An Empirical Approach to Understanding Privacy Valuation

Luc Wathieu
Harvard Business School
lwathieu@hbs.edu

Allan Friedman
Kennedy School of Government
Harvard University
allan_friedman@ksgphd.harvard.edu

ABSTRACT

The purpose of this paper is to detect the presence of sophisticated economic motives behind individual concerns for privacy. Recent theories of privacy demands in commercial contexts have assumed an economically aware and sophisticated consumer, capable of evaluating the indirect consequences of information transmission. We present evidence, from a large-scale experiment evoking a realistic context, that privacy concerns are indeed sensitive to the indirect consequences of information transmission.

1. INTRODUCTION

There are many perspectives on how privacy sentiments manifest themselves and should be addressed in society. Recently, with the rise of electronic commerce and database marketing, researchers have increasingly interpreted the demand for privacy in economic terms. Some authors and commentators suggest that consumers view targeted marketing communications as a costly annoyance [e.g., 15], while other authors posit that finely informed firms might take actions that are not indirectly decrease the consumer's welfare [e.g., 16, 27, 31]. While these models are often theoretically sound, the connection between their basic assumptions and the motives behind ordinary privacy concerns remain largely un-documented. Thus, the purpose of this paper is to suggest and test a set of behavioral hypotheses to reveal and understand the sensitivity of consumers towards indirect potential economic consequences attached to the dissemination of personal information.

Many observers have noted the existence of a "privacy paradox" in that consumers at the same time (1) routinely declare that they value their privacy highly [14], but (2) do not seem to actively incorporate privacy concerns in their transactions [25]. This paradox might suggest that consumers are too unsophisticated to envision (or even to sense) the economic consequences of transmitting their personal information. This interpretation would be bad news for economic theories of privacy demands. However, an alternative interpretation of the privacy paradox could be that, accounting for the complexity of the anticipated consequences and the lack of means at their disposal, consumers currently feel unable to enact their privacy preferences. However, to substantiate this explanation, a finer understanding of these privacy preferences needs to be provided in the first place.

In order to progress in that direction, this paper proposes an experiment based on a real-world situation involving the transmission of personal information in a commercial context. Much of the earlier research aimed at highlighting and understanding demands for privacy has been survey-based, focusing on non-situational antecedents of demand [7, 32]. An experimental approach is suitable to assess the impact of context modifications on the privacy sentiment.

There have been a limited number of experiments attempting to study privacy valuation. Huberman, Adar and Fine [17] suggest that privacy valuation is a function of perceived deviance. While this finding helps clarify the strength of some individuals' preferences, deviance cannot explain all preferences, particularly when privacy applies to data that does not fit in a normal/deviant framework (e.g., name and address on a mailing list). Rational privacy protection behavior has been isolated in a study with very explicit information on risks and rewards [9], which does not tell us whether people's natural notion of privacy usually encompass such consequences. A series of detailed, interactive surveys by Acquisti and Grossklags questioned the model of a rational privacy-protecting consumer, in an analysis that included a broad range of privacy lifestyle choices [2], but did not directly induce trade-offs between information transmission and economic benefits. Such trade-offs can be studied by actually watching user behavior [3, 6, 26] or in conjoint analysis to derive values of resolving privacy concerns [14]. While these tools provide an important understanding of privacy sentiments in a specific context, or a useful dollar value, it is difficult to apply them in a broader context where the implications of information transmissions can be complex and indirect.

2. ASSESSING THE INDIRECT VALUE OF PRIVACY

2.1. A Distinction

This paper differentiates between a *direct* privacy concern and an *indirect* privacy concern, and argues that the second form, while more subtle, is a measurable influence on ordinary privacy demands in commercial contexts. A direct privacy concern is motivated by an immediately perceived harm from information release by the offended party. For instance, fear of impersonation

fraud[1] or dislike of direct marketing represent direct disutility from lack of privacy. An indirect privacy concern, in contrast, is predicated on multiple steps between a situation in which personal information is revealed and the impact of this situation on other variables that affect the individual's well-being. These consequences could include access to low prices, product variety and quality, or even economic growth, but the effect on an individual is a function of many other variables, decisions and data beyond the collection of that individual's personal information.

Privacy concerns are often both direct and indirect in nature. For example, when thinking about government surveillance and airport security, a direct privacy concern would be the fear of a stranger intruding your intimate space, but citizens have also expressed annoyance at the delays resulting from *others* receiving such treatment, or complained from the fact that *everyone* loses rights when even a few people are unfairly treated as suspects. Interestingly, as noted by [31] such indirect privacy effects are not necessarily attached to the individual's personal information disclosure, they are related to harms incurred by the overall system when people's privacy is restricted.

In connection with marketing information, one can also highlight the distinction between direct and indirect privacy concerns, with greater doubt cast on the empirical relevance of the latter. Indeed, it should be fairly straightforward to show that individual consumers are concerned about exposing personal peccadilloes to marketers. Similarly, if revealing a telephone number or email address leads to the annoyance of telemarketing or spam, a theory claiming reluctance to reveal information should be uncontroversial. In contrast, the fear of price discrimination and other market dysfunctions that might result from consumer exposure [27, 31] requires the consumer to understand (or at least to sense) the fact that personal harms can accrue from the collection of everyone's information to gauge demand.

2.2. General Framework

We assume a general framework in which firms seek to gain consumer information in anticipation of a profitable course of action. This course of action can take many forms, including internal systems development and improvement [11], targeted marketing [20], loyalty programs [8], or maximizing profits through price discrimination [21]. In most such cases, the firm is the driving actor to collect and/or use personal information. Whether this raises a privacy problem depends on the consumer's reaction. In extreme cases, the consumer impact is obviously positive (e.g., when an emergency medical practitioner obtains life-saving information) or negative (e.g., when it results in unwanted telephone solicitation [20, 28]), and predicting consumer reaction is trivial. What is less understood are the more balanced trade-offs, with subtler benefits or less obvious cost or risk factors. In particular, indirect issues are critical for a proper understanding of privacy from an economic perspective.

Our general hypothesis is that ***consumers are capable of expressing differentiated levels of concerns in the presence of changes that suggest indirect consequences of information transmission***, at least intuitively. In other words, we suggest that there is a *homo economicus* behind the privacy concern, not simply a primal fear.

If consumers do not, in fact, have a sophisticated understanding of indirect privacy effects, then they will not be concerned with subtle factors in a given context, nor will they appreciate factors that only affect the indirect concern without triggering the direct, immediate privacy concern. The six hypotheses proposed below would not hold true unless consumers are capable of responding to non-trivial implications of information dissemination.

2.1 Specific Hypotheses

If a direct utility of privacy were driving the privacy concern, information dissemination in itself would be the critical trigger of privacy concern for consumers, with more information transmission causing a greater concern. Sophisticated actors, whose privacy concern anticipates possible indirect consequences of information dissemination should be expected to (1) be indifferent to mere information dissemination across databases when it is clearly inconsequential and (2) display a concern that varies with the likelihood of information use. The former idea provides us with a first hypothesis:

> **H1 (Indifference Towards Mere Dissemination):** Data dissemination alone has no disutility in privacy terms.

The question of how the data will be used further raises the question of data relevance. Naïve approaches to the privacy concern would seek protection of any kind of personal data (with perhaps an emphasis on personally identifying pieces of data). Hann et al even suggest that personal valuation is independent of personal context [14]. But if the privacy concern is driven by the indirect consequences of data usage by marketers, privacy demands should, ceteris paribus, be greater towards data that is more likely to be used by the firm who collects it (i.e., data that can be leveraged more profitably) when this use could be to the subject's detriment. For instance, if a consumer is worried about obtaining health insurance in a given context, then sharing family medical history should cause concern, while if the consumer is confident in his future health coverage, sharing the history is less of a concern. More generally:

> **H2 (Sensitivity to Relevance):** Situational relevance for a self-interested party increases the privacy concern.

It is commonplace nowadays to question the sophistication of consumers, their ability to anticipate the material (as opposed to "framed") consequences of their actions, and privacy is no exception [3]. Thus our general hypothesis that consumers act as if they perceived a privacy concern stemming from indirect market effects might appear somewhat surprising. The following hypothesis captures the notion that consumers will be able to produce a privacy concern reflecting indirect consequences without a need for prompts or framing:

> **H3 (Spontaneous Concern):** Consumers have a privacy concern that stems from indirect effects even in the absence of additional warnings or priming.

The most critical test of an indirect effect rests on the question of personal involvement. Under the conventional thinking regarding privacy, there should be no privacy concern if personal information is not transmitted at all. By including indirect privacy effects, a concern might exist even when the consumer's

[1] Also known as "identity theft," a term the authors believe is misleading and should be replaced by "impersonation fraud". See [12, 24]

personal information is protected. For instance, the fact that other consumers transmit their information might lead to structural changes (e.g., in monopolistic positions, or in the amount of variety available) that affect a non-transmitter and should cause a reaction in defense of privacy (as in [31]). This logic even applies in the case of a privacy concern motivated by impersonation fraud, as consumers absorb the added costs of the misuse of others' identifiers. A test of this idea could be based on the following hypothesis:

> **H4 (Privacy Externality):** Individuals may have personal privacy concerns in situations where they do not have a personal stake to directly gain or lose.

In sum, what links the above hypotheses is the notion that they detect a consumer who is able to sense indirect economic consequences of information transmission and to register a privacy demand that flows from these consequences.

Another aspect of consumer behavior that can help assess whether consumers perceive the indirect implications of personal data transmission is their approach to policy solutions in response to privacy concerns. If consumers think of privacy only in terms of direct disutility upon disclosure of information about themselves, we can expect that control levers such as the ability to opt-in and opt-out will be deemed attractive and sufficient. In contrast, a consumer's distinct call for regulation or intermediation (broadly speaking: any collective intervention to limit the transfer of data concerning a group of people) can only be understood in light of a perception of interdependence of individual (and indirect) consequences. In particular if H1 and H4 are true, perceived harm from consumer exposure can occur whether or not the individual can control his or her own participation in the data transfer. Thus the following hypothesis:

> **H5 (Limited Personal Control):** Opt-in and opt-out preferences do not completely enact privacy concerns when indirect consequences are perceived.

While personal participation preferences may not be strongly applicable in situations that suggest indirect effects, the role of the social planner (or of any representative intermediary [31]) becomes more important. If individuals are affected by the actions of the group, then individuals should sense that the solution lies with group (or intermediated) action. This gives our final hypothesis:

> **H6 (Demand for Intermediation):** When indirect threats are associated with the privacy concern, consumers are more likely to call for a collective intervention to limit data transmission.

3. RESEARCH DESIGN

To better understand how consumers treat information privacy in a complex environment and test the above hypotheses, we presented participants with a realistic scenario involving the dissemination of personal information in a commercial context, and measured their response through a brief survey. A scenario/survey-based experiment was deemed appropriate because it would allow experimental manipulations while evoking a relevant, relatively natural, relatively complex situation.

There were twelve experimental conditions, each involving a specific modification of the same baseline scenario. A manipulation check questionnaire was also performed to verify that respondents and researchers shared the same interpretation of the various scenarios.

Respondents were 647 randomly selected members of a subject pool maintained by the research center of a business school in the United States. This subject pool features over 10,000 members diverse in background and gender, including business and undergraduate students who accounts for 45% of the population. Respondents were recruited by email and participation was voluntary, with a $5 payment upon completion. The experiment was administered via a website. It was made clear to respondent that there was no right or wrong answer. The average experimental group included 54 respondents, with no group having fewer than 48 respondents. The manipulation check involved another 47 respondents.

3.1 Control Scenario

To evaluate the theoretical hypotheses presented above, we looked for a realistic and intuitive situation where consumer data were disseminated in a way that might (1) allow the consumer to access advantageous offers (2) expose the consumer to marketing hassle and (3) have likely indirect consequences in terms of the consumer's welfare.

> As a service to its members your college alumni association has negotiated a special deal with a well-known car insurance company.
>
> The insurance company will use data (including members' name and contact information) on a one-time basis to offer alumni (via a mail and phone marketing campaign) an alumni association-endorsed deal featuring first-class service levels and a 30% discount on annual insurance premiums.
>
> Based on certain parameters specified by the insurance company, data for 20% of the alumni have been transmitted to the insurance company and all of these alumni are about to be offered the deal. At this point it is still unknown whether you are among the beneficiaries of this deal.

Affinity-based direct marketing of car insurance contracts provided such a context. This marketing process, documented in a case study by Wathieu and Morris [30], uses the membership databases of trusted associations (such as alumni associations) to channel targeted deals to their members, through direct communications means that blend direct mail and telemarketing. When associations negotiate such deals, often for considerable fees, they have an interest in minimizing potential hassle for their members, and they also seek to minimize the possibility that marketers discriminate among different types of members, in order to maintain membership cohesiveness. Governments, on the other hand, monitor the impact of these arrangements on competition and the industrial structure. The control scenario, which will serve as a baseline for our analyses, is evoking one

such arrangement between an alumni association and a car insurance company.

The underlined parts of the scenarios are those privacy-sensitive aspects that will be modified in experimental conditions.

3.2 Survey Questions

The scenario itself did not explicitly offer the respondents a choice. However, after reading the scenario, respondents were asked four questions (answers were selected from 7-point Likert scales):

- How happy are you that this deal was struck between your alumni association and the car insurance company?
- In this instance, how fairly do you feel your alumni association is treating you?
- Are you fearful that this kind of activity in the insurance market might ultimately reduce your access to a low-premium contract?
- This is an example of a situation in which I am concerned about privacy.
- Alumni should be given an opportunity to opt-out (withdraw) from this program before their data is transmitted.
- Alumni should be included in this program only if they specifically sign up before their data is transmitted.
- I would like this kind of initiative to be reviewed and voted on (either banned or explicitly authorized) by the Board of Alumni.

In addition, participants were asked (yes/no) (1) whether they would opt-out of the deal if the opportunity were available to do so, (2) whether they would opt-in if assent was "necessary but easy," and (3) whether they would vote to authorize the initiative if they were on the board of alumni.

This experiment was designed to elicit honest feedback. None of the questions were asked in such a way that the respondent would be inspired to create a positive impression.

3.3 Experimental Conditions

Experimental conditions changed the baseline scenario by inserting one or more of the following five modification:

Dissemination. This modification is introduced in reference to H1. Instead of assuming that the alumni association would transmit data parsimoniously (underlined part of scenario starting with "Based on certain parameters specified by the insurance company data for 20% of the alumni have been transmitted…"), participants are told that "Data for all the alumni have been transmitted to the insurance company and, based on certain parameters certified by the insurance company, 20% of the alumni are about to be offered the deal." As a result of this manipulation the likelihood of data transmission has increased from 20% to 100%, while no other significant change is taking place,[2] in accordance with the notion of mere dissemination.

More relevant data. This modification implies increasing transmitted data to include education and occupational data, commonly viewed as relevant for an insurance company trying to assess client risk. "Name and contact information" is accordingly replaced by "name, contact information, degree obtained and year, honor student status, GPA, and current occupation." Manipulation-check respondents rated each of these elements as highly useful to predict whether a person is a safe driver or not.

More irrelevant data. This scenario modification increases transmitted data to include data that is personally meaningful, but less likely to be used by an insurance company assessing client risk: "name, contact information, membership in college associations, city of birth, and city of residence at college registration time." These additional elements were seen as least relevant as predictors of safe driving in the manipulation check questionnaire.

Priming. This modification serves to test H3. To increase the salience of a risk of discriminative practices by better-informed insurance companies, the following paragraph was inserted before the baseline scenario's last paragraph: "Some have wondered whether the premium paid by ordinary drivers can stay low if car insurance companies continue to use databases to offer special deals to consumers predicted to be 'safe drivers.'" Manipulation checks used 7-point scales to verify that respondents found this statement both clear and legitimate.

No personal benefit. In the context of a test of H4, we told some respondents that they were not beneficiaries of the deal. The last phrase of the baseline scenario was replaced by "it has become clear that you are not among the beneficiaries of the deal."

An extensive test of the effect of all these modifications and their interactions would have required 24 experimental groups: 3 (contact data, additional relevant data, additional irrelevant data) x 2 (dissemination, 20% data shared) x 2 (priming indirect concern or not) x 2 (personal benefit or not). However, for a parsimonious test of the individual impact of each modification against the control condition we only needed five experimental conditions in addition to the control. Four additional experimental conditions were added to measure, in the presence of the "dissemination" modification, the impact of each of the other four modifications. Finally, to further scrutinize H4, the condition that combined (priming indirect concern, dissemination, no personal benefit) was also included in the experiment, leading to a total of 12 experimental groups.

4. RESULTS

Table 1 gives the mean responses for each of the twelve groups, with indication of significance when the response obtained is statistically different from response in the control group. Dichotomization around an arbitrary point was sometimes used to simplify descriptive analysis by reducing the 7-point scale of relative sentiment to a simple yes/no Boolean variable. Because

[2] Responses to a manipulation check questionnaire confirmed that respondents in the target population reliably agreed with this interpretation of the manipulation.

the split point is arbitrary, however, we only use dichotomization sparingly.

Table 1: Mean Response (Privacy Concern)

CONDITIONS	CONCERNED ABOUT PRIVACY (1-7 scale)
(1) Control	4.16
(2) Dissemination	**4.86 ***
(3) More relevant data	**5.26 *****
(4) More relevant data/Dissemination	**4.95 ****
(5) More irrelevant data	4.70
(6) More irrelevant data/Dissemination	4.70
(7) Priming	4.48
(8) Priming /Dissemination	4.77
(9) No personal benefit	4.43
(10) No personal benefit/Dissemination	4.77
(11) Priming/No personal benefit	4.76
(12) Priming/No personal benefit/ Dissemination	**5.05 ****

Significant difference with respect to the control condition:
**** = (p < .01), ** = (p < .05), * = (p < .1)*

A few notes on the control group's response are in order. With 2/3 of the respondents placing their level of concern at 4 or higher out of 7 (1 meaning "Not at all concerned" and 7 "extremely concerned"), the control group already appears somewhat concerned about privacy. While the respondents were concerned, they were not dissatisfied with the offer in front of them: over 80% recorded a 6 or a 7 when asked if they were happy that a deal was struck between their alumni association and the car insurance company. The control group reveals a concerned population that is nonetheless open to making a trade-off between personal data dissemination and an opportunity to access a better deal.

4.1 Mere Dissemination

The first condition of whether or not all data were shared with the insurance company tests H1. The respondents faced the same situation, where 20% of the alumni will be offered a deal on car insurance. The first condition indicated that either only those alumni receiving the deal would have their data shared, or whether everyone's data would be shared. H1 predicts that this condition will have no effect, since the dissemination of data has no material impact on who benefits from the deal (or otherwise, as the data is assumed to be good for one-time use only). We find no significant effect on privacy concerns across any of the six categories.

Table 2 shows the difference between means of privacy concerns for each condition. The control condition comes close to significance with a p-value of .0516 (but this result is not replicated in any dichotomization of the measure). Holding everything else constant, going from a 20% chance of having ones' data disseminated to a 100% certainty of having ones data disseminated does not result in an increase of the privacy concern. This indicates support for H1: mere dissemination of data is not a driver of disutility with respect to privacy.

Table 2 (Absence of) Dissemination Impact on Privacy Concern

BASE CONDITIONS Cfr. Table 1	CHANGE IN PRIVACY CONCERN MEASURE
Control [(1) → (2)]	-0.702 (p = 0.0516)
More relevant data [(3) → (4)]	0.31 (p = 0.3777)
More irrelevant data [(5) → (6)]	0 (p = 1)
Priming [(7) → (8)]	-0.28 (p = 0.4617)
No personal benefit [(9) → (10)]	-0.33 (p = 0.3394)
Priming/No personal benefit [(11) → (12)]	-0.28 (p = 0.373)

4.2 Sensitivity to Relevance

We test H2 by varying the information passed along in two ways. The first condition submits additional information that insurers typically see as relevant to the client risk (and, thus, insurance premiums): professional and educational achievements. The second condition uses personal information that people may not particularly want exposed, but that is not relevant to the situation of auto insurance.

To assess these predictions, we compare the privacy concerns of the control group with those of groups treated by either relevant or irrelevant information. We observe a very small effect from transmission of additional irrelevant information and a large highly robust effect from the use of relevant information. The difference in means between the relevant group and the control group is significant ($p < .01$). We also find some degree of statistical significance in five out of the six dichotomizations (the irrelevant information increases the privacy concern above the control level, but it is only slightly significant in the {1-4/5-7} split). However, the importance of this finding should reflect the difference in the respective "sensitivity" of the relevant and irrelevant treatments. The former contains information such as

GPA, which many people instinctively shelter, while there is little that is as sensitive in the irrelevant treatment.

Nonetheless, these results confirm the supposition that consumers are sensitive to the type of information transmitted, particularly as it relates to the market context. Participants found the sharing of relevant data more worrisome for privacy issues than irrelevant information or no information. That the type of information is important is not a revolutionary conclusion, but it does support the vision of a more sophisticated privacy-sensitive consumer (as compared to a simple "people want privacy" view). Understanding the relevance of personal information to a given situation is a necessary step for consumers to understand indirect economic privacy threats.

4.3 Spontaneous Concern

If people are myopic and don't envision indirect information effects such as price discrimination, then priming should alter their perception and increase the privacy concern. But we find that priming has no significant effect on privacy concerns. One might argue that the respondents were unable to see or value the threat of market segmentation even with the added suggestion. The manipulation check does not support this conclusion: respondents claim to understand the priming and evaluate the highlighted concern as highly legitimate. Furthermore, the explicit measure of a fear of reduced access in the survey was not affected by the priming treatment either, confirming that the priming is "no news" for respondents. This confirms H3 and offers support the general hypothesis that indirect economic consequences have a natural influence on the feeling of a privacy concern.

4.4 Privacy Externality

Conventional discussions surrounding privacy attach privacy concerns to the existence of personal data transfer and individualized consequences. Personal privacy fears are triggered when the individual's data could be transmitted and used. The findings in 4.1 and 4.2 show that data transmission is not the binding factor here, but rather use. Accounting for the externalities arising from more informed firms, however, the data in question does not need to apply to the individual who would benefit from privacy protection. We test this sensitivity to privacy externality by telling two groups [(9) and (11)] that only those receiving the deal will have their information shared, and that they will NOT receive this deal. Ergo, their information will not be shared and they won't get called about the deal. This is the most indirect privacy story one can tell, so if a privacy concern is still registered, then consumers can be seen as sensitive to indirect situations: they reveal a perception of privacy externalities. We believe that consumers understand that they will not receive any benefit because there is an overwhelming jump in the number of participants who believe the situation is not fair. Their feelings on privacy, however, remain unchanged.

Since respondents registered a privacy concern, H4 is corroborated. The increase in privacy concern in condition (12) (priming, no personal benefit, dissemination) in a comparison with the control condition has no clear interpretation (recall that the relevant difference, between (11) and (12) is not significant, as per Table 2).

4.5 Limited Personal Control

Participants were asked, in a yes-or-no format, whether they would opt-in or opt-out from this type of arrangement.[3] One third of the control group would opt-out, while 39% would opt-in. As above, we compare the treatment groups with the control group to measure the magnitude of treatment effects. First, however, it is interesting to note the discrepancy between the number of people who would participate with an opt-out provision (66%) versus an opt-in provision (39%). While this distinction has been empirically observed elsewhere [18, 19], it is unusual to see such a strong default effect when the two questions are juxtaposed and no effort is required.

In the experiment, we find that there is generally no significant variance among the treatment groups for the personal decision to opt-in. The number of participants wishing to opt out, however, dips to 40% ($p < 0.05$) when relevant information is involved, and to 48% ($p < 0.10$) when participants will not be considered. We also note that priming as a small effect ($p < .10$).

The findings with respect to opt-in support H5: participants do not associate the decision to opt-in with variations of the privacy concern. However, the opt-out preferences seem to be a more direct function of privacy preferences. That is, even though (by H4) opting out will not completely mitigate the privacy harms identified in measures of concern, participants still would choose to exercise the option to opt out. In all, we do not find evidence to reject H5. If both opt-in and opt-out directly reflected privacy sentiments, rather than a more sophisticated solution-driven consumer, both would be strongly correlated with privacy concern. We find that only opt-out sentiments are.

4.6 Demand for Intermediation

Participants were asked to what extent they "would like this kind of initiative to be reviewed and voted on (either banned or explicitly authorized) by the Board of Alumni" and whether they would be "inclined to ban" or "inclined to authorize" the initiative if they were on the Board. The mean level of approval for review in the control group was 5.10, where 7 indicated complete agreement with review. Support for review does not indicate rejection, as only 53% of the control group would themselves vote to reject. For ease of comparison, we divide support for review into a Boolean categorization: the {1-5:6-7} split appears to align with base sentiments indicating 47% favorable towards review.

Preferences for review change under some of the conditions. The relevant data treatment increased favorability towards review to 68%, a strongly significant increase. (This treatment effect is weakly significant using the Likert data, and either strong or weak across other Boolean divisions). There was also a significant effect in both groups that were not a recipient but whose data would be transmitted anyway (with and without priming). This points towards H6, since relevant data creates an indirect harm from everyone's data, requiring review (and rejection) to mitigate. The review demanded by sharing without potential reward might

[3] On the questionnaire, the intervention question how participants felt about opt-in, opt-out or board approval regimes in each situation was offered *before* personal preferences to discourage priming. Correlation between regime and personal preferences are not unusually correlated. We present the data in this form for rhetorical clarity.

be tied to questions of fairness. Those categories registered strong sentiments about the unfairness of the deal; approval would mitigate that fairness. A full interpretation would have to disentangle that fairness, which we know is separate from privacy sentiments.

Response about approval does not lend such strong support. Only 46% would personally vote for the initiative if relevant data were used, but this is not a significant difference from the 53% of the control group. Moreover, there is no reduction in approval level if respondents did not have a chance to obtain reduced premiums. For this group, the risk of broader discrimination remain the same but the benefit disappears, so even if their information were not passed along, consumers in this group would suffer economic harms. Our theory of a sophisticated consumer would predict less support: we only find diminished declared approval when respondents were primed to think about price discrimination, when a mere 36% of those who would not receive any benefit would vote for the initiative.

We predict that respondents favor intervention mechanisms and actual intervention more because of indirect effects. Mixed support of this hypothesis does not solidly confirm it, but indicates that we cannot rule it out, either.

4.7 General Discussion

The indirect harm that applies across these conditions is the fear of price discrimination in the car insurance market. Factors such as support for the deal and feelings of fairness vary along these conditions, but the privacy concern remains constant whether the individual's personal information is involved or not. Privacy concern is heightened by market-relevant data (H2). The nature of desired intervention, in some cases, is consistent with a demand for protection against indirect harms beyond personal information dissemination. Taken together, the data suggests a concern about a privacy externality, which is not a function of data collection from individuals, but rather data use by those who collect any data.

The experimental design specifically targets this distinction between data collection and data use. We presented the participants with a tradeoff situation, rather than a generalized survey, then recorded their sentiments about privacy. Instead of attempting to directly measure the value of privacy, which would be entangled in valuations of other experiment-specific variables, we only focused on how respondents felt. Measuring the presence and relative strength of feelings across independent groups allowed us to capture feelings of utility while controlling for the anticipated benefit.

One could argue that, in showing that privacy concern did not change across conditions, we have only detected a constant, latent privacy sentiment. In such a case, distinctions may or may not exist, but the participants failed to discern the relative importance of different treatments. This could be because they did not understand the privacy issues at stake to begin with, but use of a priming condition (H3), confirmed against a manipulation check, suggests that consumers are aware of indirect ramifications. Alternatively, the experimental treatment differences were too subtle. However, treatments were administered to independent groups, and the use of other vehicles such as happiness and fairness metrics allow us to be fairly confident that we have measured valid responses for different treatment groups. While every result observed is not consistent with our set of hypotheses—only partial support was found for H5 and H6, and we cannot explain some of the more complex interactions between treatments—the experiment was designed to be simple, and the analysis of the data was fairly straightforward.

This design offers some external validity to the findings. The experiment does not use specific sets of rewards, and avoids specifying any explicit harms. Privacy sentiments are all relative, so that they can be scaled to other situations. A realistic situation was used to help prompt realistic responses, but nothing about the scenario offered implies that a similar set of incentives in a different context would produce different results.

5. CONCLUDING REMARKS

The nascent field of the economics of privacy requires more empirical information about what consumers value and why. Against a null hypothesis of a consumer that was concerned about all aspects of personal information sharing or, alternatively, was focused exclusively on direct, explicit harms from privacy violations, we proposed a set of hypotheses arguing that consumers think about context and indirect effects. We demonstrated mild or strong support for all of the above hypotheses, and found that consumers are sensitive to context and indirect effects, rather than data collection itself. There are several implications of these results.

First, the privacy concern does not revolve around unitary "atoms" of personal data. This contradicts the assumptions of some models, which assume that units of personal information have intrinsic value. If, as we show above, privacy concerns are the same whether information is shared or not, then relying on separate privacy valuations may not work. This has broader implications for privacy regulation paradigms. If the flow of personal information is not the root of how people think about privacy, then policy solutions that rely on market mechanisms to efficiently control that flow will not function properly. Moreover, the idea of a privacy externality that introduces concerns based on dissemination of other people's information means that personal use of privacy-enhancing technologies will not eliminate the privacy concern. More broadly, information protection regimes should not treat all data as equal.

In fact, the above findings suggest that focusing on the data itself does not address the source of the privacy concern: data *use*. While we can draw no conclusions based on any specific mechanisms of society-wide control, we do find evidence that there is consumer demand for some social control, and that control should focus on data use. This emphasis seems more aligned with approaches like the OECD's Guidelines, which advocate the principles of purpose specification and data limitation [22]. While such principles could be made manifest in the private market, many proposed mechanisms have fallen short [13]. Wathieu proposes market pooling mechanisms to prevent the indirect harms from segmentation [20]; strict regulatory limitations on private use of personal data could also mitigate many of the concerns discussed in this paper.

To understand and model privacy, more information is needed about consumer preferences, beyond "people want privacy."

More sophisticated privacy models require evidence of a sophisticated, economically aware consumer. We present evidence from an experiment that people do behave somewhat rationally when considering realistic privacy situations. We find evidence of a sophisticated consumer that cares about economic context and indirect economic effects.

6. REFERENCES

[1] Agre, P.E. and Rotenberg, M. Technology and privacy: the new landscape. MIT Press, 1997.

[2] Aquisti, A and Grossklags, J. Privacy and Rationality in Individual Decision-Making. IEEE Security and Privacy 3:1 2005

[3] Aquisti, A and Grossklags, J. Uncertainty, Ambiguity and Privacy. Fourth Workshop on the Economics of Information Security (WEIS05) 2005

[4] Berman, J. and Weitzner, D.J. Abundance and User Control: Renewing the Democratic Heart of the First Amendment in the Age of Interactive Media. Yale Law Journal 104:7. 1995

[5] Chellappa, R.K., " Consumers' Trust in Electronic Commerce Transactions: The Role of Perceived Privacy and Perceived Security," under submission

[6] Chellappa, R. K., and Sin, R., Personalization versus Privacy: An Empirical Examination of the Online Consumer's Dilemma. Information Technology and Management, Vol. 6, No. 2-3, 2005

[7] Cranor, L. F., Reagle, J., Ackerman, M. S. Beyond Concern: Understanding net users' attitudes about online privacy. AT&T Labs-Research Technical Report TR 99.4.3, 1999

[8] Deighton, J. Frequency Programs in Service Industries. In Handbook of Services Marketing and Management, Teresa Schwartz and Dawn Iacobucci, eds., Sage, 2000.

[9] Earp, J.B, Poindexter, J.C., Baumer, D.L. Modeling privacy values with experimental economics. Proceedings of the ACM workshop on Privacy in the electronic society(WPES), 2004.

[10] European Community. Official Journal of the European Communities of 23 November 1995 No L. 281 p. 31

[11] Google Privacy Policy. http://www.google.com/intl/en/privacy.html Version 07/01/2004

[12] Friedman, A.A. Rhetoric and the Public Policy of Security. Rump Session Presentation at Financial Cryptography (FC05), 2005

[13] Greenstadt, R, Smith, M.D., Protecting Personal Information: Obstacles and Directions. Fourth Workshop on the Economics of Information Security (WEIS05) 2005

[14] Hann, I.H., Hui, K.L., Lee, T.S. and Png, I.P.L. Online Information Privacy: Measuring the Cost-Benefit Trade-off. Proceedings of the 23rd International Conference on Information Systems, 2002

[15] Hann, I.H., Hui, K.L., Lee, T.S. and Png, I.P.L , Consumer Privacy and Marketing Avoidance. Working Paper, 2005

[16] Hermalin, B. E. and Katz, M.L. Privacy, Property Rights and Efficiency: The Economics of Privacy as Secrecy. University of California Working Paper, 2004.

[17] Huberman, B.A., Adar, E. and Fine, L.R. Valuating Privacy. To appear in IEEE Security and Privacy

[18] Johnson, E.J., Bellman, S., and Lohse, G.L. Defaults, framing and privacy: Why opting in ≠ opting out. Columbia Business School Working Paper (2000).

[19] Johnson, E.J and Goldstein, Daniel. Do Defaults Save Lives? Science 302, 2003

[20] Milne, G.R. and Rohm, A.J. Consumer Privacy and Name Removal Across Direct Marketing Channels: Exploring Opt-in and Opt-out Alternatives. Journal of Public Policy and Marketing, 19:2, 2000.

[21] Odlyzko, A.M. Privacy, economics, and price discrimination on the Internet. in Economics of Information Security, L. Jean Camp and S. Lewis, eds., Kluwer, 2004.

[22] OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. 1980.

[23] Regan, P M. Legislating Privacy. North Carolina Press, Chapel Hill NC, 1995.

[24] Schneier, B. Mitigating Identity Theft. Cnet News, April 14, 2005. Available at http://news.com.com/Mitigating+identity+theft/2010-1071_3-5669408.html

[25] Shostak, A. and Syverson, P. What Price Privacy (and why identity theft is about neither identity nor theft). in Economics of Information Security, L. Jean Camp and S. Lewis, eds., Kluwer, 2004.

[26] Spiekermann, S., Grossklags, J. and Berendt, B. E-privacy in 2nd generation E-commerce: privacy preferences versus actual behavior, Proceedings of the 3rd ACM conference on Electronic Commerce, 2001.

[27] Taylor, C. R., Consumer Privacy and the Market for Customer Information. Rand Journal of Economics 35:4, 2004.

[28] Tversky, A., & Kahneman, D.. Judgment under uncertainty: Heuristics and biases. Science, 185, 1974

[29] Varian, H. R Wallenberg, F. Woroch, G. The Demographics of the Do-Not-Call List. IEEE Security and Privacy 3:1 2005

[30] Wathieu, L. and K. Morris, Meloche Monnex. Harvard Business School Case Nr. 504 008.

[31] Wathieu, L. Marketing and the Privacy Concern. Marketing Sciences, under revision.

[32] Westin, A. Privacy concerns & consumer choice. Technical report, Louis Harris & Associates, Dec. 1998.