more common is for plaintiffs to ask the court to enjoin (i.e., prevent) the ex-employee from continuing to work at the rival firm. Cases can take months or years to work through the legal system, so plaintiffs routinely ask for a *preliminary* injunction, which requires the ex-employee to step away from the new job immediately while the

case proceeds through the legal system during a period of several months or even multiple years.

Absent a formal lawsuit, non-competes may still discourage workers from leaving their employer to found a rival. The lawsuit threat, whether written (e.g., a cease-and-desist letter) or verbal, may be enough to discourage the worker from breaching the non-compete. Even absent an explicit threat, workers may fear a potential lawsuit and decide to either remain with the firm or take a “career detour,” leaving the industry when they leave their job (Marx 2011). Thus, non-compete agreements make it more difficult to leverage one’s professional experience in an industry or to hire others with industry experience.

Non-competes Increase the Risk of Entrepreneurship

If entrepreneurs can improve their chances of success by leveraging industry experience, and if non-competes make it more difficult to found ventures in the same industry or to hire talent from within the industry, then non-compete agreements make entrepreneurship a more risky enterprise than otherwise. Table 1 sketches the relative risk of founding in four scenarios. Each cell considers constraints on the founder leveraging the founder’s own relevant expertise (if any) in the new venture as well as whether the founder can hire employees with relevant expertise. The columns consider weaker versus stricter non-compete policies; the rows consider founding a venture in a different industry versus the same industry.

Looking at the first row of Table 1, founding in a different industry is, of course, disadvantageous because the founder lacks relevant expertise to apply to the new venture. And, if founding in a strict non-compete regime (upper left-hand quadrant), it is difficult for the founder to compensate for the founder’s own lack of industry experience by hiring workers from within the industry. Even if founding in a weak NC regime (upper right-hand quadrant), although the founder can access talent with relevant expertise, the founder’s network connections to such workers may be weaker than when founding in the same industry.

Proceeding clockwise to the lower right-hand quadrant, the most advantageous entrepreneurial scenario is to launch a venture in the same industry but under a weak non-compete regime. The founder can leverage the founder’s own industry expertise and network to hire others with relevant expertise. But, in a strict non-compete regime (lower left-hand quadrant), founding a firm in the same industry exposes the founder to the greatest legal risk. Not only may the startup be sued for facilitating breach of contract if it tries to hire workers away from rivals, but the founder may be sued for starting the firm in violation of the founder’s own non-compete. Litigation is costly, and

Table 1. Non-competes, Industry Selection, and Ability to Leverage Expertise

	Stricter non-competes	Weaker non-competes
Founding in a different industry	FOUNDER EXPERTISE: N/A, as founder lacks industry experience	FOUNDER EXPERTISE: N/A, as founder lacks industry experience
	EMPLOYEE EXPERTISE: founder risks lawsuit by hiring from within the industry, and lacks networks	EMPLOYEE EXPERTISE: founder can hire those with expertise, though lacks networks
Founding in the same industry	FOUNDER EXPERTISE: founder has expertise but risks a lawsuit by founding in the same industry	FOUNDER EXPERTISE: founder has experience in the same industry and can leverage it
	EMPLOYEE EXPERTISE: founder risks lawsuit by hiring from within the industry	EMPLOYEE EXPERTISE: founder can hire those with expertise, possibly drawing on own networks

startups often lack resources, so the legal risk of being sued for violating a non-compete to start a rival also brings financial risk.

Moreover, as mentioned, even if lawsuits are never actually filed, potential hires from within the same industry may be reluctant to join because they fear a *potential* lawsuit (Marx 2011, Starr et al. 2020a). Thus, the founder may be less eager to even try to hire talent from within the industry given both the difficulty and the risk of legal action. The inability to attract talent from within the industry reduces the chances of success and makes the entrepreneurial enterprise more risky.

Increased Risk of Entrepreneurship Under Non-competes May Disproportionately Discourage Women

As noted, starting a new venture when subject to a strict non-compete policy increases risk in three ways. First, founding a business in the same industry as one’s former employer may incur legal risk if the ex-employer sues for breach of contract. Even if the case is ultimately decided in favor of the founder defendant, a prolonged period of time under a temporary injunction may hurt the performance of the firm. Second, the costs of defending oneself against a lawsuit from an established, well-resourced plaintiff creates financial risks, especially because early stage ventures lack resources. Third, the difficulty of hiring talent from within the industry—not only because of the aforementioned legal and financial risks, but also the reluctance of potential employees to breach their employment agreement—increases the risk that the venture will fail. Taken together, these increased risks of founding a new venture in the same industry when subject to stricter non-compete enforcement may dissuade would-be founders from proceeding. Indeed, prior scholarship finds that within-industry entrepreneurial entry decreases in stricter non-compete

regimes (Stuart and Sorenson 2003, Starr et al. 2017). Hence, those more averse to assuming risks are especially discouraged from proceeding with entrepreneurial entry under strict non-competes.

An extensive literature in psychology and economics provides evidence on gender differences in tolerance for risk (Eckel and Grossman 2008). Experimental results using gambles show higher risk aversion among women: Charness and Gneezy (2012) review dozens of studies *not* designed to study gender differences (and, thus, do not suffer from publication bias) that are nonetheless consistent with this finding. Although laboratory experiments tend to be conducted among college students, Dohmen et al. (2005) demonstrate gender differences in risk aversion among 22,000 German adults (and replicate a laboratory study with a subset of 450 of the respondents). Observational studies of investment patterns (Jianakoplos and Bernasek 1998), securities trading (Barber and Odean 2001), and white-collar crime (Steffensmeier et al. 2013) are moreover consistent with these findings. Although some have expressed skepticism at the magnitude of the findings in certain of these studies (Nelson 2014), the volume of experimental and observational evidence suggests that risk aversion is indeed generally higher among women.

Thus, women may be less eager than men to run the risk of incurring a lawsuit from their ex-employer by violating the postemployment restriction and starting a rival firm. Beyond the litigation risk, the stress associated with a potential lawsuit may exacerbate the effects of risk aversion. Mather and Lighthall (2012) confirm using brain scans that risk aversion is amplified more among women when subjected to stress. Important to note is that the risk-averse response does not require that the worker violating a non-compete is actually sued; as documented by Marx (2011) and Starr et al. (2020a), merely the potential for litigation is

sufficient to compel many workers not to violate their non-compete.

Another reason heightened risk aversion among women might result in a differential response to non-competes is relate to the fact that women are more severely penalized for failure. For example, Egan et al. (2017) report that female financial advisors found guilty of misconduct are more likely to be fired as a consequence. Financial misconduct is a rather crisp example of failure, but this penalty occurs even when the nature of “failure” is less clear. For example, if a surgeon loses a patient, it is often beyond the surgeon’s control to prevent the death. But Sarsons (2020) shows that female cardiologists receive fewer referrals following the death of a patient, controlling precisely for risk factors that contribute to the difficulty of avoiding a fatality. Failure is the modal outcome of the entrepreneurial process, so factors that make failure more likely may particularly discourage women from founding firms. Given the importance of financial capital, for example, the established biases of professional investors against female founders (Brooks et al. 2014, Ewens and Townsend 2020) discourage women from even founding the firm in the first place and contribute to the gender gap in entrepreneurship.

The state sanction of non-competes is established as an institutional factor affecting entrepreneurship. Indeed, researchers establish a negative relationship between non-competes and entrepreneurship. At the firm level, Stuart and Sorenson (2003) find that fewer biotech startups emerge following the completion of an IPO or acquisition when non-competes are more strictly observed. At the regional level, Samila and Sorenson (2011) report that investments by venture capitalists yield fewer startups where non-competes are sanctioned. At the team level, Starr et al. (2017) show that less intraindustry entrepreneurship occurs in states with tighter non-compete policies. However, I am unaware of any evidence regarding whether women are more or less likely to engage in entrepreneurship than men because of non-competes. One reason for this omission may be that nothing in the history of non-competes—which date back to the year 1414—suggests that these employment contracts are intended to target women. Nor are there substantial differences in the rates at women men versus women are asked to sign non-competes.

Small and Pager (2020) claim that institutional discrimination can occur *even despite any intentionality on the part of individuals*, when organizational/legal practices systemically disfavor particular groups. This can be true even when the practice is not designed to discriminate. As an example, they note that, during layoffs, many organizations reduce their liability exposure by using nonarbitrary criteria, such as rank or tenure, in deciding who stays and who goes. The intention of

such programs is not to discriminate against women and/or racial minorities, but because, in practice, white males typically have longer tenure and higher rank, the practice effectively discriminates despite zero intention to do so. Similarly, even though non-competes are designed to block *all* workers from leaving to start rival ventures, differences in risk aversion may lead women to be more likely to honor the employment agreement, not found a firm that leverages their professional experience, and thus contribute to the gender gap.

Empirical Approach

Assessing the connection between non-compete agreements and entrepreneurial activity among women places at least three demands on the data. First, I need to observe workers throughout their careers. Second, the data must contain demographic information on individuals. Third, I need to observe the formation of new firms and whether these are in the same industry as the worker’s prior employer. These requirements rule out the use of cross-sectional survey data, which has been successfully employed in prior work at the intersection of institutions, gender, and entrepreneurship (Elam and Terjesen 2010, Thébaud 2015) and also in prior studies of non-compete agreements (Marx 2011, Starr et al. 2020b). Examining a wide variety of occupations precludes the analysis of patent inventors, which have frequently been used in job-mobility studies but are limited to scientists and engineers and moreover lack reliable data on firm foundings and dissolutions (Marx et al. 2009).

Instead, I use data on the universe of firms and employees from the U.S. Census Bureau (Abowd et al. 2006). The Longitudinal Business Database (LBD) contains all U.S. employers from 1976 to 2014, including founding and dissolution dates. I link the LBD to the Longitudinal Employer Household Dynamics (LEHD), which contains quarterly wages for all U.S.-based employees of firms in the LBD. (For computational tractability, quarterly records are collapsed to the annual level.) Each worker in the LEHD has a unique identifier so that individual careers can be tracked across firms, states, and years. Moreover, each worker’s age, country of birth, gender, and level of education is available.

The LBD and LEHD are complete in their coverage of the full population of firms and workers, but these data also come with restrictions. First, their confidential nature entail that observation counts must be rounded and that summary statistics cannot be revealed for individual states. Second, because the LEHD is available to me only for 25 states and the District of Columbia,² I must account for the possibility that workers move to states I cannot observe. Fortunately, the 2014 edition of the LEHD contains a quarter-by-quarter file of every

worker who was paid wages *somewhere* in the United States even if it is a state for which I lack access. Using this file, I am able to eliminate any worker who was ever paid wages in a quarter when the worker does not appear in the states available to me. Doing so restricts my analysis to approximately 5,762,000 workers whose careers I can observe without interruption.

It is also essential to know who founds a startup. Although the LEHD has information on every worker and firm, it does not report firm ownership. Hence, I follow Kerr and Kerr (2017) as well as Azoulay et al. (2020) in inferring founder status from the LEHD as follows. For each newly founded firm reported in the LBD, I record all workers who had earnings during the quarter when the firm first paid wages. I then label a firm's top three earners in that first wage-paying quarter as "founders." More than 85% of firms had five or fewer employees during their first wage-paying quarter; results are robust to this subset. Results are moreover robust to considering not just the top three but top five earners as founders. If a firm reports more than 50 employees during its first quarter, I do not consider it a startup.

Variation in the State-Level Enforceability of Non-competes

In the United States, the enforceability of non-competes is determined at the state level and not by any federal statute. I follow Hausman and Lavetti (2020) by constructing a year-by-year index of non-compete enforceability at the state level. The full procedure to replicate my index is described in Online Appendix A, but I provide a summary here. I start with Bishara's (2011) 1991 and 2009 state-by-state indexes and determine the values between 1992 and 2008 and 2010 and 2014 by applying the state-level non-compete policy shifts reported by Garmaise (2011) and Ewens and Marx (2017). Bishara's (2011) scale ranges from 0 (North Dakota) to 470 (Florida), which I normalize to the [0, 1] interval.

The LEHD does not contain information on whether individual workers have signed non-competes. Ideally, I would know year-by-year who had signed a non-compete, but panel data on non-competes is not available to my knowledge. To address this limitation, I follow Starr (2019) in differentiating between workers in fields in which non-competes are used widely under the assumption that state-level policies should bind such workers more tightly. Specifically, Starr et al. (2020b) conduct a nationwide survey in which they ask respondents two among many other questions: (a) what industry do you work in, and (b) are you currently subject to a non-compete? The industry categories presented in their survey correspond to two-digit North American Industry Classification System (NAICS) codes. For each NAICS code, they calculate the percentage

of workers who are subject to a non-compete. Starr (2019) presents in table 2 an ordering of these industries by the percentage of workers who were subject to non-competes and separates fields into high versus low use of non-competes. I use their data to create a similar variable for the LEHD, leveraging the fact that they have used NAICS codes for their survey.

Empirical Specification

I estimate a linear probability model on the likelihood of a given worker founding a firm in the same six-digit NAICS as the worker's prior employer in a given year with a host of controls as well as the interaction terms necessary to evaluate the relationship between the state sanction of non-compete agreements and entrepreneurship among women. Standard errors are clustered at the state level. The specification is

$$\begin{aligned} \Pr(\text{FoundRival}_{it}) = & \beta_0 + \beta_1 \text{NCenforce}_{st} + \beta_2 \text{Female}_i \\ & + \beta_3 \text{NChighuse}_{it} + \beta_4 \text{NCenforce}_{st} \\ & \times \text{Female}_i + \beta_5 \text{NChighuse}_{it} \\ & \times \text{Female}_i + \beta_6 \text{NCenforce}_{st} \\ & \times \text{NChighuse}_{it} \times \text{Female}_i \\ & + X_{it} + F_{it+\gamma_t} + \vartheta_{sn} + \varepsilon_{ti}, \end{aligned} \quad (1)$$

where i , s , t , and n index the worker, state, year, and industry, respectively. The dependent variable FoundRival_{it} indicates whether worker i founded a firm in the same six-digit NAICS industry as the prior employer in year t . NCenforce_{st} is the index of non-compete enforceability in state s in year t . NChighuse_{it} indicates whether worker i was employed in a high-use non-compete field in year t . X_i is a vector of demographics, including gender, immigrant, and completion of college (non-time-varying) as well as age, wages, quarters at the employer, and quarters in the six-digit NAICS industry (time varying). The variable F_{it} captures the size of worker i 's employer in year t . γ_t and ϑ_{sn} represent year and state \times industry fixed effects, respectively. ε is the error term. The coefficient of interest in Equation (1) is β_6 . Given the state \times industry fixed effects, the effect of non-competes is identified via within-state changes in enforceability.

Not represented in Equation (1) are interactions of female \times age and female \times college degree as well as interactions of Female_i , NCenforce_{st} , and NChighuse_{it} with (respectively) immigrant status, completion of college degree, age, and prior-year wages. Some specifications involve some of these interaction terms, and others do not.

Main Results

Descriptive statistics for approximately 5,762,000 workers I can observe without interruption are shown Table 2. Slightly less than one third are female, and

Table 2. Descriptive Statistics

	Mean	Median	Standard deviation
Female	0.2968	0	0.4569
Immigrant	0.1296	0	0.3359
Age	46.31	46.5	9.002
No college degree	0.5849	1	0.4927
Prior-year wage (ln)	10.59	10.54	0.6276
NCenforce	0.5569	0.6679	0.2729
NChighuse	0.4607	0	0.4984
Firm tenure (quarters)	26.33	20	23.05
Industry tenure (six-digit NAICS, quarters)	27.78	20	23.86
Founded rival (in same six-digit NAICS)	0.0007	0	0.0272

Notes. Observations are worker-years, collected for approximately 5,762,000 workers whose entire careers occurred in the following 25 states and the District of Columbia: AR, CA, CO, DC, DE, GA, HI, IL, IN, IA, LA, ME, MD, MO, NV, OK, OR, PA, RI, SC, TN, TX, UT, VT, WA. *N* is approximately 81,170,000. Census disclosure rules mandate rounding of observations to the nearest 10,000 and the truncation of reported statistics to four significant figures. Neither minimum nor maximum values can be reported. Median is the mean of observations in the 49th, 50th, and 51st percentiles as per Census disclosure guidelines.

more than half lack a college degree. Workers are 46 years old on average and worked at their employer for 6.5 years, earning on average \$39,735 (1982 deflated). About 1% of workers start a company in the same six-digit NAICS as their former employers.³

All models of Table 2 include year and state \times industry fixed effects as well as interactions of female with age, college degree, firm size, tenure at the firm, and tenure within the industry. Column (1) shows that women are less likely than men to leave their employers to start a rival business, independent of any non-compete considerations. Column (2) adds terms for *NCenforce* and its interaction with female. The negative and statistically significant estimated coefficient on *NCenforce* \times *Female* suggests a relationship between non-competes and gender, but I interpret this model cautiously as it fails to account for the prevalence of non-competes in a worker's field. The remainder of Table 3 interacts this term with *NChighfield*, which captures whether the focal worker is in a field with a high use of non-competes.

Column (3) adds interaction terms for the prevalence of non-competes in the worker's field, including our parameter of interest: *NCenforce* \times *NChighfield* \times *Female*. Adding these terms clarifies that the relationship between higher NC enforceability, gender, and entrepreneurship is largely obtained in high-use non-compete fields. In particular, the coefficient on *NCenforce* \times *NChighfield* \times *Female* is statistically significant at the 0.01% level. Note that column (3) of Table 3 includes not only interactions with female but also age and immigrant (all determined at birth, thus exogenous). In column (4), I include interactions with two possibly endogenous characteristics: (1) whether workers completed a college degree and (2) prior-year earnings. Coefficient estimates are robust to including these factors.

A unit (or one standard deviation) increase in the strictness of the state-level NC policy (*NCenforce*)

reduces the rate of founding a firm in the same six-digit NAICS industry among men in high-non-compete fields by 0.18 percentage points (*NCenforce* \times *NChighfield*).⁴ Relative to men, a unit increase in the state-level NC policy decreases the likelihood for women by 0.02 percentage points. In other words, the impact of non-competes on would-be founders who are female is about 15% stronger. A one-standard-deviation in NC policy (from Table 2, 0.2729) is roughly equivalent to the difference between Florida (the maximum) and Minnesota. A shift from no enforcement to maximum enforcement (e.g., California or North Dakota to Florida) would have about three and one half times the impact.

This finding is moreover robust to including fixed effects for the worker's prior employer in column (5). Doing so accounts for the possibility that firms are idiosyncratic in their treatment of employees who have signed a non-compete as some employers may have a history of litigation or may employ scare tactics to discourage would-be entrepreneurs in ways that are impossible for me to observe in the Census data. Entering firm fixed effects in column (5) yields results generally consistent with column (4) with the estimated coefficient on *NCenforce* \times *NChighfield* \times *Female* statistically significant at the 1% level and of somewhat smaller magnitude. I conclude that the effect of non-competes on women's participation in entrepreneurship is not driven primarily by firm-level practices and return to this theme when examining patterns of non-compete lawsuits.

The remainder of Table 3 contains placebo tests. First, given that non-compete agreements are particular and not general restraints, they should be more enforceable in narrowly defined industries than in broad sectors. The result from columns (3)–(5) show the impact of non-competes on forming rival ventures in the same six-digit NAICS industry, which include very specific fields, such as “flour milling” and “knit fabric mills.” If non-competes are truly responsible for these

findings, we should find stronger effects in narrowly defined industries and weaker effects (if any) in broadly defined industries.

I begin in column (6) by setting the dependent variable to capture *any* instance of entrepreneurship regardless of whether the startup is in the same industry as the founder’s prior employer. Results resemble those of the first four columns of Table 3, which could indicate that non-competes discourage entrepreneurial entry regardless of how close the startup is to the former employer. This would cast doubt on the assertion that non-competes are responsible for these findings. On the other hand, it could be that this result is driven primarily by the instances of within-industry entrepreneurship captured in columns (1)–(4).

Column (7) attempts to resolve these two possible explanations by constructing a dependent variable that captures founding a firm in the same two-digit NAICS as one’s former employer. Examples of two-digit NAICS categories include “manufacturing,” “information,” and “utilities.” Compared with narrowly defined, six-digit NAICS industries, non-competes should have little or no impact in blocking

entrepreneurial area in such broad sectors. (To be clear, this is measuring entrepreneurship that is in the same two-digit NAICS *but not at any finer NAICS level.*) Indeed, the estimated coefficient on $NCenforce \times NChighfield \times Female$ in column (7) is not only much smaller in magnitude but also statistically insignificant. Likewise, the coefficient on $NCenforce \times NChighfield$ is imprecisely estimated. The failure of this placebo test indicates that the effect of non-competes on women and entrepreneurship is limited to the founding of *rival* startups when non-competes should bind most tightly.

Mechanisms

In this section, I evaluate four potential mechanisms underlying Table 3. First, I show that stricter non-competes make it more difficult to hire workers with industry experience, increasing risk and discouraging potential female founders in a manner consistent with heightened risk aversion. Second, I show that the “screening” effect of non-competes applies differently to women (Starr et al. 2017) discouraging the founding

Table 3. Employee Non-compete Policy and the Transition to Entrepreneurship

	Founded startup in same six-digit NAICS					Founded startup in any NAICS	Founded startup in same two-digit NAICS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female	-0.1439*** (0.009317)	-0.1359*** (0.006781)	-0.1402*** (0.006858)	-0.05176+ (0.02598)	0.005555 (0.03483)	-0.3122*** (0.01825)	-0.0975*** (0.0074)
NCenforce		-0.03208 (0.04317)	0.09481* (0.04152)	0.1651 (0.1039)	-0.04801 (0.1161)	0.142 (0.2447)	-0.0001 (0.0001)
NCenforce × Female		-0.01608** (0.005005)	-0.002951 (0.006253)	-0.004038 (0.005829)	0.000081 (0.007124)	0.002536 (0.01447)	-0.00001 (0.00001)
NCenforce × NChighfield			-0.1062*** (0.02645)	-0.1773** (0.05357)	-0.08143+ (0.04702)	-0.6795*** (0.1295)	-0.0001 (0.0001)
NChighfield × Female			0.005305 (0.003312)	0.006709+ (0.003594)	0.002131 (0.003516)	-0.006384+ (0.003468)	-0.0019 (0.0033)
NCenforce × NChighfield × Female			-0.02848*** (0.004669)	-0.02845*** (0.005094)	-0.01945** (0.006821)	-0.02056** (0.007020)	-0.000015 (0.00001)
Year fixed effects, state × industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Female × {Age, No college, firm size, firm tenure, industry tenure}	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NCenforce × NChighuse × {Age, Immigrant}	Yes	Yes	Yes	Yes	Yes	Yes	Yes
NCenforce × NChighuse × {No college degree, Prior-year wages}	No	No	No	Yes	Yes	No	No
Prior-employer fixed effects	No	No	No	No	Yes	No	No

Notes. Observations are worker-years, collected for approximately 5,762,000 workers whose entire careers occurred in the following 25 states and the District of Columbia: AR, CA, CO, DC, DE, GA, HI, IL, IN, IA, LA, ME, MD, MO, NV, OK, OR, PA, RI, SC, TN, TX, UT, VT, WA. *N* is approximately 81,170,000. Census disclosure rules mandate rounding of observations to the nearest 10,000 and the truncation of reported statistics to four significant figures. Models are estimated via ordinary least squares with standard errors clustered at the state level. Coefficients are multiplied by 100 for readability.

+*p* < 0.1; **p* < 0.05; ***p* < 0.01; ****p* < 0.001.

Table 4. Employee Non-compete Policy and Attracting Employees with Industry Experience

DV	% employees with industry experience		% employees with industry experience and had worked with founder	
	all startups		FoundedRival = 1	
	(1)	(2)	(3)	(4)
Sample				
FoundedRival	0.463*** (0.013)	0.404*** (0.012)		
NCenforce		-0.161* (0.072)	-0.111+ (0.059)	0.022 (0.039)
FoundedRival × NCenforce		0.119*** (0.024)		
FoundedRival × NChighfield		0.006 (0.006)		
NCenforce × NChighfield		-0.021 (0.075)	-0.006 (0.077)	0.09 (0.067)
FoundedRival × NCenforce × NChighfield		-0.040*** (0.009)		
Female			0.010** (0.003)	-0.018*** (0.002)
Female × NCenforce			-0.018+ (0.009)	0.004 (0.006)
Female × NChighfield			0.012** (0.004)	0.005 (0.004)
Female × NCenforce × NChighfield			-0.004 (0.011)	-0.024* (0.010)
Mean of DV	0.572	0.572	0.799	0.144
Year fixed effects, state × industry fixed effects	Yes	Yes	Yes	Yes

Notes. Observations are startups founded by 5,762,000 workers whose entire careers occurred in the following 25 states and the District of Columbia: AR, CA, CO, DC, DE, GA, HI, IL, IN, IA, LA, ME, MD, MO, NV, OK, OR, PA, RI, SC, TN, TX, UT, VT, WA. *N* is approximately 70,000 for columns (1) and (2) (all startups regardless of industry overlap but that hired employees beyond the founders) and 40,000 for columns (3) and (4) (startups in the same industry as the founder(s) and that hired employees beyond the founders). Models are estimated via ordinary least squares with standard errors clustered at the state level.

+*p* < 0.1; **p* < 0.05; ***p* < 0.01; ****p* < 0.001.

not of weaker startups but of *riskier* startups. Third, I provide suggestive evidence that, independent of risk aversion, women may anticipate higher relative costs in either defending against non-compete lawsuits or returning to paid employment after abandoning the startup. Fourth, I rule out the possibility that my findings can be explained by firms targeting women in non-compete lawsuits.

Increased Risk of Founding Given Difficulty of Hiring Talent with Industry Experience

I theorize in Table 1 that founders are discouraged from founding in the same industry by stricter non-competes, not only given the personal risk of a lawsuit for breaking their own non-compete, but also given the increased difficulty of hiring talent from within the industry. If, as I argue, non-competes make it harder to hire talent with industry experience, then founding a rival firm when subject to stricter non-competes becomes a riskier endeavor and may be less attractive to women if, as reported in prior literature, women have, on average, lower tolerance for risk.

Table 4 estimates the impact of non-competes on the ability of startups to attract talent with industry experience. Whether a startup employee has prior experience in the same six-digit industry as the startup is determined by reviewing the employee's entire employment history before joining the focal startup. This per-worker information is then collapsed to the startup level, creating as a dependent variable the percentage of employees at a given startup with industry experience. (Note that this variable is calculated only for startups that hired workers beyond the founders; otherwise, the DV is undefined.) State × industry and year fixed effects are again included as in Table 3 with standard errors clustered at the state level.

Column (1) of Table 4 shows that founders hire a higher percentage of employees with industry expertise when founding in an industry in which they have prior experience. The variable *FoundedRival* captures whether the focal startup was founded in the same six-digit NAICS industry as the founders. It resembles the dependent variable from the first five columns of Table 3 except calculated at the firm level.⁵ It is

perhaps not surprising that within-industry startups hire a higher percentage of employees with industry experience. Within-industry founders would understand the value of having relevant human capital and, thus, want to hire employees with such experience.

When subject to stricter non-competes, however, same-industry founders hire a *lower* percentage of employees with industry experience. This reversal is visible in the estimated coefficient on $FoundedRival \times NCenforce \times NChighfield$, which is negative and statistically significant at the 0.1% level. Given that within-industry founders hire more employees with industry experience when they are able to, column (2) indicates that—consistent with the lower left-hand corner of Table 1—stricter non-competes prevent founders from mitigating startup risk by hiring workers with relevant human capital. Coupled with the difficulty of a founder leveraging the founder's own experience in a rival startup, founding when subject to non-competes increases risk. If women tend to be more risk averse, non-competes may, therefore, discourage them from founding rival firms given the difficulty of acquiring talent with industry experience.

In the remainder of Table 4, I explore whether female entrepreneurs are less successful in acquiring experienced employees. I start in column (3), revisiting column (2) by gender. To avoid quadruple interactions, I execute this test on the subsample in which $FoundedRival = 1$. The estimated coefficient on $Female \times NCenforce \times NChighfield$ in column (3) is negative, perhaps suggestive that women are less likely than men to hire talent with industry experience, but it is imprecisely estimated. I dig deeper by examining whether women are less likely to utilize their professional networks to hire experienced talent when subject to non-competes. In column (4), I define the dependent variable as the percentage of startup employees who worked in the same industry at some prior point in their career *and* who had worked with the founder previously (i.e., both were at the same firm in the same year). The estimated coefficient on $Female \times NCenforce \times NChighfield$ in column (4) indicates that non-competes discourage women from recruiting their prior coworkers who have industry experience. One possible explanation for this distinction is that women are less eager to put their colleagues at risk for a non-compete lawsuit.

In sum, the analysis of Table 4 confirms the theorizing of Table 1 that non-competes make it more difficult to hire employees with industry experience. Coupled with the risk founders run of starting a company in the same industry as their former employer, non-competes raise the risk involved with entrepreneurship. To the extent that women are more risk averse than men, this heightened risk helps to explain why women subject to strict non-competes are less likely to start rival businesses.

Screening out of Higher Potential, Higher Risk Ventures Among Women

Additional evidence consistent with role of risk aversion in the effect of non-competes on women in entrepreneurship is found in screening (Starr et al. 2017). Focusing on startups with multiple founders from the same prior employer, they find that same-industry startups founded under stricter non-compete enforcement have higher initial size and achieve a larger eventual size.⁶ They interpret these results to mean that weaker or less promising startups are “screened out” because founders are less willing to take the risk for a likely lower payoff. I find that non-competes have a different effect among women, screening out not startups that are smaller or weaker but rather those that are at once higher potential and higher risk.

The contrast in screening behavior is visible in Table 5. In doing this comparison, it is key to control for industry as men and women may select into industries with widely varying average size and/or potential and, thus, include industry \times state and year fixed effects as in Table 5. Column (1) of Table 5 focuses on initial size (i.e., logged number of employees) for the within-industry startups from columns (3) and (4) of Table 4. If it were the case that the results in Table 3 were driven by increased screening of weaker startups that would otherwise been founded by women, one would anticipate a positive and statistically significant estimated coefficient on $Female \times NCenforce \times NChighfield$. Although positive, this coefficient is very imprecisely estimated, making it difficult to conclude that the screening effect reported by Starr et al. (2017) is exacerbated among women.

In fact, columns (2) and (3) indicate that the opposite may be the case: what would have been some of the *strongest* female-founded startups are screened out by non-competes. Column (2) of Table 5 shows that the size ultimately achieved by women-founded startups is materially lower, not higher, under strict non-compete enforcement. Note that this difference cannot be attributed to differences in initial size of male- versus female-founded startups under stricter non-competes, which is imprecisely estimated in column (1). Moreover, the difference in size achieved is recovered when controlling for initial size in column (3).

A more compelling explanation may lie in the results of Table 4, which demonstrate that founders are less likely to hire employees with industry experience when subject to non-competes and that women are especially less likely to leverage their prior professional networks to attract workers with industry experience when subject to stricter non-competes. This could explain the results in columns (2) and (3) of Table 5 in two ways. First, the difficulty of hiring talent from the same industry under stricter non-competes could discourage women from founding high-potential rival

Table 5. Employee Non-compete Policy and Possible Screening of Female-Founded Startups

DV	Ln initial size (1)	Ln final size (2)	Ln final size (3)	Years survived (4)
<i>Female</i>	−0.056*** (0.012)	0.008 (0.008)	0.036*** (0.008)	0.031 (0.063)
<i>NCenforce</i>	1.510** (0.505)	1.132* (0.443)	0.367 (0.364)	1.736** (0.539)
<i>Female</i> × <i>NCenforce</i>	0.058 (0.034)	0.103** (0.028)	0.074** (0.022)	−0.413*** (0.096)
<i>Female</i> × <i>NChighfield</i>	−0.061*** (0.016)	0.033 ⁺ (0.019)	0.064*** (0.017)	−0.263** (0.078)
<i>NCenforce</i> × <i>NChighfield</i>	0.164 (0.289)	−0.036 (0.150)	−0.119 (0.103)	−1.605 (1.104)
<i>Female</i> × <i>NCenforce</i> × <i>NChighfield</i>	0.023 (0.050)	−0.232*** (0.042)	−0.244*** (0.042)	0.993*** (0.211)
Ln initial size			0.506*** (0.027)	
Constant	1.349*** (0.242)	1.697*** (0.225)	1.014*** (0.241)	9.747*** (0.411)
Year fixed effects, state × industry fixed effects	Yes	Yes	Yes	Yes

Notes. Observations are startups founded by 5,762,000 workers whose entire careers occurred in the following 25 states and the District of Columbia: AR, CA, CO, DC, DE, GA, HI, IL, IN, IA, LA, ME, MD, MO, NV, OK, OR, PA, RI, SC, TN, TX, UT, VT, WA. *N* is approximately 40,000, including startups in the same industry as the founder(s) and that hired employees beyond the founders. Models are estimated via ordinary least squares with standard errors clustered at the state level.

⁺*p* < 0.1; **p* < 0.05; ***p* < 0.01; ****p* < 0.001.

ventures that depend crucially on such expertise. Second, even if women found such ventures, they may find themselves disproportionately limited in their ability to grow these ventures. Although I am not able to distinguish reliably between these two mechanisms, it is possible that either or both is a contributing factor.

Finally, column (4) examines survival. The estimated coefficient on *Female* × *NCenforce* × *NChighfield* is positive and statistically significant, showing that female-founded firms subject to stronger NCs survive longer. Again, this result is not driven by initial size, which does not differ by gender.

Taken together, the results of Table 5 indicate that non-competes push women toward founding businesses that do not fail as soon but also do not grow as large. Thus, when considering the effect of non-competes on women, it is not weaker startups that are screened out, but perhaps *riskier* startups: those that have greater growth potential but are also more likely to fail. This finding is consistent with the notion that higher risk aversion among women is responsible for the lower levels of entrepreneurship among women subject to stricter non-competes. Again returning to Table 4, one reason women may be averse to pursuing higher potential, higher risk startups when subject to stricter non-competes is it is more difficult to mitigate startup risk by hiring talent with industry experience.⁷ This may disproportionately discourage would-be female founders from proceeding with high-potential, high-risk ventures.

Higher Relative Costs of Non-competes for Female Entrepreneurs

The foregoing mechanisms rely on the assumption that women are more risk averse than men, but it is possible that non-competes discourage women from founding firms even independent of risk aversion. For example, even if the expense of mounting a legal defense against a non-compete does not differ for men versus women, the *relative* costs are higher if women have fewer financial resources available. As shown in columns (1) and (2) of Table 6, men have higher cumulative earnings before starting their first firm even when controlling for age, education, the number of years before starting a company, and when adding fixed effects at the six-digit NAICS level in column (2). The differential in pre-founding earnings may be explained by women having less experience and in lower paying roles (Loscocco et al. 1991).

Moreover, women may face higher relative costs upon returning to paid employment if they abandon their startup—for example, after losing or declining to contest a non-compete lawsuit. Prior studies show that entrepreneurs suffer a wage penalty upon returning to paid employment (Campbell 2013) though not along gender lines. If women are penalized more for failure (Egan et al. 1997, Sarsons 2020), women who abandon startups might be less well compensated upon returning to paid employment. The sample analyzed in Table 6 does not include entrepreneurs whose startups were acquired; rather, the founder either shut down the firm or left it in the hands of

Table 6. Pre-entrepreneurial vs. Postentrepreneurial Earnings

	Cumulative earnings before startup		Annual earnings, poststartup versus prestartup	
	(1)	(2)	(3)	(4)
Female	-0.2679*** (0.0102)	-0.3361*** (0.0111)	-0.1144*** (0.0242)	-0.1773*** (0.0256)
Age	0.0176*** (0.0006)	0.0173*** (0.0006)	-0.0312*** (0.0018)	-0.0318*** (0.0018)
No college degree	-0.3374*** (0.0086)	-0.2555*** (0.0089)	-0.2817*** (0.0264)	-0.1966*** (0.0268)
Immigrant	0.0117 (0.0134)	-0.0072 (0.0134)	-0.0708 (0.0428)	-0.0892 (0.0455)
Years worked before founding	0.2024*** (0.0012)	0.2023*** (0.0012)	-0.0749*** (0.0044)	-0.0748*** (0.0044)
Year fixed effects; state fixed effects	Yes	Yes	Yes	Yes
Six-digit NAICS industry fixed effects	No	Yes	No	Yes

Notes. Observations are approximately 20,000 entrepreneurs who held jobs at established companies both before and after their startup and who only worked in the following 25 states and District of Columbia: AR, CA, CO, DC, DE, GA, HI, IL, IN, IA, LA, ME, MD, MO, NV, OK, OR, PA, RI, SC, TN, TX, UT, VT, WA. Models are estimated via ordinary least squares with standard errors clustered at the state level. Earnings are calculated as 1982 deflated wages recorded in the LEHD. For columns (1) and (2), the DV is the sum of all reported earnings (if any) prior to the worker founding her first entrepreneurial venture. For columns (3) and (4), the DV is the worker’s average annual earnings prior to founding the worker’s first entrepreneurial venture divided by earnings in the first year after abandoning their venture and returning to paid employment with an existing firm. Startups that were acquired are excluded.

* $p < 0.1$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

others. In columns (3) and (4) of Table 6, I estimate the ratio of these former entrepreneurs’ earnings in the first year after abandoning the startup—that is, upon their return to paid employment—divided by their average earnings prior to founding. Column (3) shows that this earnings ratio is lower for women versus men, suggesting that women may be more heavily penalized for returning to the labor force following entrepreneurial failure. The industry fixed effects in column (4) confirm this is not a result of selection into different industries. That said, it is unclear whether women would be aware of such a penalty, so I treat this only as suggestive evidence.

Do Firms Target Women with Non-compete Lawsuits?

I theorize earlier that heightened risk aversion among women leads them to less often violate a non-compete by setting up a rival firm. But perhaps women face a materially higher hazard of being sued for breaking a non-compete because firms target them disproportionately in lawsuits. Consider a scenario in which someone founds a firm but is sued soon thereafter and perhaps even before the startup begins to pay wages. In such a case, the founding event would not be reflected in the LEHD data and would not influence the estimates in Table 3. This alternative mechanism seems particularly plausible given women’s higher relative costs of legal defense reported earlier. Additionally, to the degree that women are less likely to negotiate the terms of a non-compete, they may also be subject to more stringent non-compete terms regarding duration or the field of service (see also Babcock and Laschever

2009 for evidence that women are generally less likely to negotiate employment agreements). Stricter terms might lead ex-employers to believe that they will prevail in court.

To examine whether women are disproportionately targeted in non-compete lawsuits, I compare the gender of defendants in non-compete lawsuits versus the expected gender distribution of lawsuits given the occupational profiles of defendants. Doing so is difficult using legal databases such as Westlaw or Lexis as these contain only published court decisions and not every case filed. As noted, it is common for a judge to issue a preliminary injunction in a non-compete case. Many cases settle, so the number of published decisions may dramatically understate the actual number of lawsuits. However, the Courthouse News Service (CNS) curates a database of all cases *filed* in a large subset of district courts across the United States.

I extracted 7,931 lawsuits from January 2003 through August of 2017 that reference “non-compete,” “covenant not to compete,” “postemployment restraints,” or similar terms. The 11,975 defendants had 2,459 unique given names, 87% of which could be reliably classified according to genderize.io, a website that generates country-specific probabilities. I assigned names to a particular gender with a confidence score of >80%. (Classifications are available from the author.) For the remaining 415 defendants, I searched for profiles on LinkedIn, Facebook, and company websites, assigning gender only when (a) the defendant’s name and company could both be confirmed and (b) personal pronouns, suffixes (i.e., “Jr.”), or photos provided clues. All but 59 defendants could be classified.

Table 7. Expected Gender Distribution of Defendants in Employee Non-compete Lawsuits

Occupation	Number of CNS defendants	Percentage of of CNS defendants	BLS estimates of percentage females in occupation	Projected percentage of female defendants	Non-compete usage by occupation (mean = 1)	Projected percentage of female defendants (adjusted)
Sales	727	34.7	49.0	17.0	0.8	14.0
Manager	554	26.4	39.1	10.3	1.6	17.0
Engineer	214	10.2	14.2	1.4	1.9	2.8
Doctor	182	8.7	38.2	3.3	1.0	3.4
Salon	124	5.9	92.4	5.5	1.0	5.6
Broker	83	4.0	52.4	2.1	1.2	2.6
CEO	39	1.9	27.3	0.5	1.6	0.8
Coach	34	1.6	34.2	0.6	1.0	0.6
Real estate	37	1.8	55.5	1.0	1.0	1.0
Construction	39	1.9	3.0	0.1	0.6	0.0
Financial advisor	37	1.8	52.5	0.9	1.2	1.1
Physical therapy	15	0.7	69.3	0.5	1.4	0.7
Customer care	8	0.4	65.0	0.2	1.0	0.3
Artist	3	0.1	56.6	0.1	1.2	0.1
Total	2,096			43.5		50.2

Notes. Table shows the occupations in which employee non-compete lawsuits could be classified between 2003 and 2017 by the Courthouse News Service (CNS). The number and percentage of defendants who are female is then shown for all lawsuits reported by CNS. The percentage of workers in each occupation that are female (from the Current Population Survey) is presented and multiplied by the percentage of lawsuits in each occupation and non-compete usage by occupation to predict the final column, expected percentage of female defendants in each occupation. The italicized totals in the final row sum the values for all occupations in the rows above.

Only 26.7% of defendants were classified as female. Whether 26.7% is higher or lower than one would expect depends on the occupations in which non-compete lawsuits occur. For example, if non-competes were used only in the construction industry, nearly all defendants should be male as the Current Population Survey (CPS) reports a 93% male workforce. I, therefore, classified occupation for the 2,096 defendants for which CNS reported this field as shown in Table 7.

The next step in constructing an expected gender distribution of non-compete lawsuit defendants was to input the gender distribution for defendants' occupations. For example, CPS reports that 49% of salespeople are female, whereas that figure is 14.2% of engineers and 92.4% of hairstylists. I multiply the percentage of lawsuits in each occupation by the percentage of workers in each industry who are female (per the CPS) and then sum these to yield an overall expected percentage of female defendants. This figure is 43.5% as shown at the bottom of the fourth column of Table 7. Finally, I adjust this prediction given the prevalence of non-competes per occupation from Starr et al. (2020b), raising the expected percentage of female defendants to 50.2%.⁸ Given that only 26.7% of actual CNS defendants appear to be female, this indicates nearly a two times underrepresentation of women as defendants in non-compete lawsuits.

In sum, it does not appear that firms target women with non-compete lawsuits. In fact, the severe underrepresentation of women in non-compete lawsuits underscores that non-competes may have a stronger

chilling effect on women—consistent with the mechanisms of higher relative costs for defending against a lawsuit and a higher wage penalty upon returning to paid employment.

Discussion and Conclusion

I interpret these results cautiously for several reasons. First, because the LEHD data available to me cover only 25 states and the District of Columbia, it is possible that dynamics for those who worked the remaining 25 states differ meaningfully. Another limitation, as discussed, is that the Census data do not indicate who signed a non-compete (indeed, no panel data to my knowledge include this information, only point-in-time surveys). My approach to addressing this limitation, following prior work (Starr 2019), is to compare fields in which non-competes are more frequently used versus in which they are less common. But it is possible that the central tendencies of this variable are somehow unrepresentative.

Notwithstanding these limitations, this paper provides evidence that women are less likely than men to start rival businesses when they are subject to employee noncompetition agreements. This finding is shown using register data on all employees in 25 states and D.C. from 1990 to 2014 whose careers were entirely captured in those states. Although it has long been known that non-competes discourage entrepreneurial activity, this is, to my knowledge, the first evidence that the effect of non-competes on entrepreneurship varies by gender. If these findings are correct, then

one step to reducing the gender gap in entrepreneurship would be to curtail the use of non-compete agreements. One attractive aspect of addressing institutional discrimination is that it could be implemented quickly (via judicial fiat or legislative vote), whereas the attitudes that underpin individual discrimination can be slow to change.

I show that heightened risk aversion among women (Eckel and Grossman 2008, Bertrand 2011, Charness and Gneezy 2012) discourages them from violating a non-compete by setting up a rival firm. I establish two mechanisms underlying this finding. First, I show that, although same-industry founders prefer to hire workers with industry experience, they are unable to do so when subject to strict non-competes. Women subject to stricter non-competes are particularly unlikely to hire workers with industry experience from their own networks, increasing the chances of failure and discouraging them from founding in the first place. Second, non-competes redirect women away from high-growth, high-risk ventures toward startups that stay in business but do not grow. Moreover, I am able to rule out the alternative mechanism that firms target women in non-compete lawsuits. Thus, non-competes not only discourage entrepreneurial activity among women in general, but in particular, depress the launch of high-potential ventures as non-competes make risky businesses even riskier.

This paper contributes to the nascent literature on formal institutions, gender, and entrepreneurship (Elam and Terjesen 2010, Thébaud 2015, Castellaneta et al. 2020). Legislative efforts focus on implementing programs targeted at ameliorating frictions that may hold women back from pursuing entrepreneurial ventures. My findings complement this work by showing that ostensibly gender-neutral policies may nonetheless have substantially different outcomes by gender. This result also adds to a growing body of work (Small and Pager 2020) on institutional discrimination driven by organizational practices not originally intended to discriminate but nonetheless yield differential outcomes.

These results also have implications for public policy. Although non-compete statutes often vary by occupation or wage, I am unaware of any state that has implemented gender-specific non-compete policies. Many states have pursued reforms in non-compete policy, so understanding that non-competes disproportionately affect women may further galvanize these efforts. Moreover, the fact that non-compete effects on women are not driven by actual lawsuits, but instead by the chilling effect of potential lawsuits suggests a different approach to reform. Legislators often focus on courtroom outcomes, including whether the non-compete will be enforced by a judge. My results reinforce previous findings (Marx 2011, Starr et al. 2020a) that even the *possibility* of a lawsuit may have a chilling effect on

worker behavior. This raises the question of whether employers should be allowed to ask their employees to sign such contracts in the first place. As of January 1, 2022, Illinois will become the first state with a law that restricts not just the enforceability but the use of non-competes with certain workers: “No employer shall enter into a covenant not to compete with any employee unless the employee’s annualized rate of earnings exceeds \$75,000 per year” (Illinois General Assembly, 2021). Similar approaches in other states may help to shrink the gap.

Acknowledgments

The author is grateful to Lauren Aydinliyim, John Bauer Esq., Russell Beck Esq., Norm Bishara, Esq. Niles Eldredge, Michael Ewens, Benjamin Fink Esq., Lee Fleming, Jessica Jeffers, Aleksandra Kacperczyk, Sarah Kaplan, David Pardue Esq., Olav Sorenson, and Timothy Simcoe for feedback as well as to seminar participants at the NBER Summer Institute, the Kauffman Foundation Entrepreneurship Scholars Conference, Baruch College, Yeshiva University, Cornell University, and Copenhagen Business School. Any views expressed are those of the author and not those of the U.S. Census Bureau. The Census Bureau’s Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 1182. (CBDRB-FY21-P1182-R9045). Mistakes are the author’s.

Endnotes

¹ See https://obamawhitehouse.archives.gov/sites/default/files/docs/eleven_facts_about_family_and_work_final.pdf.

² Coverage of LEHD data available to me includes MD/CO/IL/IN/LA/MO/WA since 1990, CA/OR/PA since 1991, GA since 1994, NM/RI/TX since 1995, HI/ME since 1996, DE/IA/NV/SC/TN since 1998, UT since 1999, OK/VT since 2000, DC since 2002, and AR since 2003.

³ From Table 2, there are about 57,000 startups in this sample. This does not resemble the total annual number of U.S. startups because both this figure is limited to same-industry rivals and count is only for the approximately six million workers whose entire careers can be observed in the LEHD states available to me.

⁴ For purposes of calibration, Starr et al. (2017) find that a unit increase in NC policy is associated with a 0.13 percentage point decrease in within-industry entrepreneurship.

⁵ For the firm-level analysis in Tables 4 and 5, *NCenforce* and *NChighfield* are calculated as the maximum values for any founder. Unreported results using instead mean or median are similar.

⁶ The long-term size results in Starr et al. (2018) are tempered by controlling for initial size.

⁷ An alternative interpretation of Table 5 is that non-competes do not dissuade women from *founding* high-potential ventures, but that these ventures simply fail to grow (for example, because of the difficulty of hiring talent with relevant industry experience). If this were the case, we would expect an increase in failure, but that is not what column (4) indicates.

[‡] Given that occupation is available only for about one fifth of CNS cases, I reestimate given 755 NC-related decisions in Lexis (2003–2006), in which every decision includes defendant occupation. The distribution of occupations is generally similar though lawsuits involving CEOs are overrepresented in Lexis. Using Lexis instead of CNS yields an estimated 49.1% of defendants being female—quite similar to the 50.2% obtained using CNS.

References

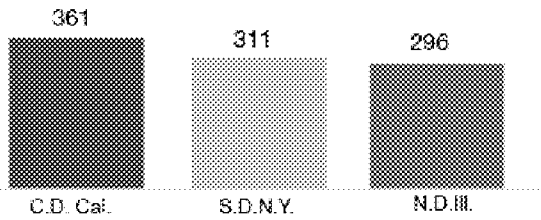
- Abowd J, Stephens B, Vilhuber L, Andersson F, McKinney K, Roemer M, Woodcock S (2006) The LEHD infrastructure files and the creation of the Quarterly Workforce Indicators. U.S. Census Bureau, LEHD Program, Technical Paper No. TP-2006-01. Suitland, MD.
- Acs ZJ, Desai S, Hessels J (2008) Entrepreneurship, economic development and institutions. *Small Bus. Econom.* 31(3):219–234.
- Agarwal R, Echambadi R, Franco AM, Sarkar MB (2004) Knowledge transfer through inheritance: Spin-out generation, development, and survival. *Acad. Management J.* 47(4):501–522.
- Ahmed S (1985) nAch, risk-taking propensity, locus of control and entrepreneurship. *Personality Individual Differences* 6(6):781–782.
- Aidis R, Estrin S, Mickiewicz T (2008) Institutions and entrepreneurship development in Russia: A comparative perspective. *J. Bus. Venturing* 23(6):656–672.
- Azoulay P, Jones B, Kim JD, Miranda J (2020) Age and high growth entrepreneurship. *Amer. Econom. Rev. Insights* 2(1):65–82.
- Babcock L, Laschever S (2009) *Women Don't Ask: Negotiation and the Gender Divide* (Princeton University Press, Princeton, NJ).
- Barber BM, Odean T (2001) Boys will be boys: Gender, overconfidence, and common stock investment. *Quart. J. Econom.* 116(1):261–292.
- Belcourt M, Burke R, Lee-Gosselin H (1991) *The Glass Box: Women Business Owners in Canada* (Canadian Advisory Council on the Status of Women).
- Bhide A (2000) *The Origin and Evolution of New Businesses* (Oxford University Press, Oxford, England).
- Bertrand M (2011) New perspectives on gender. Ashenfelter O, Card D, eds. *Handbook of Labor Economics*, vol. 4 (Elsevier, North-Holland, Netherlands), 1543–1590.
- Bishara ND (2011) Fifty ways to leave your employer: Relative enforcement of covenants not to compete, trends, and implications for employee mobility policy. *Univ. Pennsylvania J. Bus. Law* 13:751–795.
- Brooks AW, Huang L, Kearney SW, Murray FE (2014) Investors prefer entrepreneurial ventures pitched by attractive men. *Proc. Natl. Acad. Sci. USA* 111(12):4427–4431.
- Brush C, Greene P, Balachandra L, Davis A, Blank AM (2014) *Women Entrepreneurs 2014: Bridging the Gender Gap in Venture Capital* (Arthur M. Blank Center for Entrepreneurship Babson College), 1–26.
- Campbell BA (2013) Earnings effects of entrepreneurial experience: Evidence from the semiconductor industry. *Management Sci.* 59(2):286–304.
- Castellaneta F, Conti R, Kacperczyk A (2020) The (un)intended consequences of institutions lower barriers to entrepreneurship: The impact on female workers. *Strategic Management J.* 41(7):1274–1304.
- Charness G, Gneezy U (2012) Strong evidence for gender differences in risk taking. *J. Econom. Behav. Organ.* 83(1):50–58.
- Chatterji AK (2009) Spawned with a silver spoon? Entrepreneurial performance and innovation in the medical device industry. *Strategic Management J.* 30(2):185–206.
- Coleman S, Robb A (2009) A comparison of new firm financing by gender: Evidence from the Kauffman Firm Survey data. *Small Bus. Econom.* 33(4):397–411.
- Dohmen TJ, et al. (2005) Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey. Mimeo.
- Eckel CC, Grossman PJ (2008) Men, women and risk aversion: Experimental evidence. Plott D, Smith V, eds. *Handbook of Experimental Economics Results*, vol. 1, 1061–1073.
- Egan ML, Matvos G, Seru A (2017) When Harry fired Sally: The double standard in punishing misconduct. NBER Working Paper No. 23242, National Bureau of Economic Research, Cambridge, MA.
- Elam A, Terjesen S (2010) Gendered institutions and cross-national patterns of business creation for men and women. *Eur. J. Development Res.* 22:331–348.
- Ewens M, Marx M (2017) Founder replacement and startup performance. *Rev. Financial Stud.* 31(4):1532–1565.
- Ewens M, Townsend RR (2020) Are early-stage investors biased against women? *J. Financial Econom.* 135(3):653–677.
- Garmaise M (2011) Ties that truly bind: Noncompetition agreements, executive compensation, and firm investment. *J. Law Econom. Organ.* 27(2):376–425.
- Global Entrepreneurship Monitor (2017) Global Report. Global Entrepreneurship Research Association (GERA, London Business School, London).
- Goldin C, Katz LF, Kuziemko I (2006) The homecoming of American college women: The reversal of the college gender gap. *J. Econom. Perspect.* 20(4):133–156.
- Gupta V, Bhawe N (2007) The influence of proactive personality and stereotype threat on women's entrepreneurial intentions. *J. Leadership Organ. Stud.* 13(4):73–85.
- Guzman J, Kacperczyk AO (2019) Gender gap in entrepreneurship. *Res. Policy* 48(7):1666–1680.
- Hall RE, Woodward SE (2010) The burden of the nondiversifiable risk of entrepreneurship. *Amer. Econom. Rev.* 100(3):1163–1194.
- Hausman N, Lavetti K (2020) Physician concentration and negotiated prices: Evidence from state law changes. *Amer. Econom. J. Appl. Econom.* 13(2):258–296.
- Henrekson M, Rosenberg N (2001) Designing efficient institutions for science-based entrepreneurship: Lesson from the US and Sweden. *J. Tech. Transfer* 26(3):207–231.
- Hvide H, Panos G (2014) Risk tolerance and entrepreneurship. *J. Financial Econom.* 111(1):200–223.
- Illinois General Assembly (2021) SB0672 Enrolled: An Act Concerning Business. Accessed August 7, 2021, <https://www.ilga.gov/legislation/102/SB/PDF/10200SB0672lv.pdf>.
- Jianakoplos NA, Bernasek A (1998) Are women more risk averse? *Econom. Inquiry* 36(4):620–630.
- Kanze D, Huang L, Conley MA, Higgins ET (2018) We ask men to win and women not to lose: Closing the gender gap in startup funding. *Acad. Management J.* 61(2):586–614.
- Kacperczyk A, Younkin P (2017) The paradox of breadth: The tension between experience and legitimacy in the transition to entrepreneurship. *Admin. Sci. Quart.* 62(4):731–764.
- Kerr S, Kerr W (2017) Immigrant Entrepreneurship. Haltiwanger J, Hurst E, Miranda J, Schoar AA, eds. *Measuring Entrepreneurial Businesses: Current Knowledge and Challenges*. (University of Chicago Press, Chicago), 187–252.
- Klepper S (2009) Spinoffs: A review and synthesis. *Eur. Management Rev.* 6(3):159–171.
- Loscocco KA, Robinson J, Hall RH, Allen JK (1991) Gender and small business success: An inquiry into women's relative disadvantage. *Soc. Forces* 70(1):65–85.
- Markussen S, Roed K (2017) The gender gap in entrepreneurship—The role of peer effects. *J. Econom. Behav. Organ.* 34:356–373.
- Marx M (2011) The firm strikes back: Non-compete agreements and the mobility of technical professionals. *Amer. Sociol. Rev.* 76(5):695–712.

- Marx M, Strumsky D, Fleming L (2009) Mobility, skills, and the Michigan non-compete experiment. *Management Sci.* 55(6): 875–889.
- Mather M, Lighthall NR (2012) Risk and reward are processed differently in decisions made under stress. *Current Directions Psych. Sci.* 21(1):36–41.
- Minniti M, Arenius P (2003) Women in entrepreneurship. *Entrepreneurial Advantage Nations First Annual Global Entrepreneurship Sympos.*, vol. 29.
- Miric M, Yin P (2020) Population-level evidence of the gender gap in technology entrepreneurship. Mimeo.
- Nanda R, Sørensen J (2010) Workplace peers and entrepreneurship. *Management Sci.* 56(7):1116–1126.
- Nelson J (2014) Are women really more risk-averse than men? A re-analysis of the literature using expanded methods. *J. Econom. Surveys* 2(3):1–20.
- Organisation for Economic Co-operation and Development (2016) Policy brief on women’s entrepreneurship. Accessed March 18, 2020, <http://www.oecd.org/cfe/smes/Policy-Brief-on-Women-s-Entrepreneurship.pdf>.
- Organisation for Economic Co-operation and Development (2017) The missing entrepreneurs. Accessed November 30, 2018, <http://www.oecd.org/industry/the-missing-entrepreneurs-2017-9789264283602-en.htm>.
- Sarsons H (2020) Interpreting signals in the labor market. Mimeo.
- Samila S, Sorenson O (2011) Venture capital, entrepreneurship, and economic growth. *Rev. Econom. Statist.* 93(1):338–349.
- Small M, Pager D (2020) Sociological perspectives on racial discrimination. *J. Econom. Perspect.* 34(2):59–67.
- Steffensmeier DJ, Schwartz J, Roche M (2013) Gender and twenty-first-century corporate crime: Female involvement and the gender gap in Enron-era corporate frauds. *Amer. Sociol. Rev.* 78(3): 448–476.
- Starr E (2019) Consider this: Training, wages, and the enforceability of covenants not to compete. *Indust. Labor Relations Rev.* 72(4): 783–817.
- Starr E, Balasubramanian N, Sakakibara M (2018) Screening spinouts? How non-compete enforceability affects the creation, growth, and survival of new firms. *Management Sci.* 64(2):552–572.
- Starr E, Prescott, JJ, Bishara N (2020a) The behavioral effects of (unenforceable) contracts. *J. Law Econom. Organ.* 36(3):633–687.
- Starr E, Prescott JJ, Bishara N (2020b) Non-competes in the U.S. labor force. *J. Law Econom.* 64(1):53–84.
- Stuart T, Sorenson O (2003) Liquidity events and the geographic distribution of entrepreneurial activity. *Admin. Sci. Quart.* 48(2): 175–201.
- Thébaud S (2010) Gender and entrepreneurship as a career choice: Do self-assessments on ability matter? *Soc. Psych. Quart.* 73(2): 288–304.
- Thébaud S (2015) Business as plan B: Institutional foundations of gender inequality in entrepreneurship across 24 industrialized countries. *Admin. Sci. Quart.* 60(4):671–711.

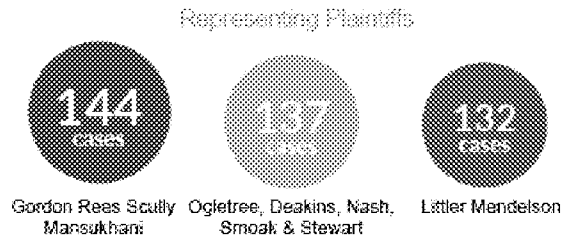
Matt Marx is the Bruce F. Failing Sr. Chair in Entrepreneurship and professor of management and organizations at Cornell University. His research interests include the commercialization of science and technology, especially by new ventures.

TRADE SECRET LITIGATION REPORT 2023

Districts with Most Cases Filed 2018-2022

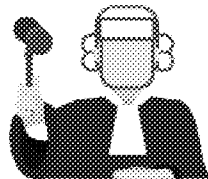


Most Active Law Firms by Cases 2018-2022



Most Active Judges 2018-2022

Judge Pitman of the Western District of Texas heard the highest number of Trade Secret cases: 50 cases.



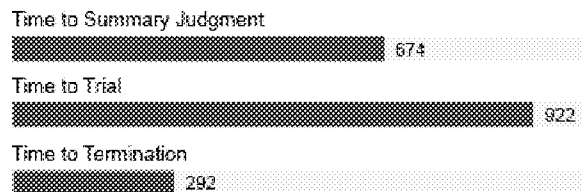
Representing Defendants



Most Active Plaintiffs by Cases 2018-2022

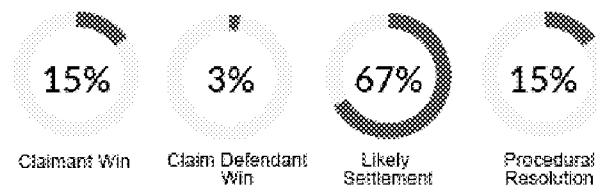


Median Time to Judgment in Days for Cases Terminated 2018-2022



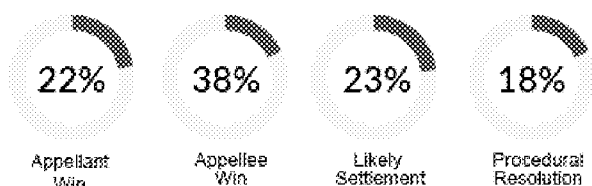
Case Resolutions for Cases Terminated 2018-2022

In Federal District Court

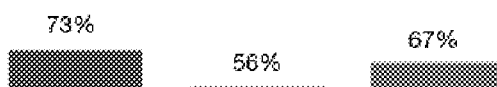


Case Resolutions for Cases Terminated 2018-2022

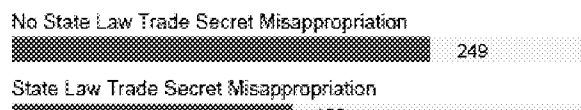
In Federal Appellate Court

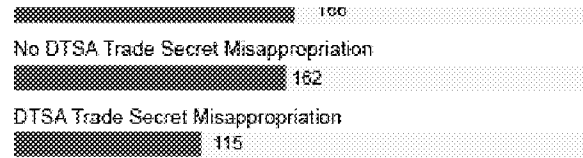
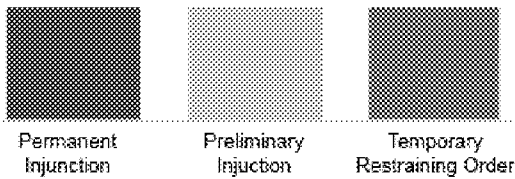


Injunctions % Granted by Judgment on the Merits for Cases Terminated 2018-2022



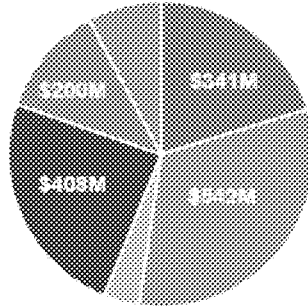
Findings in Cases Terminated 2018-2022





Damages Breakdown 2018-2022

- Actual Damages/Lost Profits: \$341M
- Punitive/Willfulness Damages: \$542M
- Reasonable Royalty: \$64M
- Other/Mixed Damages: \$408M
- Prejudgment Interest: \$200M
- Attorneys' Fees/Costs: \$132M



Total Damages Awarded in 2022

\$85M

Cases Awarded Damages in 2022

24 cases



[Click here to request the full 2023 Trade Secret Litigation Report](#)

Product:
Legal Analytics
Platform
Quick Tools
How We're Different
Federal Practice Areas
State Courts
All Areas of Coverage

Business Type:
Law Firms
Companies
Consultants
Public Interest

About Us
News & Events
Resource Library
Careers
Blog

Lex Machina
1010 Doyle Street,
Suite 200
Menlo Park, CA 94025



[REQUEST A LIVE DEMO](#)

Lex Machina is a LexisNexis Company | [Privacy Policy](#) | [Cookie Settings](#) | [Ad Choices](#) | [Terms of Use](#) | © 2015-2024



An Examination of Firm-Manager Match Quality in the Executive Labor Market

Matthew Ma
College of Business and Analytics
Southern Illinois University
lei.ma@siu.edu

Jing Pan
Cox School of Business
Southern Methodist University
jingp@smu.edu

Xue Wang
Fisher College of Business
Ohio State University
wang.4821@osu.edu

August 2021

Abstract

We use a three way mixed-effects model to quantify firm-manager match effects in executive compensation, and find that unobservable match heterogeneity explains a considerable proportion of the compensation variation. Firms compensate managers for productivity generated by the efficient match, so we propose a measure of firm-manager match quality based on the match effects in executive pay. We validate the proposed measure by showing that it captures systematic firm-manager complementarities and that it is positively and significantly associated with firm operating performance. We also find significant negative stock market reactions to the news of sudden deaths of CEOs with higher match quality, which helps address the concern that our match quality measure captures managerial rent extraction. Further, we show that match quality is an economically significant factor in CEO turnover decisions, an aspect of the executive labor market that features matching consideration prominently.

JEL Classification: G30, G32, J24, J31, J33, C23

Keywords: executive compensation, firm-manager match quality, match effects, mixed-effects model

We appreciate helpful comments from Hank Bessembinder, Brian Cadman, Ted Christensen, Jeff Coles, John Graham, Rachel Hayes, Kevin Murphy, Yihui Pan, Jiaping Qiu, James Schallheim, Jason Smith, Keke Song, Chris Stanton, Steve Stubben, Mohammad Uddin, Mike Weisbach, Feng Zhang and participants at the University of Utah and the 2019 Lone Star Accounting Research Conference. We are particularly grateful for C. Kirabo Jackson for his insights on the application of mixed-effects model. All remaining errors are our own.

1. Introduction

The quality of the match between firms and managers plays a central role in theoretical models of executive compensation and turnover. Gabaix and Landier (2008) and Tervio (2008) model the level of compensation as a competitive outcome when firms compete for scarce managerial talent. Assuming that talented managers generate greater productivity at larger firms, their models imply that the most talented managers match with the largest firms and earn the highest pay.¹ Likewise, Eisfeldt and Kuhnen (2013) develop a competitive assignment model in which the match between a firm and a CEO on multiple dimensions explains CEO turnover. Management consultants also emphasize that each unique business situation requires executives with a specific set of management skills and characteristics (Gerstein and Reisman 1983).

Despite the importance of match effects for understanding the executive labor market, there is little direct empirical investigation into the match quality between firms and managers. A major empirical challenge in this area of research is to measure the match quality between firms and managers. It is challenging because (1) most firm-manager complementarities are not readily observable, and (2) quantifying matching complementarities is hampered by a lack of satisfactory measurements of manager traits such as talent, risk preference, and personality. As an illustration, Graham et al. (2012) document that time-invariant firm and manager fixed effects explain a significant proportion of the variation in top executive pay.² However, they acknowledge that their

¹ Pan (2015) extends the one-dimensional assignment model to a multi-dimensional model. In her model, two additional complementarities between firms and managers affect the level of pay: (1) matching diversified firms to CEOs with cross-industry experience, and (2) matching R&D intensive firms to CEOs with high innovation propensity.

² An extensive body of research investigates the determinants of executive pay. To date, existing studies have examined a wide variety of observable firm and managerial characteristics that affect executive pay (e.g., Rose and Shepard 1997; Core et al. 1999; Gabaix and Landier 2008; Frydman and Jenter 2010; Custodio et al. 2013; Falato et al. 2015; Ellahie et al. 2016).

empirical methods cannot fully address the matching issue if firms and managers are matched on unobservable firm and manager characteristics.³

In this paper, we take a first step toward addressing this empirical challenge. Applying recent methodology advances in labor economics, we use a three way mixed-effects model to quantify firm-manager match effects in top executive compensation. The model includes time-invariant firm and manager random effects and an interaction (random) effect between the firm and the manager. The interaction effect captures time-invariant, unobservable match effects in compensation.⁴ The three way mixed-effects model has been used extensively in the labor economics literature to capture match effects. For example, Woodcock (2011) employs this method to quantify firm-worker match effects in determining worker wages; Jackson (2013) uses it to investigate the importance of match quality between teachers and schools for student achievement; and Lazear et al. (2015) apply it to study how much bosses enhance worker productivity. This mixed-effects model allows us to demonstrate the importance of the match between firms and managers in that it captures unobservable complementarities and overcomes the limitations of using indirect measures such as managerial traits.⁵

Using ExecuComp firm-executive linked data from 1992 to 2015, we construct a connectedness sample containing all the managers who have ever worked for a firm that has hired at least one other manager who previously worked for other firms. We find that firm-manager

³ In a recent paper, Liu, Mao and Tian (2017) use the method in Abowd, Kramarz, and Margolis (1999, AKM) to show the importance of human capital in enhancing firm innovation. While Liu et al. (2017) apply similar empirical approaches as in Graham et al. (2012) to show the robustness of their results, they too provide a caveat that they cannot completely rule out the matching concern.

⁴ While researchers could identify manager, firm, and match effects by either a three-way fixed effects model or a three way mixed-effects model, the latter is more suitable for most compensation studies for two reasons. First, the fixed effects model suffers from “limited mobility bias,” which leads to imprecise estimation of fixed effects (Andrews et al. 2008). Second, to identify match effects, the three-way fixed effects model assumes match effects are orthogonal to person and firm effects. This assumption is almost always violated in economic data.

⁵ By “unobservable complementarities,” we mean the firm-manager complementarities that are difficult to quantify or unavailable to financial economists.

match effects are an important determinant of executive pay level. Specifically, firm-manager match effects explain 12% of the variation in executive pay, compared to 7.6% of the variation in executive pay attributed to manager effects. Importantly, the result shows that the power of manager effects in explaining the variation in executive pay reduces by about 59% (from 18.7% to 7.6%) when match effects are included in the mixed-effects model. This finding indicates that a substantial portion of what is considered a manager effect is not portable across firms.

The above results supporting the importance of match effects in executive compensation might not be surprising. Firms compensate managers for productivity generated by the efficient match between firms and managers. As such, we propose a measure of firm-manager match quality based on the extent to which match effects explain executive pay. This approach could open the door to research studies that previously were challenging because of the difficulty to isolate match effects from firm and manager effects (Graham et al. 2012; Liu et al. 2017). We conduct a number of empirical analyses in an attempt to establish the validity of the measure and to illustrate the use of the measure in understanding the role of match quality in the executive labor market.

First, we examine the information contents of the match quality measure. We show that compensation match effects are positively associated with two complementarities documented in prior work (e.g., Gabaix and Landier 2008; Pan 2015): (1) firm size and managerial talent, and (2) diversification degree and general management skills. These findings suggest that the proposed match quality measure can capture complementarities between firms and managers.

Second, the job match theory assumes heterogeneity in the productivity of firm-manager matches, and predicts that high quality firm-manager matches are characterized by better firm performance relative to low quality matches (Garen 1988; Bishop 1990; Jovanovic 1979). We find

that the proposed measure, the match effects in compensation, is positively and significantly correlated with firm operating performance over the manager's tenure.

Next, we address one concern about the proposed match quality measure that it might capture managerial rent extraction. Beginning with the work of Jensen and Murphy (1990), a stream of literature attributes the high level of executive pay to managerial entrenchment (e.g., Borokhovich et al. 1997; Bebchuk and Fried 2003). If entrenched managers match with poorly governed firms and extract rents in the form of high pay, then compensation match effects might capture rent extraction. To investigate this possibility, we analyze changes in shareholder value around CEO sudden deaths. When an incumbent CEO matches efficiently with the firm, the sudden death of the incumbent CEO should be associated with a subsequent sub-optimal firm-CEO match, which decreases shareholder value. If the proposed match quality measure captures value created by firm-CEO pairs, we expect to observe more negative market reactions to the sudden deaths of CEOs with higher match quality. Using a sample of 71 CEO sudden deaths, we find that the average three-day cumulative abnormal announcement returns around deaths of CEOs with high match effects (above median) are -1.94% and statistically significant, but are insignificant around deaths of CEOs with low match effects (below median). When we conduct regressions analyses, the negative association between market reactions to CEO sudden deaths and match effects remains robust to the inclusion of a number of firm and CEO characteristics. These results indicate that the proxy of compensation match effects is more likely to capture match-specific productivity rather than rent extraction.

In a final step, we illustrate the usefulness of the measure by investigating CEO turnover decisions, another aspect of the executive labor market that features matching consideration prominently (Allgood and Farrell 2003). CEO turnover can be driven by the board's conclusion

about the suitability of the match. If the CEO's skills are not well matched with the firm's needs, then the board is more likely to fire the CEO. While prior research documents a robust relation between CEO turnover and poor firm performance, there is no direct empirical evidence supporting the role of matching consideration.⁶ In an influential discussion, Brickley (2003) argues that a large amount of variation of CEO turnover decisions remains unexplained even after controlling for firm and executive characteristics. We employ the proposed match quality measure to examine whether match quality can in part account for the unexplained variation in CEO turnover decisions. The results show that CEO turnover is more likely when the quality of the firm-CEO match is poor after we control for a number of firm and CEO characteristics in the empirical analyses, including market and accounting performance measures. Specifically, one standard deviation increase in the proxy of match quality is associated with a decrease of 6.4% (3.3%) of CEO turnover (forced turnover) likelihood, which is about 77% (143%) of the mean turnover (forced turnover) probability.

Our study makes three primary contributions. First, our study advances understanding of executive compensation and turnover decisions. To the best of our knowledge, this paper is the first empirical study to quantify firm-manager match effects in executive compensation. Building on a competitive assignment model, the literature emphasizes the importance of firm-manager match in understanding the executive labor market (e.g., Gabaix and Landier 2008; Tervio 2008; Pan 2015). Our results suggest that time-invariant, unobservable match effects explain a sizable

⁶ Allgood and Farrell (2003) document that a high percentage of CEO turnovers in the early years of tenure, and interpret it as supporting that the job match theories apply to the CEO labor market.

proportion of the variation in executive compensation, and that match effects are an economically significant determinant of executive turnover decisions.⁷

Second, we introduce a firm-manager match quality dimension to empirical studies in the executive labor market. We propose a measure of firm-manager match quality that is based on easily obtainable compensation data and available for a broad cross-section of firms. Researchers might apply this measure to study the role of match quality in a wide array of economics, finance and accounting settings, which were previously difficult to implement.

Third, our study contributes to a growing literature exploring the role of multi-dimensional match between firm characteristics and manager attributes in shaping corporate outcomes (e.g., Bandiera et al. 2012; Graham et al. 2013; Pan et al. 2017; Liu et al. 2017). Extending these studies, we document that unobservable heterogeneity in firm-manager match is an important determinant of firm performance. Furthermore, our findings suggest the importance of considering firm-manager match effects if the goal is to study how unobserved managerial attributes influence corporate policies and accounting practices (e.g., Bertrand and Schoar 2003; Bamber et al. 2010; Bushman et al. 2020).

2. Empirical Methodology of Identifying Match Effects in Compensation

2.1. A Model of Firm-Manager Match Effects and Executive Pay

In this section, we describe the empirical methodology of identifying match effects in compensation. Early labor economics literature decomposes worker wages into a worker effect and a firm effect (Abowd et al. 2004). We follow Jackson (2013) and use his model to illustrate

⁷ While the quality of the match can change over time as a firm evolves or as industry and market conditions change, we note that our proxy of firm-manager match quality is time invariant and does not capture changing match quality. It is difficult to account for changing match quality because of the difficulty of identifying when the changes occur.

the role of firm-manager match effects in executive pay.

Suppose that (1) manager i has productive characteristics (e.g., ability, education, human capital, and other “portable” determinants of productivity) indexed by L_i ; (2) firm j has productivity characteristics (e.g., technology and capital intensity) indexed by K_j . When a manager’s productive attributes and a firm’s productive attributes are complementary, certain pairings of managers and firms are particularly productive.⁸ We denote M_{ij} as the complementarities between the productive characteristics of manager i and firm j . Output Q_{ij} is given by the Cobb-Douglas function:

$$Q_{ij} = L_i^\theta K_j^\psi M_{ij}^\phi \quad (1)$$

Suppose that firm j faces price P_j for its output, and manager i ’s compensation is share γ_{ij} of firm j ’s output. The log of the manager’s compensation W_{ij} is given by equation (2).

$$\ln W_{ij} = \ln P_j + \ln \gamma_{ij} Q_{ij} = \theta \ln L_i + \ln P_j + \psi \ln K_j + \ln \gamma_{ij} M_{ij}^\phi \quad (2)$$

In Equation (2), $\ln W_{ij}$ is determined by three separable components: the manager productivity component $\theta \ln L_i$; the firm productivity component $\ln P_j + \psi \ln K_j$; and the firm-manager match-specific productivity component $\ln \gamma_{ij} M_{ij}^\phi$.⁹

2.2. Empirical Methodology

Consistent with Equation (2), our empirical model of executive compensation includes a firm component, a manager component, and a firm-manager match component. We decompose firm and manager components into observable and unobservable parts. Given the inherent difficulty of empirically measuring match-related characteristics, we rely on a time-invariant

⁸ Rosen (1981, 1982) suggests that an efficient match between firms and managers creates economic surplus.

⁹ In a competitive assignment model framework (e.g., Gabaix and Landier 2008; Tervio 2008), the output is shared between a firm and a manager. The sharing rule is an equilibrium outcome of the competition for scarce managerial talent and is determined by the distribution of the manager’s and the firm’s productive characteristics.

match component to capture firm-manager complementarities.⁵ Specifically, we use the following regression model of executive compensation:

$$\begin{aligned} \ln(\text{Total compensation})_{ijt} = & \mu_0 + \beta_1 \ln(\text{Assets})_{jt-1} + \beta_2 \text{MTB}_{jt-1} + \beta_3 \text{Return}_{jt} + \beta_4 \text{Return}_{jt-1} + \\ & \beta_5 \text{ROA}_{jt} + \beta_6 \text{ROA}_{jt-1} + \beta_7 \sigma(\text{Return})_{jt} + \beta_8 \text{Duality}_{jt} + \beta_9 \ln(\text{Tenure})_{it} + \beta_{10} \text{CEO}_{it} + \\ & \beta_{11} \text{FirstYear}_{it} + \text{Year Fixed Effects} + \theta_i + \psi_j + \phi_{ij} + \varepsilon_{ijt} \end{aligned} \quad (3)$$

where i indexes managers, j indexes firms, and t indexes years. θ_i is a time-invariant portable manager effect, which captures unobservable managerial skills transferable across companies; ψ_j is a time-invariant firm effect, which captures unobservable firm-specific characteristics or compensation policies; ϕ_{ij} is a match effect, which captures the value of productive complementarities between manager i and firm j .

We include a set of variables to control for observable firm and manager characteristics that determine the level of executive compensation. Following Core, Holthausen and Larcker (1999), we use the natural log of total assets ($\ln(\text{Assets})$) to proxy for firm size; market-to-book ratio (MTB) to proxy for firm's investment opportunity; stock returns (Return) and return on assets (ROA) to proxy for firm performance; stock return volatility ($\sigma(\text{Return})$) to proxy for firm risk; the CEO/board chair indicator variable to proxy for board independence (Duality); and the natural log of job tenure ($\ln(\text{Tenure})$) to proxy for human capital investment. We also include an indicator variable for CEO (CEO) to capture compensation differences between CEOs and non-CEO top executives. Many firms offer abnormally large equity awards and signing bonuses when an outsider first joins a firm, so we include an indicator variable for the first year when an executive

⁵ Gabaix and Landier (2008) and Tervio (2008) suggest that the match between talented managers and large firms generates match-specific productivity. Without a good measure of managerial talents, it is difficult to quantify the compensation consequences of match-specific productivity by regressing executive compensation on match-related characteristics. Match effects driven by firm-manager complementarities are likely to evolve slowly over time. Hence, we rely on a time-invariant match component to capture the average observable and unobservable match effects in compensation for each firm-manager pair. Given that this method does not identify match effects that are time varying, it might underestimate match effects.

joins a firm (*FirstYear*). The inclusion of *FirstYear* also helps mitigate concerns that sign-up awards might reflect premiums for the uncertainty of the match when a manager joins a new firm (Carter, Franco and Tuna 2019). We additionally include year fixed effects to capture differences in compensation across years. Lastly, we adjust standard errors at the firm level to correct for within firm correlation. Appendix 1 provides detailed definitions of all the variables used in the empirical analyses.

The assumption of using Equation (3) to identify firm, manager, and match effects is that the error terms (ε_{ijt}) have a zero conditional mean:

$$E(\varepsilon_{ijt} | \theta_i, \psi_j, \phi_{ij}, x_{ijt}) = 0 \quad (4)$$

This identification condition is weaker than that required by the person (worker) and firm effects model in Abowd, Kramarz and Margolis (1999; AKM) and Graham et al. (2012), $E(\varepsilon_{ijt} | \theta_i, \psi_j, x_{ijt}) = 0$. Adding match effects allows the mobility of managers to depend on unobserved match-specific heterogeneity (ϕ_{ij}). If the mobility of managers depends on unobservable match effects (e.g., a manager moves to a firm because of a better match at this firm), omitting match effects leads to biased and inconsistent estimations of manager and firm effects.¹⁰

By observing compensation data for multiple managers who switched across firms, we can separate match effects from firm and manager effects. Woodcock (2011) proves that the group connectedness of individuals across firms is the necessary and sufficient condition to separate firm,

¹⁰ We note that ignoring match effects predominantly biases manager effects not firm effects. This is due to the nature of the firm-manager linked data. In such data, a typical firm matches with a number of managers, but a typical manager only matches with a few firms. Among the managers who have worked for a particular firm, some of them have positive match effects and others have negative match effects. On average, firm effects contain small measurement errors because positive and negative omitted match effects tend to cancel out over many matches. However, there is no such effect for managers who only match with a few firms. As such, manager effects estimated by the AKM (1999) method could be biased.

person, and match effects, which is the same as the identification conditions of separating firm and person effects in AKM (1999).

For mover managers who are observed in multiple firms, the mixed-effects model identifies firm effects based on the common component of compensation among managers, manager effects based on conditional covariation of a manager's compensation at different firms, and match effects based on conditional covariation in compensation within the firm-manager match that is not explained by firm and manager effects.

To illustrate the intuition, consider two firms A and B, and two managers C and D. The difference in Manager C's compensation between Firm A and Firm B captures the difference in firm effects, as well as the difference in match effects for Manager C between the two firms. If there are no firm-manager match effects, the difference in a manager's compensation between Firm A and Firm B reflects only the difference in firm effects, which is the same for any manager. In contrast, when match effects exist, the difference in a manager's compensation between Firm A and Firm B will vary with the manager, which allows us to separate match effects from firm and manager effects.¹¹

At the same time, the mixed-effects model allows the identification of match effects of non-mover managers based on an empirical Bayesian approach (Jackson, 2013). The mixed-effects model estimates distributional information (the variances of manager, firm, and match effects) based on mover managers. The high compensation of a non-mover manager (after controlling for observable manager characteristics and firm effect) could be explained by (1) a large positive manager effect and/or (2) the manager is a good match with the firm. If the variance

¹¹ Mathematically, under the assumption of the mixed-effects model that the mean of the match effects is equal to 0 in expectation for each firm and for each manager, we have four equations and four unknowns about match effects between firms A and B, and managers C and D for mover managers: $\phi_{AC} + \phi_{AD} = 0$, $\phi_{BC} + \phi_{BD} = 0$, $\phi_{AC} + \phi_{BC} = 0$, and $\phi_{AD} + \phi_{BD} = 0$. As a result, there is a unique solution for the values of the match effects.

of match effects is larger than that of manager effects, then the manager is more likely to draw a large match effect rather than a large manager effect. Under this circumstance, the mixed-effects model attributes a larger proportion of the compensation unexplained by manager and firm effects to match effects.¹² As demonstrated above, the mixed-effects model uses distributional information of the mover managers to generate estimators of manager and match effects rather than mechanically attributing the unexplained variation to managers, as in the fixed-effects model.

Following AKM (1999), we construct a connectedness sample that contains all the managers who have ever worked for a firm that has hired at least one mover manager, i.e., a manager previously worked for other firms. The procedure goes as follows: we begin with an arbitrary manager and include all the firms for which she has ever worked. Next, we include all the managers who have ever worked for those firms, and continue to add other firms for which any of these managers have ever worked until we could not add more managers or firms to the current group. We repeat the above steps for the next group and continue until all data are exhausted. The connectedness sample contains both mover and non-mover managers as long as they work in firms that have hired at least one mover manager.

Equation (3) can be estimated by either a fixed-effects model or a mixed-effects model. In a fixed-effects model, time-invariant firm, manager, and match effects are identified as fixed effects. The firm fixed effect is the common component of compensation for all managers who have worked for the same firm. The manager fixed effect is the common component of a manager's pay at each firm in her employment history. To separate match effects from firm and manager effects, the three way fixed-effects model assumes that match effects are orthogonal to firm and manager effects (Woodcock 2011). The orthogonality assumption introduces two undesirable

¹² We thank C. Kirabo Jackson for confirming the mechanisms of using a mixed-effects model to estimate match effects for the non-mover managers.

features to fixed-effects estimators. First, this approach attributes the proportion of match effects correlated with firm or manager effects to firm or manager effects.⁶ Second, the orthogonality assumption restricts the mean of match effects to zero for each manager and firm. For the manager who has only worked for one firm during her employment history, the match effect is zero by construction.¹³

Alternatively, we can estimate Equation (3) using a mixed-effects model, in which the unobservable match heterogeneities are identified as random effects (Woodcock 2011; Jackson 2013; Lazear et al. 2015). We describe in detail the estimation procedure in Appendix 2 of the paper. The mixed-effects model offers several advantages relative to the fixed-effects model. First, the two-step estimation approach, also referred as the “hybrid” mixed-effects model, does not assume zero correlation between random effects and observable characteristics (Woodcock 2011).⁸ By relaxing the zero-correlation assumption, the (hybrid) mixed-effects model generates consistent and unbiased estimates of the parameters on observable characteristics (Cameron and Trivedi 2005).

Second, as mentioned earlier, the fixed-effects model mechanically assigns match effects to zero for non-mover managers and loads match effects that correlate with firm or manager effects to firm or manager effects for mover managers. In contrast, the mixed-effects model loads the

⁶ The orthogonality assumption is always violated in economic data. For example, Gabaix and Landier (2008) suggest that larger firms tend to match with more talented CEOs. If high match quality arises from such complementarity, match effects might correlate with CEO effects that capture latent managerial talent.

¹³ Prior studies also raise concerns about the “limited mobility bias” suffered by both the two way and three way fixed-effects models. Abowd et al. (2004) and Andrews et al. (2008) show that when the worker mobility in the connectedness sample is limited, the estimated fixed effects are imprecise and the standard deviations of worker and firm fixed effects are biased upward.

⁸ The conventional mixed-effects model assumes that firm, CEO, and match effects do not correlate with the observable firm and CEO characteristics. This assumption is too restrictive. For example, if the length of a CEO’s job tenure is related to her latent ability, the variable $\ln(Tenure)$ is correlated to unobservable person effects.

variation in compensation, which is unexplained by observable characteristics, to firm, manager, and match random effects based on the distributional information $(\hat{\sigma}_\theta^2, \hat{\sigma}_\psi^2, \text{ and } \hat{\sigma}_\phi^2)$.

Third, the mixed-effects approach does not suffer from the “limited mobility bias.” As the number of firms, managers, and matches are sufficiently large in the sample, the estimated random effects and estimated variance of random effects are consistent (Jiang 1996, 1998). Further, the mixed-effects model allows manager, firm, and match effects to correlate with each other in the sample, and does not mechanically impose the restriction that match effects are orthogonal to firm and manager effects as in the fixed-effects model.

3. Data and Sample

3.1 Sample

We construct a firm-manager matched panel dataset from 1992 to 2015. ExecuComp allows us to track through time the top five highest-paid managers of S&P 1,500 firms. As described in Section 2, we construct a connectedness sample that contains all the managers who have ever worked for a firm that has hired at least one manager who previously worked for other firms. We obtain financial accounting data from Compustat, and stock return data from CRSP. We remove observations with missing accounting and stock return variables. The resulting sample contains 153,833 firm-manager-year observations.

3.2 Measurements of Top Executive Compensation

We use ExecuComp data item TDC1 to measure total compensation for fiscal years before 2006. The Securities and Exchange Commission (SEC) adopted new reporting requirements for executive compensation for fiscal years ending after December 15, 2006. To ensure that we measure the compensation variables consistently over our sample period, we follow the formula

of TDC1 to compute total compensation for observations after fiscal year 2006. Specifically, we first calculate the Black-Scholes value of stock options granted using the modified Black-Scholes formula,⁹ then we compute total compensation as the sum of salary, bonus, non-equity incentives, the Black-Scholes value of stock options, the grant date fair value of stock grants, and other compensation.

3.3 *Executive Mobility*

Because the separation of manager, firm, and match effects depends on the mobility of managers, we summarize the mobility structure of the sample in Table 1. The sample contains 32,420 managers and 2,385 firms. Panel A shows that 90.19% of managers are non-movers who have worked as top executives for only one firm, and that 9.81% of managers have worked for at least two firms in the sample.¹⁴ Panel B shows that 94.34% of the firms have hired more than five top managers during the sample period. In particular, 28.51% of the firms have hired 6 to 10 managers, 23.61% of them have hired 11 to 15 managers, 31.45% of them have hired 15 to 20 managers, and 10.78% of them have hired more than 25 managers. In general, a typical firm matches with multiple managers.

[Insert Table 1 about here]

3.4 *Summary Statistics*

Table 2 reports summary statistics for the sample. The average (median) total executive compensation is \$2.39 (\$1.26) million, suggesting that the distribution of executive pay is skewed

⁹ Following the methodology of ExecuComp, we make the following assumptions about the inputs of the modified Black-Scholes formula: (1) the exercise price per share is the market price at the time of grant; (2) the grant date is July 1st of the grant year; (3) the term of grant equals 0.7 times the period between the grant date and the expiration date; (4) the risk-free rate is the annual yield on a seven-year U.S. Treasury bond; (5) the estimated stock volatility during the time of grant is the standard deviation of the past 60 months' stock return; (6) the estimated dividend yield during the time of grant equals the average dividend yields over a three-year period.

¹⁴ Note that the statistics of 90% non-mover managers in our sample is comparable to that of teachers in Jackson's (2013) sample in which 80% of teachers only work for one school.

to the right side. The average tenure of top executives is 11.41 years, and 18% of them are CEOs. Among firm characteristics, the mean (median) of total assets is \$10,982 (\$1,986) million in the connectedness sample, indicating skewness in the distribution of firm size. On average, the sample firms have positive ROA and positive stock returns.

[Insert Table 2 about here]

4. Match Effects in Executive Compensation

We start empirical analyses by demonstrating match effects in executive compensation. Panel A of Table 3 reports the estimation results of four different model specifications of Equation (3). Column (1) is a pooled OLS regression without firm or manager fixed effects; column (2) is an OLS regression with firm and manager fixed effects; column (3) is a two way mixed-effects regression with firm and manager random effects; and column (4) is a three way mixed-effects regression with firm, manager, and match random effects. Across all columns, we find that managers earn higher pay in firms with larger size, higher growth opportunities, better performance, and higher risk, consistent with findings from prior studies (Core et al. 1999; Core et al. 2008; Graham et al. 2012). The results also show that CEOs on average earn higher pay than non-CEO top executives, and that executives earn higher pay in their first year with a firm.¹⁵

[Insert Table 3 about here]

We employ two tests to assess the importance of match effects in executive pay. First, we use a REML likelihood ratio test (REMLRTs) based on the log-likelihoods of specifications with

¹⁵ We find that the coefficient estimates on most observable firm and manager variables are similar in magnitudes in columns (2) - (4). This result is not surprising for the following reasons. First, as discussed in Appendix 2, we estimate the coefficients on observable firm and manager characteristics from the first-step (or “spell fixed effects” approach) in the mixed-effects models. Thus, the coefficient estimates in columns (3) - (4) are identical. Second, Graham et al. (2012) also report that the coefficient estimates in the “spell fixed effects” approach are similar to those in the fixed-effects regression.

and without match effects to assess the existence of match effects.¹¹ In the last row of column (4), we report the result of the likelihood ratio test which rejects the null hypothesis of no match effects at a conventional significance level ($p < 0.00001$).

Second, following Graham et al. (2012) and Coles and Li (2013), we use proportional variance decomposition to determine the proportion of the variance of $\ln(\text{Total compensation})$ explained by each component to capture the relative economic importance of match effects. The decomposition is based on the following Equation:

$$1 = \frac{\text{Var}(y_{ijt})}{\text{Var}(y_{ijt})} = \frac{\text{Cov}(y_{ijt}, \hat{y}_{ijt})}{\text{Var}(y_{ijt})} + \frac{\text{Cov}(y_{ijt}, e_{ijt})}{\text{Var}(y_{ijt})} = \frac{\text{Cov}(y_{ijt}, X_{ij}\hat{\beta} + W_{jt}\hat{\gamma} + \hat{\mu}_t + \hat{\theta}_i + \hat{\psi}_j + \hat{\phi}_{ij})}{\text{Var}(y_{ijt})} + \frac{\text{Cov}(y_{ijt}, e_{ijt})}{\text{Var}(y_{ijt})} \quad (5)$$

$$= \frac{\text{Cov}(y_{ijt}, X_{ij}\hat{\beta})}{\text{Var}(y_{ijt})} + \frac{\text{Cov}(y_{ijt}, W_{jt}\hat{\gamma})}{\text{Var}(y_{ijt})} + \frac{\text{Cov}(y_{ijt}, \hat{\mu}_t)}{\text{Var}(y_{ijt})} + \frac{\text{Cov}(y_{ijt}, \hat{\theta}_i)}{\text{Var}(y_{ijt})} + \frac{\text{Cov}(y_{ijt}, \hat{\psi}_j)}{\text{Var}(y_{ijt})} + \frac{\text{Cov}(y_{ijt}, \hat{\phi}_{ij})}{\text{Var}(y_{ijt})} + \frac{\text{Cov}(y_{ijt}, e_{ijt})}{\text{Var}(y_{ijt})}$$

where i indexes managers, j indexes firms, and t indexes years. y_{ijt} is the dependent variable $\ln(\text{Total compensation})$; X_{it} is a vector of time-variant observable manager characteristics; W_{jt} is a vector of time-variant observable firm characteristics; $\hat{\mu}_t$ captures year effects; $\hat{\theta}_i$ captures manager effects; $\hat{\psi}_j$ captures firm effects; $\hat{\phi}_{ij}$ captures firm-manager match effects, and e_{ijt} is the corresponding residual. The proportion explained by each component equals the covariance between $\ln(\text{Total compensation})$ and each component, scaled by the variance of $\ln(\text{Total compensation})$.

In Panel B of Table 3, column (2) shows that firm and manager fixed effects explain 8.9% and 35.4% of the variation in $\ln(\text{Total compensation})$ in an OLS regression. Column (3) demonstrates that, in the two way mixed-effects model, manager random effects explain 18.7% of

¹¹ REMLRTs is the likelihood ratio test based on REML estimators of Equation (3). We test the null hypothesis of no match effects ($H_0: \sigma_{\phi}^2 = 0$ versus $H_a: \sigma_{\phi}^2 > 0$). In particular, we compare the log likelihood of the null model to that of the alternative model. The null model is Equation (3) without the match effects and the alternative model is Equation (3). The test statistic is $\chi^2 = -2(\log\text{likelihood}(\text{null model}) - \log\text{likelihood}(\text{alternative model}))$. Abowd, Kramarz and Woodcock (2006) provide detailed discussion of the likelihood ratio test.

the variation in $\ln(\text{Total compensation})$ while the explanatory power of firm random effects increases to 20.5%. The results suggest that the fixed-effects model is more likely to over-estimate (under-estimate) the explanatory power of time-invariant manager (firm) effects in executive pay relative to the mixed-effects model.

Turning to column (4), the three way mixed-effects model estimation results indicate that firm, manager, and match effects explain 21%, 7.6%, and 11.5% of the variation in $\ln(\text{Total compensation})$, respectively. Compared to the results in column (3), the explanatory power of manager effects decreases by 59% (from 18.7% to 7.6%) after controlling for match effects. However, the explanatory power of firm effects is similar, regardless of whether we control for match effects or not. As discussed in Section 2, due to the firm-manager linked data structure, omitted match effects are likely to be absorbed in manager effects rather than in firm effects. Thus, omitting match effects inflates the explanatory power of manager effects but not of firm effects.

We use the summary statistics in Panel B to demonstrate the economic impact of match effects. Column (4) shows that the standard deviations of firm, manager, and match effects are 0.42, 0.16, and 0.23, respectively. Thus, if match effects increase one standard deviation, the average $\ln(\text{Total compensation})$ increases from 7.2 to 7.43 ($=7.2+0.23$), corresponding to an increase of \$0.35 million in total compensation ($=\$1.69 \text{ million} - \1.34 million). As a benchmark, if firm (manager) effects increase one standard deviation, $\ln(\text{Total compensation})$ increases 0.42 (0.16), equivalent to an increase of \$0.7 (\$0.23) million in total compensation. Overall, match effects appear to have a significant economic impact in determining executive pay.

One concern about our empirical research methodology is the inclusion of the first-year observation of a manager in the sample. Sign-up awards during the first year might reflect premiums for the uncertainty of the match when a manager joins a new firm, confounding the

estimation of the firm-manager match effects during a manager's tenure (Carter et al. 2019). To address this issue, we estimate match effects after removing the first year observations from the sample. The results are qualitatively similar to those presented in Table 3.¹⁶

Overall, the results presented thus far suggest that unobservable time-invariant firm-manager match heterogeneity explains a sizeable proportion of the variation in executive pay, and that the inclusion of match effects substantially reduces the explanatory power of time-invariant manager effects.

5. Match Effects in Executive Compensation as a Proxy of Match Quality

The findings supporting the importance of match effects in executive compensation might not be surprising given the theoretical predictions about the role of match in the executive labor market. Firms compensate managers for productivity generated by the efficient match between firms and managers, so we propose using the extent to which the match effects explain executive pay as a proxy of firm-manager match quality. For the remaining of the study, we conduct a series of empirical tests in an attempt to establish the validity of the proposed measure and to illustrate the use of the measure in understanding the role of match quality in the executive labor market with a focus on CEOs.

5.1. The Information Content of the Match Quality Measure

As a first step, we assess whether the proposed match quality measure captures systematic complementarities between particular managers and particular firms. We examine two firm-manager complementarities documented in prior studies. One is the well-established complementarity between firm size and managerial talent (Gabaix and Landier 2008 and Tervio

¹⁶ We note that this similarity is probably not surprising because we include an indicator variable of the first year of a manager's tenure in Model (3). The regression results are not tabulated for brevity and are available upon request.

2008). The other is the complementarity between diversification and general management ability. Economies of scope suggest that the skills of managers with past working experience in conglomerates are more suited for more diversified firms (Pan 2015). Consistent with this expectation, Maksimovic and Philips (2002) document that managers in more diversified firms have a higher level of general managerial ability than those in less diversified firms.

We test whether the proposed match quality measure is positively associated with the two firm-manager complementarities. To assess complementarity between firm size and managerial talent, we measure firm size using the logarithm of total assets ($Ln(Assets)$), and managerial talent ($Talent$) using the manager efficiency score developed in Demerjian et al. (2012).¹⁷ Following Pan (2015), we allocate the efficiency score to a manager based on the proportion of the whole management team's compensation earned by the manager. To capture complementarity between diversification and CEO general management ability, we measure diversification (DIV) using a segment sales-based Herfindahl Index, and general management skills (GAI) using the General Ability Index of CEOs in Custodio et al. (2013).¹⁸ The GAI is standardized to have a zero mean and a standard deviation of one.

Given that the estimated compensation match effects are time-invariant, we average firm and manager attributes over the job tenure for each manager. We then calculate interaction terms between averaged firm and managerial attributes at the firm-manager level. We estimate the following cross-sectional regressions:

$$Match\ Effect_{ij} = \beta_0 + \beta_1 Ln(Assets)_j + \beta_2 Talent_i + \beta_3 Ln(Assets)_j \times Talent_i + \varepsilon_{ij} \quad (6.1)$$

¹⁷ The efficiency score is based on the technique of input-oriented data envelopment analysis (DEA). Demerjian et al. (2012) show that the efficiency score is a more precise measure of managerial talent than existing ability measures such as firm size, ROA, and media mentions.

¹⁸ Custódio et al. (2013) develop the GAI for CEOs of S&P 1,500 firms from 1992 through 2012. Cláudia Custódio generously shared with us the updated GAI data to 2016. Thus, our regression analysis is conducted over the period from 1992 to 2015.

$$\begin{aligned}
Match\ Effect_CEO_{ij} = & \beta_0 + \beta_1 Ln(Assets)_j + \beta_2 Talent_i + \beta_3 Ln(Assets)_j \times Talent_i + \beta_4 DIV_j \\
& + \beta_5 GAI_i + \beta_6 DIV_j \times GAI_i + \varepsilon_{ij}
\end{aligned}
\tag{6.2}$$

where i indexes managers and j indexes firms. *Match Effect* is the manager compensation match effects, and *Match Effect_CEO* is the CEO compensation match effects, both estimated from Equation (3).

[Insert Table 4 about here]

Table 4 reports the results. Column (1) presents the results on complementarity between firm size and managerial talent of managers (equation 6.1); column (2) presents the results on complementarity between diversification and general management ability of CEOs (equation 6.2); and column (3) presents the results after including both complementarities for CEOs (equation 6.2). The results show that the interaction terms— $Ln(Assets) \times Talent$ and $DIV \times GAI$ —are both positive and statistically significant across all columns. These findings suggest that firms compensate managers for match-specific productivity derived from size/talent and diversification/general management ability complementarities.

5.2. Match Quality and Firm Performance

Next, we investigate whether firm performance varies predictably with the proposed measure of firm-manager match quality. Match theory proposes that workers differ in their productivity because of the heterogeneity in the quality of job matches across firms. As such, high quality matches are more productive than poor quality matches. Applying this insight to the executive labor market, we expect a positive association between firm-manager match quality and firm performance. We focus on CEOs because their primary responsibility is to lead the firm and should be responsible for a firm's performance. To test this implication, we run a regression of firm performance on the measure of firm-CEO match quality as follows:

$$Performance_{ijt} = \mu_0 + \beta_1 Match\ Effect_CEO_{ij} + \beta_2 Ln(Assets)_{jt} + \beta_3 MTB_{jt} + \beta_4 Leverage_{jt} + \beta_5 \sigma(Return)_{jt-1} + Year\ FEs + \varepsilon_{ijt} \quad (7)$$

where i indexes CEOs, j indexes firms, and t indexes years. We employ two operating performance measures, ROA and ROE. The variable *Match Effect_CEO* denotes the compensation match effects for CEOs estimated from Equation (3).

[Insert Table 5 about here]

Table 5 column (1) shows the regression results when the performance measure is ROA, and column (2) reports the results when the performance measure is ROE. Regardless of the performance measures used, we find a positive and significant coefficient on *Match Effect_CEO*, supporting that high quality firm-CEO match has a positive impact on firm performance.

5.3. *Match Effects and Managerial Rent Extraction*

One concern about the proposed match quality measure is that the estimated compensation match effects might capture managerial rent extraction. Prior literature documents that managerial entrenchment or poor corporate governance allows CEOs to extract rents from shareholders (e.g., Jensen and Murphy 1990; Borokhovich et al., 1997; Bebchuk and Fried, 2003; Acharya and Volpin 2010; Dicks 2012). If entrenched CEOs match with poorly governed firms and extract private benefits in the form of excessive pay, high compensation match effects might reflect rent extraction. To investigate this possibility, we examine the association between the proposed match quality measure and the stock market reactions to CEO *sudden deaths*. Following existing literature (Johnson et al. 1985; Bennedsen et al. 2007; Nguyen and Nielsen 2014), we use the stock price reactions to CEO sudden deaths to proxy for the CEO's expected contribution to shareholder value.

Under the competitive assignment framework of Gabaix and Landier (2008) and Tervio (2008), the current CEO creates more firm value than the next candidate because otherwise, the

firm would have replaced the current CEO. That is, in a competitive and frictionless world, a CEO sudden death decreases firm value. If the proposed match quality measure reflects value created by the firm-CEO pairs, we should observe that the stock market reacts more negatively to the sudden deaths of CEOs with higher match effects. In contrast, if the proposed match quality measure captures disproportionately rent extraction, i.e., the incumbent CEOs extract more compensation than the value they generate, we should observe that the stock market reacts more positively to the sudden deaths of CEOs with higher match effects.

Identifying value created by the CEOs requires that the deaths be sudden and unexpected by the stock market. Following Fee et al. (2013), we search Factiva and Capital IQ to identify CEO sudden deaths (e.g., heart attack, stroke, and other unspecified causes) and collect the announcement dates of CEO deaths.¹⁹ We are able to identify 71 CEO sudden deaths during the period from 1992 to 2012. Panel A of Table 6 shows that the average CEO age is around 59 for both CEOs with high (above median) match effects and those with low (below median) match effects.

[Insert Table 6 about here]

We calculate abnormal returns using the market model, and estimate the expected returns over the trading window [-240, -30] prior to the event. Panel B presents cumulative abnormal returns (CAR) for three windows starting up to two trading days before the announcement, and ending up to five trading days after the announcement. We find that on average, the stock market responds negatively to CEO sudden deaths. CARs are significantly negative for sudden deaths of CEOs with high match effects, but are insignificant for sudden deaths of CEOs with low match effects. For example, the average CARs over a three-day window, $CAR[-1, 1]$, is -0.194 and

¹⁹ We thank Edward Fee, Charles Hadlock, and Joshua Pierce for kindly providing us with their classification of CEO turnovers over the period from 1992 to 2007.

statistically significant (p-value<0.05) for sudden deaths of CEOs with high match effects, while $CAR[-1, 1]$ is 0.13 and not statistically significant for sudden deaths of CEOs with low match effects.

We additionally conduct a regression analysis to address the possibility that firm and CEO characteristics might confound the interpretation of the univariate statistics. For example, Jenter, et al. (2018) document that CEO characteristics such as age and tenure correlate strongly with stock market reactions to CEO deaths. We run an OLS regression of CARs around CEO sudden deaths on the measure of firm-CEO match quality controlling for several firm and CEO characteristics.²⁰

$$CAR_{ij} = \mu_0 + \beta_1 Match\ Effect_CEO_{ij} + \beta_2 CEO\ Effect_i + \beta_3 Ln(Temure)_{ij} + \beta_4 Ln(Assets)_j + \beta_6 Adj_Return_j + \beta_7 Adj_EBIT_j + \beta_{10} Ln(Age)_i + \varepsilon_{ij} \quad (8)$$

where i indexes CEOs and j indexes firms. We measure both firm and CEO characteristics at the year of CEO death. Panel C of Table 6 reports the regression results. The dependent variable is CAR around CEO sudden deaths measured over three different windows: $[-1, 1]$ in column (1), $[-2, 2]$ in column (2), and $[-2, 5]$ in column (3). We find a negative and statistically significant coefficient on $Match\ Effect_CEO$ across all columns, suggesting more negative market reactions to the sudden deaths of CEOs with higher match effects.

We interpret the overall evidence of stock market reactions to CEO sudden deaths as suggesting that investors view sudden deaths of CEOs with high match effects as bad news. The findings support that the proposed match quality measure captures match-specific productivity rather than rent extraction.

5.4. Match Quality and CEO Turnover

²⁰ The definitions of all the explanatory variables are available in Appendix 1.

Finally, we use CEO turnover as a setting to demonstrate the usefulness of our proposed match quality measure in understanding the role of match in the executive labor market. In an influential discussion, Brickley (2003) concludes that a large amount of variation of CEO turnover decisions remains unexplained partly due to the narrow focus of CEO turnover research on the inverse relation between turnover and firm performance. He proposes that researchers examine other less-explored issues in CEO turnover decisions. In this section, we use the proposed match quality measure to investigate whether match effects can in part account for unexplained variation in CEO turnover decisions.²¹

CEO turnover is an aspect of the executive labor market that features matching consideration prominently (Allgood and Farrell 2003). CEO turnover can be driven by the board's conclusion about the suitability of the match. If the CEO's skills are not well matched with the firm's needs, then the board is more likely to fire the CEO. Prior research uses firm performance to infer the matching considerations. However, poor performance also reflects CEO ability and efforts. Therefore, it is difficult to use firm performance to disentangle the incentive from the matching consideration. To the extent that the estimated compensation match effects capture the underlying firm-CEO match quality, we would expect a higher turnover likelihood for a CEO with lower match effects. Our proposed measure of match quality also allows us to assess the economic significance of matching consideration in CEO turnover decisions.

We use ExecucComp to identify CEO turnovers. Prior studies (e.g., Warner et al. 1988 or DeFond and Park 1999) suggest that it is not always possible to determine whether a CEO turnover was forced, so we examine both turnover and forced turnover. We obtain CEO forced turnover

²¹ Along those lines, the job match theory assumes heterogeneity in the productivity of firm-manager matches and predicts that high quality firm-manager matches are characterized by longer tenure of managers relative to low quality matches (Garen 1988; Bishop 1990; Jovanovic 1979). Consistent with this theory, we find that match effects in manager compensation is positively and significantly correlated with the tenure of the managers.

data from Peters and Wagner (2014). The forced turnover data cover the period from 2002 to 2015, and we use the same period for the identification of CEO turnovers. For each firm, we compare the CEO names from one year to the next, and identify a CEO turnover if the CEO listed in the sample has changed from one year to the next. We employ the following probit regression to investigate the role of match effects in CEO turnover decisions:

$$\begin{aligned} Turn_{ijt} / Forced_{ijt} = & \mu_0 + \beta_1 Match\ Effect_CEO_{ij} + \beta_2 CEO\ Effect_t + \beta_3 Ln(Temure)_{ijt-1} + \\ & \beta_4 Ln(Assets)_{jt-1} + \beta_5 MTB_{jt-1} + \beta_6 Adj_Return_{jt-1} + \beta_7 Adj_EBIT_{jt-1} + \beta_8 \sigma(Return)_{jt-1} + \\ & \beta_9 Duality_{jt-1} + \beta_{10} Ln(Age)_{it-1} + \beta_{11} Retire_{it-1} + Year\ FEs + Industry\ FEs + \varepsilon_{ijt} \end{aligned} \quad (9)$$

where i indexes CEOs, j indexes firms, and t indexes years. We use two turnover measures in our tests: *Turn*, an indicator variable equals to one for firm-years where there is a CEO turnover and zero otherwise; and *Forced*, an indicator variable equals to one for firm-years where there is a forced CEO turnover and zero otherwise. The variable of interest is *Match Effect_CEO*, which is CEO compensation match effects estimated from Equation (3). Following prior literature on CEO turnover (e.g., Murphy 1999; Engel et al. 2003), we include a set of firm and CEO variables in the empirical analyses.²² We also control for year and industry fixed effects. After imposing requirements for non-missing control variables, we arrive at a final sample of 16,389 observations. The final sample includes 1,363 CEO turnover events, 349 of which are forced turnover events.

[Insert Table 7 about here]

Panel A of Table 7 presents summary statistics for the variables in Equation (9) for the Turnover, Forced Turnover, and Control subsamples. The average *Match Effect_CEO* in the Control subsample is higher than that in the Turnover and Forced Turnover subsamples, suggesting that, on average, continuing CEOs have better match quality than CEOs that depart the firm.

²² The definitions of all the explanatory variables are available in Appendix 1.

Panel B of Table 7 presents the results of the probit regression. For each independent variable, we report the coefficient estimates, z-statistics, and marginal effects. Columns (1) and (2) show the results when the dependent variable is *Turn*, and columns (3) and (4) demonstrate the results when the dependent variable is *Forced*. Across all columns, we find that the coefficient on *Match Effect_CEO* is negative and statistically significant. CEO turnover is more likely when the CEO-firm match quality is poor after we control for a number of firm and CEO characteristics including market and accounting performance measures. The results are also economically significant. We note that the average probability of CEO turnover (forced turnover) is 8.3% (2.3%) over the sample period. The results suggest that one standard deviation increase in the proxy of matching quality is associated with a decrease of 6.4% (3.3%) of the CEO turnover (forced turnover) likelihood, which is about 77% (143%) of the mean turnover (forced turnover) probability. The economic effects of match quality in explaining CEO turnover is comparable to that of poor stock and accounting performance.²³ Decades of research in CEO turnover provides robust evidence that boards consider firm performance when making CEO retention decisions. However, Brickley (2003) argues that CEO research might have reached a point of diminishing returns in estimating regressions focusing on the relation between turnover and performance. Our exploration of the role of matching consideration in CEO turnover decisions answers his call to consider less-explored issues to expand our understanding of CEO turnover decisions.

5.5. *Match Quality based on Output Data*

²³ We note here the economic effects of poor firm performance in explaining CEO turnover. First, one standard deviation decrease in adjusted stock returns is associated with an increase of 4.3% (3.7%) of the CEO turnover (forced turnover) likelihood, which is about 52% (161%) of the mean turnover (forced turnover) probability. Second, one standard deviation decrease in adjusted ROA is associated with an increase of 4.6% (4.8%) of the CEO turnover (forced turnover) likelihood, which is about 55% (209%) of the mean turnover (forced turnover) probability.

Before we conclude our paper, we investigate an alternative firm-manager match quality measure based on output data. Recent studies in labor economics use output data to capture match quality. For example, Jackson (2013) employs a data set of student test scores linked to teachers to estimate school-teacher match quality and the implications of match quality for student achievement. The use of output data to estimate match quality is particularly appealing in the Jackson's (2013) setting because teacher pay might vary across teachers and schools for reasons unrelated to productivity. In the executive labor market setting, we expect that firms reward managers for productivity generated by the efficient match between firms and managers, so our proposed firm-manager match quality measure based on compensation data might be less likely affected by the issue in the school setting. This expectation also implies a positive association between compensation match effects and match effects in firm performance. To investigate this possibility, we apply a similar research methodology to estimate match effects in firm performance. We again use two operating performance measures: ROA and ROE. As mentioned earlier, CEOs should be responsible for overall firm performance, so we focus on the firm-CEO match. We estimate the following OLS regression:

$$Match\ Effect_CEO_{ij} = \alpha + \beta Match(Performance)_{ij} + \varepsilon_{ij} \quad (10)$$

where i indexes managers and j indexes firms. $Match\ Effect_CEO$ denotes for the compensation match effects estimated for CEOs from Equation (3), and $Match(Performance)$ captures match effects in one performance measure. Similar to $Match\ Effect_CEO$, $Match(Performance)$ is estimated by a three way mixed-effects model.¹⁵

[Insert Table 8 about here]

¹⁵ Following Bertrand and Schoar (2003), we regress performance measures on the lagged logarithm of total assets, year fixed effects, firm effects, CEO effects, and match effects.

Panel A of Table 8 reports the likelihood ratio tests on the existence of match effects in firm performance measures. The tests reject the null hypothesis of no match effects in firm performance. Panel B of Table 8 reports the regression results of Equation (10). Following Bertrand and Schoar (2003) and Graham et al. (2012), we interpret the estimated coefficients as correlations rather than causal relationships. The results show that the coefficients on the match effects based on performance measures are positive and statistically significant, providing further support that firms compensate managers for match-specific productivity.

6. Conclusion

Although the quality of the match between firms and managers plays a central role in understanding the executive labor market, there is little direct empirical investigation into match quality between firms and managers. In this study, we take a first step toward addressing the empirical challenge of measuring match quality between firms and managers.

The innovation of our research methodology is to use a three way mixed-effects model to quantify firm-manager match effects in executive compensation, which has been applied to capture match effects in the field of labor economics. Using this approach, we find that unobservable match heterogeneity explains a considerable proportion of the variation in executive compensation. The results also demonstrate that the lack of consideration of match effects is associated with an overstatement of the importance of manager effects in explaining executive compensation.

Firms reward managers for productivity generated by the efficient match between firms and managers. As such, we propose using the extent to which match effects explain executive pay as a measure of firm-manager match quality. We validate the proposed match quality measure by showing that it captures systematic complementarities between particular managers and particular

firms, and that it is positively and significantly associated with firm operating performance. We also find significant negative stock market reactions to the news of sudden deaths of CEOs with high match effects, a result contradicting the prediction of managerial rent extraction. Finally, we illustrate the usefulness of the measure by showing that match quality is an economically significant factor in CEO turnover decisions, an aspect of the executive labor market that features matching consideration prominently.

Despite the appealing features of the three way mixed-effects model, we caveat that this method only identifies time-invariant match effects, but not match effects that change over time. Nevertheless, the availability of a valid match quality measure could expand our knowledge of matching consideration in the executive labor market. We believe that the empirical approach developed in this study has potential applications in the areas of economics, finance, and accounting research where the separation of manager, firm, and match effects is desirable and meaningful.

References:

- Abowd, J. M., Kramarz, F., and Margolis, D. N., 1999, High wage workers and high wage firms. *Econometrica* 67(2), 251-334.
- Abowd, J. M., Kramarz, F., Lengermann, P., and Perez-Duarte, S., 2004, Are good workers employed by good firms? A test of a simple assortative matching model for France and the United States, *Working paper*.
- Abowd, J. M., Kramarz, F., and Woodcock, S., 2006, Econometric analysis of linked employer-employee data, Working Paper.
- Acharya, V., and Volpin, P., 2010, Corporate governance externalities, *Review of Finance* 14 (1), 1-33.
- Allgood, S., and Farrell, K., 2003, The match between CEO and firm, *Journal of Business* 76, 317-341.
- Andrews, M., Gill, L., Schank, T., and Upward, R., 2008, High wage workers and low wage firms: negative assortative matching or limited mobility bias, *Journal of Royal Statistical Society* 171, 673-697.
- Bamber, L., Jiang, J., and Wang, I., 2010, What's my style? The influence of top managers and their personal backgrounds on voluntary corporate financial disclosure, *The Accounting Review* 85 (4): 1131-62.
- Bandiera, O., Guiso, L., Prat, A., and Sadun, R., 2012, Matching firms, managers, and incentives, *Working paper*.
- Bebchuk, L., and Fried, J., 2003, Executive compensation as an agency problem, *Journal of Economic Perspectives* 17, 71-92.
- Bennedsen, M., Nielsen, K., Perez-Gonzalez, F., and Wolfenzon, D., 2007, Inside the family firms: The role of families in succession decisions and performance, *Quarterly Journal of Economics* 122 (2), 647-691.
- Bertrand, M. and Schoar, A., 2003, Managing with style: the effect of managers on corporate policy, *Quarterly Journal of Economics*, 118 (4), 1169-1208.
- Bishop, J., 1990, Job performance, turnover, and wage growth, *Journal of Labor Economics* 8 (July), 363-386.
- Brickley, J., 2003, Empirical research on CEO turnover and firm-performance: a discussion, *Journal of Accounting and Economics* 36, 227-233.
- Borokhovich, K., Brunarski, K., and Parrino, R., 1997, CEO contracting and antitakeover amendments, *The Journal of Finance* 52, 1495-1517.
- Bushman, R., Gao, J., Martin, X. and Pacelli, J., 2020. The Influence of Loan Officers on Loan Contract Design and Performance. *Journal of Accounting and Economics*, 101384.
- Cameron, A. C. and Trivedi, P. K., 2005, *Microeconometrics: Methods and Applications*, Cambridge University Press.
- Carter, M. E., Franco, F., and Tuna, I., 2019, Matching premiums in the executive labor market, *The Accounting Review*, 94 (6), 109-136.

- Coles, J.L. and Li, Z.F., 2013, Managerial attributes, incentives, and performance, Working paper.
- Core, J. L., Guay, W. and D. F. Larcker, D. F., 1999, Corporate governance, chief executive officer compensation, and firm performance, *Journal of Financial Economics*, 51, 371-406.
- Core, J. L., Guay, W. and D. F. Larcker, D. F., 2008, The power of the pen and executive compensation, *Journal of Financial Economics*, 88, 1-25.
- Custodio, C., Ferreira, M. A., and Matos, P., 2013, Generalists versus specialists: lifetime work experience and CEO pay, *Journal of Financial Economics* 108, 471-492.
- DeFond, M., and Park, C., 1999, The effect of competition on CEO turnover, *Journal of Accounting and Economics* 27, 35-56.
- Demerjian, Peter, Baruch Lev, Sarah McVay, 2012, Quantifying managerial ability: A new measure and validity tests, *Management Science* 58, 1229-1248.
- Dicks, D., 2012, Executive compensation and the role of corporate governance regulation, *Review of Financial Studies* 25 (6), 1971-2004.
- Dyreng, S., Hanlon, M. and Maydew, E., 2010, The effects of executives on corporate tax avoidance, *The Accounting Review* 85 (4), 1163-1189.
- Eisfeldt, A., and Kuhnen, C., 2013, CEO turnover in a competitive assignment framework, *Journal of Financial Economics*, Vol. 109, Issue 2, 351-372.
- Ellahie, A., Tahoun, A. and Tuna, A., 2016. Do common inherited beliefs and values influence CEO pay?, Working paper.
- Engel, E., Hayes, R. M., and Wang, X., 2003, CEO turnover and properties of accounting information, *Journal of Accounting and Economics*, 36, 197-226.
- Frydman, C., and Jenter, D., 2010, CEO compensation, *Annual Review of Financial Economics* 2, 75-102.
- Falato, A., Li, D. and Milbourn, T., 2015, Which skills matter in the market for CEOs? Evidence from pay for CEO credentials, *Management Science* 61(12), 2845-2869.
- Fee, C.E., Hadlock, C.J. and Pierce, J.R., 2013, Managers with and without style: Evidence using exogenous variation, *The Review of Financial Studies* 26(3), 567-601.
- Gabaix, X., and Landier, A., 2008, Why has CEO pay increased so much? *Quarterly Journal of Economics*, Vol. 123, Issue 1, 49-100.
- Garen, J., 1988, Empirical studies of the job matching hypothesis, In R.G. Ehrenberg (ed.), *Research in Labor Economics*, Vol. 9, pp. 187-224. Connecticut: JAI Press.
- Ge, W., Matsumoto, D. and Zhang, J.L., 2011, Do CFOs have style? An empirical investigation of the effect of individual CFOs on accounting practices, *Contemporary Accounting Research* 28(4), 1141-1179.
- Gerstein, M., and Reisman, H., 1983, Strategic selection: Matching executives to business conditions, *Sloan Management Review* 24 (Winter), 33-49.
- Graham, J., Harvey, C., and Puri, M., 2013, Managerial attitudes and corporate actions, *Journal of Financial Economics* 109, 103-121.

- Graham, J. R., Li, S., and Qiu, J., 2012, Managerial attributes and Executive Compensation, *The Review of Financial Studies*, Vol. 25, Issue 1, 144-186.
- Henderson, C. R., and Kempthorne, O., Searle, S. R., and Von Krosigk, C. M. 1959, The estimation of environmental and genetic trends from records subject to culling, *Biometrics* 15, 192-218.
- Jackson, K., 2013, Match quality, worker productivity, and worker mobility: direct evidence from teachers, *Review of Economics and Statistics* 95, 1096-1116.
- Jensen, M., and Murphy, K., 1990, Performance pay and top-management incentives, *Journal of Political Economy* 98, 225-264.
- Jenter, D., Matveyev, E., and Roth, L., 2018, Good and bad CEOs. Working paper.
- Johnson, W., MaGee, R., and Nagarajan, N., 1985, An analysis of the stock price reaction to sudden executive deaths, *Journal of Accounting and Economics* 7, 151-174.
- Jovanovic, B., 1979, Firm-specific capital and turnover, *Journal of Political Economy* 87, 1246-1260.
- Jiang, J., 1996, REML estimation: asymptotic behavior and related topics, *The Annals of Statistics* 24, 255-286.
- Jiang, J., 1998, Asymptotic properties of the empirical BLUP and BLUE in mixed linear models, *Statistica Sinica* 8, 861-885.
- Lazear, E., Shaw, K., and Stanton, C., 2015, The value of bosses, *Journal of Labor Economics* 33, 823-862.
- Maksimovic, V., and Philips, G., 2002, Do conglomerate firms allocate resources inefficiently across industries? Theory and evidence, *Journal of Finance* 57, 721-767.
- Murphy, K.J., 1999, Executive compensation, *Handbook of Labor Economics* 3b, 2485–2563.
- Nguyen, B., and Nielsen, K., 2014, What death can tell: are executives paid for their contribution to firm value? *Management Science*, Vol. 60, Issue 12, 2994-3010.
- Pan, Y., 2015, The determinants and impact of executive-firm matches, *Management Science* 63, 185-200.
- Pan, Y., Siegel, S., and Wang, T. Y., 2017, Corporate risk culture, *Journal of Financial and Quantitative Analysis*, 52(6), 2327-2367.
- Peters, F., and Wagner, A., 2014, The executive turnover risk premium, *Journal of Finance* 69 (4), 1529-1563.
- Rose, N., and Shepard, A., 1997, Firm diversification and CEO compensation: Managerial ability or executive entrenchment, *Rand Journal of Economics* 28, 489-514.
- Rosen, S., 1981, The economics of superstars. *American Economic Review*, 71(5), 845-858.
- Rosen, S., 1982, Authority, control, and the distribution of earnings, *The Rand Journal of Economics*, 13(2), 311-323.
- Tervio, M., 2008, The Difference that CEOs make: An assignment model approach. *The American Economic Review*, Vol.98, No.3, 642-668.

- Warner, J.B., Watts, R.L. and Wruck, K.H., 1988, Stock prices and top management changes, *The Journal of Financial Economics* 20, 461-492.
- Woodcock, S., 2008, Wage differential in the presence of unobserved worker, firm, and match heterogeneity, *Labour Economics*, 15, 772-794.
- Woodcock, S., 2011, Match effect, Working paper.

Appendix 1. Variable Definitions

Variable	
<i>Advertising</i>	Advertising expense (Compustat data item XAD) scaled by sales (Compustat data item SALE).
<i>Adj_EBIT</i>	Industry-median adjusted EBIT. EBIT is defined as earnings before interest, taxes, and minority (Compustat data item EBIT), scaled by total assets at the beginning of period (Compustat data item AT). Industries are defined based 2-digit SIC codes.
<i>Adj_Return</i>	Industry-median adjusted annualized stock returns. Industries are defined based 2-digit SIC codes.
<i>Age</i>	CEO age.
<i>CapitalExp</i>	Capital expenditure (Compustat data item CAPX) scaled by the total assets at the beginning of period (Compustat data item AT).
<i>Cash</i>	Cash and short-term investments (Compustat data item CHE) scaled by non-cash total assets (Compustat data item AT minus Compustat data item CHE).
<i>CEO</i>	An indicator variable that equals one if the manager is the CEO of the firm for all or most of the fiscal year, and zero otherwise.
<i>DIV</i>	One minus the sum of the squares of each segment's sales over total sales at the firm level for each fiscal year.
<i>Dividend</i>	An indicator variable that equals one if dividend payout (Compustat data item DVC) is positive, and zero otherwise.
<i>Duality</i>	An indicator variable that equals one if the CEO of a company is also the board chair, and zero otherwise.
<i>EduTech</i>	5 for CEOs who earned a doctoral degree, 4 for a master's degree in science, 3 for a bachelor's degree in science, 2 for a master's degree outside of science, and 1 for a bachelor's degree outside of science.
<i>FirstYear</i>	An indicator variable that equals one for the year that the manager joined a firm.
<i>GAI</i>	General Ability Index (GAI) for CEOs of S&1500 firms from 1992 through 2007, as developed in Custodio et al. (2013).
<i>Ln(Total compensation)</i>	Natural log of total compensation. For fiscal years ending before December 15 2006, total compensation is measured by ExecuComp data item TDC1. For fiscal years ending after December 15 2006, total compensation is calculated as the sum of salary, bonus, non-equity incentive, the Black-Scholes value of option, the grant date fair value of stock grants (restricted stock and stock awarded under Long-term incentive plan), and the other compensation. Total compensation is measured in \$thousands.
<i>Ln(Total wealth)</i>	Natural log of total wealth. Total wealth is the sum of the value of the stock and option portfolio held by the executive.
<i>Ln(Assets)</i>	Natural log of total assets (Compustat Data item AT). Assets is measured in \$millions.
<i>Ln(Tenure)</i>	Natural log of manager tenure. Tenure is measured as the difference between the year the manager joined the firm and the year of the observation.
<i>Leverage</i>	Total debt (Compustat data item DLTT plus data item DLC) divided by total assets (Compustat data item AT).
<i>Match Effect</i>	Time-invariant firm-manager match effects in compensation estimated from Equation (3).
<i>Match Effect_CEO</i>	Time-invariant firm-CEO match effects in compensation estimated from Equation (3).
<i>CEO Effect</i>	Time-invariant CEO effects in compensation estimated from Equation (3).

<i>MTB</i>	Market value of equity plus the book value of debt (Compustat data item PRCC_F \times CSHO + AT- CEQ) divided by total assets (Compustat data item AT).
<i>R&D</i>	R&D expense (Compustat data item XRD) scaled by sales (Compustat data item SALE).
<i>ROA</i>	Net income (Compustat data item IB) scaled by total assets at the beginning of period (Compustat data item AT).
<i>ROE</i>	Net income (Compustat data item IB) scaled by book value of equity at the beginning of period (Compustat Data item CEQ).
<i>SG&A</i>	Selling, general and administrative expense (Compustat data item XSGA) scaled by sales (Compustat data item SALE).
<i>Retire</i>	An indicator variable for CEOs between 64 and 66 years of age.
<i>Return</i>	Annualized monthly stock returns.
<i>σ(Return)</i>	Standard deviation of monthly stock returns over the past 60 months.
<i>Talent</i>	The productive efficiency score of the manager by weighting the management team score with the executive's share of pay received for the whole management team. The efficiency score of the management team is developed in Demerjian et al. (2012).
<i>Turnover</i>	An indicator variable equals to one for all firm-years where there is a CEO turnover and zero otherwise,
<i>Forced Turnover</i>	An indicator variable equals to one for all firm-years where there is a forced CEO turnover and zero otherwise.

Appendix 2. The Estimation of The Mixed Effect Model

The estimation of the mixed effect model follows a two-step approach. First, we calculate the within-match estimators of β in Equation (3) by estimating the following Equation:

$$y_{ijt} = x'_{ijt}\beta + v_{ij} + \varepsilon_{ijt} \quad (i)$$

where v_{ij} is an indicator variable for each unique combination of manager i and firm j . x_{ijt} denotes the intercept term, observable characteristics, and year effects in Equation (3). The OLS estimator $\hat{\beta}$ from Equation (i) is the within-match estimator, which is equivalent to the estimator from the “spell fixed effects” approach in Graham et al. (2012).

Second, we estimate the following Equation to decompose $y_{ijt} - x'_{ijt}\hat{\beta}$ (e.g., the firm-by-manager effect v_{ij} and the error term ε_{ijt}) into firm, manager, and match random effects.

$$y_{ijt} - x'_{ijt}\hat{\beta} = \theta_i + \psi_j + \phi_{ij} + e_{ijt} \quad (ii)$$

where $\theta_i \sim N[0, \sigma_\theta^2]$, $\psi_j \sim N[0, \sigma_\psi^2]$, and $\phi_{ij} \sim N[0, \sigma_\phi^2]$. Following Woodcock (2011) and Jackson (2013), we normalize the population means of firm effects, manager effects, and match effects to zero. We estimate the variance of these random effects by employing the Restricted Maximum Likelihood (REML) approach. The REML approach estimates $\hat{\sigma}_\theta^2$, $\hat{\sigma}_\psi^2$, and $\hat{\sigma}_\phi^2$ simultaneously in a way that minimizes the variance of e_{ijt} . Then, we recover the best linear unbiased predictors (BLUPs) of firm, manager, and, match random effects by solving the Henderson equations.⁷ The REML approach is discussed in detail in Abowd et al. (2006).

⁷ The BLUPs are calculated by solving the Henderson equation system (e.g., Henderson et al. 1959). Let the total number of observations be N^* , the number of managers be N , the number of firms be J , and the number of firm-manager matches (pairs) be M . Stacking all the observations, we can write the linear model (6) in matrix form as: $y - x'\hat{\beta} = D\theta + F\psi + G\phi + e$, where $\theta \sim N[0, \sigma_\theta^2 I_N]$, $\psi \sim N[0, \sigma_\psi^2 I_J]$, and $\phi \sim N[0, \sigma_\phi^2 I_M]$. D is the N^* by N matrix of manager indicator variables. F is N^* by J matrix of firm indicator variables, and G is N^* by M matrix of manager-CEO pair

indicator variables. Define the matrix of variance components as $H = \begin{bmatrix} \sigma_\theta^2 I_N & 0 & 0 \\ 0 & \sigma_\psi^2 I_J & 0 \\ 0 & 0 & \sigma_\phi^2 I_M \end{bmatrix}$ and the matrix of

variance of error terms as $R = \sigma_e^2 I_{N^*}$. The BLUPs are the solutions to the Henderson equation system:

$$\begin{bmatrix} D' \\ F' \\ G' \end{bmatrix} \widehat{R}^{-1} [D \quad F \quad G] + \widehat{H}^{-1} \begin{bmatrix} \widehat{\theta} \\ \widehat{\psi} \\ \widehat{\phi} \end{bmatrix} = \begin{bmatrix} D' \\ F' \\ G' \end{bmatrix} \widehat{R}^{-1} (y - x'\hat{\beta})$$

Table 1. Structure of the Connectedness Sample

This table provides information about the structure of the connectedness sample. The sample period starts in 1992 and ends in 2015. In the connectedness sample, all the managers have worked for the firms that have hired at least one mover manager. The connectedness sample includes 32,420 managers and 2,385 firms.

Panel A: Number of firms for which managers have worked

Number of firms for which a manager has worked	Number of managers	Percentage	Cumulative percentage
1	29,238	90.19%	90.19%
2	2,804	8.65%	98.83%
3	324	1.00%	99.83%
4	48	0.15%	99.98%
5	4	0.01%	99.99%
6	2	0.01%	100.00%
Total number of managers	32,420		

Panel B: Number of managers who have worked for the same firm

Number of managers who have worked for the same firm	Number of firms	Percentage	Cumulative percentage
1-5	135	5.66%	5.66%
6-10	680	28.51%	34.17%
11-15	563	23.61%	57.78%
16-25	750	31.45%	89.22%
>25	257	10.78%	100.00%
Total number of firms	2,385		

Table 2. Summary Statistics

This table provides summary statistics of the main variables in the connectedness sample. The sample period starts in 1992 and ends in 2015. The definitions for the variables are available in Appendix 1.

Variable	N	Mean	Median	STD	P25	P75
<i>Total compensation_t</i> (\$Thousands)	153,833	2,393.12	1,261.25	3,192.15	623.28	2,686.42
<i>Assets_{t-1}</i> (\$Millions)	153,833	10,982.83	1,985.70	28,516.04	614.34	7,364.29
<i>MTB_{t-1}</i>	153,833	1.92	1.47	1.28	1.14	2.15
<i>Return_t</i>	153,833	0.15	0.10	0.47	-0.13	0.34
$\sigma(\text{Return})_t$	153,833	0.41	0.36	0.20	0.26	0.49
<i>ROA_t</i>	153,833	0.05	0.04	0.10	0.01	0.09
<i>Tenure_t</i>	153,833	11.41	10.01	7.33	6.59	14.70
<i>Duality_t</i>	153,833	0.45	0.00	0.50	0.00	1.00
<i>CEO_t</i>	153,833	0.18	0.00	0.38	0.00	0.00
<i>FirstYear_t</i>	153,833	0.03	0.00	0.18	0.00	0.00

Table 3. Determinants of the Level of Total Executive Compensation

Panel A of Table 3 reports the regression results for the determinants of total executive compensation, using the connectedness sample.

$$\begin{aligned} \ln(\text{Total compensation})_{ijt} = & \mu_0 + \beta_1 \ln(\text{Assets})_{jt-1} + \beta_2 \text{MTB}_{jt-1} + \beta_3 \text{Return}_{jt} + \beta_4 \text{Return}_{jt-1} + \beta_5 \text{ROA}_{jt} + \\ & \beta_6 \text{ROA}_{jt-1} + \beta_7 \sigma(\text{Return})_{jt} + \beta_8 \text{Duality}_{jt} + \beta_8 \ln(\text{Tenure})_{it} + \beta_9 \text{CEO}_{it} + \beta_{10} \text{FirstYear}_{it} + \text{Year Fixed} \\ & \text{Effects} + \theta_i + \psi_j + \phi_{ij} + \varepsilon_{ijt} \end{aligned}$$

The sample period starts in 1992 and ends in 2015. Column (1) is a pooled OLS regression without firm or manager fixed effects. Column (2) is an OLS regression with firm and manager fixed effects. Column (3) is a two way mixed-effects regression model with firm and manager random effects. Column (4) is a three way mixed-effects regression model with firm, manager, and match random effects. The definitions for all the variables are available in Appendix 1. The standard errors are clustered by firms. We report t-statistics in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively. In the last row, we report the likelihood ratio test based on the log-likelihoods of specifications with and without match effects. Panel B reports the summary statistics of observable and unobservable components that determine $\ln(\text{Total compensation})_{ijt}$ and presents the decomposition of the variance of $\ln(\text{Total compensation})_{ijt}$, using the estimation results from columns (2) – (4) of Panel A. The observable time-variant characteristics component includes $\ln(\text{Assets})_{jt-1}$, MTB_{jt-1} , Return_{jt} , Return_{jt-1} , ROA_{jt} , ROA_{jt-1} , $\sigma(\text{Return})_{jt}$, Duality_{jt} , $\ln(\text{Tenure})_{it}$, CEO_{it} , and FirstYear_{it} . The percentage of the variance of total compensation attributable to particular components equals the covariance between each component and $\ln(\text{Total compensation})_{ijt}$ scaled by the variance of $\ln(\text{Total compensation})_{ijt}$.

Panel A: Regression Results

	Pooled OLS (1)	Firm and manager fixed effects (2)	Firm and manager mixed effects (3)	Firm, manager and match mixed effects (4)
$Ln(Assets)_{jt-1}$	0.363*** (59.85)	0.226*** (16.74)	0.221*** (15.86)	0.221*** (15.86)
MTB_{jt-1}	0.170*** (19.80)	0.098*** (12.00)	0.098*** (11.68)	0.098*** (11.68)
$Return_{jt}$	0.222*** (22.44)	0.173*** (18.86)	0.172*** (18.55)	0.172*** (18.55)
$Return_{jt-1}$	0.068*** (7.25)	0.087*** (11.37)	0.085*** (11.01)	0.085*** (11.01)
ROA_{jt}	0.381*** (5.74)	0.361*** (5.61)	0.368*** (5.66)	0.368*** (5.66)
ROA_{jt-1}	0.350*** (5.94)	0.251*** (4.63)	0.257*** (4.70)	0.257*** (4.70)
$\sigma(Return)_{jt}$	0.620*** (12.80)	0.086* (1.81)	0.092* (1.87)	0.092* (1.87)
$Duality_{jt}$	0.085*** (5.77)	0.005 (0.46)	0.004 (0.38)	0.004 (0.38)
$Ln(Tenure)_{it}$	0.105*** (11.52)	0.144*** (17.12)	0.149*** (16.35)	0.149*** (16.35)
CEO_{it}	0.956*** (109.21)	0.423*** (33.39)	0.414*** (31.08)	0.414*** (31.08)
$FirstYear_{it}$	0.513*** (21.41)	0.365*** (14.83)	0.356*** (14.18)	0.356*** (14.18)
Year Effects	Yes	Yes	Yes	Yes
Firm Effects	No	Yes	Yes	Yes
Manager Effects	No	Yes	Yes	Yes
Match Effects	No	No	No	Yes
# Observations	153,833	153,833	153,833	153,833
R ²	0.533	0.831	N/A	N/A
H_0 : No match effects (p-value)				2,166(< 0.000)

Panel B: Relative Importance of Different Components in Determining Compensation

		Firm and manager fixed effects	Firm and manager mixed effects	Firm, manager and match mixed effects
		(1)	(2)	(3)
	<i>Ln(Total compensation)</i>	1.05	1.05	1.05
	Observable time-variant characteristics ($x\beta$)	0.45	0.44	0.44
S.D.	Firm Effects	0.80	0.41	0.42
	Manager Effects	0.93	0.38	0.16
	Match Effects	N/A	N/A	0.23
	Year Effects	0.35	0.35	0.35
	Residuals	0.43	0.46	0.45
% of the variance of total compensation attributable to particular components	Observable time-variant characteristics ($x\beta$)	29.4%	28.9%	28.9%
	Firm Effects	8.9%	20.5%	21.0%
	Manager Effects	35.4%	18.7%	7.6%
	Match Effects	N/A	N/A	11.5%
	Year Effects	9.5%	9.6%	9.6%
	Residuals	16.9%	22.4%	21.5%

Table 4. The Relation between Compensation Match Effects and Firm-Manager Complementarities

This table reports the relation between compensation match effects and two firm-manager complementarities: (1) between firm size and managerial talent and (2) between diversification degree and CEO general management skills. The sample period starts in 1992 and ends in 2015.

$$Match\ Effect_{ij} / Match\ Effect_CEO_{ij} = \beta_0 + \beta_1 Ln(Assets)_j + \beta_2 Talent_i + \beta_3 Ln(Assets)_j \times Talent_i + \beta_4 DIV_j + \beta_5 GAI_i + \beta_6 DIV_j \times GAI_i + \varepsilon_{ij}$$

Match Effect denotes for manager compensation match effects and *Match Effect_CEO* denotes for CEO compensation match effects estimated from Equation (3). $Ln(Assets)_j \times Talent_i$ is the interaction term between size and talent averaged over the job tenure for each manager, $DIV_j \times GAI_i$ is the interaction term between *DIV* and *GAI* averaged over the job tenure for each CEO. The definitions of all the variables are available in Appendix 1. We report t-statistics in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

	<i>Match Effect</i> _{ij}	<i>Match Effect_CEO</i> _{ij}	
	(1)	(2)	(3)
<i>Ln(Assets)</i> _j	-0.012*** (-3.15)		-0.031** (-2.08)
<i>Talent</i> _i	0.392*** (2.89)		-0.061 (-0.58)
<i>Ln(Assets)</i> _j × <i>Talent</i> _i	0.334*** (4.43)		0.314** (2.18)
<i>DIV</i> _j		0.056*** (6.28)	0.059*** (6.05)
<i>GAI</i> _i		0.044*** (6.26)	0.047*** (6.42)
<i>DIV</i> _j × <i>GAI</i> _i		0.005** (2.49)	0.005** (2.44)
# Observations	28,746	2,880	2,880
R ²	0.005	0.042	0.046

Table 5. Compensation Match Effects and Firm Performance

This table reports the results from the OLS regression of firm performance on CEO compensation match effects and other control variables. The sample period starts in 1992 and ends in 2015.

$$ROA_{jt} / ROE_{jt} = \beta_0 + \beta_1 Match\ Effect_CEO_{ij} + \beta_2 Ln(Assets)_{jt-1} + \beta_3 MTB_{jt-1} + \beta_4 Leverage_{jt} + \beta_5 \sigma(Return)_{jt} + Year\ FEs + \varepsilon_{ijt}$$

Match Effect_CEO denotes for CEO compensation match effects estimated from Equation (3). The definitions of all the variables are available in Appendix 1. We report t-statistics in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

	<i>ROA_{jt}</i>	<i>ROE_{jt}</i>
	(1)	(2)
<i>Match Effect_CEO_{ij}</i>	0.010** (2.03)	0.046*** (2.63)
<i>Ln(Assets)_{jt}</i>	0.001 (0.74)	0.019*** (5.27)
<i>MTB_{jt}</i>	0.019*** (7.13)	0.034*** (7.24)
<i>Leverage_{jt}</i>	-0.086*** (-6.18)	0.055 (1.48)
<i>σ(Return)_{jt}</i>	-0.156*** (-19.14)	-0.236*** (-7.33)
Year FEs	Yes	Yes
# Observations	26,685	26,685
Adjusted R ²	0.214	0.0428

Table 6. Compensation Match Effects and Cumulative Abnormal Returns around CEO Sudden Deaths

This table reports empirical results on cumulative abnormal returns (CARs) around the announcement date for firms with a CEO sudden deaths over the period from 1992 to 2012. We partition the sample of CEO sudden deaths into two groups based on whether the CEO's compensation match effect is above the median or not. Panel A presents CEO characteristics for two groups: high and low match effects. Panels B presents CARs over three different horizons ([-1, 1], [-2, 2], and [-2, 5]) around the announcement of CEO sudden deaths for two groups: high and low match effects. CARs are the sum of market-model adjusted abnormal returns. Panels C presents results from regressions of CARs around the announcement of CEO sudden deaths on CEO compensation match effects and other control variables.

$$CAR_{ij} = \mu_0 + \beta_1 Match\ Effect_CEO_{ij} + \beta_2 CEO\ Effect_i + \beta_3 Ln(Tenure)_{ij} + \beta_4 Ln(Assets)_j + \beta_6 Adj_Return_j + \beta_7 Adj_EBIT_j + \beta_{10} Ln(Age)_i + \varepsilon_{ij}$$

We measure both firm and CEO characteristics at the year of CEO death. Robust standard errors are used to calculate test statistics for means. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Panel A: CEO Characteristics

CEO characteristics	Full sample	High match	Low match	High - Low	
	(N=71)	(N=35)	(N=36)	Mean	t-stat
<i>Match Effect_CEO</i>	0.22	0.42	0.02	0.40***	6.57
<i>Age</i>	59.02	59.12	58.97	0.15	0.64
<i>Tenure</i>	8.50	10.17	6.64	3.53*	1.76

Panel B: Cumulative Abnormal Returns around CEO Sudden Deaths

Window	Full sample		High match		Low match		High - Low	
	(N=71)		(N=35)		(N=36)		Mean	t-stat
[-1, +1]	-1.02*	-1.81	-1.94**	-2.01	0.13	0.22	-2.07**	-1.97
[-2, +2]	-1.23*	-1.80	-2.40***	-2.98	0.10	0.16	-2.50**	-2.18
[-2, +5]	-1.52*	-1.93	-2.42***	-2.68	0.64	0.98	-3.06**	-2.43

Panel C: Multivariate Regression Results

	$CAR[-1, 1]_{ij}$	$CAR[-2, +2]_{ij}$	$CAR[-2, +5]_{ij}$
	(1)	(2)	(3)
<i>Match Effect_CEO_{ij}</i>	-0.540*	-2.199*	-5.315***
	(-1.73)	(-1.77)	(-2.84)
<i>CEO Effect_i</i>	-0.703	-1.176	-2.618
	(-1.39)	(-1.21)	(-1.60)
<i>Ln(Temure)_{ij}</i>	-0.095*	-0.119**	-0.181*
	(-1.79)	(-2.13)	(-1.89)
<i>Ln(Assets)_j</i>	0.387	0.624	0.567
	(0.76)	(1.25)	(1.27)
<i>Adj_Return_j</i>	-0.774	-0.664	-0.652
	(-0.29)	(-0.29)	(-0.30)
<i>Adj_EBIT_j</i>	-2.302**	-6.204	-3.701
	(-2.23)	(-1.56)	(-1.44)
<i>Ln(Age)_i</i>	0.032	0.083	0.067
	(0.25)	(0.67)	(0.61)
# Observations	71	71	71
R ²	0.075	0.092	0.149

Table 7. Compensation Match Effects and CEO Turnover

This table reports empirical results about the relation between compensation match effects and CEO turnover. Panel A reports summary statistics of the Turnover, Forced Turnover, and Control subsamples during the period from 1992 to 2015. The Turnover subsample includes the firm-year observations where CEO changed. The Forced Turnover subsample includes firm-year observations where the CEO turnover is classified as forced based on Peters and Wagner (2014). The Control subsample includes the firm-year observations where no CEO turnover occurs. Panel B presents results from the probit regression of CEO turnover /forced turnover on CEO compensation match effects and other control variables.

$$Turn_{ijt} / Forced_{ijt} = \mu_0 + \beta_1 Match\ Effect_CEO_{ij} + \beta_2 CEO\ Effect_i + \beta_3 Ln(Tenure)_{it-1} + \beta_4 Ln(Assets)_{jt-1} + \beta_5 MTB_{jt-1} + \beta_6 Adj_Return_{jt-1} + \beta_7 Adj_EBIT_{jt-1} + \beta_8 \sigma(Return)_{jt-1} + \beta_9 Duality_{jt-1} + \beta_{10} Ln(Age)_{it-1} + \beta_{11} Retire_{it-1} + Year\ FEs + Industry\ FEs + \varepsilon_{ijt}$$

Match Effect_CEO denotes for CEO compensation match effects estimated from Equation (3). For each independent variable, we report the coefficient estimate, z-statistics (in parentheses), and the marginal effect. The definitions of all the variables are available in Appendix 1. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Panel A: Summary Statistics

	Turnover (N=1,363)		Forced Turnover (N=349)		Control (N=15,026)	
	Mean	STD	Mean	STD	Mean	STD
<i>Match Effect_CEO_{ij}</i>	0.09	0.19	0.09	0.19	0.11	0.18
<i>CEO Effect_i</i>	0.05	0.10	0.05	0.10	0.06	0.10
<i>Tenure_{it-1}</i>	16.19	8.56	14.55	7.56	15.27	8.14
<i>Assets_{jt-1}</i>	14,290.67	34,217.49	15,402.25	40,469.04	13,381.89	33,713.34
<i>MTB_{jt-1}</i>	1.74	1.28	1.52	0.83	1.82	1.17
<i>Adj_Return_{jt-1}</i>	-0.03	0.36	-0.15	0.41	0.04	0.37
<i>Adj_EBIT_{jt-1}</i>	0.00	0.09	-0.02	0.10	0.01	0.08
<i>σ(Return)_{jt-1}</i>	0.42	0.21	0.49	0.23	0.42	0.21
<i>Duality_{jt-1}</i>	0.32	0.47	0.20	0.40	0.31	0.46
<i>Age_{it-1}</i>	59.34	7.13	54.07	6.05	55.39	6.83
<i>Retire_{it-1}</i>	0.14	0.35	0.01	0.11	0.05	0.22

Panel B: Probit Regression Results

	<i>Turn_{ijt}</i>		<i>Forced_{ijt}</i>	
	Coefficients	Marginal effects	Coefficients	Marginal effects
<i>Match Effect_CEO_{ij}</i>	-0.420*** (-4.43)	-0.064*** (-4.46)	-0.615** (-2.17)	-0.033** (-2.17)
<i>CEO Effect_{ij}</i>	0.126 (0.85)	0.019 (0.85)	0.524 (1.01)	0.030 (1.01)
<i>Ln(Tenure)_{it-1}</i>	-0.001 (-0.03)	-0.000 (-0.03)	-0.132*** (-3.20)	-0.007*** (-3.19)
<i>Ln(Assets)_{jt-1}</i>	0.046*** (4.90)	0.007*** (4.90)	0.039* (1.75)	0.002* (1.75)
<i>MTB_{jt-1}</i>	0.008 (0.52)	0.001 (0.52)	-0.097** (-2.40)	-0.005** (-2.40)
<i>Adj_Return_{jt-1}</i>	-0.302*** (-7.32)	-0.043*** (-7.39)	-0.655*** (-7.87)	-0.037*** (-8.04)
<i>Adj_EBIT_{jt-1}</i>	-0.324* (-1.76)	-0.046* (-1.76)	-0.845** (-2.24)	-0.048** (-2.25)
<i>σ(Return)_{jt-1}</i>	0.273*** (3.52)	0.039*** (3.51)	0.796*** (5.94)	0.045*** (6.04)
<i>Duality_{it-1}</i>	0.187*** (4.22)	0.026*** (4.22)	0.130** (2.36)	0.007** (2.36)
<i>Ln(Age)_{it-1}</i>	0.039*** (12.58)	0.005*** (13.31)	-0.008* (-1.93)	-0.000* (-1.93)
<i>Retire_{it-1}</i>	0.338*** (6.01)	0.048*** (6.12)	-0.572** (-2.44)	-0.032** (-2.45)
Year FEs	Yes	Yes	Yes	Yes
Industry FEs	Yes	Yes	Yes	Yes
N (Turn or Forced)	1,363		349	
N (Control)	15,026		15,026	
# observations	16,389		15,375	
Pseudo R ²	0.078		0.111	

Table 8. The Relation between Compensation Match Effects and Match Effects in Firm Performance

This table presents the relation between compensation match effects and match effect in firm performance. The sample period starts in 1992 and ends in 2015. Panel A reports chi-square statistics and p-value of the likelihood ratio test based on the log-likelihoods of specifications with and without match effects to test the existence of match effects in performance measures, including ROA and ROE. Panel B presents the results from the following regressions.

$$Match\ Effect_CEO_{ij} = \alpha + \beta Match(Performance)_{ij} + \varepsilon_{ij}$$

In each regression, we regress compensation match effects on match effects in one particular performance measure. *Match Effect_CEO* denotes for CEO compensation match effects estimated from Equation (3). *Match(Performance)* represents match effects in one particular performance measure. We report t-statistics in parentheses. ***, **, and * denote significance at the 1, 5, and 10 percent levels, respectively.

Panel A: Likelihood Ratio Tests on Match Effects in Firm Performance Measures

Performance measure	Likelihood ratio tests	
	Chi-square stat.	p-value
<i>ROA</i>	490.10	<0.000
<i>ROE</i>	25.39	<0.000

Panel B: Univariate Regression Results

	<i>Match Effect_CEO_{ij}</i>	
	(1)	(2)
<i>Match(ROA)_{ij}</i>	0.294*** (4.91)	
<i>Match(ROE)_{ij}</i>		0.227*** (3.62)
# Observations	28,561	28,561
R ²	0.002	0.001



(https://www.nsf.gov/)



(https://nces.nsf.gov)

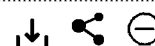
National Center for Science and Engineering Statistics
(https://nces.nsf.gov)



InfoChart

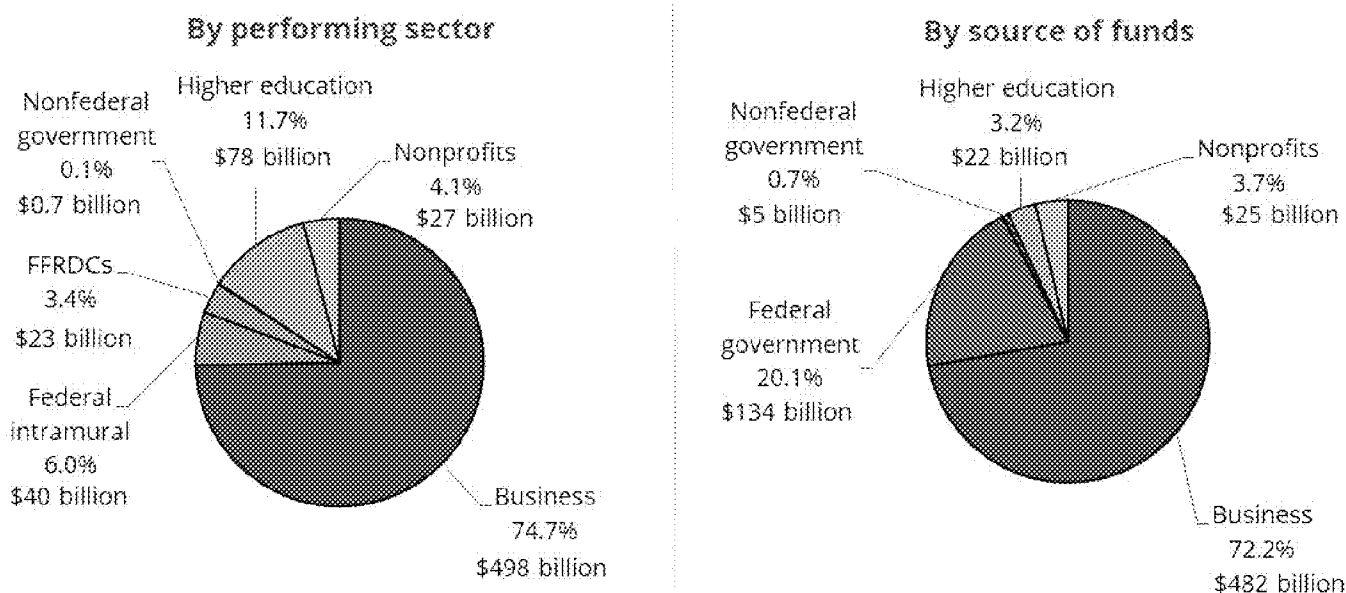
New Data on U.S. R&D: Summary Statistics from the 2019–20 Edition of National Patterns of R&D Resources

Figure 1



U.S. R&D expenditures, by performing sector and source of funds: 2019

U.S. total of R&D expenditures in 2019: \$667 billion



Data View

FFRDCs = federally funded research and development centers.

Source(s):

National Center for Science and Engineering Statistics, National Patterns of R&D Resources (annual series).

New data on research and experimental development (R&D) expenditures indicate that U.S. R&D performance totaled \$667 billion in 2019 (figure 1). The business sector accounted for 75% of the U.S. R&D performance total; the other larger performers were higher education (12%) and the federal

government (federal intramural facilities and federally funded research and development centers)



source of funds, the largest funders were the business sector (72%) and the federal government (20%). The \$667 billion U.S. R&D total in 2019 compares with \$554 billion in 2017 and \$505 billion in 2016. These increases in 2018 and 2019—\$51 billion and \$62 billion, respectively—

were large compared with recent history. (The average annual increase in 2010–17 was \$21 billion).

InfoChart

The total for 2020 is \$708 billion—a \$41 billion increase over the 2019 level. Data are from the *National Patterns of R&D Resources* series from the National Center for Science and Engineering Statistics (NCSES). Further detail will be available in NCSES reports to follow.

Suggested Citation: Boroush M; (mailto:mboroush@nsf.gov) National Center for Science and Engineering Statistics (NCSES). 2021. *New Data on U.S. R&D: Summary Statistics from the 2019–20 Edition of National Patterns of R&D Resources*. NSF 22-314. Alexandria, VA: National Science Foundation. Available at <https://ncses.nsf.gov/pubs/nsf22314> (<https://ncses.nsf.gov/pubs/nsf22314>).

Authors

Mark Boroush

Source Data & Analysis

National Patterns of R&D Resources: 2019–20 Data Update (<https://ncses.nsf.gov/pubs/nsf22320>)

Can't Find What You Are Looking For?

Search NCSES (/search) to find data and reports.



(<https://www.facebook.com/US.NSF>)



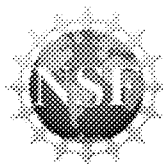
(<https://twitter.com/NCSESGov>)



(<https://www.instagram.com/nsfgov/>)



(<https://www.youtube.com/user/VideosatNSF>)



(<https://www.nsf.gov/>)



(<https://ncses.nsf.gov/>)

Website Policies (<https://nsf.gov/policies/>) | Budget and Performance (<https://nsf.gov/about/performance/>) |
Inspector General (<https://nsf.gov/oig/>) | Privacy (<https://nsf.gov/policies/privacy.jsp>) |
FOIA (<https://nsf.gov/policies/foia.jsp>) | No FEAR Act (https://nsf.gov/od/oecr/nofearact_notice.jsp) |

USA.gov (<https://www.usa.gov/>) | Accessibility (<https://nces.nsf.gov/accessibility>) |

Language (https://nsf.gov/policies/nsf_plain_language.jsp) |

<https://www.nsf.gov/help/contact.jsp>) | Help (<https://www.nsf.gov/help/>)

National Science Foundation (<https://nces.nsf.gov>), Alexandria, Virginia 22314, USA Tel: (703) 292-5111, FIRS:
(800) 877-8339 | TDD: (800) 281-8749

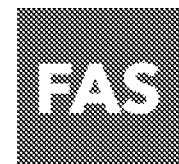
InfoChart

DAY ONE PROJECT

Supporting Market Accountability, Workplace Equity, and Fair Competition by Reining in Non-Disclosure Agreements

**Rachel Arnow-Richman
Gretchen Carlson
Orly Lobel
Julie Roginsky
Jodi Short
Evan Starr**

January 2022



DAY ONE PROJECT

Summary

Overuse of non-disclosure agreements (NDAs) is a pervasive problem in the United States. Companies apply these silencing tools to prevent their workers from sharing critical information with one another and the public. This in turn threatens economic growth, limits competition, and inhibits workplace equity. Workers need reliable information about corporate practices to assess job quality, ensure personal safety, and obtain pay commensurate with their worth. The public needs information about corporate practices to decide how to use their investment and purchasing power. Yet existing laws give companies enormous latitude to designate information as confidential, allowing them to impose NDAs and other contract clauses and internal policies that prevent workers from sharing information with those who need to know.

It is time for government to rein in corporate secrecy. The #MeToo movement revealed how NDAs enable and perpetuate misconduct at work, prompting public outrage and support for legislative action. New empirical evidence has exposed just how widely NDAs are being used in the corporate world: researchers estimate that between 33% and 57% of U.S. workers are constrained by an NDA or similar mechanism.^{1,2} At recent hearings and public events, regulators have signaled their concern about the anti-competitive effects of restrictive employment agreements.³ Policymakers should seize this moment of support to pursue a comprehensive legislative and multi-agency agenda limiting inappropriate use of NDAs. A strong action plan should include proactive enforcement of existing laws governing NDAs; new legislation prohibiting the most harmful uses of NDAs; and interagency collaboration to educate the public, collect data, and support research on impacts of corporate secrecy practices. Together, these efforts to limit NDA abuse will promote market accountability, workplace equity, and fair competition.

Challenge and Opportunity

NDAs are contracts in which parties agree not to disclose any information designated confidential by the agreement. In some cases, NDAs may be used appropriately to protect valuable trade secrets or other intellectual property. But employers often draft these agreements broadly to conceal many other types of information, sometimes in ways that overstep existing legal bounds. For instance, the Weinstein Company required employees to sign NDAs that prohibit disclosure of “any confidential, private, and/or non-public information obtained by Employee during Employee’s employment with the Company concerning the personal, social, or business activities of the Company, the Co-Chairmen, or the executives, principals, officers, directors, agents, employees of, or contracting parties (including, but not limited to artists) with,

¹ Starr, E.P.; Prescott, J.J.; Bishara, N.D. (2020). Noncompete Agreements in the US Labor Force. *The Journal of Law and Economics*, 64(1).

² Balasubramnian, N.; Starr, E.; Yamaguchi, S. (2021). Bundling Employment Restrictions and Value Capture from Employees. Available at <http://dx.doi.org/10.2139/ssrn.3814403>.

³ Federal Trade Commission. (2021). [Making Competition Work: Promoting Competition in Labor Markets](#).

DAY ONE PROJECT

the Company.”⁴ Some companies require employees to sign non-disparagement agreements. These particularly broad NDAs prohibit employees from disclosing any information that might, as a non-disparagement agreement for employees of Task Rabbit reads, “disparage the Company, and the Company’s officers, directors, employees, investors and agents, in any manner likely to be harmful to them or their business, business reputation or personal reputation.”⁵ NDAs and non-disparagement agreements often purport to apply indefinitely, preventing workers from sharing information long after they have left employment.

NDAs are imposed on workers at various points during the employment relationship. They are regularly included as part of a bundle of mandatory HR forms that new hires must sign as a condition of employment. They can also be imposed and enforced through confidentiality policies contained in personnel manuals or codes of conduct that prevent employees from sharing information about the company with outsiders and sometimes even with co-workers. They are also routinely included in standardized as well as negotiated severance agreements that workers sign when ending their employment. Lastly, they are also often included in settlement agreements that resolve workplace disputes and in agreements that force employees to arbitrate disputes in secret. By preventing workers from disclosing information on everything from workplace harassment and abuse to compensation practices and safety conditions, NDAs stifle competition, limit the free flow of ideas,⁶ and allow toxic workplace conditions to fester.^{7,8,9}

Prevalence of NDAs

Though researchers estimate that between 33% and 57% of U.S. workers are constrained by an NDA or similar mechanism,^{10,11} it is difficult to precisely determine how many employees are silenced by NDAs because NDAs are designed to conceal information. In fact, NDAs often provide that the mere existence of the agreement is itself a secret. Lawyers regularly encourage firms¹² to use broad NDAs as a condition of employment—not only to protect trade secrets, but also to discourage employees from revealing bad employment experiences.¹³ Prevalence of NDAs also varies by sectors. For instance, 73% of workers in “computer or mathematical jobs” report having an NDA with their employer.^{14,15}

⁴ Sockin, J.; Sojourner, A.; Starr, E. (2021). Non-Disclosure Agreements and Externalities from Silence. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3900285.

⁵ Ibid.

⁶ Lobel, O. (2015). The New Cognitive Property: Human Capital Law and the Reach of Intellectual Property. *Texas Law Review*, 93(4): 789–854.

⁷ Lobel, O. (2018). NDAs are Out of Control. Here's What Needs to Change. *Harvard Business Review*, January 30.

⁸ Hemel, D. (2017). How Nondisclosure Agreements Protect Sexual Predators. *Vox*, October 9.

⁹ Birnbaum, E. (2020). A wall of silence holding back racial progress in tech: NDAs. *Protocol*, July 1.

¹⁰ Starr, E.P.; et al. (2020).

¹¹ Balasubramnian, N.; et al. (2021).

¹² Gensing-Pophal, L. (2019). How to Protect Your Employer Brand on Glassdoor. SHRM, October 25.

¹³ Sockin, J.; Sojourner, A.; Starr, E. (2021). Non-Disclosure Agreements and Externalities from Silence. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3900285.

¹⁴ Balasubramnian, N.; et al. (2021).

¹⁵ Balasubramnian, N.; et al. (2021).

DAY ONE PROJECT

How NDAs Hurt Workers, the Public, and the Economy

The overuse of broad NDAs can have harmful economic and social effects. Depending on how they are drafted and enforced, NDAs may undermine law enforcement and regulatory compliance, distort labor and investment markets, constrain fair competition, allow toxic workplace conditions such as harassment and discrimination to persist, and undercut efforts to make workplaces more diverse and equitable.

Interference with Law Enforcement and Regulatory Compliance

Social, psychological, and economic disincentives already discourage employees from blowing the whistle on harmful and illegal corporate behavior.¹⁶ NDAs add another barrier preventing this critical information from reaching regulators and the public. NDAs have been used by companies to cover up illegal behavior. They have been used to silence whistleblowers who disclose information about products that threaten public health and safety.¹⁷ They have even been used to prevent employees from disclosing illegal conduct to government regulators despite countervailing law. A complaint filed by the California Department of Fair Employment and Housing (DFEH) against gaming company Activision Blizzard alleges that, contrary to law, the company pressured employees to sign contracts waiving their right to speak to investigators and requiring them to notify the company before disclosing information to DFEH.¹⁸ Some companies have required employees to agree to secrecy about corporate pay practices and diversity statistics, thereby depriving regulators of vital information about companies' compliance with pay equity and anti-discrimination laws.¹⁹ The dangers of overly aggressive NDAs have become especially clear during the COVID-19 pandemic, when it is vital for the public to know if companies are disregarding essential health and safety guidelines designed to reduce virus spread.

Market Distortion

NDAs deprive individuals of information they need to assess competing job offers and make informed decisions about where to work. They also degrade the reliability of employer reviews that workers post to online job platforms. This is because workers subject to broad NDAs are more likely to censor themselves and withhold negative information. New research shows that on Glassdoor, workers in states with more stringent limits on NDAs are 16% more likely to give a one-star review, write 8% more about the “cons” of working at the firm, and discuss harassment at work 22% more often.²⁰ That same research also shows that states with more stringent limits on NDAs increase reporting of sexual harassment and safety violations to federal agencies. NDAs hence remove an important check on corporate behavior, since companies

¹⁶ Feldman, Y.; Lobel, O. (2010). The Incentives Matrix: A Study of the Comparative Effectiveness of Monetary Rewards as Compliance Systems. *Texas Law Review*, 88(6).

¹⁷ Short, J.L. (1999). Killing the Messenger: The Use of Nondisclosure Agreements to Silence Whistleblowers. *University of Pittsburgh Law Review*, 60(4): 1207–1234.

¹⁸ Amended Complaint & Demand for Jury Trial, Department of Fair Employment and Housing v. Activision, No. 21STCV26571 (L.A. Cty. Super. Ct., August 23, 2021). Available at

<https://www.documentcloud.org/documents/21048497-activision-amended-complaint-82321>

¹⁹ Bowman Williams, J. (2019). Diversity as a Trade Secret. *Georgetown Law Journal*, 107(6): 1685–1732.

²⁰ Sockin, J.; et al. (2021).

DAY ONE PROJECT

have been shown to improve their practices in response to negative job reviews and investigations into their practices.²¹ NDAs thus enable bad employers to hide their flaws and make it difficult for good employers to distinguish themselves in the market.

Accurate information about workplace conditions is also valuable to investors, who have increasingly come to recognize that the ways companies treat their workers impact corporate financial performance.²² A nonprofit investment group recently called on the Securities and Exchange Commission to develop a standardized set of workplace-practice metrics as part of a comprehensive framework for evaluating socially responsible corporate governance.²³ NDAs can hide information about workplace conditions that investors value.

Constraints on Fair Competition

Broad NDAs can impede fair competition. Research has demonstrated that non-compete agreements—which prohibit departing workers from joining competitors—impede worker mobility, economic growth, and new firm entry. Broad NDAs pose some of the same competitive risks as non-competes because they limit workers' ability to share and apply knowledge gained through on-the-job experience. This in turn diminishes workers' human capital and makes them less competitive in the labor market.²⁴ Indeed, employers in states that ban non-competes have illegally attempted to use broad NDAs as an alternative mechanism to impede employee mobility.²⁵

Harassment and Discrimination

NDAs conceal harassment, discrimination, and abuse in the workplace. As the #MeToo movement showed, perpetrators of harassment and discrimination are often repeat offenders.^{26,27} NDAs may prevent victims of harassment and discrimination from warning co-workers and prospective employees about a company's toxic workplace environment, leaving others at risk. NDAs may also prohibit or inhibit employees from disclosing information to government agencies, shielding offenders from outside investigation. By limiting what employees can share, NDAs allow harmful and abusive behavior to persist.

²¹ Dube, S.; Zu, C. (2021). The Disciplinary Effect of Social Media: Evidence from Firms' Responses to Glassdoor Reviews. *Journal of Accounting Research*, 59(5): 1783-1825. See also Johnson, Matthew S. "Regulation by shaming: Deterrence effects of publicizing violations of workplace safety and health laws." *American economic review* 110, no. 6 (2020): 1866-1904.

²² Mahoney, C. (2021). Companies that Have the Most Workers Earning a Living Wage Have Higher Return on Assets. *Just Capital*, February 21.

²³ Whittaker, M. (2021). CEO JUST Capital, Letter filed in response to SEC request for comments on ESG disclosures. Available at <https://www.sec.gov/comments/climate-disclosure/cl112-8922517-245118.pdf>. June 14.

²⁴ See *TLS Management & Marketing Services v. Rodriguez-Toledo*, 966 F.3d 46 (1st Cir. 2020). The Court of Appeals found that "overly broad nondisclosure agreements, while not specifically prohibiting an employee from entering into competition with the former employer, raise the same policy concerns about restraining competition as noncompete clauses where...they have the effect of preventing the defendant from competing with the plaintiff."

²⁵ *Brown v. TGS Management Co., LLC*, 57 Cal.App.5th 303 (2020).

²⁶ Ayres, I. (2018). Targeting Repeat Offender NDAs. *Stanford Law Review Online*, 71.

²⁷ Lobel, O. (2016). The Prisoner's Dilemma in Airing Fox's Corporate Culture. *Fortune*, July 28.

DAY ONE PROJECT

Diversity, Equity, and Inclusion

Restrictions on employee disclosure of harassment and discrimination undermine the goal of achieving diverse and equitable workplaces. Workers of color, women, and LGBTQ+ workers are disproportionately likely to suffer harassment and discrimination in the workplace. Such adverse experiences can have significant psychological and professional consequences, including driving workers out of certain jobs and even out of certain industries.^{28,29} NDAs exacerbate these harms by suppressing information about systemic workplace inequities and by denying workers a forum to expose and discuss harassment and discrimination.

Corporate-secrecy practices shrouding employee compensation similarly undermine efforts of diverse employees to achieve pay equity. Contrary to law, some NDAs and confidentiality policies prohibit employees from discussing their compensation, which makes it challenging for those employees to negotiate fair salary terms commensurate with their value.³⁰ Studies have found that states that adopted anti-secrecy pay laws increased gender wage equality relative to states that did not.^{31,32}

National Leadership Is Needed

As described above, overly broad NDAs and the organizational secrecy practices they support pose serious risks to our economy and our society. Yet absent government intervention, these challenges will persist. Individual firms have incentives to maintain their reputations using corporate-secrecy tactics despite the social costs of such behavior. Many of those who value the information concealed by NDAs lack the capacity and power to pressure companies to change. Policymakers have an imperative to use the levers of government to curb NDA abuse.

A minority of states, including California, Illinois, New Jersey, New York and Washington, passed legislation in the wake of #MeToo regulating some uses of NDAs. But these laws comprise an inconsistent and incomplete regulatory patchwork. State laws differ in scope of coverage and impose different compliance standards, making it difficult for employees and companies to determine what employee disclosures are legally protected where. Moreover, the harms caused by NDAs do not stop at state lines. In fact, uneven regulation of NDAs further distorts markets by making it easier for companies to conceal information and restrict competition in some states than in others. Multi-state firms can use choice of law and choice of forum provisions to exploit inter-state legislative discrepancies, i.e., to apply the most lenient state-level secrecy laws to the entirety of a multi-state workforce.

²⁸ Williams, J.; Short, J.; Brooks, M.; Hardcastle, H.; Ellis, T.; Saron, R. What's Reasonable Now? Sexual Harassment Law after the Norm Cascade. *Michigan State Law Review* 2019(1): 139-224 (2019).

²⁹ Pina, A.; Gannon, T.A. (2010). An Overview of the Literature on Antecedents, Perceptions and Behavioural Consequences of Sexual Harassment. *Journal of Sexual Aggression*, 18(2): 209-232.

³⁰ Lobel, O. (2020). Knowledge Pays: Reversing Information Flows and the Future of Pay Equity. *Columbia Law Review*, 120(3): 547-612.

³¹ Fetisova-Canas, O. (2014). Effects of Anti-Secrecy Pay Laws on the Gender Wage Gap. University of Maryland, May.

³² Kim, M. (2015). Pay Secrecy and the Gender Wage Gap in the United States. *Industrial Relations*, 54(4): 648-667.

DAY ONE PROJECT

The upshot is clear. National leadership is the only way to support market accountability, workplace equity, and fair competition by reining in non-disclosure agreements.

Plan of Action

Multiple policy interventions could curtail NDA misuse. Select options are presented below.

Better Enforce Existing Laws

Existing laws restrict some of the harmful uses of NDAs. But laws must be enforced to be effective. Research shows that some employers include unlawful non-compete clauses in their employment contracts, capitalizing on workers' ignorance of the law and fears of being sued. Employers may similarly use NDAs in ways that violate existing law.^{33,34} Ensuring that employers are following laws that protect certain disclosures and forms of communication is a common-sense place to start when it comes to curbing NDA abuse.³⁵

The Federal Trade Commission (FTC) has an important role to play in enforcement. The FTC has broad authority to punish companies engaging in unfair or deceptive practices that harm consumers or competitors, as NDA misuse often does.³⁶ Unfair practices include practices that offend public policy as established by state statutes and common law,³⁷ which already restrict use of overly broad NDAs as well as NDA misuse to silence disclosures of employer wrongdoing. Stronger enforcement by the FTC would give these laws some needed teeth and would help establish norms governing responsible NDA use. The FTC could also work with companies to develop standards and best practices around NDA use and to encourage companies to engage in robust self-regulation and police one another.³⁸

Stronger enforcement of existing laws could also come from the various federal agencies that help oversee labor and employment in the United States. For example, the National Labor Relations Act protects workers who make common cause in seeking to discuss the terms and conditions of their employment. Regional offices of the National Labor Relations Board (NLRB) have the authority to investigate employers' use of policies that discourage this type of communication, and to file unfair labor practice charges against employers acting unlawfully. The NLRB can and should exercise this authority more forcefully. Similarly, the Occupational Safety and Health Administration (OSHA) could better use its power to enforce whistleblower laws protecting employees who report unlawful behavior. The Equal Employment

³³ Starr, E.; Prescott, J.J.; Bishara, N. (2020). The behavioral effects of (unenforceable) contracts. *The Journal of Law, Economics, and Organization*, 36(3): 633–687.

³⁴ Prescott, J.J.; Starr, E. (2021). Subjective Beliefs about Contract Enforceability. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3873638.

³⁵ Drange, M. (2021). Apple told the SEC it doesn't silence employees regarding workplace harassment or discrimination. New whistleblower documents show that isn't true. *Business Insider*, November 23.

³⁶ Section 5 of the Federal Trade Commission Act.

³⁷ *Federal Trade Commission v. Wyndham Worldwide Corporation*, 799 F.3d 236 (3d Cir. 2015).

³⁸ Majoras, D.P. (2005). Self-Regulatory Organizations and the FTC.