THE UNIVERSITY OF WARWICK

June 2013 No.137

Pre-Reformation Roots of the Protestant Ethic

Thomas Barnebeck Andersen, Jeanet Bentzen,
Carl-Johan Dalgaard and Paul Sharp

WORKING PAPER SERIES

Centre for Competitive Advantage in the Global Economy

Department of Economics

CAGE
Competitive Advantage in the Global Economy

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

Pre-Reformation Roots of the Protestant Ethic*

Thomas Barnebeck Andersen† Jeanet Bentzen‡ Carl-Johan Dalgaard

Paul Sharp

June 4, 2013

Abstract

We hypothesize that cultural appreciation of hard work and thrift, the Protestant ethic according to Max Weber, had a pre-Reformation origin. The proximate source of these values was, according to the proposed theory, the Catholic Order of Cistercians. In support, we first document an impact from the Order on growth within the epicenter of the industrial revolution; English counties that were more exposed to Cistercian monasteries experienced faster productivity growth from the 13th century onwards. Consistent with a cultural influence, this impact is also found after the monasteries were dissolved in the 1530s. Second, we find that the values emphasized by Weber are relatively more pervasive in European regions where Cistercian monasteries were located historically, and that the legacy of the Cistercians can be detected in present-day employment rates across European sub-regions.

Keywords: Cultural values; Protestant ethic; Economic development

JEL Classification codes: N13; O11; Z12

*This paper was previously circulated under the title "Religious Orders and Growth through Cultural Change in Pre-Industrial England". We thank Daron Acemoglu, Philippe Aghion, Quamrul Ashraf, Sascha Becker, Steve Broadberry, Roland Benabou, Francesco Caselli, Matteo Cervellati, Matthias Doepke, Raquel Fernández, Oded Galor, Peter Sandholt Jensen, Andros Kourtellos, Thomas Markussen, Stelios Michalopoulos, Omer Moav, Kalle Moene, Ola Olsson, Enrico Perotti, Karl Gunnar Persson, Christoffer Sonne-Smith, Holger Strulik, Uwe Sunde, Joachim Voth, Fabrizio Zilibotti as well as seminar participants at the AEA meeting in Denver, the CEPR/CAGE conference on unified growth theory at Warwick, the conference on "Cultural Change and Economic Growth" in Münich, on "Intergenerational Transmission of Entrepreneurship, Occupations and Cultural Traits in the process of long-run economic growth" in Naples, DEGIT XV in Frankfurt, Gothenburg University, University of London Royal Holloway, the Macro and Income distribution meeting at the NBER Summer Institute, University of Oslo, University of Cyprus, and University of Zürich for comments and suggestions. Errors are ours.

†Andersen and Sharp are affiliated with the Department of Business and Economics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M. Contact: Thomas Barnebeck Andersen (barnebeck@sam.sdu.dk) and Paul Sharp (pauls@sam.sdu.dk).

‡Bentzen and Dalgaard are affiliated with the Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, DK-1353 Copenhagen, Denmark. Contact: Jeanet Bentzen (jeanet.bentzen@econ.ku.dk) and Carl-Johan Dalgaard (carl.johan.dalgaard@econ.ku.dk).

1 Introduction

In what is surely one of the most famous works in all of social science, Max Weber (1905) argued that the Protestant Reformation was instrumental in facilitating the rise of capitalism in Western Europe. More specifically, Weber argued that Protestantism, in contrast to Catholicism, commends the virtues of hard work and thrift. These values, which Weber famously referred to as the "Protestant ethic", laid the foundation for the eventual rise of modern capitalism. Despite its prominence Weber's hypothesis nevertheless remains controversial.

The central hypothesis advanced in the present study is that the cultural virtues emphasized by Weber had a pre-Reformation origin in the Order of Cistercians, a Catholic order which spread across Europe as of the 11th century, and that this monastic order served to stimulate growth during the second millennium by encouraging cultural change in local populations. That is, we argue that the Cistercians encouraged growth by instigating the kind of cultural change that Weber attributed to Protestantism.

The Cistercian order, a Benedictine offshoot, was established in France in 1098 as a reformist movement with the aim of returning to the literal observance of the "Rule of St. Benedict". They rejected the developments the Benedictines had undergone and tried to reproduce life exactly as it had been in St. Benedict's time; in fact, they often ventured beyond it in austerity. The salient feature in the reform was a return to hard manual labor and the restraint from consumption (Kieser 1987). This meant that within the walls of the Cistercian monasteries one would find cultural values similar to those which, promulgated by the Protestant Reformation centuries later, is thought to have assisted the rise of capitalism outside the monastic walls. Several scholars have noted that the simplicity of the Order's lifestyle and their pursuit of wealth were in fact early manifestations of "the Protestant ethic" (e.g., Baumol 1990, p. 906; Collins 1986, p. 54; Kieser 1987, p. 116). Weber (1958, p. 118-119) himself singled out the Cistercians as encompassing values with a clear antecedent to the Protestant ethic.

Accordingly, we hypothesize that cultural values associated with the Protestant movement started to spread long before Martin Luther posted his theses on the door of the All Saints' Church in Wittenberg. Of course, the cultural influence from the Cistercians was not immediate. Initially, the Cistercians may only have "convinced" a (potentially very) small group of people to "adopt" their attitudes towards hard work and thrift. But prior to the fertility transition, in an era where Malthusian forces are at play, work ethic and thrift translates into economic success and thus reproductive success. To the extent that cultural values carry over from parent to offspring, a cumulative process of growth through cultural change can be envisioned. If the pervasiveness of "Protestant-type" cultural values increases, this will stimulate work effort, investments and technological change; in turn, this works to encourage population growth and, as a consequence of selection,

cultural change.[1]

We construct a simple model that illustrates this cumulative process. To fix ideas, we focus on how Cistercians may have influenced the attitude towards hard work and thereby macroeconomic development. Using the model, we demonstrate that an initially small group of dynasties featuring a relatively strong preference for work effort could plausibly have come to dominate the population within the span of 500 years. Moreover, we show that small differences in the initial rate of "conversion" to a high work ethic could result in considerable variations in cultural values in the course of centuries. Finally, we derive an estimable equation from the model.

As a historical testing ground for the theory we use cross-county data from England, where the Cistercians arrived early in the 12th century. England is of particular interest as it centuries later turned out to be the epicenter of the industrial revolution. Moreover, an advantage of examining England is that high quality *regional* population data is available from the 13th century onwards, which we employ as our measure of productivity in keeping with the predictions of the theoretical model.[2] In order to proxy the initial cultural influence from the Cistercians on local populations, we employ information on the historic location of English Cistercian abbeys at the county level. With this data in hand, we proceed to document that the intensity of Cistercian presence left an important imprint on comparative development across English counties until 1801; that is, long after the Dissolution of the Monasteries, which took place between 1536 and 1540.[3] Specifically, we show that, conditional on relevant exogenous controls, English counties with more Cistercian monasteries experienced faster population growth during the period 1377-1801.[4]

Correlations should be interpreted with care. We cannot rule out that some omitted factor is driving the link between Cistercian presence and long-run population growth. But we do examine a rich set of potential confounders. In addition, we provide IV estimates of the Cistercian/population growth nexus, where we draw on the work of historians to produce a plausible instrument for the location of Cistercian monasteries in England. The IV estimates corroborate our OLS findings that the Cistercians had an impact on growth, even after the monasteries were dissolved. Hence, the weight of the evidence suggests a causal effect running from Cistercian presence to long-run comparative development in England.

We believe the most plausible interpretation of this finding is that the Cistercians influenced local cultural

[1] The fundamental influence of parents on children in terms of transmitting cultural values is well established; see Bisin and Verdier (2000, 2001) and Dohmen et al. (2011). Observe, however, that one may well imagine values gradually spreading across dynasties, which would work to speed up the process of cultural change (see Dohmen et al. 2011). For evidence on the relevance of Malthusian dynamics during pre-industrial times, see Ashraf and Galor (2011).

[2] See e.g. Ashraf and Galor (2011; 2013) for a similar empirical strategy in a Malthusian setting.

[3] During 1536-1540 England went through her own version of the Protestant Reformation, which entailed the dissolution of all monasteries.

[4] By 1377 most of the Cistercians were settled; only a few additional monasteries emerged after that year. Hence, by selecting 1377 we can treat Cistercian presence as pre-determined. 1801 is chosen to permit the longest possible window of observation while at the same time ending before the fertility transition in England occurs. After the fertility transition population growth is no longer a sensible marker of productivity growth.

values, which subsequently took hold in the population. These new values in turn stimulated growth through attendant changes in work effort, investment behavior and technological progress. If indeed values changed, as hypothesized in the present study, it would provide a reasonable explanation for why the influence of the Cistercians appears to extend itself beyond the Dissolution of the Monasteries.

In order to test this account further, we turn to the present day where it is possible to obtain information on cultural values. While it is possible to study comparative cultural differences across England we do not only follow this track. Instead, we broaden the scope of the analysis to Europe in its entirety. The key advantages of this approach are that it enables us both to expand the number of observations substantially and to examine the influence of the Cistercians on sub-samples of individuals that are Catholic today. The latter check is useful in that current values may well have been influenced by the Reformation as well as by the Cistercians, which could prevent a clean test of the proposed hypothesis if the Reformation served to spread "Protestant ethics" across Protestant Europe at large, thereby muting the early influence from the Cistercians on cultural *differences* across individuals living in Protestant regions. So if the hypothesized data pattern fails to materialize in Catholic sub-samples this cannot be dismissed as the result of a confounding influence from the Reformation and thus constitutes a good opportunity to reject the hypothesis. However, we find in fact that the historical presence of the Cistercians predicts contemporary values regarding the importance of "hard work" and (to a far lesser extent) thrift across European Catholics. These results carry over if we study English citizens, or individuals living in Western Europe more broadly, as befits the hypothesis. We also find that the Cistercians appear to have left a long-run legacy on contemporary employment rates across European sub-regions, consistent with a productivity enhancing effect of the Order in the presence of a mobile labor force.

To be sure, it is impossible to establish definitively that our results with regard to contemporary or historical economic outcomes are solely attributable to a cultural impact of the Order. For instance, the Cistercians were highly innovative and fostered early industrial developments, as explained below. If the pace of technology diffusion was sufficiently slow across English counties during the second millennium, this may also have influenced growth beyond the period where Cistercians were active in England, and it may also account for the impact on employment that we detect across European sub-regions. Yet our analysis of the nexus between the Cistercians and values today makes probable that cultural change very likely is *part* of the story.

The present research is related to the literature which examines the influence from religious values on economic activity (e.g., Landes 1999; Barro and McCleary 2003; Guiso et al. 2006; McCleary and Barro 2006; Becker and Woessmann 2009; Cantoni 2009). Whereas most studies explore the "Weberian transmission mechanism", Landes (1999) and Becker and Woessmann (2009) propose that the Protestant Reformation led to a higher appreciation of literacy due to the new religious dogma, which required Protestants to be

able to read the Bible in their own language. While Landes also admits an important role for the Weber mechanism, Becker and Woessmann (2009) find little evidence of an influence from what Weber called "the Protestant ethic" for comparative development across Prussia.

More broadly, our theory is related to studies that propose that changes in the composition of the population affect long-run development in fundamental ways; whether such changes were cultural (e.g., Clark 2007; Doepke and Zilibotti 2008) or of a genetic nature (Galor and Moav 2002; Ashraf and Galor, 2013). We differ from these contributions in emphasizing a *shock* to cultural values, viz. the settlement of the Cistercians. This allows us to test our argument statistically.[5]

The rest of the paper is organized as follows: Section 2 provides background on the Order of the Cistercians and develops a model that shows how Cistercian values could spread beyond the Order itself thereby influencing productivity and population growth in a Malthusian environment. Section 3 contains our empirical analysis of historical England, where we demonstrate that the Cistercians appear to have influenced productivity growth in a manner consistent with the proposed hypothesis. Section 4 then takes a step further and examines whether the legacy of the Cistercians can be detected across contemporary Europe. In this regard, we examine both whether the Cistercians appear to have influenced the pervasiveness of values hitherto regarded as being of Protestant origin, i.e., thrift and "hard work", as well as whether the historical location of Cistercian monasteries predicts contemporary economic outcomes. Finally, Section 5 concludes the paper.

2 Theory: The Cistercians and Why they Mattered

2.1 The Order of Cistercians

The Cistercian order was founded in 1098 in France; the first Cistercian monastery in England was founded in 1128 (Cooke 1893; Donkin 1963). During the 12th century the Order spread rapidly across England, cf. Figure 1. By the end of the 14th century the expansion of the Order had essentially ceased. Hence from the perspective of our regression analysis below, which mainly involves the time period from 1377 onwards, we can treat Cistercian settlements as predetermined.

There is little doubt that the Cistercians held beliefs which were later to be associated with the Protestant ethic. By seeking to return to a literal interpretation of the Rule of St. Benedict, the small book written in the sixth century by its namesake, they stressed the trinity of prayer, work and study, as well as the values of practicality, adaptability, simplicity and moderation (Hill 1968, p. 3). The Exordium Cistercii, written

[5] Hence, in this latter respect our work is related to Nunn and Wantchekon (2011), who tests whether historical slave trade - an external shock from the point of view of the individual - has had a lasting impact on cultural values across Africans in ways of reduced trust.

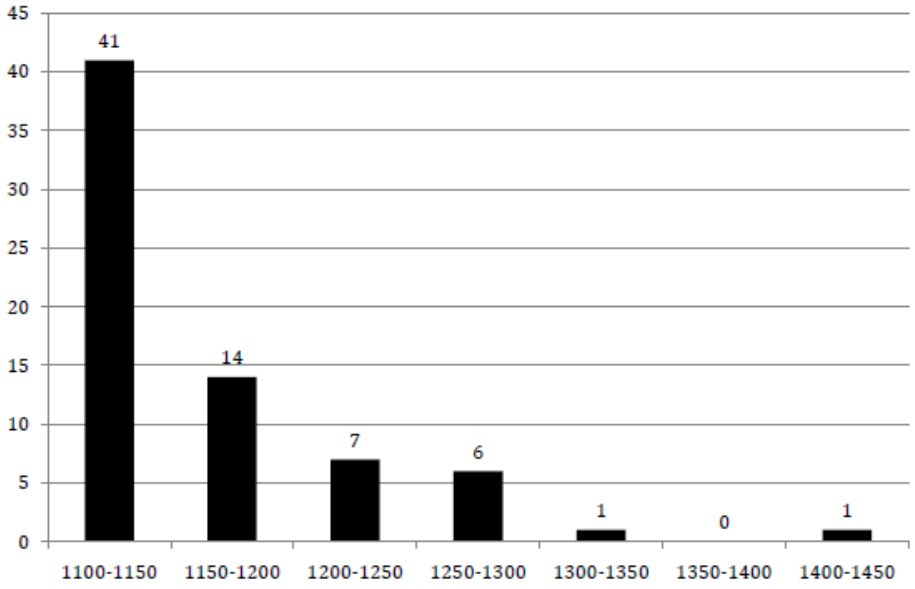

Figure 1: Frequency of founding years of Cistercian monasteries in England.

in the 1120s, and the statutes promulgated at the general chapter of 1134, stated that the monks were to work hard and live "*from the labour of their own hands, from cultivation and from their flocks*". They were also to live frugally, and were not permitted to have any possessions "*contrary to monastic purity*" such as parish churches, the tithes of other men's labor, dependent peasants, mills, ovens, or other income sources attached to the land. Hence it is no surprise that Baumol (1990, p. 906) suggests that the monks of the Order of Cistercians may have embodied an earlier "Protestant ethic": "*Puritanical, at least in the earlier years, in their self-proclaimed adherence to simplicity in personal lifestyle while engaged in dedicated pursuit of wealth, they may perhaps represent an early manifestation of elements of 'the Protestant ethic'*". Collins (1986, p. 54) is slightly more direct when he notes that the Cistercians: "*had the Protestant ethic without Protestantism*".[6]

The simplicity of the Cistercians was thus only a liturgical simplicity, replacing long days of ritual with short prayers that could be said in pauses from labor (Bouchard 1991; Hill 1968). Moreover, "useless" labor, such as painting pictures, decorating books, breeding useless animals, etc. was banned (Kieser 1987). Some have suggested that they were attempting to reduce the need for manual labor in order to leave more time for prayer (Bloch 1935; Gimpel 1976; Ovitt 1986; Landes 1999). Whatever the case, from the very beginning the Cistercians were involved in the rapidly developing economic practices of the 12th century, and were in

[6] Kiefer (1987, p. 116) makes the same observation.

some cases initiators of these practices. Moreover, the monks' asceticism, by keeping down consumption, drove up levels of investment (Kiefer 1987; Baumol 1990).

Kaelber (1998) points out that Weber himself saw monastic asceticism as a clear precursor to ascetic Protestantism, the key driving force behind European capitalism according to Weber. More specifically, as argued by Weber (1958, p. 118-19): "*In the rules of St. Benedict, even more so in the case of the monks of Cluny and the Cistercians...[Christian asceticism] has become a systematically developed method of rational life conduct, with the goal to overcome the status naturae, to free man from the power of irrational impulses and his dependence on the world and on nature...It attempted to subject man under the supremacy of purposive will, to bring his action under constant self-control with a careful consideration of their ethical consequences. Thus it trained the monk, objectively, as a worker in the service of the Kingdom of God, and thereby further, subjectively, assured the salvation of his soul...[T]he end of this asceticism was to be able to lead an alert, intelligent life: the most urgent task the destruction of spontaneous, impulsive enjoyment, the most important means was to bring order into the conduct of its adherents. All these important points are emphasized in the rules of Catholic monasticism as strongly as in the principles of conduct of the Calvinists.*" Hence the idea that the Cistercians held values close to those promulgated by the Protestant Reformation has a long and distinguished tradition.[7]

The emphasis on hard work and thrift made the Cistercians entrepreneurial and ultimately very successful economically (Baumol 1990). They contributed much as agriculturists and as horse and cattle breeders. Their major contribution was the introduction of the grange system, whereby land was held in compact blocks, in contrast to the usual fragmented and unenclosed village holdings (Donkin 1963). Another contribution seems to have been advanced irrigation techniques, thus predating Rowland Vaughan's famous popularization of these methods by centuries.[8] Moreover, their high level of agricultural technology was matched by their industrial technology. Every monastery had a model factory, often as large as the church, with waterpower to drive the machinery (Gimpel 1976). This power was used for crushing wheat, sieving flour, fulling cloth and tanning (Baumol 1990). The Cistercians are also known to have been skilled metallurgists (Gimpel 1976).

The Cistercian monastic system was one based on the principle of kinship, and thus Cistercian work practices and technology seem to have spread easily from house to house (Donkin 1978). These values in turn spread into the local area partly due to the Cistercian practice of incorporating illiterate peasant lay brothers (known as conversi) for agricultural labor (Berman 2000). Lay brothers were bound by vows of chastity and obedience to their abbot, but were otherwise permitted to follow a less demanding form of

[7] As Weber points out, similar values were found among the Cluniacs. The impact of the Cluny order has received scant attention in the literature in comparison with the Cistercians. Yet, as we shall see, they do not seem to have left a mark on pre-industrial growth in England.

[8] Vaughan's Golden Valley was actually located in an area where the Cistercians had held extensive estates prior to the Dissolution (Cook, Stearne and Williamson 2003).

Cistercian life. Work on Cistercian granges was also carried out by various classes of secular laborers. These included servi (servants), mercenarii (hired laborers), familiares (workers with intermediate status between hired workmen and lay brothers) and donate or oblate (pious laymen exchanging work for support). The exact fraction of lay brothers to these other types of labor is difficult to determine, but the latter seem to have become increasingly important at the turn of the 13th century (Noell 2006). Another important group of settlers in the abbeys were the corrodians, who spent their years of retirement there. Moreover, settled communities, including shopkeepers, formed outside the monasteries (Williams 1970). In this manner, the ways of the Cistercians spread beyond the Order itself; by power of demonstration, by word of mouth, or both.

If indeed the Cistercians influenced the values of local populations it is easy to envision how the process would become cumulative. Up until the fertility transition, which occurs in England around 1880 (e.g., Hatton and Martin 2010), households with greater earnings capabilities would proliferate at a greater rate (e.g., Clark 2007; Ashraf and Galor 2011). As a consequence, families valuing hard work and thrift should be expected to have more offspring. Provided cultural values are transmitted from generation to generation (e.g., Dohmen et al. 2011), the share of the population featuring the new values would gradually rise. As the fraction of the population with greater earnings capabilities increased there would be a positive feedback to overall population growth. In this manner, the initial cultural influence from the Cistercians would eventually have a macroeconomic impact on population density.

2.2 A Model of Growth through Cultural Change

In order to think more formally about how the ways of the Cistercians spread beyond the Order itself, and the ensuing macroeconomic impact, we construct a growth model designed to elucidate the long-run consequences of cultural change with respect to work ethics in a Malthusian environment. The model can be seen as a slight extension of the basic Malthusian framework (Ashraf and Galor 2011), to a situation where labor supply is endogenous. This extension allows us to have a simple representation of greater work ethics: reduced utility value of leisure. As a result, we can study the process of cultural change that is unleashed by a work-inducing preference shock to a subset of the population. Eventually the shock leads to greater aggregate population density, as a consequence of selection of individuals with a greater appreciation of work. By focusing on the impact of changes in the attitude to work we suppress changes in cultural attitudes towards saving and investment; i.e., thrift. It is worth observing, however, that while the model focuses on the work ethic of individuals, similar results would arise if we instead examined thrift. As long as thrift implies a greater earnings potential, groups with high thriftiness will be selected in the Malthusian

setting thereby propelling aggregate population density.[9]

2.2.1 Individuals' optimization

We are considering a closed OLG economy with one good. Time is discrete and extends to infinity, $t = 0, 1, 2...\infty$. Individuals live for two periods. During the first period of life individuals are children and live off their parents' consumption. During their second period individuals decide on labor supply and use the proceeds to rear offspring and for consumption. For simplicity, individuals are assumed to reproduce asexually; in the model there is therefore no distinction between a household and an individual.

Individuals derive utility from consumption, c, off-spring, n, and leisure, $1 - e$. The utility function, $u(c_t, n_t, e_t)$, is quasi-linear:

$$u(c_t, n_t, e_t) = \log(c_t) + \delta \log(n_t) + \eta(1 - e_t), \tag{1}$$

where δ and η are positive parameters.[10] The individual's budget constraint is

$$e_t y_t = c_t + n_t \tau, \tag{2}$$

where y is potential income and τ is the output cost of a child. Optimal fertility and effort is given by[11]

$$n_t = (\delta/\tau\eta) y_t, \tag{3}$$

$$e_t = e = (1 + \delta)/\eta. \tag{4}$$

Hence, if the taste for leisure, η, declines, labor supply expands, and fertility increases. The latter is caused by the fact that as income goes up some of the gain is used for consumption and some is converted into larger families.

2.2.2 Production

Output in the economy, Y, is produced using technology, A, labor, L, and land, X: $Y_t = AL_t^{\alpha} X^{1-\alpha}$. Following Galor and Weil (2000) we assume that individuals' income is given by the average product of labor

[9] Becker (1980) explores a dynamic economy where agents differ in terms of the rate of time preference; i.e., in terms of "thrift". In the long run the most patient dynasty ends up "owning" the economy. Below we demonstrate a similar result in that dynasties with greater work ethics will end up dominating the population.

[10] The quasi-linearity of the utility function ensures time invariance of e_t, and thus that increases in income are converted into larger families on a one-to-one basis. Accordingly, these preferences allow us to lay out the logic of the argument in a particularly simple and transparent way.

[11] In the interest of brevity we suppress the solution for consumption in the text. It is straightforward to solve for optimal c by inserting the solutions for n and e into the budget constraint.

so that potential income of the individual is $y_t \equiv Y_t/L_t = A\left(X/L_t\right)^{1-\alpha}$.[12]

2.2.3 Macro Dynamics: Cultural homogeneity

The size of population at $t+1$ is given by the number of individuals at time t multiplied by their number of offspring, $L_{t+1} = n_t L_t$. Inserting the solution for optimal family size, n, and the expression for potential income, we obtain

$$L_{t+1} = \left(\delta/\tau\eta\right) A L_t^{\alpha} X^{1-\alpha} \equiv F\left(L_t\right), \ L_0 \text{ given.}$$

It is straightforward to demonstrate that the model allows for a unique and globally stable steady state, L^*, with L^* thus fulfilling $L_{t+1} = L_t = L^*$ and $L^* = F\left(L^*\right)$. In the steady state population size is given by

$$L^* = \left(\frac{\delta A}{\tau\eta}\right)^{1/(1-\alpha)} X. \tag{5}$$

In the steady state we obtain the standard comparative statics with respect to δ, τ, A, and X (see Ashraf and Galor 2011). In addition, the present model contains the prediction that if preferences for leisure declines (i.e., η declines) population density rises in the long run; greater preference for labor supply allows for greater fertility and thus a larger population in the long run.

2.2.4 Macro Dynamics: Cultural heterogeneity

In order to study cultural change in the population at large in the present setting we need a shock to ignite the process. Accordingly, we assume that an arbitrarily small subset of society develops a greater taste for work. That is, a preference alteration emerges such that for a group of citizens η declines to $\tilde{\eta} < \eta$. The source of the change is left unspecified but it could be thought of as resulting from religious persuasion, which we think of as an exogenous shock from the point of view of the individual.

The economy subsequently nests two types of individuals; people with relatively low valuation of leisure, $\tilde{L}$, and the rest, $\hat{L}_t$. Otherwise the two groups are identical. We assume preferences are passed on from one generation to the next ensuring the initial one-off shock persists.[13]

After the shock to preferences the relative size of the two groups will evolve in accordance with[14]

$$\frac{\hat{L}_{t+1}}{\tilde{L}_{t+1}} = \frac{\tilde{\eta}\hat{L}_t}{\eta\tilde{L}_t}.$$

If we define $z_t \equiv \hat{L}_t/\tilde{L}_t$ the above equation can be viewed as a linear first order difference equation in z_t,

[12] Since individuals only supply e units of time, the actual income of an individual is $y = Ae^{\alpha}\left(X/L\right)^{\alpha-1}$.
[13] See Bisin and Verdier (2000, 2001) and Dohmen et al. (2011) for evidence on the "inheritability" of values.
[14] For each group, $L_{t+1} = \left(\delta/\eta\tau\right) y_t L_t$ holds. The equation in the text emerges by dividing them while recalling only η differs.

which can be solved so as to yield

$$z_t = z_0 \left(\frac{\tilde{\eta}}{\eta}\right)^t.$$

Obviously, $\lim_{t\to\infty} z_t = 0$ since $\tilde{\eta} < \eta$. By implication, the share of the total population, which belongs to the high effort group, evolves in accordance with

$$\frac{\tilde{L}_t}{L_t} \equiv \frac{\tilde{L}_t}{\tilde{L}_t + \hat{L}_t} = \frac{1}{1 + z_0 \left(\frac{\tilde{\eta}}{\eta}\right)^t}. \tag{6}$$

As is readily seen, $\lim_{t\to\infty} \frac{\tilde{L}_t}{L_t} = 1$. Since one group has a reproductive advantage over the other, the former will eventually dominate the population in its totality.

Turning to aggregate dynamics, the total population evolves in accordance with:

$$L_{t+1} = (\delta/\tau\tilde{\eta})\, y_t \tilde{L}_t + (\delta/\tau\eta)\, y_t \left(L_t - \tilde{L}_t\right),$$

which can be rewritten, using equation (6), as the following law of motion for population:

$$L_{t+1} = \theta_t A L_t^{\alpha} X^{1-\alpha} \equiv G\left(L_t, t\right), \; L_0 \text{ given}, \tag{7}$$

where

$$\theta_t \equiv \frac{\delta}{\tau} \frac{(1/\tilde{\eta}) - (1/\eta)}{1 + z_0 \left(\frac{\tilde{\eta}}{\eta}\right)^t} + \frac{\delta}{\tau\eta}, z_0 \text{ given}.$$

Hence, after the shock the law of motion is affected by the time autonomous factor θ_t, which reflects the influence from cultural change. As the fraction of the population with greater preference for work increases, θ shifts upwards over time. Asymptotically, when $z_t \to 0$, the law of motion will only reflect the preferences of the $\tilde{L}$ type.

The phase diagram for the economy, which is depicted in Figure 2, illustrates the adjustment dynamics after a small group in society changes values at time $t = 0$. Since the new group works harder, its income is greater. This works to increase population density relative to the initial situation where all individuals held the same preference for leisure, η. However, initially the high work ethic group may be very small, for which reason the *immediate* impact on aggregate population size could be miniscule. But since the hard working group holds a reproductive advantage, the group's population share gradually rises over time, thereby increasingly stimulating aggregate population size. The rise in importance of the new cultural values is reflected in the upward shifts in the law of motion for population size depicted in Figure 2. Eventually, the group with greater work ethics will dominate the population, and the economy convergences to a steady

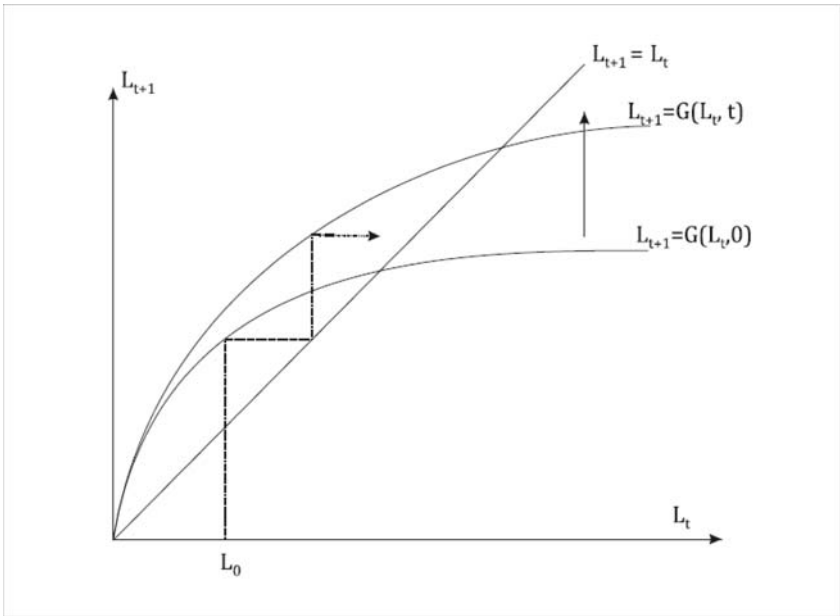

Figure 2: Phase diagram for the model with cultural heterogeneity.

state where population size reflects the preferences of the high work ethic group. In the steady state

$$L^* = \left(\frac{\delta A}{\tau \tilde{\eta}}\right)^{1/(1-\alpha)} X.$$

Since $\tilde{\eta} < \eta$ it follows that the impact of the cultural change has been to elevate population density (cf. equation (5)).

The model thus shows how a change in a certain cultural attitude in a small subset of the population may rise in importance due to selective pressures and eventually influence the macroeconomy. The source of the change of preferences is left unexplained by the model. But it seems plausible that the Cistercians have influenced county populations in this manner, as argued in Section 2.1. Accordingly, our hypothesis is that Cistercians planted the seeds of change by affecting the cultural attitudes; or, more appropriately, the work ethic of a (in principle arbitrarily) small part of the county population. By so doing, they instigated a process of growth through cultural change.

2.2.5 Speed of Diffusion

A question of some relevance is how fast the cultural diffusion process played out if it only emanates from differential population growth rates across dynasties with different values. Naturally, the process would conceivably occur at a faster rate than what we find below if values gradually diffuse *across* dynasties as

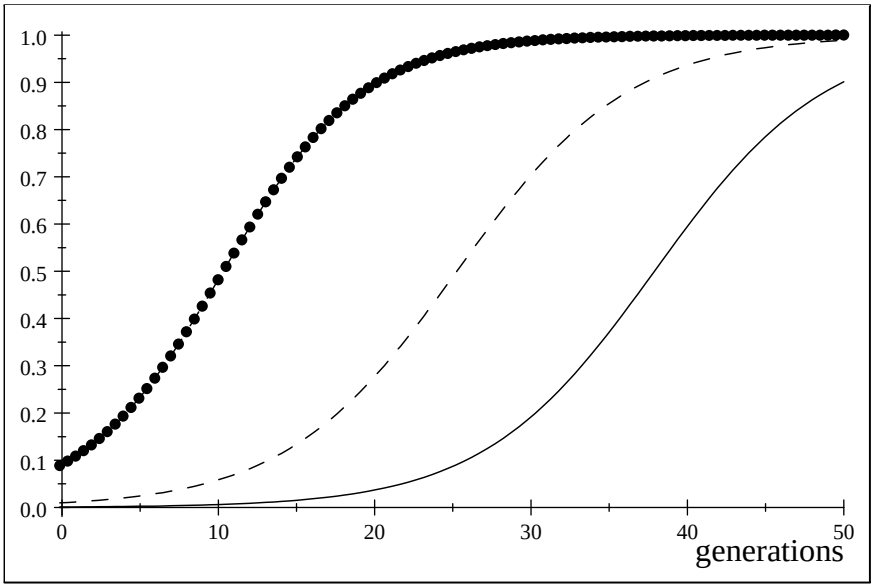

Figure 3: The rise of new cultural values in the population. *Assumptions*: (a) 20% higher work effort among individuals with "high" work effort". (b) Initial "infection" rate: 1/1000 (solid black), 1/100 (dashed), 1/10 (dotted).

well, following the initial shock to a select group of dynasties. In practice, both mechanisms may have been at work.

In order to examine the speed of population-growth driven cultural change, observe that the fraction of individuals with high work ethic at time t is

$$\pi_t \equiv \frac{\tilde{L}_t}{L_t} = \frac{1}{1 + z_0 \left(\frac{\tilde{\eta}}{\eta}\right)^t}. \tag{8}$$

The speed at which $\tilde{\eta}$ becomes dominant in society depends on how much more effort the high work ethic group exerts, $\frac{\tilde{\eta}}{\eta} = \frac{\hat{e}}{\bar{e}}$, as well as how many individuals were "persuaded" to change their values as of time $t = 0$. The ratio of η's is hard to pin down in any precise manner. But suppose the high work-ethic group works 20% more than the other group.[15] In this case Figure 3 shows how the new cultural values grow in significance over time for different assumptions about the initial degree of cultural change; that is, $\pi_0 = 0.1\%$, 1% and 10% of the population, respectively.

[15] Clark and Van der Werf (1998) estimate that the number of days worked per year (standard deviation in parenthesis) rose in England from 266 (4.8) in 1560-99 to 280 (12.9) in 1771. Suppose this increase is attributable to the rise of the Protestant work ethic, resulting from the Cistercian presence and the Reformation. Then the estimated increase over time in work days provides a crude guesstimate for the ratio $\tilde{\eta}/\eta$. Factoring in the statistical uncertainty we may note that working days in 1771 may have been between 5% lower and 23% higher than in 1560, with a mean around +10%. Hence, assuming a 20% higher work effort may not be unrealistic.

The spread of the new cultural values follows an S-shaped trajectory: the process is slow to begin with but accelerates over time and ultimately levels off. Consider the curve in the middle, associated with an initial degree of cultural change of 1%. The first 10 generations only raise the fraction with strong work ethic modestly (to about 6%), the next 10 generations increase the share to 30% of the population, and another five to nearly 50%. If a generation is about 20 years, 25 generations (what it takes to go from 1% to 50%) is about 500 years. The point is that, within the window of observation available to us (about 500 years), it is possible for a small (initial) cultural shock from the Cistercians to accumulate into a major aggregate impact on the composition of the population solely by way of selective pressure.

Another point worth emphasizing is the implied comparative differences in cultural values that seemingly small initial differences translates into. With an initial infection rate, π_0, of one percent, 50 percent of the population holds a high work ethic after 25 generations; but only eight percent have high work ethic after 25 generations if the initial infection rate is 1/10th of a percentage point. This implies that, by affecting π_0, variations in the intensity of Cistercian presence may have generated substantial comparative differences in cultural values across English counties over the period in question. It may therefore be possible to detect the legacy of the Cistercians on population dynamics over the period 1377-1801, which we examine below.

3 The Impact of the Cistercians on Productivity across Historical England

This section proceeds in a series of steps. We begin by deriving an empirical model based on the theoretical model from the previous section. Subsequently, in Section 3.2 we present our data, Section 3.3 contains our OLS regression results, whereas Section 3.4 reports our IV results.

3.1 Empirical Specification

The theory predicts that changes in population can be written (taking logs in equation (7)):

$$\log\left(L_{t+1}\right) - \log\left(L_t\right) = \left(\alpha - 1\right)\log\left(L_t/X\right) + \log\left(A\right) + \log\left(\theta_t\right),$$

where we may write:

$$\theta_t \equiv \frac{\delta}{\tau}\left[\left(1/\tilde{\eta}\right) - \left(1/\eta\right)\right]\pi_t + \frac{\delta}{\tau\eta},$$

with π_t, recall, denoting the proportion of individuals with high work ethics in the population. If we linearize $\log\left(\theta_t\right)$ around $\pi_t = 0$ we obtain $\log\theta_t \approx \log\left(\frac{\delta}{\tau\eta}\right) + \eta\left[\left(1/\tilde{\eta}\right) - \left(1/\eta\right)\right]\pi_t$, which can be reinserted into the

equation for population growth so as to yield

$$\log(L_{t+1}) - \log(L_t) \approx (\alpha - 1)\log(L_t/X) + \log(A) + \log\left(\frac{\delta}{\tau\eta}\right) + \eta\left[(1/\tilde{\eta}) - (1/\eta)\right]\pi_t.$$

Finally, denoting a county by i and adding an error term, we arrive at the empirical model

$$\Delta\log(L_{it+1}) = a_0 + a_1\log(L_{it}/X_i) + a_2\pi_{it} + \mathbf{Z}_i'\mathbf{a} + \epsilon_i,$$

where $\Delta\log(L_{it+1}) \equiv \log(L_{t+1}) - \log(L_t)$ and $\mathbf{Z}_i$ contains time-invariant controls for productivity (A).

Naturally, we do not have data on π_{it}. But, according to the theory, we may proxy it using some measure of Cistercian presence in the county, as it should influence π_{i0}, and thereby π_{it} (see equation (8)). We define this intensity as the *Cistercian presence relative to other moral influences*. Since the Church was the principal authority in matters of moral in medieval times, we construct π as the ratio of Cistercian monasteries, M_c, to all religious houses; i.e. $\pi = M_c/M$. However, the counterfactual we are interested in is that of changing the *composition* of moral influences while at the same time holding constant its *level*. This dictates that we also control for the total number of religious houses, M, separately.

Hence, our main empirical specification reads as follows:

$$\Delta\log(L_{it+1}) = a_0 + a_1\log(L_{it}/X_i) + a_2\left(M_c/M\right)_i + a_3M_i + \mathbf{Z}_i'\mathbf{a} + \varepsilon_i. \tag{9}$$

Ceteris paribus, areas with more Cistercians saw a larger fraction of the population initially being "persuaded" by the Cistercian work ethic. As seen from Figure 3 this should imply a higher π_{it} at any given point in time, stimulating growth according to the model. As a result, we expect a_2 to come out with a positive sign. In addition, theory predicts that $a_1 < 0$, capturing convergence effects. The coefficient a_3 is a priori indeterminate.

There are clearly other ways in which one could introduce the influence from the Cistercians into the empirical model, aside from the choice made in specifying equation (9). For instance, Cistercian monasteries could enter linearly (i.e., $a_2M_c + a_3M$), or one could introduce a dummy variable (present/not present) in order to gauge their impact. We explore such alternatives below, with little effect on the results.

Before we turn to a description of our data two remarks on the testing strategy are warranted. First, when examining the proposed hypothesis we are studying the period 1377-1801. More specifically, we have county-level data on population density at several points in time: in 1377 (right after the Cistercians had completed their settlement in England); 1600 (shortly after the Dissolution of the Monasteries); and in 1801. It is obviously important that this period, in its entirety, is a period during which English population growth is likely to be a sensible marker of productivity growth. We believe this is a plausible assumption as England

did not go through the fertility transition until around 1880 (e.g., Hatton and Martin 2010). To be clear, the fact that the industrial revolution (may have) occurred earlier in England is immaterial to the present empirical analysis, as long as the productivity gains it brought about resulted in faster population growth, which it should have done until the onset of the fertility transition.

Second, we have made no mention of migration in the discussion above. Yet productivity gains in one county could plausibly attract immigrants from lagging counties. This is observationally equivalent to population growth arising from higher fertility. While we cannot distinguish between these two alternatives, a positive influence from the Cistercians on population growth will in any case testify to a productivity enhancing effect from this particular religious order.[16]

3.2 Data

3.2.1 Population density

Our dependent variable is population density. We obtained data on population density for the year 1377 from Campbell (2008). Campbell also provides the area of the counties; we transformed them from square miles into square kilometers. The distribution of the population in 1377 is based on 1.38 million adult males and females who contributed to the poll tax of 1377.[17] The level of the population is based on an estimate by Campbell (2000) of a total population of 4 million. Campbell only reports population numbers for the aggregate of London and Middlesex, not for the two counties separately. In order to match the data, all data on all variables is aggregated in this way. Yet we end up excluding London and Middlesex in all regressions, since it is an outlier. We note for completeness, however, that including London and Middlesex makes no difference to our results. We also have data on population density in 1600, which is taken from Broadberry et al. (2010). Finally, population density in 1801 is from Wrigley (2007). The latter data are based on registered marriages, which were more completely recorded than baptisms and burials on which previous population estimates were based (Rickman, 1802).[18]

3.2.2 Religious Houses

In controlling for Cistercian presence, as well as of other religious orders, we rely on the English Monastic Archive (EMA), which has been constructed by researchers at University College London. The database

[16] In order to distinguish between the two cases we would need county level data on income per capita, which does not exist. If productivity induces migration, income per capita should increase as a result from Cistercian presence; the same is not true if fertility is the driver. See Ashraf and Galor (2011) for further discussion and tests on cross-country data.

[17] These numbers are available in Dobson (1983).

[18] Campbell (2008) also reports population data for 1290 based on taxable wealth. But since about 10% of Cistercian settlement occurred around that time, the risk of reverse causality tainting our estimations would be enhanced if we used 1290 as our initial year. As a result we stick with 1377 as the initial date in our main analysis. However, in the IV setting we do explore the consequence of extending the period of observation to 1290-1801.

involves 776 religious houses in England, which date from the 10th to the 16th century. We gathered these data into one dataset, which we then used to calculate the number of religious houses in each county (*religious houses*) and the number of Cistercian monasteries as a share of total religious houses in each county (*Cistercian share*). We also construct the share of Benedictine monasteries (*Benedictine share*), as well as shares associated with Augustinians, Cluny, and the Premonstratensian Monastic Orders.

Taken together these monastic orders accounted for 3/4 of the 776 religious houses in England as recorded by the EMA. The Benedictine Order was the largest, accounting for roughly 30% of all religious houses. The Cistercians accounted for circa 10% or about the same as the Premonstratensians and Cluniacs taken together. Finally the Augustinians accounted for about 26% of the total religious houses.

We made one correction to the data with respect to the city of York, which was listed in EMA as a county. York was (is) a walled city situated in North Yorkshire. To be able to match the data with the data on population density, we re-coded it as part of the county of North Yorkshire.

3.2.3 Other controls

In order to control for potential steady state determinants of productivity, we control for a range of plausible correlates with "A". Specifically, we control for *land quality* in order to capture productivity in a predominantly agrarian society such as England during most of the period in question. In addition, we control for *access to ocean*, as well as to *rivers*. As *coal* arguably played a key role in the industrialization process (Allen 2009; Pomeranz 2000), occurring at the end of our observation window, we also include a control for physical availability of this resource across counties. Moreover, we employ a full set of *regional fixed effects* to control for unobserved heterogeneity. Details on this data are found in the Appendix.

Table 1 provides summary statistics and a correlation matrix of key variables discussed above**.**

[Table1 about here]

As a prelude for things to come, it is worth observing from Table 1 that M_c/M (*Cistercian share*) is negatively correlated with population density in 1377, yet positively correlated with population density in 1801; both correlations are significant at a ten percent level of confidence (p-values of 0.08 and 0.09, respectively). In the middle of the period, in 1600, the correlation is essentially nil. As explained in Section 3.4, the Cistercians had a preference for locating in sparsely populated areas, which likely explains the negative correlation in 1377. And yet, the correlation changes markedly during the ensuing roughly 500 years, consistent with a productivity enhancing influence from the Order beyond the Dissolution of the Monasteries in the 1530s. One may also observe that a similar time-varying correlation is not found between population density and any other religious order.

Figure 4 provides a complementary perspective. The figure shows the evolution of average population

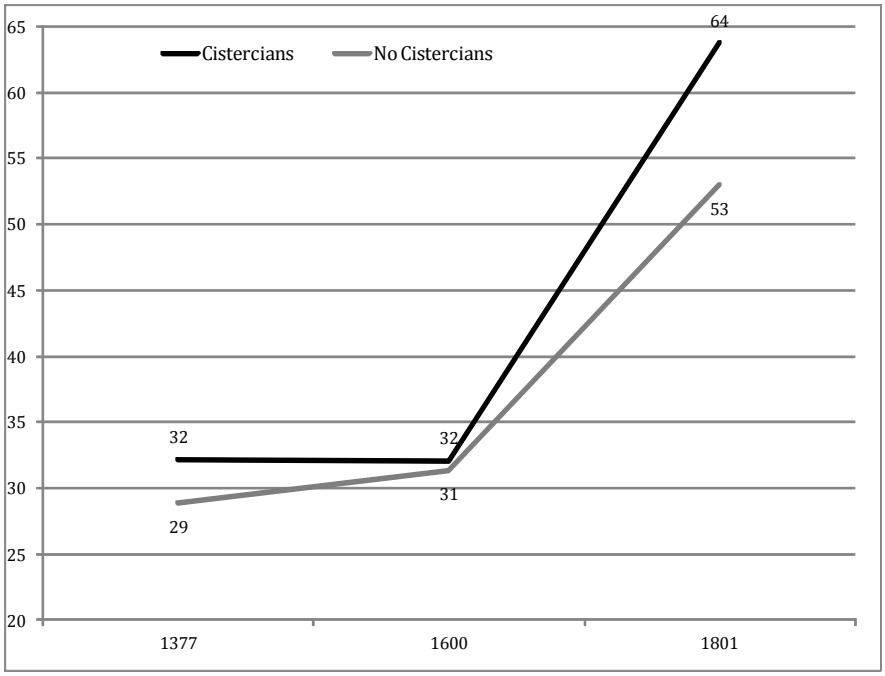

Figure 4: The figure compares average population density (persons/km2) in areas where at least one Cistercian monastery was found to areas without Cistercian monasteries.

density in areas that were "treated" by Cistercians (i.e., areas that hosted at least one Cistercian monastery) and those that were not.[19]

It is interesting to note that prior to the Dissolution of the Monasteries, there seems to be little difference between areas "treated" by Cistercians and those that were not (if anything, population growth is faster in the latter areas). However, after the Dissolution divergence seems to take place. This is consistent with (though, obviously, not proof of) a gradually accelerating impact from cultural change, as predicted by the model. At the same time, these patterns seem more difficult to square with a hypothesis according to which the Cistercians impacted growth via technological change. In this case one would surely expect to see an impact on population dynamics while the monasteries were physically in place. Still, many factors might simultaneously impact on cross-county population growth. Hence, in the remaining we resort to regression analysis in order to elicit information about the partial impact from the Order of Cistercians.

[19] A total of eight counties were left "untreated": Berkshire, Cambridgeshire, Cornwall, Derbyshire, Durham, Hertfordshire, Rutland and Westmorland.

3.3 OLS Results

Table 2 reports our baseline model. In all columns of the table we control for initial population density, the total number of religious houses, the share of all religious houses which are Cistercian, and the most obvious productivity control: agricultural land quality. The regression in column 1 shows that these variables collectively hold significant explanatory power with respect to population growth over the period 1377-1801. The baseline model explains almost two thirds of the variation in the dependent variable.

In column 2 we add the regional fixed effects. The null hypothesis that the regional dummies are jointly zero cannot be rejected at conventional levels (p-value = 0.49). This means that, conditional on our baseline controls, regional effects appear unimportant. Columns 3-6 add controls one by one; column 7 estimates the full model, which involves our baseline controls along with the additional confounders that were found to be statistically significant in columns 3 to 6. In this latter specification we can account for about three quarters of the variation in population growth from 1377-1801.

Several features of the results are noteworthy. First, the share of Cistercians stays statistically significant in all columns. This means that the composition of religious houses seems to matter, with a larger share of Cistercians being associated with greater population growth. In addition, Cistercian point estimates are fairly stable, always situated in the interval [1.7, 2.0]. Second, while initial population density displays the expected conditional convergence feature, it is surprising that agricultural land quality has a negative impact on population growth. A simple explanation might be that population density in 1377 is measured with error, and that land quality partly serves to correct for it. This is certainly possible as land quality is correlated with population density in 1377, cf. Table 1b. Thus the point estimate might simply reflect convergence: places with greater initial density (good soil conditions) would be expected to grow at a slower rate. Third, land area adds significant explanatory power, whereas the physical infrastructure of rivers does not seem to matter for population growth, perhaps suggesting that neither irrigation nor water-based transportation were significant binding constraints to growth. Fourth, our coal control variable is significant. This is consistent with Allen (2009) and Pomeranz (2000), who both argue that proximity to coal production was critical for British industrialization because it supplied an inexhaustible supply of cheap energy, which may have stimulated growth.

[Table 2 about here]

Figure 5 provides a visual depiction of the relationship between the share of Cistercians and the growth rate in population as estimated by column 7, Table 2. The partial correlation does not appear to be driven by outliers.

As explained in Section 3.1, the information we wish to elicit pertains to changing the composition of religious houses, holding constant its level. Other variations of the theme are obviously possible beyond the

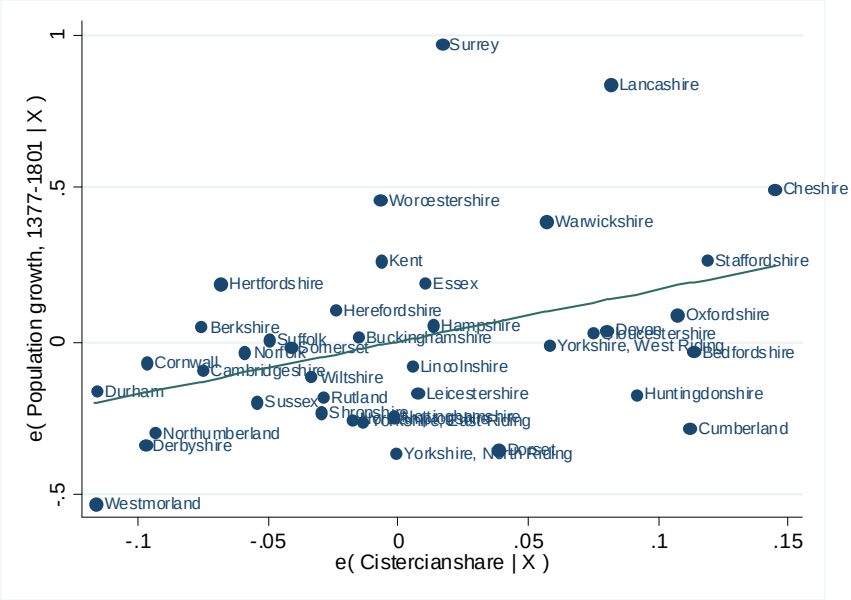

Figure 5: The partial correlation between Cistercians and population growth, 1377-1801. Notes: The plot illustrates the regression results reported in Table 2, column 7.

introduction of M_c/M. For instance, one option would be to use the total number of Cistercian monasteries, in which case the partial effect is assumed to be constant. Another option would be to simply use a dummy variable indicating Cistercian presence. Yet another variation we will consider is the number of Cistercian houses per unit area. This variable will become useful in Section 4 when we examine the impact of the Cistercians across Europe, an area where we lack data on total religious houses.

Table 3 reports the results using these alternative indicators of Cistercian influence. Columns 1 to 4 show that in the baseline model little is changed in terms of statistical significance when using the different Cistercian variables. However, using our preferred variable, *cistercianshare*, leads to a higher R^2 in the baseline model. The same is true for the augmented model, as immediately seen upon inspecting columns 5 to 8. In the full model, however, the dummy variable is a borderline statistically insignificant predictor of population density growth (p-value of 0.102).

[Table 3 about here]

While the Cistercians (however measured) appear to be positively correlated with population growth 1377-1801, one may reasonably wonder if this apparent influence is unique to this particular monastic order. In an effort to learn the answer we re-ran the regressions in Table 2, replacing the Order of Cistercians with the share of the total number of monasteries that were *Benedictine*; the order from which the Cistercians originated. In stark contrast to the Cistercians, the Benedictine order does not seem to be associated with

faster population growth over the period. The point estimate is close to zero, sometimes even featuring the "wrong" sign, and it is always statistically insignificant. The details of the results can be found in Appendix Table A1, which documents that the Order of Cistercians convey different information about population growth than the order from which they originated. In Appendix Table A2 we further examine whether the Cistercians hold a differential impact on population growth compared to other religious orders. In these exercises we include the Cistercians alongside the Augustinians, the Cluniacs, the Premonstratensians as well as the Benedictines. We once again find that only the Cistercians appear to be correlated with population growth over the period.

Accordingly, a positive correlation between the intensity of Cistercian presence and population growth 1377-1801 appears reasonably robust, and statistically different from the link between population growth and any other major religious order present at the time. Yet a legitimate concern is whether the positive correlation reflects a causal influence from the Cistercians or perhaps simply the influence from omitted factors. We address this concern next.

3.4 Endogenous Location of Cistercian Monasteries: IV Results

An objection to the preceding results is that they could be spurious. That is, perhaps the Cistercians simply chose to locate in areas with a pronounced productive potential?

Based on the historical evidence, however, this seems unlikely. The Order had a stated preference for situating their monasteries in remote, even devastated locations (Cooke 1893; Donkin 1963). Indeed, it has long been accepted by scholars that the Cistercians acted as transformers of wastelands into fertile farms, as mirrored in the poet Wordsworth's *Cistercian Monastery.*[20] The fact that Cistercian presence is negatively correlated with initial population density (see Table 1b) provides some formal corroboration of these assessments. Nevertheless, in order to check we further scrutinize the Cistercian/population growth nexus using instrumental variables estimation.

The Cistercians had a particular preference for locating in secluded and sparsely populated areas, as explained above. At the time of arrival the most secluded areas may well have been the forests owned by the Crown: royal forests.[21] As Donkin (1963, p. 184) observes: "*..there is a really significant connection with the Royal Forests; one-third of all the English [Cistercian] houses lay at first within or very near their bounds [...]. In these areas there was a good deal of land of low value for endowments; nonroyal landowners were gravely hampered by the forest laws; and, as elsewhere, prospective founders undoubtedly responded to the willingness of the early generations of monks to exploit rough, undeveloped country.*" Thus, there may well

[20] "Where'er they rise, the sylvan waste retires, And aery harvests crown the fertile lea."

[21] The concept of a royal forest was introduced in England by the Normans in the 11th century. They were protected areas of land (not necessarily woodland) where the king had privileged hunting rights under the "forest law", which offered strict penalties to anyone using these areas for hunting or farming.

have been a double coincidence of wants. Nonroyal landowners, wanting to save their souls, had an interest in allowing Cistercians to settle at or near royal forests, which were of limited value beyond the occasional hunt with the monarch. At the same time, this location satisfied the ascetic needs of the Cistercian settlers. Finally, the monarch may also have had an incentive to encourage the practise. Madden (1963) notes that the king likely granted rights of pasture over wide tracts of the royal lands and forests because the Cistercians were willing to pay for this service using revenue from sale of wool; wool which derived from sheep using the royal lands for grazing. Hence the presence of a royal forest in a county could be a potentially viable instrument for Cistercian settlements.

We obtained data on the location of royal forests in the 13th century from Bazeley (1921). Based on the maps constructed by Bazeley, we constructed a dummy variable: *Rforest*, which is equal to one if a royal forest were to be found in the county in the 13th century. Accordingly, we expect to find a positive partial effect of royal forest on the intensity of Cistercian settlements.

One potential problem with the use of *Rforest* as an instrument for the intensity of Cistercian presence is that it might capture resource growth. The royal forest system was at its height in the late 12th and early 13th century. But already in 1215 Magna Carta laid down limits to the power of the monarchy in the forests, and the Great Perambulation of 1300 reduced the scale of the forests. Hence, counties with royal forests may have experienced growth in agricultural land area, as the importance of royal forests receded.[22]

To alleviate this cause for concern we add a new control variable, based on Bazelay's map, which measures the size of the county area that was covered by royal forest in the 13th century as a share of the total county area: *forest share*. Needless to say, places with greater forest area should be places where the scope for growth in land area is greater once the royal forests start to recede. Thus, adding *forest share* to the control set should make the excludability of *Rforest* in the second stage plausible.

However, another potential problem is that if *Rforest* predicts the location of the Cistercians, it might also predict the location of *other* religious houses. This too might jeopardize identification. Of course, our OLS results give us no particular reason to expect that, say, the Benedictine order influenced growth. But since the OLS results could be biased these findings are hardly definitive. As a result, one might legitimately worry about identification if *Rforest* predicts *other* religious orders.

In order to gauge the likely severity of this concern we ran regressions that correspond to the first stage results reported below, but exchanging Cistercians for other religious houses: Augustinians, Premonstratensians, Cluniacs and Benedictines. As seen from Appendix Table A3, *Rforest* is never a significant predictor of the intensity (or presence) of any of these orders. Hence, the exclusion restriction is unlikely to be violated

[22] We have admittedly been unable to find examples of historical writings hypothesizing that land expansion, prompted by deforestation, had an important impact on population growth. Still, it does seem to be the case that forest areas receded particularly markedly from the 16th century onwards (e.g., Young 1978). In this light it would appear reasonable to regard expansion of agricultural land expansion as a potential problem for identification.

on account of *Rforest* picking up the influence from other religious orders.[23]

Our main IV results are reported in Table 4. We focus on two basic specifications: our baseline specification and the "full specification" (cf. column 7, Table 2). Cistercian influence is measured in two different ways; as the *Cistercian share* and as the *Cistercian presence,* a binary indicator of whether Cistercians are present or not in the county. In this manner we can assess the impact of Cistercians both along the extensive margin (Cistercian presence) and the intensive margin (Cistercian share), where in the latter case we use the measure featuring the highest R^2 in the OLS setting. Moreover, we look at two periods. The full period from 1377-1801 (columns 1-4) and the post-dissolution period from 1600-1801 (columns 5-8). The motivation for the period split is to gauge whether the Cistercians have influenced the growth process beyond the period where they were physically present.

[Table 4 about here]

Several observations are worthwhile. *Rforest* is a strong instrument for whether Cistercians were *present* in a county or not, cf. the Kleibergen-Paap statistic. In some cases when the *Cistercian share* indicator is used the instrument turns weak. But, as established via the Anderson-Rubin test, the Cistercians do appear to exert a causal impact on population growth both before and after the dissolution of the Monasteries, and this being regardless of the way their influence is actually measured. To gauge the robustness of the IV results above we have examined the consequences of widening the window of observation to the period 1290-1801; i.e. by nearly an additional century, which allows us to take an initial year before the plague hit England. The results, which are reported in Appendix Table A4, are very similar to those reported in Table 4.

As can be seen upon comparing the results in Tables 2-4, IV estimates exceed OLS counterparts by roughly a factor of two. One possible explanation is that OLS suffers from a negative bias due to omitted confounders. Our preferred interpretation, however, is attenuation bias resulting from the obvious fact that our indicators of Cistercian presence are imperfect indicators of the fraction of the population with "Protestant ethics".

The impact of the Cistercians appears economically significant. To see this, observe from Table 4 that countries with at least one Cistercian monastery had about 0.5 log points faster growth in population density between 1377 and 1801 compared to counties that were "untreated". This is equivalent to an acceleration in average annual population growth of about 1/10th of a percentage point. During the period in question, average cross-county population growth was 0.16 percent per year (with a standard deviation of 0.13 percent

[23] For completeness we also considered the possibility that *total* religious houses is an endogenous regressor. As demonstrated in Table A5, our instrument - Royal Forest - is not correlated with total religious houses, conditional on our choice of the controls. Table A6 reports the second stage results (Table A5 contain first stage results), where total religious houses are *omitted* from the control set. Evidently, with a full set of controls, Cistercian presence remains a significant determinant of population growth, regardless of the choice of period (1290-; 1377-; 1600-).

per year). Hence the impact of the Cistercians is clearly economically significant.

In sum, we believe a strong case can be made that the Cistercian order had a causal impact on population growth in England during the pre-demographic transition period, consistent with an impact on productivity. The effect is found both before and after the Dissolution of the Monasteries and it appears to be economically significant.

4 The Legacy of the Cistercians: Values and Economic Outcomes Across Europe

The above analysis makes probable that Cistercian monasteries left a lingering impact on county-level productivity in England. Yet so far we have not narrowed down the mechanism. It could be that the Cistercians simply managed to provide some areas with a technological lead, which was expanded after the Dissolution of the Monasteries. To be sure, this is a viable candidate explanation which could potentially motivate the results above, with little or no mentioning of cultural change. If indeed the Cistercians had an impact on cultural values, and in light of the likely persistence of cultural values, perhaps we can detect a lingering impact on present-day cultural values and economic outcomes influenced by the selfsame cultural values? This is the question to which we now turn.

4.1 Cistercians and Contemporary Values

In order to examine the potential influence of the Cistercians on values we estimate the following individual-level regression:

$$v_{i,s,c} = a_0 + a_1 \left(M_c/X\right)_{s,c} + a_2 X_{s,c} + \mathbf{b}'\mathbf{W}_{i,s,c} + \boldsymbol{\gamma}'\mathbf{W}_{s,c} + c_c + \varepsilon_{i,s,c}, \tag{10}$$

where $v_{i,s,c}$ refers to cultural values (work ethic or thrift, respectively) of individual i who is residing in sub-region s of country c. As described below, we can observe the location of individuals in Europe, which is partitioned into so-called NUTS2 sub-regions (s). We observe Cistercian presence at the NUTS2 level, and we measure it chiefly as Cistercian density, $\left(M_c/X\right)_{s,c}$, since we do not have data on all religious houses across Europe; other variations are also employed, however, as explained below. $\mathbf{W}_{i,s,c}$ is a vector of individual-level co-variates: age, age squared, sex of the respondent, educational attainment, marital status, and the religious denomination of the respondent. $\mathbf{W}_{s,c}$ captures geographical information at the NUTS2 level, latitude and longitude, and c_c represents a country fixed effect. Finally, $\varepsilon_{i,s,c}$ is a noise term. The parameter of interest is a_1, which represents the link between the intensity of the Cistercian historical presence and individual-level values with regards to thrift and hard work. We estimate equation (10) as a logistic regression model. In

order to check the robustness of our results we further examine the Cistercians/values link by aggregating the responses to the NUTS2 level. In these regressions the left hand side variable becomes the fraction of respondents that find hard work and thrift to be important values to pass on to their children. We return to these results below.

We measure Protestant values according to whether the individual believes hard work and thrift, respectively, are important traits for children to learn at home; this is similar to McCleary and Barro (2006), who examine Weber's hypothesis at the country level. The data derives from the European Values survey (EVS), and the latest wave (2008-10) provides information about the location of individuals at the NUTS2 level. More specifically, we have information about where the respondent lived when he or she was 14 years old. We code the individual as belonging to this particular region. The rationale is that values are predominantly formed during childhood. In total we have access to data for 56,227 individuals (See Appendix for further details).

In order to measure the Cistercian influence we employ data on the location of European Cistercian monasteries from Donkin (1978). Donkin's map is reproduced in Figure 6. Using GIS software and a shapefile of European NUTS regions, we construct a variable measuring the number of Cistercian monasteries per NUTS2 region across Europe. To make sure that we use only the regions included in Donkin's map, we restrict our sample to regions with a centroid between longitudes -10 and 26 and latitudes 37 to 63. This leaves us with a total of 32,641 respondents from the EVS that we were able to match with the data on Cistercian monasteries. To capture Cistercian density we divide the number of Cistercian monasteries by the size of the geographical area of the NUTS2 region. We exclude three NUTS2 regions from the analysis throughout as our analysis revealed they constitute outliers: Inner London, Outer London, and Brabant Wallon in Belgium**.** If individuals from these areas are included in the analysis the link between the cultural values and Cistercian presence is *strengthened*. With these sample restrictions we are left with a sample of 32,358 individuals who grew up in 242 different NUTS2 regions across Europe.

Table 5 reports the results from estimating equation (10). The first four columns focus on hard work, whereas the last four examine thrift. In columns 1 and 5 we examine a baseline specification without individual-level controls, whereas they are added in columns 2 and 6. In the remaining columns we examine the interaction with Protestantism in two different ways: In columns 3 and 7 we control for the religious denomination of the respondent, and in columns 4 and 8 we only consider respondents that report they are Catholic.

[Table 5 about here]

In all specifications greater Cistercian density appears to elevate the likelihood that the respondent values hard work. It is perhaps revealing to observe that once we only examine Catholics, the influence from the Cistercians increases considerably. This is consistent with the idea that the Reformation led to the diffusion

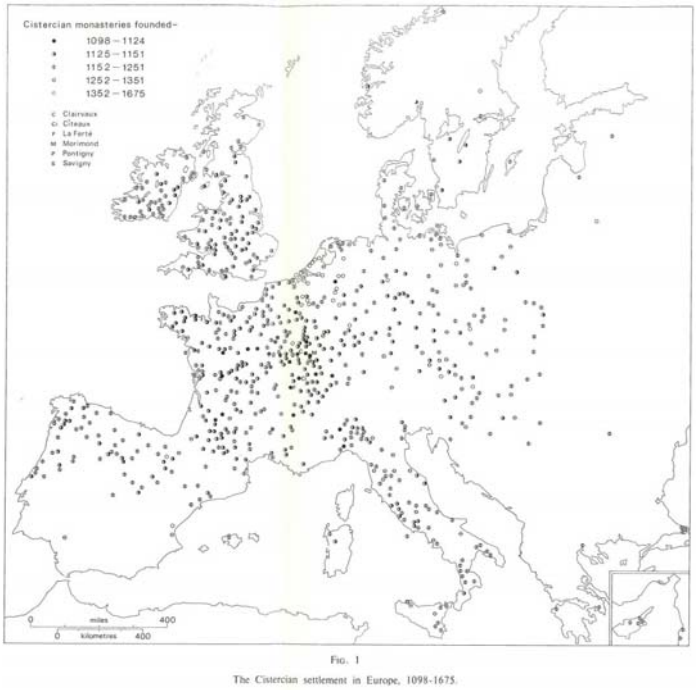

Figure 6: Map of the historic location of Cistercian monasteries across Europe. *Source*: Donkin (1978).

of ideas similar to those promulgated by the Cistercians, thereby serving to mute the historical influence from the latter on contemporary outcomes in Protestant regions. In contrast, however, we find little evidence of an impact of the Cistercians with regards to thrift; in all settings the correlation (albeit positive, as expected) is insignificant.

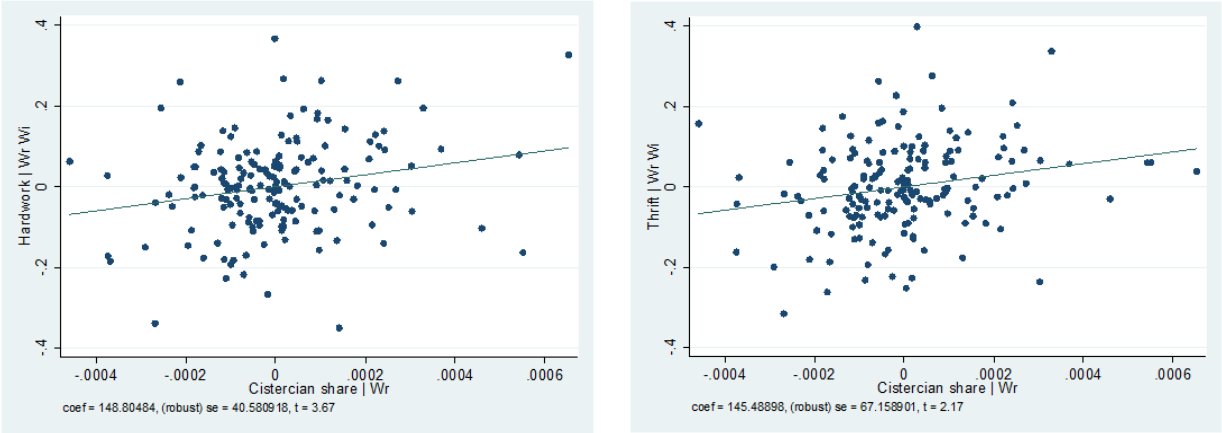

Figure 7: Cistercian density versus values across NUTS2 regions of Europe. The left figure shows the link to hard work and the right figure shows the link to thrift. Variation in hard work and thrift caused by the individual-level baseline

controls ($\mathbf{W}_i$) is first removed in the individual sample, then aggregated to the regional level, and then regressed on Cistercian share and the remaining regional-level controls. The partial correlations correspond to columns 4 and 8 in Appendix Table A7.

In Table 6 we attempt to gauge the robustness of the link between the Cistercians and contemporary values in two ways. First, we measure the influence from the Cistercians in several alternative ways in addition to Cistercian density. We include total Cistercian monasteries linearly, log transformed, and as a 0/1 indicator for whether the individual NUTS2 sub-region hosted at least one Cistercian monastery. Second, we vary the sample. That is, we first look at all European regions, then we focus on Catholics, and finally we restrict the sample to England. The latter, of course, in an effort to see if the Cistercians indeed left a cultural imprint on the English population, consistent with our results from Section 3.

[Table 6 about here]

As can be seen from Table 6, the link between Cistercian presence and hard work appears rather robust. It is generally found no matter how we measure the Cistercian influence, and in all sub-samples. In particular, in the non-Protestant sample every indicator carries significant explanatory power. These results are consistent with a cultural impact of the Cistercians in terms of work ethics across Europe in general as well as within England specifically. The latter results further support our interpretation of the significant impact from the Cistercians on population growth in England that we documented above.

The results with regards to thrift are less strong, and only in one case do we obtain a significant partial correlation between Cistercian historical locations and thrift.

Things change, however, when we aggregate to the NUTS2 level and examine the determinants of the *fraction* of respondents at the NUTS2 level valuing hard work and thrift (results are reported in Appendix Table A7). For hard work we find results that are qualitatively similar to those pertaining to the individual-level: Areas with greater Cistercian density harbor a larger fraction of respondents who find hard work to be a value worth passing on to their children. For thrift the results are now somewhat stronger. The partial correlation between M_c/M and the fraction of respondents who find thrift to be a value worth passing on to their children is positive and significant at conventional levels in all but one column. Figure 7 depicts the partial correlation between Cistercian density and hard work and thrift, respectively. It is clear that no particular region or group of regions seems to be driving the results in either case.

Overall we conclude that the Cistercians do appear to have had a lasting impact on cultural values, though most strongly with regards to hard work. Interestingly, these results appear to be economically stronger in the context of non-Protestants, consistent with the view that Protestantism also brought similar values to bear albeit considerably later in history.

4.2 The Cistercians and Contemporary Economic Outcomes

If indeed the Cistercians left a cultural imprint on the values of European citizens, it is of interest to see if their influence is also detectable on present-day economic *outcomes*. The historical analysis for England suggests the Cistercians led to higher productivity as reflected in greater population density. In a modern day context it is inappropriate to use population density as a marker for productivity, for which reason we use *employment* at the sub-national level as the key outcome variable in the context of contemporary Europe.

The logic is simple. If the Cistercians eventually led to higher productivity in some regions compared to others, labor should be attracted to the former thereby raising the ratio of employment to population size. To investigate whether this is the case we estimate the following model:

$$\log\left(E_{t,s,r}\right) = a_0 + a_1\left(M_c/X\right)_s + a_2 X_s + \mathbf{b}'\mathbf{W}_{s,r} + c_r + \varepsilon_{t,s,r}, \tag{11}$$

where $\log\left(E_{t,s,r}\right)$ is the (log of) total regional employment at time t at the NUTS2 sub-regional level s in region r. In practice we look at the year 2007 (i.e., $t = 2007$), which is the year before the financial crisis.[24] In addition to Cistercian density, $\left(M_c/X\right)_s$, we control for a set of variables that vary across NUTS2 sub-regions ($\mathbf{W}_{s,r}$): *geography*, in the sense of latitude and longitude, and *demography*, in the form of the average age of the local population. The (log) size of the population is also included in $\mathbf{W}_{s,r}$, which means that a significant coefficient for a_1, the link between Cistercian density and contemporary employment, speaks to a higher employment *rate* in sub-region s. As seen from equation (11) we also include a full set of NUTS1 fixed effects, c_r.

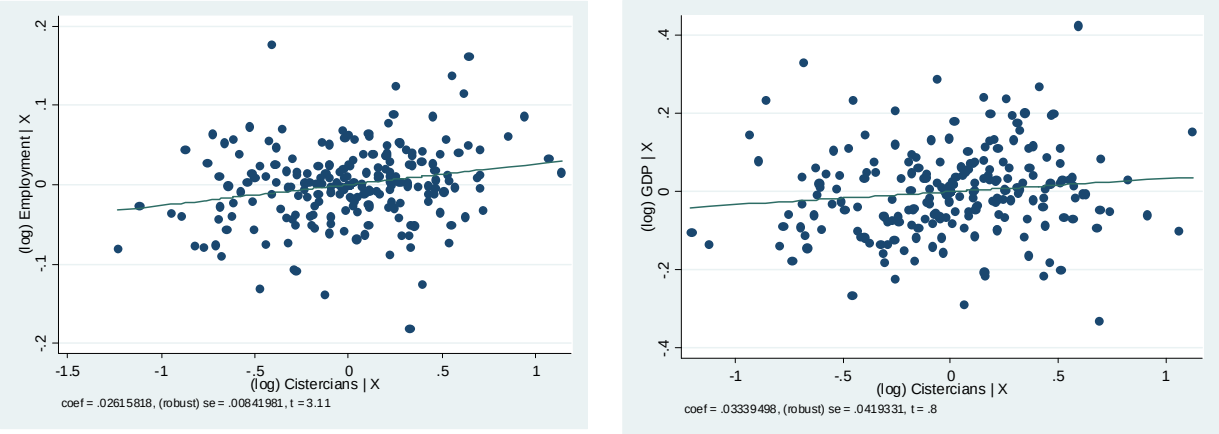

Figure 8: (log) Cistercian monasteries versus (log) employment and (log) GDP, respectively. Corresponds to columns 3 and 7 of Table 7, respectively.

[24] The results are qualitatively similar if we examine 2005, 2006, 2008, or 2009, though significance is slightly reduced in the years 2005 and 2006, presumably because of fewer observations.

The link between productivity and total employment is theoretically mediated by labor mobility. But if labor is (fully) mobile, another proposition follows. Namely that high productivity and high employment regions should not distinguish themselves in terms of *labor* productivity. Incipient labor productivity differences, and thus wage differences, is what theoretically drives mobility, leading to a reallocation of employment. In this process, labor productivity is reduced in the high productivity regions due to diminishing returns. Hence, in the absence of large frictions to labor mobility at the NUTS2 sub-regional level we would not expect to see a link between Cistercian historical influence and current *labor productivity* (i.e., GDP per employed). In order to check, we therefore also examine the impact of the Cistercians on GDP per employed. The specification mirrors equation (11) though with (log) GDP as left hand side variable, and with employment on the right hand side in place of total population. The sources for our data on employment, population and GDP at the NUTS2 level are laid out in the Appendix; summary statistics are found in Table A8.

Table 7 reports the results. In the first four columns we focus on employment (rates), whereas the last four columns concern GDP (per employed). Cistercian presence is measured in four different ways; as density, $(M_c/X)_s$; as presence (0/1); as total Cistercian monasteries, Cistercians; and as log(1 + Cistercians).

Regardless of how we measure Cistercian presence, we find a statistically significant correlation with employment rates, as seen from columns 1-4, consistent with the hypothesis under scrutiny. Economically, the effect is non-trivial: regions "treated" with Cistercian presence have on average 2.5% higher employment, cf. column 4.

Columns 5-8 turns to GDP (per employed). Consistent with high labor mobility at the sub-regional level across Europe, we find no statistically significant impact of the Cistercians as implied by the reallocation argument discussed above: High productivity sub-regions draw in labor, which serves simultaneously to elevate the employment rate and eliminate labor productivity differences.

[Table 7 about here]

Figure 8 shows the partial correlation between the Cistercians and employment and GDP, respectively, conditional on the co-variates mentioned above. Specifically, Figure 8 depicts the estimates reported in columns 3 and 7. The results do not seem to be driven by any particular regions.

These results, and those from the previous section, paint a coherent picture. When we examine the link between Cistercian historical presence and contemporary values related to work ethics in particular, we find systematically that the latter are more pervasive in places with more Cistercian monasteries. To the extent that work ethics influence productivity, one would expect to see higher employment rates in areas that were exposed to Cistercians if labor mobility is sufficiently high. This is exactly what we see in the data. These results provide strong evidence in favor of the hypothesis under examination: the Order of Cistercians instigated changes in cultural values in local populations, which worked to influence growth and

development in the long run.

5 Concluding Remarks

The present paper documents that the Order of Cistercians left a lingering imprint on long-run comparative development across English counties during the pre-industrial era. In counties with greater Cistercian presence population growth was faster during the period 1377-1801, suggesting that the Cistercians stimulated productivity. The Catholic monasteries were dissolved by 1540 in England. Hence our results suggest that the regional impact of the Cistercian order was felt more than 250 years after the Order's discontinuation in England. This finding is robust and IV estimates suggest that the correlation can be given a causal interpretation.

We have also offered a mechanism behind our finding, namely that the Cistercians ignited a process of growth through cultural change. That is, a gradual change in local populations in terms of taste for hard work and thrift, much like Max Weber suggested was the end result of the Protestant reformation.

We think this explanation is plausible for the following reasons. First, a cultural concordance between the Cistercians and the Protestants in the dimensions of work ethics and thrift has already been observed by several scholars, including Weber himself. Second, the cultural explanation has the virtue of being able to account for the lingering Cistercian influence on growth. Third, consistent with the cultural mechanism, using data from the European Values Survey we find that Catholic regions in Europe, which historically were influenced relatively more by the Cistercians, tend to have populations with greater taste for hard work and, to a lesser extent, thrift today. These results carry over to England as well as to Europe more broadly. Fourth, given free mobility in the labor market, one would expect to see sub-regions with higher productivity to also feature higher employment rates. This suggests that regions that were "treated" by Cistercian monasteries historically should feature higher employment rates today, if indeed there is a lingering impact on productivity. This is exactly what we find. Hence, the Cistercian order seems to have had a lasting impact on economic development in Europe.

Overall, our research suggests that Weber was right in stressing the importance of a cultural appreciation of hard work and thrift but quite likely wrong in tracing the origins of these values to the Protestant reformation.

A Data Appendix

A.1 Regional analysis for England

Religious Houses in England. The data on religious houses is available from: http://www.ucl.ac.uk/history2/englishmonasticarchives/religioushouses/index.php.

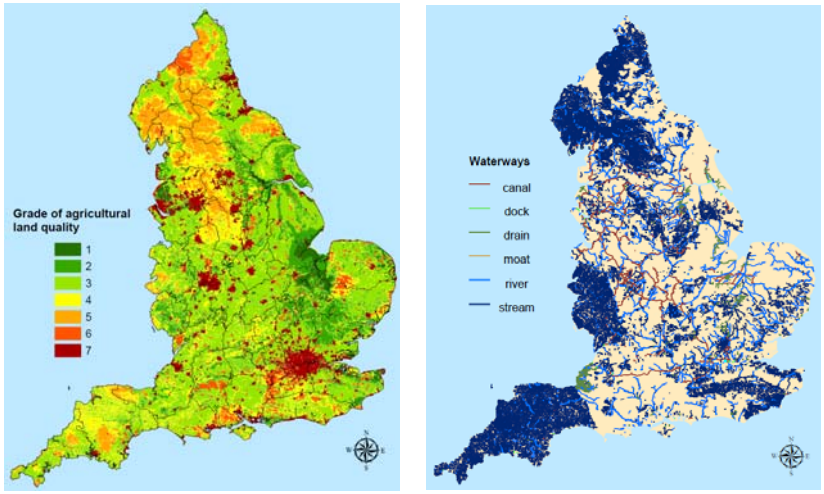

Figure A1. The figure shows the data on land quality (left panel) and on waterways (right panel) discussed below.

Land quality. Natural England provides a measure of agricultural land classified into five grades plus classifications for non-agricultural and urban land. Grade one is best quality and grade five is poorest quality, grade six is non-agricultural land and grade seven is urban. The measure is calculated by Natural England using information on climate (temperature, rainfall, aspect, exposure, frost risk), site (gradient, micro-relief, flood risk) and soil (depth, structure, texture, chemicals, stoniness). The source of the data is Raster Digital mapping with a scale of 1:250,000, which is available online at: http://www.gis.naturalengland.org.uk/pubs/gis/gis_register.asp.[25] The data was gathered with coordinate precision of 1 meter. We used these data to create a measure of agricultural land quality within each county. The earliest digital map of English counties is from 1851. These data were kindly provided to us by the University of Portsmouth and the Great Britain Historical GIS Project. Combining the shapefile including the agricultural land quality and the shapefile including English county borders, we were able to create measures of the area in a county with agricultural land of quality level 1-5, each as a share of total county area; the total county area was here calculated by summing over the land quality variable, since this variable spans the entire area. Our variable

[25] Additional Data description is also available online at: http://www.magic.gov.uk/datadoc/metadata.asp?dataset=2&x=16&y=10 and http://naturalengland.etraderstores.com/NaturalEnglandShop/product.aspx?ProductID=88ff926a-3177-4090-aecb-00e6c9030b29.

"land quality" is the combination of qualities 1 and 2.

Access to Water. The German company Geofabrik freely provides shapefiles on various geographic features.[26] Of our interest is their data on waterways in Great Britain, where waterways are divided into canal, dock, drain, moat, river, and stream. These data are available online at: http://download.geofabrik.de/osm/europe/great_britain/. As with the data on agricultural land quality, we merge the shapefile describing waterways with the shapefile describing the county borders of England. The outcome of interest from this procedure is the total length of rivers as a share of the total area in a county (*rivershare*). In addition *oceandummy* equals one if the county borders the ocean, zero otherwise.

Coal. Allen (2009) and Pomeranz (2000) argue that proximity to coal production was critical for British industrialization because it supplied an inexhaustible supply of cheap energy. We therefore construct a variable called *coalshare*, measured as the surface area of coalfields to total area in 1871.[27] The map of coalfields is taken from Redmayne (1903).

Regional fixed effects. The regional classification that we employ is based on Government Office regions: East Midlands, East of England, London, North East, North West, South East, South West, West Midlands, and Yorkshire and the Humber.

A.2 Values and Outcomes across Europe.

Values and individual level controls. The data derives from the European Values Survey, which is available online at http://www.europeanvaluesstudy.eu. We focus on the 2008-10 wave as detailed information of the place of residency at age 14 (NUTS2 level) is available. We use whether respondents indicate that they think that valuing "hard work" is an important trait for children to learn at home (variable a030 in EVS) and whether they think "thrift, saving money and things" is an important trait for children to learn at home (variable a038 in EVS). In Appendix tables we also aggregate to the NUTS2 level, which means the variable becomes the fraction of respondents (appropriately weighted) that subscribe to thrift and hard work. The EVS is also the source of the individual level controls highlighted in the text.

Cistercian presence. Derives from Donkin (1978). Shapefiles for NUTS regions were obtained from eurostat.com.

Employment, Population and GDP per employed. For the outcomes regressions, employment, population, and total GDP is measured at the NUTS2 level and provided by EuroStat (ec.europa.eu/eurostat). Employed persons is the total number of employed persons aged 15-64. Population is total persons living in the NUTS2 region. GDP is total regional gross domestic product in million PPS.

[26] These shapefiles are based on maps created by the OpenStreetMap project using data from portable GPS devices, aerial photography, other free sources, or simply from local knowledge.

[27] A coalfield is an area of certain uniform characteristics where coal is mined.

B Additional Results Appendix

[Insert Tables A1 - A8]

References

[1] Ashraf, Q. and O. Galor, 2011. Dynamics and Stagnation in the Malthusian Epoch: Theory and Evidence. *American Economic Review* 101, 2003-2041.

[2] Ashraf, Q. and O. Galor, 2013. The Out of Africa Hypothesis, Human Genetic Diversity and Comparative Development. *American Economic Review* 103, 1-46

[3] Allen, R. 2009. *The British industrial revolution in global perspective.* Cambridge University Press, Cambridge.

[4] Barro, R. and R. McCleary, 2003. Religion and Economic Growth across Countries. *American Sociological Review*, 68, 760-781

[5] Baumol, W., 1990. Entrepreneurship: Productive, Unproductive, and Destructive. *Journal of Political Economy*, 98, 893-921

[6] Bazeley, M.L., 1921. The Extent of the English Forest in the Thirteenth Century. Transactions of the Royal Historical Society, 4 , pp. 140-172.

[7] Becker, R.A., 1980. On the long-run steady state in a simple dynamic model of equilibrium with heterogeneous households. *Quarterly Journal of Economics*, 95, 375–82

[8] Becker, S. and L. Woessmann, 2009. Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *Quarterly Journal of Economics*, 124, 531-596

[9] Berman, C.H., 2000, *The Cistercian Evolution.* Penn Press

[10] Bisin, A. and T. Verdier, 2000. Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits. Quarterly Journal of Economics 115, 955-88.

[11] Bisin, A. and T. Verdier, 2001. The Economics of Cultural Transmission and the Dynamics of Preferences. *Journal of Economic Theory* 97, 298-319.

[12] Bloch, M., 1935. *Avènement et conquêtes du moulin a eau. Annales d'Histoire Èconomique et Sociale,* 7, 538–563

[13] Bouchard, C., 1991. *Holy Entrepreneurs.* Cornell University Press

[14] Broadberry S., B. Cambell and Leeuwen B., 2010. English Medieval Population: Reconciling time series and cross sectional evidence. Working Paper (Warwick)

[15] Campbell, B., 2000. *English seigniorial agriculture, 1250-1450.* Cambridge University Press

[16] Campbell, B., 2008. Benchmarking medieval economic development: England, Wales, Scotland and Ireland, c. 1290. *Economic History Review*, 61, 896-945

[17] Cantoni, D., 2009. The Economic Effects of the Protestant Reformation: Testing the Weber Hypothesis in the German Lands. Manuscript (Harvard University)

[18] Chamberlain, G. 1980. Analysis of covariance with qualitative data. *Review of Economic Studies* 47, 225-238.

[19] Clark, G., 2007. *A Farewell to Alms*. Princeton University Press

[20] Clark, G. and Y. Van Der Werf, 1998. Work in progress? the industrious revolution, *Journal of Economic History*, 58, 830–43

[21] Collins, R., 1986. *Weberian Sociological Theory*. Cambridge University Press

[22] Cook, H., Stearne, K., and T. Williamson, 2003. The origins of water meadows in England. *Agricultural History Review*, 51, 155-162

[23] Cooke, A., 1893. The Settlement of the Cistercians in England. *English Historical Review*, 8, 625-676

[24] Dobson, R., 1983. *The Peasants' Revolt*. Palgrave Macmillan

[25] Doepke, M. and F. Zilibotti, 2008. Occupational choice and the spirit of capitalism. *Quarterly Journal of Economics*, 747-93

[26] Dohmen, T., A. Falk, D. Huffman and U. Sunde, 2011. The Intergenerational Transmission of Risk and Trust Attitudes. Forthcoming: *Review of Economic Studies*.

[27] Donkin, R., 1963. The Cistercian Order in Medieval England: Some Conclusions. *Transactions and Papers (Institute of British Geographers)*, 33, 181–198

[28] Donkin, R., 1978. *The Cistercians: Studies in the geography of medieval England and Wales*. Toronto: Pontifical Institute of Medieval Studies

[29] Galor, O. and D. Weil, 2000. Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. *American Economic Review*, 806–828.

[30] Galor, O. and O. Moav, 2002. Natural Selection and the Origin of Economic Growth. *Quarterly Journal of Economics*, 117, 1133–91

[31] Gimpel, J., 1976. *The medieval machine*. Penguin

[32] Guiso, L. and P. Sapienza, L. Zingales, 2006. Does Culture Affect Economic Outcomes? *Journal of Economic Perspectives* 20, 23–48.

[33] Hatton, T. and R.M. Martin, 2010. Fertility decline and the heights of children in Britain, 1886–1938. *Explorations in Economic History* 47, 505-19.

[34] Hill, B.D., 1968. *English Cistercian monasteries and their patrons in the twelfth century.* University of Illinois Press

[35] Kaelber L., 1998. *Schools of Asceticism: Ideology and Organization in Medieval Religious Communities.* Penn State Press

[36] Kieser, A., 1987. From Asceticism to Administration of Wealth. Medieval Monasteries and the Pitfalls of Rationalization. *Organization Studies*, 8, 103-123

[37] Landes, D., 1999. The Wealth and Poverty of Nations. Abacus

[38] Madden, J.E. (1963). Business Monks, Banker Monks, Bankrupt Monks: The English Cistercians in the Thirteenth Century. *The Catholic Historical Review*, Vol. 49, No. 3, pp. 341-364

[39] McCleary, R. and R. Barro, 2006. Religion and Economy. *Journal of Economic Perspectives*, 20, 49-72

[40] Neyman, J. and S., Elizabeth L, 1948. Consistent estimates based on partially consistent observations. *Econometrica*, 1–32.

[41] Noell, B., 2006. Expectation and unrest among Cistercian lay brothers in the twelfth and thirteenth centuries. *Journal of Medieval History*, 32, 253-274.

[42] Nunn N, and L. Wantchekon, 2011. The Slave Trade and the Origins of Mistrust in Africa. *American Economic Review.* 101, 3221-3252.

[43] Ovitt Jr, G., 1986. The Cultural Context of Western Technology: Early Christian Attitudes toward Manual Labor. *Technology and Culture*, 27, 477–500

[44] Pommeranz, K. 2000. *The great divergence: China, Europe, and the making of the modern world economy.* Princeton University Press.

[45] Redmayne, R. A. S. (1903) *The Coal-Mining Industry of the United Kingdom,* in Factory and Industrial Management vol 26, 1903.

[46] Rickman, J., 1802. *Observations on the Results of the Population Act,* 41 Geo. III. Gregg.

[47] Young, C.R. (1978). 'Conservation Policies in the Royal Forests of Medieval England'. *Albion: A Quarterly Journal Concerned with British Studies*, 10, 95-103

[48] Weber, M., 1905. *The Protestant Ethic and the Spirit of Capitalism*. Allen & Unwin

[49] Weber, M., 1958. *The Protestant Ethic and the Spirit of Capitalism*. Charles Scribner's Sons

[50] Williams, D.H., 1970. *The Welsh Cistercians: Aspects of Their Economic History*. The Griffin Press

[51] Wrigley, E., 2007. English county populations in the later eighteenth century. *Economic History Review*, 60, 35-69

Table 1a: Summary statistics	**Obs**	**Mean**	**Std.**	**Min**	**Max**
Cistercian share	40	0.09	0.07	0.00	0.25
Religious houses	40	19.03	12.93	2.00	73.00
Pop dens 1377	40	31.55	11.83	8.98	52.98
Pop dens 1600	40	29.99	6.46	13.97	43.33
Pop dens 1801	40	60.45	24.82	20.92	143.77
Augustinian share	40	0.28	0.13	0.00	0.62
Benedictine share	40	0.31	0.16	0.00	0.67
Cluniac share	40	0.04	0.05	0.00	0.15
Premonstratensian share	40	0.05	0.09	0.00	0.50

Table 1b: Correlation matrix									
Cistercian share	1								
Religious houses	0.01	1							
Pop dens 1377	-0.28	0.31	1						
Pop dens 1600	-0.01	0.01	0.45	1					
Pop dens 1801	0.32	-0.16	-0.06	0.57	1				
Augustinian share	-0.10	-0.12	0.35	0.23	0.30	1			
Benedictine share	-0.15	0.04	0.01	0.06	-0.05	-0.27	1		
Cluniac share	-0.09	0.07	-0.01	0.34	0.26	0.18	-0.15	1	
Premonstratensian share	-0.18	-0.07	-0.33	-0.24	-0.14	-0.22	-0.46	0.02	1

Table 2. Baseline Results: The Cistercians and Cross-County Development

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Population growth 1377-1801						
Cistercian share	1.934**	1.977*	1.931**	1.751**	1.808**	1.921**	1.702**
	(0.887)	(1.078)	(0.916)	(0.798)	(0.670)	(0.876)	(0.644)
Pop dens 1377 (log)	-0.614***	-0.482**	-0.606***	-0.438**	-0.474***	-0.601***	-0.374**
	(0.171)	(0.235)	(0.186)	(0.173)	(0.138)	(0.192)	(0.142)
Religious houses	-0.00682*	-0.00615*	-0.00683*	-0.0170***	-0.00541	-0.00740*	-0.0124***
	(0.00376)	(0.00359)	(0.00376)	(0.00400)	(0.00377)	(0.00417)	(0.00355)
Land quality	-0.634*	-0.682*	-0.634*	-0.536*	-0.540	-0.635*	-0.486
	(0.313)	(0.350)	(0.315)	(0.297)	(0.325)	(0.319)	(0.312)
rivershare			0.267				
			(1.940)				
Area (log)				0.303**			0.203**
				(0.113)			(0.0967)
Coal					1.587***		1.391**
					(0.535)		(0.517)
Ocean						0.0279	
						(0.128)	
Regional FE	NO	YES	NO	NO	NO	NO	NO
Observations	40	40	40	40	40	40	40
R-squared	0.641	0.678	0.641	0.676	0.712	0.641	0.727

Notes. (i) Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. (ii) All regressions contain a constant.

Table 3: Alternative indicators of Cistercian influence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES	Population growth 1377-1801							
Cistercian share	1.934**				1.702**			
	(0.887)				(0.644)			
Cistercians		0.0840**				0.0711**		
		(0.0408)				(0.0296)		
Cistercian presence			0.262*				0.223	
			(0.153)				(0.133)	
Cistercian density (/km2)				198.4*				188.6**
				(106.8)				(84.62)
Area				0.000114*				-3.92e-05
				(6.48e-05)				(4.95e-05)
Pop dens 1377 (log)	-0.614***	-0.640***	-0.715***	-0.533***	-0.374**	-0.379**	-0.467***	-0.448***
	(0.171)	(0.181)	(0.188)	(0.194)	(0.142)	(0.155)	(0.161)	(0.163)
Religious houses	-0.00682*	-0.0130**	-0.00891**	-0.0188***	-0.0124***	-0.0190***	-0.0133***	-0.0138***
	(0.00376)	(0.00535)	(0.00362)	(0.00667)	(0.00355)	(0.00543)	(0.00387)	(0.00451)
Land quality	-0.634*	-0.605**	-0.573*	-0.563*	-0.486	-0.451	-0.436	-0.473
	(0.313)	(0.298)	(0.320)	(0.323)	(0.312)	(0.308)	(0.305)	(0.306)
Area (log)					0.203**	0.245**	0.180*	0.362***
					(0.0967)	(0.110)	(0.104)	(0.0923)
Coal					1.391**	1.334**	1.498**	1.432**
					(0.517)	(0.565)	(0.571)	(0.563)
Observations	40	40	40	40	40	40	40	40
R-squared	0.641	0.611	0.609	0.629	0.727	0.702	0.700	0.705

Notes. (i) Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1. (ii) All regressions contain a constant. (iii) "Cistercian share" measures total cistercian houses relative to all religious houses in the county; "Cistercians" is total Cistercian houses; "Cistercian presence" is a dummy taking on the value 1 if just one Cistercian house is found, and "Cistercian density" is the number of cistercian houses per square kilometer.

Table 4. IV results: Full period and the "post-dissolution" period

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Second stage				
Dependent variable:	Population growth 1377-1801				Population growth 1600-1801			
Cistercian share	3.399*	2.752**			3.485*	3.883**		
	(1.804)	(1.203)			(1.793)	(1.829)		
Cistercian presence (0/1)			0.515*	0.444*			0.469*	0.501**
			(0.305)	(0.240)			(0.241)	(0.234)
Pop density 1377 (log)	-0.505**	-0.341**	-0.686***	-0.505***				
	(0.197)	(0.136)	(0.174)	(0.156)				
Pop density 1600 (log)					0.123	0.142	0.076	0.137
					(0.265)	(0.246)	(0.166)	(0.152)
Religious houses	-0.009***	-0.012***	-0.014***	-0.014***	-0.004	0.001	-0.010***	-0.005*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.006)	(0.003)	(0.003)
Land quality	-0.608**	-0.491*	-0.485	-0.392	-0.421	-0.373	-0.496**	-0.360
	(0.287)	(0.290)	(0.321)	(0.297)	(0.268)	(0.259)	(0.240)	(0.233)
Area (log)		0.147		0.085		-0.147		-0.082
		(0.111)		(0.142)		(0.175)		(0.131)
Coal		1.332***		1.531***		1.079**		1.464***
		(0.484)		(0.536)		(0.497)		(0.492)
Forest share	-0.391*	-0.222	-0.369*	-0.228	-0.286	-0.302	-0.295*	-0.240
	(0.223)	(0.231)	(0.216)	(0.246)	(0.187)	(0.258)	(0.173)	(0.215)
Observations	40	40	40	40	40	40	40	40
Kleibergen-Paap	16.63	12.25	40.99	20.91	16.65	10.98	43.08	25.30
Anderson-Rubin (p-value)	0.0931	0.0425	0.0931	0.0425	0.0427	0.0137	0.0427	0.0137
				First stage				
Dependent variable:	Cistercian share		Cistercian Presence		Cistercian Share		Cistercian Presence	
Rforest	0.102***	0.0980***	0.672***	0.608***	0.0872***	0.0778***	0.648***	0.602***
	(0.0250)	(0.0280)	(0.105)	(0.133)	(0.0214)	(0.0235)	(0.0987)	(0.120)
Pop density 1377 (log)	-0.0690*	-0.0589	-0.104	0.00500				
	(0.0344)	(0.0408)	(0.107)	(0.123)				
Pop density 1600 (log)					-0.0139	0.00688	-0.00251	0.0630
					(0.0507)	(0.0515)	(0.162)	(0.148)
Religious houses	0.00178**	0.00149	0.0200***	0.0129**	0.00107	-1.44e-05	0.0189***	0.0128**
	(0.000816)	(0.00136)	(0.00428)	(0.00626)	(0.000747)	(0.00134)	(0.00371)	(0.00539)
Land quality	-0.0161	-0.0102	-0.345	-0.288	-0.0820	-0.0348	-0.450	-0.296
	(0.0687)	(0.0696)	(0.481)	(0.518)	(0.0612)	(0.0651)	(0.448)	(0.517)
Area (log)		0.00949		0.199		0.0429		0.203
		(0.0310)		(0.174)		(0.0308)		(0.155)
Coal		0.0475		-0.152		0.0807		-0.144
		(0.119)		(0.616)		(0.128)		(0.583)
Forest share	0.0343	0.0416	0.183	0.271	0.0188	0.0516	0.159	0.274
	(0.0464)	(0.0511)	(0.266)	(0.284)	(0.0489)	(0.0529)	(0.267)	(0.283)
Observations	40	40	40	40	40	40	40	40
R-squared (First stage)	0.302	0.308	0.459	0.482	0.168	0.248	0.449	0.483

Table 5. Logit of values on Cistercian density across individuals in Europe

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	hardwork				thrift			
Cistercians as share of area	290.991**	314.301**	316.153*	530.757***	75.268	75.938	127.636	314.192
	(133.282)	(130.718)	(171.451)	(192.081)	(152.100)	(134.900)	(179.991)	(196.735)
Area, km2	0.000	-0.000	-0.000	-0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	32,358	31,953	22,230	16,282	31,945	31,542	21,880	15,984
Pseudo R-squared	0.251	0.261	0.282	0.223	0.0477	0.0615	0.0713	0.0468
Country FE	Y	Y	Y	Y	Y	Y	Y	Y
Baseline controls	N	Y	Y	Y	N	Y	Y	Y
Religion dummies	N	N	Y	N	N	N	Y	N
Sample	full	full	full	cath	full	full	full	cath

Logit estimates. Robust standard errors in paranthesis, clustered at the country level. ***,**,* indicates significance at 1, 5, and 10 percent, respectively. All regressions contain a constant term. Baseline controls are age, age squared, a dummy for males, married, educational attainment, absolute latitude, and longitude. Religion dummies refers to whether or not dummies for Protestantism, Roman Catholicism, Orthodox Catholicism, Islam, Buddhism, Free Church, and Judaism are included. Sample refers to whether the full sample is included or whether the sample is restricted to only Catholics. The latter is the case in columns 4 and 8.

Table 6. Values in Europe, alternative specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(3)	(8)	(9)	(10)
Dependent variable:			hardwork					thrift		
					Panel A: Full sample					
Cistercian share	316.153*					127.636				
	(171.451)					(179.991)				
Cistercians		0.020**	0.026**				0.003	0.016*		
		(0.010)	(0.011)				(0.013)	(0.009)		
Cistercians (log)				0.125**					-0.002	
				(0.057)					(0.059)	
Cistercian presence					0.097					-0.036
					(0.079)					(0.080)
Area, km2	-0.000	-0.000	-0.000			0.000	0.000	-0.000		
	(0.000)	(0.000)	(0.000)			(0.000)	(0.000)	(0.000)		
Area (log)				-0.112**	-0.082				0.097*	0.103**
				(0.045)	(0.051)				(0.052)	(0.049)
Observations	22,230	22,230	22,084	22,230	22,230	21,880	21,880	21,738	21,880	21,880
R-squared	0.282	0.282	0.284	0.282	0.282	0.0713	0.0713	0.0721	0.0718	0.0718
					Panel B. Catholics only					
Cistercian share	533.956***					314.733				
	(192.375)					(196.204)				
Cistercians		0.025**	0.034***				-0.001	0.014		
		(0.011)	(0.013)				(0.015)	(0.010)		
Cistercians (log)				0.163***					-0.015	
				(0.059)					(0.066)	
Cistercian presence					0.153**					-0.019
					(0.074)					(0.094)
Area, km2	-0.000	-0.000	-0.000			0.000	0.000	-0.000		
	(0.000)	(0.000)	(0.000)			(0.000)	(0.000)	(0.000)		
Area (log)				-0.124**	-0.089				0.109	0.107
				(0.048)	(0.056)				(0.076)	(0.070)
Observations	16,282	16,282	16,148	16,282	16,282	15,984	15,984	15,851	15,984	15,984
R-squared	0.224	0.223	0.226	0.224	0.223	0.0468	0.0465	0.0474	0.0471	0.0471
Country FE	Y		Y	Y	Y	Y		Y	Y	Y
Baseline controls	Y		Y	Y	Y	Y		Y	Y	Y
Religion dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sample	full		ex Midland	full	full	full		ex Midland	full	full
					Panel C. England					
Cistercian share	439.086***					110.691				
	(122.939)					(282.894)				
Cistercians		0.104***					-0.006			
		(0.035)					(0.054)			
(Log) Cistercians				0.347**					0.063	
				(0.144)					(0.189)	
Cistercian presence					0.196					-0.089
					(0.217)					(0.345)
Area, km2	-0.000***	-0.000***				0.000*	0.000			
	(0.000)	(0.000)				(0.000)	(0.000)			
(Log) Area				-0.362***	-0.221*				0.104	0.181
				(0.120)	(0.116)				(0.172)	(0.178)
Observations	1,085	1,085		1,085	1,085	1,088	1,088		1,088	1,088
R-squared	0.0144	0.0144		0.0131	0.0110	0.0571	0.0570		0.0558	0.0558
Baseline controls	Y	Y		Y	Y	Y	Y		Y	Y
Sample	full	full		full	full	full	full		full	full

Logit estimates. Robust standard errors in paranthesis, clustered at the country level in panels A and B, clustered at the nuts2 level in Panel C. ***,**,* indicates significance at 1, 5, and 10 percent, respectively. All regressions contain a constant term. Sample refers to whether the full sample is included or whether the sample excludes the region of "Border, Midland, and Westland" in Ireland, when this is an outlier. Baseline controls are age, age squared, a dummy for males, married, educational attainment, absolute latitude, and longitude.

Table 7. OLS of outcomes on Cistercians across European regions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	(Log) employment 2007				(Log) Gross Regional Product 2007			
Cistercian share	55.158**				42.563			
	(22.917)				(58.000)			
Cistercians		0.004*				0.009		
		(0.002)				(0.007)		
Cistercians (log)			0.026***				0.033	
			(0.008)				(0.042)	
Cistercian presence				0.025*				-0.011
				(0.013)				(0.057)
Population 2007 (log)	0.986***	0.984***	0.984***	0.989***				
	(0.012)	(0.012)	(0.012)	(0.013)				
Employment 2007 (log)					1.087***	1.078***	1.096***	1.108***
					(0.037)	(0.040)	(0.032)	(0.029)
Area, km2	0.000	0.000			-0.000**	-0.000**		
	(0.000)	(0.000)			(0.000)	(0.000)		
Area (log)			-0.011	-0.005			-0.166***	-0.149***
			(0.017)	(0.017)			(0.036)	(0.032)
Observations	234	234	234	234	234	234	234	234
R-squared	0.996	0.996	0.996	0.996	0.978	0.978	0.983	0.983
NUTS1 FE	Y	Y	Y	Y	Y	Y	Y	Y
Baseline controls	Y	Y	Y	Y	Y	Y	Y	Y
No scale effects p value	0.269	0.219	0.204	0.428	0.0253	0.0616	0.00579	0.00104

OLS estimates. Robust standard errors in paranthesis, clustered at the country level. Observations are nuts2 regions. ***,**,* indicates significance at 1, 5, and 10 percent, respectively. All regressions contain a constant term. Baseline controls are average age, latitude and longitude. No scale effects is the p-value of the test of (log) Population = 0 in columns 1-4 and (log) Employment = 0 in columns 5-10.

Table A1. Population growth and the Benetictine Order

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Population growth 1377-1801						
Benedictine share	0.0148	-0.246	0.0301	0.152	0.131	-0.0351	0.213
	(0.368)	(0.489)	(0.353)	(0.374)	(0.350)	(0.344)	(0.347)
Pop dens 1377 (log)	-0.709***	-0.464*	-0.695***	-0.492**	-0.555***	-0.678***	-0.421**
	(0.193)	(0.239)	(0.210)	(0.194)	(0.160)	(0.211)	(0.160)
Religious houses	-0.00560	-0.00421	-0.00563	-0.0179***	-0.00422	-0.00683	-0.0133***
	(0.00390)	(0.00302)	(0.00385)	(0.00500)	(0.00398)	(0.00427)	(0.00381)
Land quality	-0.651**	-0.712**	-0.650**	-0.529*	-0.547*	-0.653**	-0.474
	(0.282)	(0.308)	(0.284)	(0.311)	(0.305)	(0.285)	(0.320)
rivershare			0.463				
			(2.114)				
Area (log)				0.362**			0.261**
				(0.152)			(0.115)
Coal					1.699**		1.457**
					(0.707)		(0.633)
Ocean						0.0596	
						(0.150)	
Regional FE	No	Yes	No	No	No	No	No
Observations	40	40	40	40	40	40	40
R-squared	0.577	0.631	0.578	0.627	0.658	0.579	0.682

Notes. (i) Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. (ii) All regressions contain a constant.

Table A2. Population growth, Cistercians and other religious orders

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Population growth 1377-1801					
Pop dens 1377 (log)	-0.374**	-0.360**	-0.446**	-0.439***	-0.420**	-0.485**
	(0.142)	(0.133)	(0.178)	(0.144)	(0.156)	(0.236)
Cistercian share	1.702**	1.789***	1.700**	1.839**	1.531**	2.134**
	(0.644)	(0.647)	(0.656)	(0.691)	(0.695)	(0.967)
Religious houses	-0.0124***	-0.0132***	-0.0116***	-0.0105***	-0.0119***	-0.0113**
	(0.00355)	(0.00349)	(0.00361)	(0.00330)	(0.00370)	(0.00465)
Land quality	-0.486	-0.469	-0.442	-0.427	-0.517	-0.306
	(0.312)	(0.332)	(0.280)	(0.285)	(0.333)	(0.379)
Area (log)	0.203**	0.224**	0.205*	0.127	0.196*	0.175
	(0.0967)	(0.103)	(0.109)	(0.123)	(0.104)	(0.159)
Coal	1.391**	1.429***	1.227*	1.296**	1.359**	1.152*
	(0.517)	(0.497)	(0.610)	(0.552)	(0.524)	(0.599)
Benedictine share		0.316				0.614
		(0.319)				(0.517)
Augustinian share			0.446			0.661
			(0.638)			(0.614)
Cluniac share				1.494		1.497
				(1.329)		(1.288)
Premon share					-0.446	0.329
					(0.622)	(1.096)
Observations	40	40	40	40	40	40
R-squared	0.727	0.735	0.735	0.739	0.731	0.764

Notes. (i) Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. (ii) All regressions contain a constant

Table A3. Instrument falsification: Royal Forest and other Monastic Orders

Panel A: Intensive margin

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	Cluniac share		Benedictine share		Augustinian share		Premon share	
Rforest	0.0134	-0.00561	-0.120	-0.109	0.0260	0.0238	-0.0804	-0.0751
	(0.0219)	(0.0201)	(0.127)	(0.141)	(0.0650)	(0.0750)	(0.0950)	(0.0916)
Pop dens 1377 (log)	0.00819	0.0480**	0.0251	-0.00370	0.126**	0.154**	-0.0553	-0.0714
	(0.0193)	(0.0214)	(0.0573)	(0.0895)	(0.0509)	(0.0627)	(0.0354)	(0.0582)
Religious houses	0.000498	-0.00136*	-7.06e-05	0.000812	-0.00195	-0.00132	-0.000437	-0.000102
	(0.000640)	(0.000742)	(0.00237)	(0.00334)	(0.00143)	(0.00246)	(0.00152)	(0.00193)
Land quality	-0.0601	-0.0379	-0.0223	-0.0392	-0.120	-0.101	-0.0507	-0.0605
	(0.0414)	(0.0470)	(0.160)	(0.156)	(0.119)	(0.122)	(0.0428)	(0.0446)
Area (log)		0.0549***		-0.0286		-0.00769		-0.0122
		(0.0179)		(0.0622)		(0.0611)		(0.0324)
Coal		0.0662		-0.126		0.374*		-0.0968
		(0.0810)		(0.428)		(0.209)		(0.277)
Forest share	0.00970	0.0400	0.176**	0.155*	-0.00627	0.0112	-0.0136	-0.0250
	(0.0305)	(0.0303)	(0.0766)	(0.0799)	(0.0699)	(0.0745)	(0.0288)	(0.0320)
Observations	40	40	40	40	40	40	40	40
R-squared	0.051	0.229	0.088	0.099	0.192	0.261	0.235	0.248

Panel B: Extensive margin

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	Cluniac presence		Benedictine presence		Augustinian presence		Premon presence	
Rforest	0.191	-0.0578	0.210	0.204	0.207	0.204	0.0883	0.0483
	(0.257)	(0.195)	(0.191)	(0.198)	(0.198)	(0.176)	(0.244)	(0.262)
Pop dens 1377 (log)	0.155	0.616**	0.0508	0.0552	0.0104	0.0639	-0.161	-0.101
	(0.220)	(0.258)	(0.0652)	(0.129)	(0.0880)	(0.121)	(0.196)	(0.270)
Religious houses	0.00829	-0.0182**	0.00530	0.00440	0.00388	0.00520	0.0202***	0.0155*
	(0.00779)	(0.00885)	(0.00342)	(0.00402)	(0.00310)	(0.00448)	(0.00490)	(0.00882)
Land quality	-0.795	-0.547	0.0865	0.0877	0.198	0.234	-0.711*	-0.682*
	(0.496)	(0.547)	(0.114)	(0.0867)	(0.229)	(0.229)	(0.351)	(0.357)
Area (log)		0.754***		0.0226		-0.0179		0.130
		(0.179)		(0.0609)		(0.0805)		(0.207)
Coal		-0.0494		-0.107		0.728		-0.236
		(0.714)		(0.800)		(0.445)		(1.074)
Forest Share	0.145	0.507*	0.0543	0.0591	0.0172	0.0497	0.318	0.367
	(0.321)	(0.265)	(0.0634)	(0.0575)	(0.0392)	(0.0555)	(0.258)	(0.263)
Observations	40	40	40	40	40	40	40	40
R-squared	0.098	0.318	0.210	0.212	0.158	0.246	0.238	0.245

Table A4. Robustness: IV estimates 1290-1801

	(1)	(2)	(3)	(4)
Dependent variable:	Population growth 1290-1801			
	Second stage			
Cistercian share	4.041*	4.086**		
	(2.342)	(2.011)		
Cistercian presence			0.616*	0.745**
			(0.363)	(0.352)
Pop dens 1290 (log)	-0.683**	-0.646**	-0.963***	-1.062***
	(0.270)	(0.310)	(0.164)	(0.173)
land quality	-0.447	-0.389	-0.168	-0.071
	(0.349)	(0.360)	(0.388)	(0.420)
Religious houses	-0.008*	-0.007	-0.013***	-0.003
	(0.005)	(0.008)	(0.005)	(0.007)
Forest share	-0.378	-0.341	-0.324	-0.399
	(0.275)	(0.297)	(0.225)	(0.297)
Area (log)		-0.023		-0.293*
		(0.169)		(0.174)
Coal		0.825		1.317**
		(0.565)		(0.671)
Observations	40	40	40	40
K-P	12.18	13.52	32.44	25.46
A-R (p-value)	0.0858	0.0274	0.0858	0.0274
	First stage			
Dependent variable:	Cistercian share		Cistercian presence	
Rforest	0.102***	0.110***	0.667***	0.602***
	(0.0291)	(0.0299)	(0.117)	(0.119)
Pop dens 1290 (log)	-0.0854***	-0.0977***	-0.106	0.0231
	(0.0260)	(0.0338)	(0.122)	(0.125)
Land quality	0.0205	0.0233	-0.319	-0.299
	(0.0678)	(0.0660)	(0.471)	(0.504)
Religious Houses	0.00199**	0.00307**	0.0201***	0.0123*
	(0.000743)	(0.00119)	(0.00432)	(0.00633)
Forest share	0.0421	0.0355	0.188	0.273
	(0.0530)	(0.0559)	(0.270)	(0.281)
Area (log)		-0.0272		0.213
		(0.0341)		(0.178)
Coal		0.0917		-0.157
		(0.105)		(0.582)
Observations	40	40	40	40
R-squared	0.393	0.410	0.460	0.482

Table A5. Robustness: Endogeniety of total religious houses?

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
			Relgious houses (total)						Cistercian presence (0/1)			
Rforest	-11.25	-9.105	-11.69	-9.800	-10.68	-8.662	0.442*	0.490***	0.438*	0.482***	0.446*	0.491***
	(10.80)	(5.671)	(10.82)	(6.540)	(11.74)	(7.896)	(0.245)	(0.125)	(0.249)	(0.158)	(0.254)	(0.173)
Pop dens 1290 (log)	7.977**	13.06***					0.0544	0.183				
	(3.491)	(2.387)					(0.153)	(0.109)				
Pop dens 1377 (log)			8.316**	12.09***					0.0619	0.161		
			(4.065)	(2.593)					(0.132)	(0.115)		
Pop dens 1600 (log)					-0.0371	3.494					-0.00321	0.108
					(8.100)	(5.808)					(0.167)	(0.142)
land quality	9.246	3.749	10.78	8.572	21.00	18.54	-0.134	-0.253	-0.130	-0.178	-0.0528	-0.0582
	(16.25)	(9.424)	(15.48)	(9.516)	(17.92)	(12.92)	(0.625)	(0.533)	(0.615)	(0.537)	(0.573)	(0.524)
Area (log)		18.57***		17.67***		15.98***		0.440***		0.426***		0.408***
		(2.986)		(3.109)		(3.560)		(0.1000)		(0.0978)		(0.0976)
Coal		-24.58**		-20.49**		-37.94***		-0.459		-0.416		-0.630
		(10.50)		(8.011)		(10.85)		(0.604)		(0.625)		(0.529)
Forest share	-7.281	4.551	-6.978	4.864	-5.576	4.409	0.0422	0.329	0.0434	0.333	0.0538	0.331
	(10.60)	(4.317)	(9.505)	(3.718)	(9.560)	(4.348)	(0.263)	(0.264)	(0.259)	(0.268)	(0.265)	(0.271)
Observations	40	40	40	40	40	40	40	40	40	40	40	40
R-squared	0.236	0.766	0.236	0.724	0.165	0.608	0.147	0.446	0.148	0.435	0.144	0.417

Table A6. Robustness: IV estimates without control for total religious houses

	(1)	(2)	(3)	(4)	(5)	(6)
			Second stage			
Dependent variable:	Population gr, 1290-		Population gr, 1377-		Population gr, 1600-	
Cistercian presence	0.941	0.806**	0.880	0.724*	0.699*	0.591**
	(0.613)	(0.395)	(0.631)	(0.378)	(0.423)	(0.266)
Pop dens 1290 (log)	-1.083***	-1.116***				
	(0.175)	(0.155)				
Pop dens 1377 (log)			-0.822***	-0.717***		
			(0.201)	(0.177)		
Pop dens 1600 (log)					0.077	0.109
					(0.191)	(0.152)
Land quality	-0.243	-0.068	-0.586	-0.460	-0.685*	-0.449
	(0.571)	(0.445)	(0.535)	(0.426)	(0.373)	(0.276)
Forest share	-0.245	-0.434	-0.289	-0.388	-0.254	-0.292
	(0.240)	(0.322)	(0.215)	(0.312)	(0.192)	(0.237)
Area (log)		-0.380**		-0.278		-0.200
		(0.182)		(0.199)		(0.133)
Coal		1.425**		1.930***		1.714***
		(0.676)		(0.596)		(0.475)
Observations	40	40	40	40	40	40
K-P	3.252	15.42	3.089	9.289	3.079	8.091
A-R (p-value)	0.0380	0.0275	0.0241	0.00391	0.0119	0.00613

Table A7. OLS of values on Cistercians across regions in Europe

Dependent variable:	(1) hardwork	(2)	(3) hardworkcond	(4)	(5) thrift	(6)	(7) thriftcond	(8)
Cistercian share	58.507	147.329***	72.110	148.805***	61.515**	124.756**	73.352	145.489**
	(41.900)	(31.961)	(52.217)	(40.581)	(27.703)	(45.269)	(54.316)	(67.159)
Area, km2	0.000	-0.000	0.000	-0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	242	174	241	174	242	174	241	174
R-squared	0.858	0.872	0.800	0.846	0.441	0.419	0.368	0.379
Country FE	Y	Y	Y	Y	Y	Y	Y	Y
Regional controls	Y	Y	Y	Y	Y	Y	Y	Y
Sample	full	prot<50%	full	prot<50%	full	prot<50%	full	prot<50%
No. countries	29	26	29	26	29	26	29	26

Robust standard errors in paranthesis, clustered at the country level. The unit of observation is nuts2 regions. ***,**,* indicates significance at 1, 5, and 10 percent, respectively. All regressions contain a constant term. Sample refers to whether the full sample is included or whether the sample is restricted to include only regions with a maximum of 50% Protestants. Hardworkcond is the residuals aggregated up to the nuts2 level of a regression of hardwork on age, age squared, a dummy for males, married, educational attainment, and religion dummies in the individual sample. Likewise for thrift. Regional controls are latitude and longitude.

Table A8. Summary Statistics across European regions

Variable	Obs	Mean	Std. Dev.	Min	Max
Hardwork	242	0.400	0.260	0.000	1.000
Hardworkcond	242	0.382	0.259	0.000	1.000
Thrift	242	0.369	0.145	0.000	0.908
Thriftcond	242	0.460	0.145	0.000	1.000
Cistercian share of total area	242	0.000	0.000	0.000	0.001
Area	242	15323	15266	173	92961
Absolute latitude	242	48.893	5.048	37.589	62.329
Longitude	242	9.023	8.755	-9.046	25.476
Age	242	47.864	4.829	27.000	70.538
Protestants, share	241	0.294	0.356	0.000	1.000
(Log) employment 2007	241	13.383	0.710	10.931	15.462
(Log) Population 2007	242	14.205	0.708	11.735	16.266
(Log) Gross Regional Product 2007	235	10.435	0.808	8.302	13.141

Hardworkcond is the residuals aggregated up to the nuts2 level of a regression of hardwork on age, age squared, a dummy for males, married, educational attainment in the individual sample. Likewise for thrift.