Valuing Intrinsic and Instrumental Preferences for Privacy

Valuing Intrinsic and Instrumental Preferences for Privacy

Tesary Lin*

August 26, 2021

Abstract

I empirically separate two components in a consumer's privacy preference. The intrinsic component is a "taste" for privacy, a utility primitive. The instrumental component comes from the consumer's anticipated economic loss from revealing his private information to the firm, and arises endogenously from a firm's usage of consumer data. Combining an experiment and a structural model, I measure the revealed preferences separately for each component. Intrinsic preferences have seemingly small mean values, ranging from \$0.14 to \$2.37 per demographic variable. Meanwhile, they are highly heterogeneous across consumers and categories of data: The valuations of consumers at the right tail often exceed the firm's valuation of individual consumer data. Consumers' self-selection into data sharing depends on the respective magnitudes and correlation between the two preference components, and often deviate from a simplistic "low types are more willing to hide" argument. Through counterfactual analysis, I show how this more nuanced selection pattern changes what firms can infer from consumers' privacy decisions and its implication on effective data buying strategies.

Keywords: privacy, revealed preference, value of data, experiment, selection, bias

*Boston University Questrom School of Business; tesary@bu.edu. I am grateful to Pradeep Chintagunta, Sanjog Misra, Brad Shapiro, and Oleg Urminsky for their advice and support. Thanks go to Eric Anderson, Alessandro Bonatti, Michael Dinerstein, Jean-Pierre Dubé, Andrey Fradkin, Avi Goldfarb, Wes Hartmann, Yufeng Huang, Garrett Johnson, Nitin Mehta, Kanishka Misra, Sarah Moshary, Harikesh Nair, Olivia Natan, Sridhar Narayanan, Aniko Öry, Anita Rao, Robbie Sanders, Avner Shlain, Valeria Stourm, Abigail Sussman, Anna Tuchman, Dana Turjeman, Kosuke Uetake, Ken Wilbur, my two reviewers of the MSI Alden G. Clayton dissertation proposal, and participants at various seminars for their thoughtful comments. This study was approved by the Institutional Review Board at the University of Chicago (IRB18-0234).

1 Introduction

With the arrival of privacy regulations across the globe, companies increasingly need to seek consent from consumers before collecting and processing their personal data. In the EU, General Data Protection Regulation (GDPR) mandates that firms deliver transparent information and seek opt-in consent before data processing.[1] As a result, European websites lost 12.5% of their recorded traffic due to increased consumer vigilance (Aridor et al. 2020). Outside Europe, transparency and consent requirements are incorporated by privacy regulations in many other countries and local states, including California in the US, Brazil, Canada, and India.[2] They are also the backbone of recent industry self-regulation, such as Apple's iOS 14 update that mandates mobile apps to display privacy "nutrition labels" and seek consent before tracking users.[3]

When consent becomes a prerequisite for personal data processing, consumers' preferences for privacy determine what and whose data firms can collect. Consumers exhibit heterogeneity in privacy choices when informed (Goldfarb & Tucker 2012*b*, Varian et al. 2005, Johnson et al. 2020). This heterogeneity poses a potential selection problem in the data shared by consumers. Selection in data and the resulting bias in data-driven insights have been in the spotlight, with examples spanning automatic resume sorting (Cowgill et al. 2020), medical research (Al-Shahi et al. 2005), and public opinion polling.[4]

To better understand consumers' privacy preferences and how they affect the selection pattern in voluntarily shared data, I empirically distinguish between two preferences for protecting privacy. Privacy preferences can emerge because privacy itself is valued as an *intrinsic* right (Warren & Brandeis 1890). They can also arise as an *instrumental* value, the payoff of preventing their private "type" from being revealed through data (Stigler 1980, Posner 1981). Consumers can hold both types of privacy preferences. Intrinsically, people find it creepy to have smart thermostats tracking their activities at home, regardless of whether their behaviors are benign or objectionable (Pew Research Center 2015). Instrumentally, risky drivers may avoid installing telematics devices that allow an insurance firm to monitor their driving habits (Jin & Vasserman 2018, Soleymanian et al. 2019). Although the conceptual distinction of intrinsic and instrumental preferences dates back to Becker (1980), it has drawn little empirical attention thus far.

Empirically separating these preference components is crucial for two reasons. First, it allows us to learn how consumers self-select into sharing their data, and how this selection affects a firm's inferences about consumers. The reasoning "if you've got nothing to hide, you've got nothing to fear" is only valid when consumers harbor a purely instrumental preference for privacy. On the other hand, assuming consumers value privacy solely intrinsically can lead to the misleading conclusion that people who value privacy more are no different from the rest of the population.

[1]https://gdpr-info.eu/recitals/no-39/; https://gdpr-info.eu/recitals/no-32/.
[2]https://piwik.pro/blog/privacy-laws-around-globe/; https://parl.ca/DocumentViewer/en/43-2/bill/C-11/first-reading.
[3]https://developer.apple.com/app-store/user-privacy-and-data-use/
[4]https://www.nytimes.com/2019/07/02/upshot/online-polls-analyzing-reliability.html

Incorrect inferences about consumers who protect their privacy can lead to suboptimal decisions, such as under- or over-targeting, or putting these consumers under excessive scrutiny.

Second, empirically separating these two preferences enables us to understand how consumers' privacy choices respond to policy shocks or changes in firms' data analytic strategies. While the intrinsic preference is a utility primitive, the instrumental preference arises endogenously from how the firm uses consumer data. As such, instrumental preferences can shift with the purpose of data collection, the performance of a firm's model used for processing the data, and what other data the firm already obtains. Separating the instrumental preference from the intrinsic allows us to model these shifts accordingly. Accounting for the endogenous nature of instrumental preference allows us to calculate the equilibrium impact of privacy regulations and changes in firms' data analytic strategies.

To precisely define the constructs that underlie my measurement task, I start by summarizing the conceptual roots of intrinsic and instrumental preferences in the social science literature. Economists traditionally view privacy preferences as instrumentally motivated, derived from how the other party uses the private information. Other disciplines, such as philosophy and law, have long recognized people's need for privacy as an intrinsic value, the same way that they value freedom and autonomy. The intrinsic preference for privacy is a concrete value, not just fear about unforeseeable economic consequences of privacy intrusion. In other words, it is far from the "residual form of instrumental preference." Then I introduce the conceptual framework that builds on Becker's (1980) seminal paper. This conceptual model further clarifies how the two privacy preferences are different from another component in a consumer's utility function—the compensation for sharing data, that is, the "price" that a firm offers to collect data from consumers. Even in cases where data is considered to come "for free," firms still often pay consumers for their data in the form of products or services. Separating the benefits from privacy values is essential for understanding consumer privacy choices and welfare under alternative data collection arrangements.

Based on the framework, I design an experiment that measures revealed preferences for privacy in dollar terms and captures preference heterogeneity. Revealed preferences are solicited by requesting consumers to share their data with a company. To capture the heterogeneity in privacy preferences and its impact on selection, I use a novel two-stage design, which sequentially records consumers' private types and their privacy choices. This design enables me to observe the contents of personal information *even from consumers who decline to share their data*. The experiment generates variation needed to identify my model: (a) the level of instrumental incentives, which separates the two preference components; (b) the amount of compensation, which enables me to calculate the dollar values of privacy preferences. The experiment also contains a conjoint survey, which allows me to counterfactually calculate the value of personal data in the context of price targeting.

I then estimate a structural model to quantify each preference component, taking into account that instrumental preference is endogenous. Intrinsic preferences are highly heterogeneous across both consumers and categories of data. In terms of willingness to accept (WTA), the mean intrinsic valuation ranges from $0.14 to $2.37 across categories of data requested. In comparison, consumers at the 95% quantile value each personal variable from $2.19 to $5.08. To obtain a representative set of data, a firm will need to pay as high as $29.72 per consumer per demographic profile.

For the instrumental preference, consumers' beliefs in the economic consequence of revealing private information are first-order consistent with the actual payoff scheme. This belief consistency is conditional on a setting where the information about data usage is explained in plain language, as is required by most recent privacy regulations. This finding implies that when data usage information is transparently disclosed, the magnitude and heterogeneity of instrumental preferences will match those of the actual payoff scheme, even when the latter changes.

I show that the selection pattern in shared data is determined by the joint distribution of the two preference components. When intrinsic preferences are more heterogeneous and negatively correlated with instrumental preferences, high-type consumers are more likely to self-select *out of* rather than *into* data sharing. In the experiment, this pattern is manifested in the sharing of income. On the other hand, the "low types are more willing to hide" argument is more likely to hold when the two components are positively correlated or when the heterogeneity of instrumental preference dominates. This is true when I examine the sharing of purchase intent data.

By changing consumers' self-selection into data sharing, the co-presence of intrinsic and instrumental preferences changes what the firm can learn from a consumer's sharing decision and the value of acquired consumer data. Through counterfactual analysis, I explore these two aspects in the context of price targeting. The first counterfactual examines the information value of data sharing decisions. If consumers had only instrumental preferences, their decision to not share data would be informative enough about their type, such that the firm could effectively target them using privacy decisions alone. In contrast, I show that adding privacy decisions to the model only improves the firm's targeting slightly, despite its ability to debias the prediction result. Intuitively, adding privacy decisions allows the model to distinguish between sharing and non-sharing consumers, but not consumers inside the non-sharing group. The information value of privacy decisions is smaller when the difference within the non-sharing group is larger than the between-group difference, which is more likely when intrinsic preferences are more heterogeneous.

The second counterfactual conducts cost-benefit analysis for buying additional data. The scale of consumers' intrinsic preferences implies a high cost of buying additional data. The presence of instrumental preferences makes additional data more valuable to the firm by exacerbating the selection bias in existing data. However, it also makes additional data more costly by increasing consumers' total valuation for privacy. The effect of the latter dominates the former in my setting. Due to the heightened cost, the firm faces a quantity-representativeness trade-off when buying consumer data. Prioritizing representativeness is sensible when the direction and degree of consumer self-selection are ex-ante unknown, and can be made more efficient by leveraging the information externality among consumers. Through a simple calculation, I show how the firm can compare sampling and mass-buying strategies, accounting for the fact that different data analytic stages exhibit different degrees of information externality.

My paper contributes to several strands of literature. First and foremost, it builds on Becker's (1980) dual-privacy-preference framework (also see Farrell 2012, Cooper 2017, and Jin & Stiver 2017) by empirically measuring the two components. Moreover, it shows how this empirical separation can help us understand the selection in shared data and its implication on firms' data analytic and collection strategies. In doing so, my paper builds the link between consumers' privacy preferences and the quality of consumer data as firms' input.

Second, my paper builds on existing work that measures revealed privacy preference, including Goldfarb & Tucker (2012*b*), Athey et al. (2017), Kummer & Schulte (2019), as well as Acquisti et al. (2013) and Tang (2019) who provide dollar-value measures. Compared with these papers, mine separates intrinsic and instrumental preferences. Given the endogenous nature of instrumental preference, separating these two components is useful for characterizing equilibrium privacy choices and market outcomes when evaluating new privacy regulations or firms' data analytic strategies.

My paper also contributes to the literature on context-dependent privacy preferences by highlighting how instrumental preferences respond to changes in (perceived) economic consequences of sharing data, such as entities that have data access (Martin & Nissenbaum 2016) and information that changes consumer belief on data usage (John et al. 2010, Athey et al. 2017, Miller & Tucker 2017). As such, it complements the previous literature (Egelman et al. 2009, Acquisti et al. 2012, 2013, Adjerid et al. 2019, Lee 2019), which emphasizes psychological factors that generate context dependence.

Lastly, by discussing how consumers' privacy choices affect firms' inferences and resultant profits in the new policy regime, my paper adds to the research on how privacy regulations influence firms' managerial outcomes, including the effectiveness of advertising (Goldfarb & Tucker 2011, Tucker 2014), funds raised (Burtch et al. 2015), innovation activities (Goldfarb & Tucker 2012*a*, Adjerid et al. 2015), market concentration (Batikas et al. 2020, Johnson et al. 2021), and profits (Marotta et al. 2019, Aridor et al. 2020, Johnson et al. 2020, Ke & Sudhir 2020, Goldberg et al. 2021). Instead of examining the holistic impact of a particular regulation, my paper focuses on one mechanism: How consumers' self-selection into data sharing affects the quality of firms' data-driven decisions. In doing so, I am able to evaluate strategies that allow firms to address the impacts of selection.

2 The Conceptual Framework

2.1 What Are Intrinsic and Instrumental Preferences for Privacy?

Privacy preference is the preference to control information about oneself (Warren & Brandeis 1890, Westin & Ruebhausen 1967, Posner 1981). Different disciplines hold different views about the nature of privacy preference. Philosophers often see privacy as an intrinsic moral value. Privacy is considered "an aspect of human dignity," because it provides personal autonomy and independence (Bloustein 1964, Parent 1983) and enables people to have experiences with spontaneity and without shame (Gerstein 1978). Legal scholars also justify privacy protection on moral grounds (Westin & Ruebhausen 1967, Gavison 1980).

In comparison, economists often think of privacy preference as generated from the need to protect one's private information in market exchanges (Stigler 1980, Posner 1981). In this view, consumers have an instrumental need to protect their privacy because revealing their private information can lead to negative economic outcomes, such as a worse insurance rate or a higher price.

Becker (1980) is the first to propose that privacy preferences have both intrinsic and instrumental dimensions:

> Persons having their privacy invaded would protest not only because they may value their privacy as a "good" but also because privacy raises the probability that their crimes would succeed.

This dual-privacy preference framework is echoed in subsequent conceptual work, including Farrell (2012), Jin & Stivers (2017), and even Posner (2008) who contended that privacy preference is purely instrumental in his earlier work. Farrell (2012) calls the intrinsic preference a "final good" and the instrumental as an "intermediate good." These labels capture the distinct natures of the two preferences. While the former is a utility primitive, the latter is endogenously driven by how the other party uses the private information to change the transactional outcome.

To make the distinction more concrete, consider the following example. A firm wants to collect personal data from consumers in order to improve its price targeting practice. Denote data from consumer i as d_i. I refer to consumers who are more price-sensitive as the *high type*, and the less sensitive as the *low type*. In other words, high-type consumers are those who will get more generous discount offerings if the firm learns their price sensitivities. At the data request stage, the firm may offer compensation to consumers who share their data, denoted as P. Examples of compensation include perks offered to consumers who sign up for a loyalty program, or gift cards for sharing email. This compensation is offered ex ante and does not depend on consumers' private information.

Consumer i decides whether to share their data with the firm ($s_i \in \{0, 1\}$). He has an *intrinsic preference* for privacy c_i, which is a taste for protecting his data. If he shares data ($s_i = 1$), the firm will give him a discount offer conditional on the content of the shared data $T(d_i)$. For consumers who decline to share ($s_i = 0$), the firm forms belief about their average price elasticity, and offers the same discount $T(s_i = 0)$ to everyone in this group. Anticipating how the firm sets the discount offers, consumer i also has an *instrumental preference* for privacy, the expected economic gain from not revealing his type: $\Delta T(d_i) \equiv T(s = 0) - T(d_i)$. He shares data if the privacy cost is offset by the compensation that the firm provides:

$$s_i = 1 \text{ iff } - c_i - \Delta T(d_i) + P > 0. \tag{1}$$

From this example, we can see that intrinsic, instrumental utilities, and the utility from compensation are three distinct constructs. The intrinsic preference does not respond to how the firm uses data to design customized offers. By contrast, the instrumental preference is endogenously driven by how the firm uses data to deliver targeted payoff; thus, it changes with the payoff function $T(.)$. Both intrinsic and instrumental utility are privacy preferences, and they stem from the fact that the firm (or the firm's algorithm) sees the contents of the data. In comparison, utility from compensation does not depend on what the data reveals.

The dual-preference framework is not restricted to cases where high-type consumers benefit while low-types get hurt from sharing their private information. Instrumental preference can also arise from targeting that matches consumers and products. In cases like this, we can still categorize consumers who benefit more (e.g., those who have niche tastes) versus less (e.g., those who are already satisfied with mainstream products) into high and low types. The only difference here is that the instrumental preference may be negative for all consumers: They may all prefer sharing data from an instrumental standpoint.

This framework is different from other categorizations of privacy preferences. For example, Calo (2011) separates privacy harm into subjective and objective. Although the intrinsic preference can be viewed as subjective, the instrumental preference is also a consumer's subjective expectation about the objective market consequences of revealing the personal information. Characterizing this subjective expectation and its consistency with the objective outcome is crucial for understanding consumers' self-selection into data sharing, and is one major component of my empirical analysis.

Do consumers value privacy intrinsically? Writings in the public sphere suggest so. The examples given in a *Wired* article are relatable to many readers (Schneider 2006): "*We keep private journals, sing in the privacy of the shower, and write letters to secret lovers and then burn them. Privacy is a basic human need.*" Meanwhile, the economics literature enumerates various cases of instrumental preferences. Consumers may refrain from sharing their information which indicates that they are an incapable employee, a price-insensitive consumer, or a risky client to be insured. Yet to date, no one has measured how much consumers value privacy intrinsically versus instrumentally.

2.2 Why Measure Intrinsic and Instrumental Preferences?

Empirically separating intrinsic and instrumental preferences is useful for two reasons. First, it allows us to understand why and how privacy preferences change across economic scenarios. While the intrinsic preference is a utility primitive, the instrumental preference is endogenous to subsequent data usage strategies that the firm intends to adopt (including data sales). Thus, it changes with how the firm uses consumer data. Understanding this endogeneity is useful for evaluating the impact of new data usage practices, such as those triggered by new privacy regulations. Since the degree of preference endogeneity depends on the relative magnitudes of the two components, we want to measure them separately.

Recognizing the instrumental preference as endogenous also helps us reconcile previous findings that privacy preferences are context-specific. For example, the fact that privacy choices change with consumers' belief about data usage (John et al. 2010, Athey et al. 2017, Miller & Tucker 2017) is a manifestation of instrumental preference. The pattern that privacy choices and attitudes change with who collects the data and how data are used (Rainie & Duggan 2015, Prince & Wallsten 2020) also reflects the instrumental preference. By distinguishing this economic context-dependence from the psychological one (i.e., preference change with behavioral frames such as anchors), we can better project how privacy preferences change across contexts.

Second, we need to measure the heterogeneity in each preference component separately in order to characterize how consumers self-select into sharing data. Understanding the nature of this self-selection matters when the firm wants to know the characteristics of non-sharing consumers and design a marketing strategy suited to them. It also matters when voluntarily shared personal data are used as input for data-driven decisions. Biased data can lead to biased models, but understanding the nature of this selection bias allows us to design bias correction strategies. Unfortunately, a model that sees privacy preference as monolithic can offer misleading predictions about the selection pattern.

To see why this is true, we can revisit the pricing example. To set a suitable price for non-sharing consumers, the firm needs to conjecture the average price elasticity among them. Suppose poor college students are more sensitive to prices, but they also care about their privacy much more than the rest of the firm's customer base. A model with pure instrumental preference predicts that non-sharing consumers are likely more wealthy and should be charged with a higher price. This is the rationale that underlies the argument "you don't care about privacy if you have nothing to hide." However, due to the higher intrinsic preference among the poor, the non-sharing group may tilt towards more price-sensitive consumers. In other words, the act of withholding data not longer signals the low type when consumers have heterogeneous intrinsic preferences.

More generally, the "nothing to hide" argument is less likely to hold when the intrinsic preference is more heterogeneous than the instrumental, especially when high-type consumers have a high intrinsic preference. On the other hand, this argument may be true when the instrumental preference dominates (a formal characterization of this statement and proof can be found in Appendix A). To know which case is more likely, we need to characterize the heterogeneity of each component and how they correlate with each other.

2.3 Taking the Framework to Data

To proceed with the measurement task, first we need to answer a question: How will we measure the instrumental preference when it is endogenous? Although the instrumental preference itself changes with the economic outcome of revealing private information (hereafter the *instrumental payoff*), the degree to which consumers understand this instrumental payoff—their sophistication—is stable conditional on an information environment. Therefore in the estimation model, the utility primitives are the intrinsic preference and the degree of consumers' sophistication when forming the instrumental preference. In doing so, the model learns whether consumers are naive or rational directly from the data.

3 The Experiment

Empirically separating intrinsic and instrumental preferences is difficult. First, the economic incentive is usually fixed in observational settings, making it infeasible to separate instrumental preferences from the intrinsic. Second, in most observational settings, the request for personal data is bundled with product provision. As a result, the preferences for privacy are confounded with the preferences for products concurrently offered. For example, consumers who keep using Gmail after learning that Google analyzes all their email texts may have either a low preference for privacy or a high valuation of Google's email service. Lastly, both consumers' privacy choices and their private types need to be observed to identify my model; however, privacy choice is precisely the decision to reveal these private types. As long as variation in privacy decisions exists, the data available to the researcher will exhibit selection in a conventional setting.

My experiment includes three features to circumvent these challenges. First, instrumental incentives are turned on or off across treatments. I can thereby measure the intrinsic preferences directly when the instrumental incentives are off, and use the difference between treatments to measure the instrumental preferences. Second, I exclude the confound from product preference by using monetary incentives (which have known values) to compensate for data sharing. Furthermore, the amount of compensation to encourage data sharing varies across treatments, allowing me to measure the dollar values of privacy preferences. To overcome the last challenge, I adopt a two-stage design, where the first stage collects participants' private information, and the second stage solicits revealed preferences for privacy.

3.1 Experiment Design

The experiment uses a survey as an instrument to solicit revealed preference. Specifically, I insert personal questions into an otherwise standard survey. Thus, a participant's decision to share the response to a question indicates his level of privacy cost associated with this personal variable. This technique has been deployed by Acquisti et al. (2012) and Goldfarb & Tucker (2012*b*). Research shows that in the domain of privacy preferences, attitude- and behavior-based measures often disagree (Harper & Singleton 2001, Spiekermann et al. 2001). I focus on revealed preference because it is incentive compatible and more relevant for managerial decisions and policy analysis.

The experiment consists of two stages. In stage one, participants see themselves participating in a market research survey sent by the University of Chicago. The survey includes conjoint questions about smartwatch attributes and about participants' *intent to purchase* a digital device. They are followed by demographic questions, including *gender, age, education level, income, relationship status, whether they have children, zip code,* and *ethnicity*. Each personal question in the first stage includes a "prefer not to say" option. People who find the question too sensitive are thus allowed not to respond rather than forced to fabricate a response. Up to the end of the first stage, consumers are unaware that they will later be requested to share personal data with the firm, thus not actively considering privacy.

Stage one serves two roles. First, it records private information from consumers, including those choosing not to share data in the subsequent stage. This full information allows me to measure heterogeneity in privacy preferences and characterize how the interplay between intrinsic and instrumental motives determines selection in shared data. Second, the conjoint questions provide inputs for calculating the value of data to firms in a pricing context, which becomes the basis for evaluating the data collection and analysis strategies in my counterfactual analysis.

Stage two solicits privacy choices. After finishing the survey, participants navigate to a new screen. Here, they receive a request to share survey responses with a third party, a smartwatch manufacturer that wants to use the data to inform its product-design decision. Participants can choose whether to share each personal variable separately via check boxes.[5] The compensation for data sharing takes the form of a gift-card lottery. Participants are unaware of the possibility of sharing data with the third-party until they answer all questions in stage one. Once they reach the second stage, the "return" button is disabled, preventing them from changing previous responses to facilitate sharing. These two features, along with the "prefer not to say" option in Stage one, are included to ensure responses in the first stage are truthful.

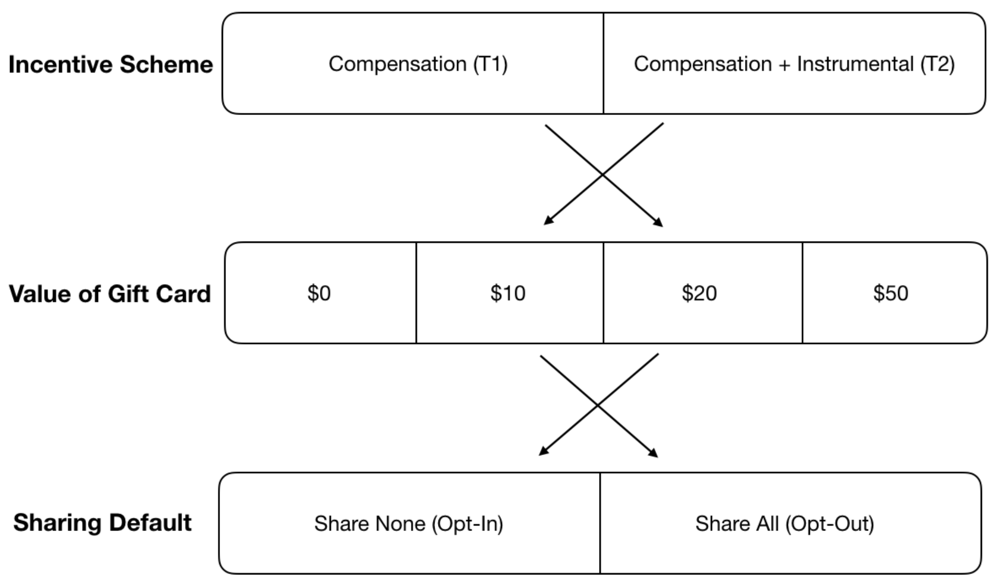
Stage two is also where treatments take place. There are three layers of treatments: the incentive scheme, the amount of compensation, and the sharing default. These treatments are orthogonal to each other (see Figure 1).[6] The first layer varies the incentive scheme:

[5] Only informative responses (i.e. other than "prefer not to say") in Stage one are allowed to be shared in Stage two.

[6] One exception: by design, participants who receive zero compensation do not receive any instrumental incentives.

- Treatment 1 (compensation): The amount of compensation increases proportionally to the amount of data shared and is common across all participants. In particular, sharing one additional personal variable increases the probability of winning the gift card by one percentage point. In other words, the price for data is the same regardless of what the firm learns about the consumer.
- Treatment 2 (compensation + instrumental incentive): A baseline level of compensation exists and takes the same form as in Treatment 1. The baseline amount is then adjusted based on whether *the company perceives* the participant to be a potential customer *from analyzing the data it obtains*. Likely customers receive higher compensation than the baseline, whereas unlikely customers get a cut in the compensated amount. Participants are told the company's target customers are high-income people who intend to buy a digital product, and therefore, they will receive more *if the shared data indicate* they fit this profile.

Figure 1: Treatment Design



Note: The experiment follows a factorial design. Treatments are assigned with equal probability in each layer.

In sum, *privacy choices in Treatment 1 alone identify intrinsic privacy preferences*. Here, the stated purpose of data collection does not imply continuous tracking or any other future interactions with consumers. Moreover, participants did not know about this company before entering the experiment; thus, they are unlikely to anticipate the instrumental consequences of sharing data from interacting with the firm in the future. By contrast, choices in Treatment 2 are motivated by both intrinsic and instrumental preferences. The instrumental preferences are induced by an incentive scheme that depends on a participant's income and product-purchase intent. These two characteristics constitute a consumer's "type" in this experiment. Therefore, the *differential responses between Treatments 1 and 2 identify instrumental preferences for privacy*.

The experiment uses type-dependent monetary compensation instead of personalized product prices to induce the instrumental incentive. Although the latter is more natural, it may not induce variations of instrumental preference in my setup. Given that participants have never interacted with the featured company (it is fictitious), they may not plan to engage in future transactions with this company. In this case, the firm's pricing practices will not matter to them.

The incentive scheme is presented in two parts (see Appendix B for the information participants see in each treatment). The main page concisely explains who collects the data and for what purpose, and how a participant's payoff will qualitatively depend on the data shared. The detailed screen shows quantitatively how the payment will be calculated, and is displayed when a participant clicks the "see details" link. Overall, the incentive scheme display is transparent and clear, similar to the format of most post-GDPR website banners.

The information environment is designed to mimic the consent requirements from recent privacy regulations. For example, both GDPR and CCPA require data controllers and processors to use clear and plain language to describe the purpose of data processing and consumer rights.[7] The "layered approach" of information display that I adopt (with the main screen giving a summary of data usage and a link to details) is recommended by recent regulations to ensure both the completeness of disclosure and clarity.[8]

The other treatments are designed as follows. The second treatment layer changes the value of the gift card (essentially cash) across participants, creating variations for measuring the dollar values of privacy preferences. The third layer varies default choice, set to either sharing all data (opt-out) or sharing none (opt-in). Treatments are assigned with equal probability within each layer.

To measure if participants understand and trust the validity of incentive treatments, I send follow-up questions to participants after they make the data-sharing choices. These questions include their perceived purpose of the study, what determines the amount of expected compensation, the reasons they choose (not) to share the survey responses, and if they prefer a sure reward with the same expected value as the gift-card lottery. Within three days after a participant completes the survey, participation fee, compensation, and instrumental incentives are all paid as described in the study.

[7]https://gdpr-info.eu/recitals/no-39/; https://oag.ca.gov/privacy/ccpa.

[8]See point 69 of the European Data Protection Board's *Guidelines 05/2020 on consent under Regulation 2016/679*.

3.2 Discussion

Using a more controlled setting for my experiment allows me to strengthen the internal validity in a way that a field setting cannot. Most importantly, having a new firm requesting consumer data allows me to measure intrinsic preferences cleanly using the control group. On the contrary, in field settings where the firm has an existing reputation, consumers may have a strong prior on how the firm uses their data, making it nearly impossible to remove instrumental preferences. Replicating my experiment in the field is possible but requires an additional assumption on consumer belief. I elaborate on field replication in Online Appendix C.

The data usage, in particular how the instrumental incentive depends on a consumer's type, is straightforward and explicitly explained in my setting. There are two reasons that an explicit explanation of data usage is necessary. First, by providing a more straightforward explanation, the design makes sure that the measured intrinsic preference is not just a vague expectation about potential instrumental consequences from sharing data. Second, a setting with transparent information about data usage is a scenario made relevant by new regulations. For example, the GDPR states that prior to seeking consent, information about data usage should "be easily understandable for the average person" and disclose all third-parties who will access data (European Data Protection Board, 2020). The California Privacy Rights Act also mandates companies to disclose the purpose of collecting sensitive consumer data. In these regulatory environments, consumers are likely to have a more informed understanding of how their data are used as well as the associated personal consequences.

One caveat of my design is that the instrumental incentive is simple. The data are used for a single purpose and do not change hands. As such, it does not reflect scenarios where consumer data can be used in multiple ways in a complex data ecosystem. In these scenarios, it may be harder for consumers to comprehend data usages compared to my setting.

Another caveat of my design comes from the two-stage setup. Although the first stage is necessary for obtaining the contents of personal data, it also creates attrition prior to treatment. One concern is that participants who stay after the first stage may care less about privacy, which can skew my measurement result downwards. In Section 4, I show evidence that first-stage attrition *caused by privacy concern* is mild.

Due to logistic constraints, the experiment uses a lottery instead of sure rewards for compensation.[9] If participants are predominantly risk-averse, their perceived gain from the gift-card lottery will be lower than its objective expected value, and the estimated dollar value of privacy preferences will be an upper bound of their true valuation. In the follow-up survey, I ask participants whether they prefer the lottery or a sure reward with the same expected value. The response distribution shows that my sample consists of both risk-loving and risk-averse participants, with

[9] The gift cards have fixed face values that are less granular than the amounts of compensation I give.

35% of them preferring the lottery and the rest opting for the sure reward. However, since the question only solicits a yes-or-no answer, I am unable to quantify their risk preferences.

4 Data and Descriptive Evidence

In what follows, I describe the data source and sample characteristics, then present model-free patterns of intrinsic and instrumental preferences. The main analysis focuses on privacy choices in the opt-in regime given its prominent empirical relevance.[10] Data show how consumers share some data while protecting others, how the instrumental incentive changes the composition of consumers that share data, and how this composition shift changes the quality of data shared.

4.1 Data Source and Cleaning

Participants of the experiment come from Qualtrics Panels. Existing work finds that Qualtrics panel is more representative of the population than alternative online panels and students (Heen et al. 2014, Boas et al. 2018), which are popular sample sources for other choice experiments. To further alleviate discrepancies, I apply stratified sampling so that the demographics of participants entering the survey resemble the distribution given by the 2018 US Census. Nevertheless, stratified sampling only accounts for observable differences. To the extent that Qualtrics panel members are more willing to share personal information without anticipating any instrumental consequences, my measurement result provides a lower bound for the population-level intrinsic preferences. Qualtrics provides three demographic variables on the back end, including income, age, and ethnicity. I use these data to validate the truthfulness of responses in the first stage. Not all demographic variables I intend to collect are available through Qualtrics. Therefore, having the first stage is still necessary.

A total of 4,142 participants enter the survey; 3,406 of them proceed to the second stage when treatment occurs. For people who leave the survey upon seeing the request, I code their choices as sharing nothing, regardless of the default condition. Figure D.1 shows the participant attrition throughout the experiment. Among the 18.4% of participants who leave the survey before seeing the treatment, 91% exit before or during the conjoint survey. This pattern indicates that attrition is mainly caused by a lack of interest in the conjoint questions rather than a reluctance to share personal data in the first stage. To prevent treatment contamination, I deduplicate the respondents by IP address. I also exclude respondents whose time spent on the survey or time spent responding to the data-sharing request is in the lowest decile. The cleaned data include 2,583 participants.

[10] Regulations differ in what default action is allowed when firms seek consent. EU laws such as GDPR and ePrivacy Regulation require opt-in consent, while the US adopts a mixed approach, with opt-in consent required for sensitive data and data sales. Regardless of the regulatory requirement, requests effectively operate in an opt-in condition for data not generated or tracked by default, such as survey responses or test results. I compare privacy choices in different consent regimes in Online Appendix H.1.

4.2 Sample Characteristics

Attrition and sample cleaning can change the characteristics of the final sample. Table 1 summarizes the demographics of survey participants in the cleaned sample, and compares them with the 2018 Current Population Survey (CPS) whenever similar statistics are available. Some discrepancies come from differences in counting. For example, the mean age provided by CPS includes juniors (ages 15–18), whereas my sample contains only adults; "black" in my sample includes mixed-race groups, while CPS's definition excludes it. Another difference comes from the fact that some participants choose not to share all demographics during the first stage. As a result, the percentages of different income levels do not sum up to 1, whereas in the census, the disclosure is complete. Compared with the population, participants who finish the survey tend to be female, less educated, and have lower income.

Table 1: Demographics of Experiment Participants (Cleaned Sample)

	Variables	Experiment Sample	2018 Census
	Female	65.31%	50.80%
	Married	47.39%	51.16%
	Have young child	24.78%	–
	Mean age	47.60 (16.89)	45.9 (–)
Education	High school degree or less	47.00%	39.93%
	College degree	40.65%	48.67%
	Master's degree or higher	11.39%	11.40%
Race	White	71.27%	76.60%
	Black	15.37%	13.40%
Annual Household Income	$25,000 or less	21.99%	20.23%
	$25,000 to $50,000	29.54%	21.55%
	$50,000 to $100,000	30.12%	28.97%
	$100,000 or more	13.55%	29.25%
No. Observations		2,583	–

Source of the census data: U.S. Census Bureau, Current Population Survey, 2018 Annual Social and Economic Supplement. "–" indicates that no corresponding statistics are available.
Note: For discrete variables, I collapse their values into larger groups to facilitate the exhibition. Numbers corresponding to the same category may not sum to 1, given that smaller groups are left out and that some participants choose not to respond in the first stage. For continuous variables, I report their mean values with standard deviation in parenthesis.

Purchase intent is one of the consumer types in the instrumental-incentive treatment. It is calculated based on participants' responses to two questions in the first stage: (A) "How likely will you buy a new smartwatch within the next 3 months?" (B) "How likely will you buy any other digital devices within the next 3 months?" Each question uses a 5-point Likert scale. Different answers are then given different scores. For example, "extremely likely" is scored 2, while "extremely unlikely" is scored -2. Purchase intent is then constructed by summing up these two scores; a higher value indicates higher purchase intent. Across participants, the mean purchase-intent score is -0.17, with a standard deviation of 1.72.

4.3 Intrinsic Preferences

Table 2 shows how the inclination to share data varies with the category of data requested and compensation in Treatment 1, where privacy preference is purely intrinsic. Consumers do not want to share personal data when not compensated: The percentages of data shared are all at or below 50% (the indifference benchmark) in the first column of Table 2.

Table 2: Frequency of Data Shared with Intrinsic Utility

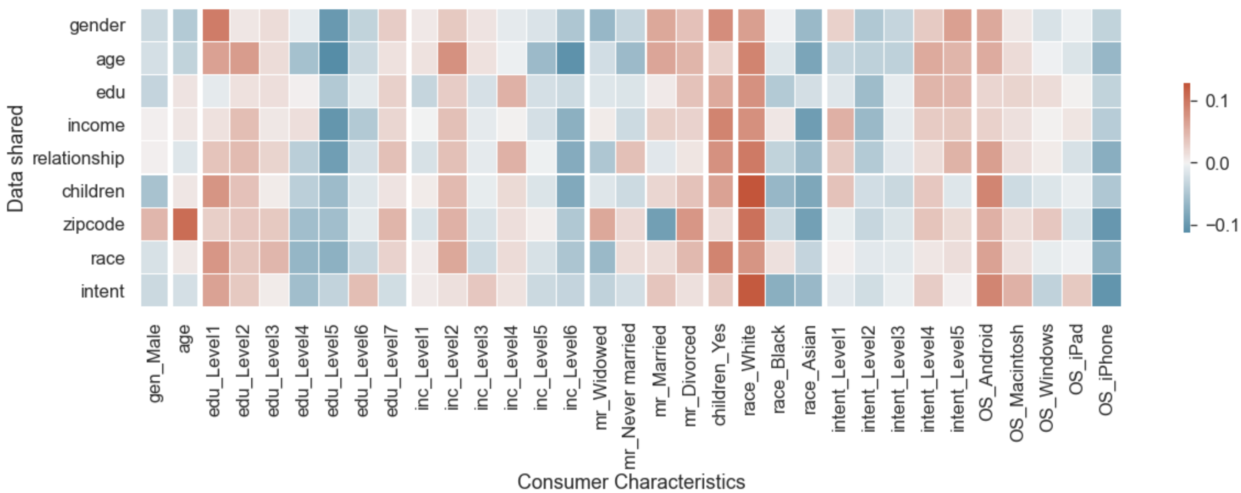
	Category of Data								
Compensation	Gender	Age	Education	Income	Relationship	Children	Zipcode	Race	Purchase Intent
= 0	0.50	0.47	0.43	0.36	0.46	0.29	0.41	0.42	0.43
> 0	0.70	0.68	0.62	0.56	0.66	0.53	0.63	0.63	0.54

Note: "Relationship" corresponds to their responses about marital status. "Children" corresponds to responses to the number of children they have. Among the compensated groups, the value of gift card is $33 on average, with a 1% increase in the possibility of winning for each variable shared.

Compensation is effective in shifting privacy decisions. An average price of 33 cents per variable increases the probability of sharing by about 20% across variables. Data about household income and their children are valued the most, whereas gender is considered the least sensitive. The sensitivity ranking across personal variables remains largely unperturbed regardless of whether data sharing is compensated. Overall, the table shows that participants make attentive trade-offs in the experiment and value various data differently.

Figure 2 unpacks how the intrinsic preferences vary across demographic strata. In the heatmap, a positive correlation (red) indicates people in this strata are more likely to share their data in Treatment 1, which means they have lower intrinsic preferences; a negative correlation (blue) indicates the opposite. Consumers who value privacy intrinsically tend to be more educated and wealthier. This pattern suggests that intrinsic preference is a learned value and, to some extent, a "luxury good." They are also more likely to use iPhone than Android devices, which is correlated with their income status. Across racial strata, minorities are less willing to share their data compared to White participants. This pattern is consistent with the notion that intrinsic preference comes from the fact that privacy ensures personal autonomy, as "autonomy is the counterweight to the 'group rights' already enjoyed by majority populations" (Wright 1999). We caution that the patterns are all correlational, and speculations about causal mechanisms should be taken with a grain of salt.

Figure 2: Correlation between Sharing Decisions and Consumer Characteristics in Control Group



Note: Attribute labels follow the "attribute-value" convention. To minimize clutter, attributes values with a natural order are coded as "Level k," where a lower level indicates a lower value.

While the intrinsic preferences are different across multiple characteristic dimensions, the magnitudes of correlation coefficients are small (between -0.11 and 0.13). As we will see later in estimation results, even for the dimension that shows the biggest differences across strata (e.g., across racial groups), their preference distribution still overlaps with each other a lot. This pattern suggests that the preference heterogeneity is not fully captured by any single dimension. I return to this point in Section 6.

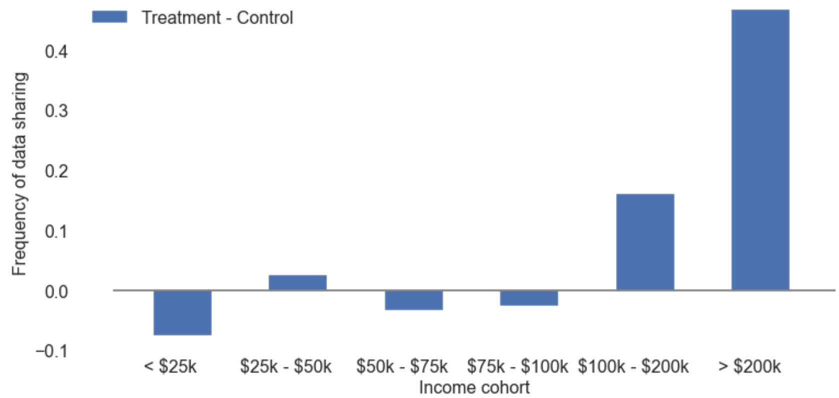
4.4 Instrumental Preferences

The treatment group introduces the instrumental incentive: Participants benefit more if they are perceived as wealthy or intend to buy digital products in the short term (hereafter *high types*). Figure 3 shows how the presence of instrumental incentives influences privacy choices. High-type consumers are more willing to share their data in the treatment than in the control group, whereas the opposite pattern holds for low-type consumers. This pattern indicates that participants are attentive to instrumental incentives. For income sharing, the behavioral differences between the treatment and control groups are overall insignificant. The lack of behavioral difference may be caused by a greater heterogeneity in intrinsic preferences, which makes the utility variation caused by instrumental preference less effective in shifting choices.

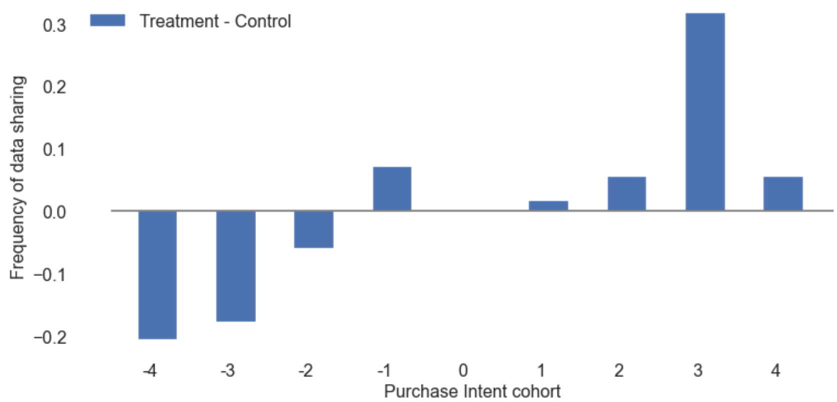
Figure 4 shows how the privacy preferences vary across consumer characteristics, this time focusing on the comparison between intrinsic and instrumental. For income data, high-income people have lower instrumental incentives but greater intrinsic preferences (see the region inside the rectangle in Panel (a)). On the other hand, Panel (b) indicates that the correlations between purchase intent types and each preference component are both positive. Due to correlation across

Figure 3: Difference in Frequency of Data Shared across Incentive Treatments

(a) Income Sharing



(b) Purchase-Intent Sharing



Note: Frequencies in each treatment group are calculated as the proportion of participants who share their data within each cohort (not across).

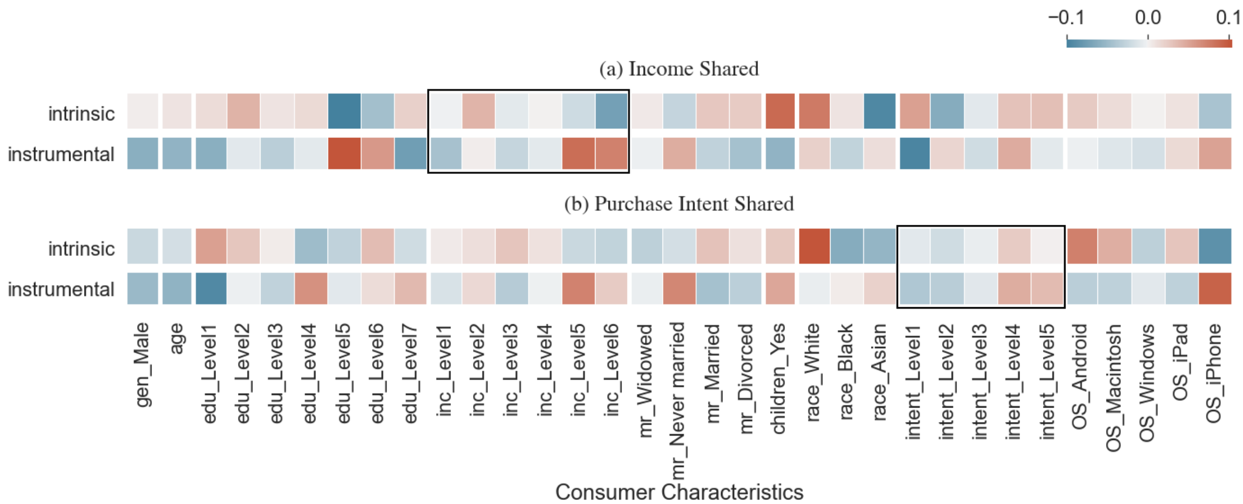
attributes, the heterogeneity in instrumental preferences is also present in other dimensions. For example, more educated participants and iPhone users also have higher instrumental preferences for income.

4.5 Dual Privacy Preferences and Selection in Shared Data

To understand the combined effect of the two preferences on selection, I compare the mean value of consumer attributes reflected by the shared data (from Stage 2) and the true mean among participants (from Stage 1). The comparison is conducted separately for the control and treatment group participants. The former reflects the impact of intrinsic preference alone, while the latter reflects the aggregate impact from both components.

Table 3 displays the t-test statistics for this comparison. For purchase intent sharing, the instrumental incentive makes the shared data feature more high-types than the true data has (see

Figure 4: Correlation between Sharing Decisions and Consumer Characteristics across Treatments



Note: The instrumental preference is reflected by the choice difference between the two treatment groups. Correlation on the second row of each panel is thus calculated as $Corr(Y_t - Y_c, X) = (sd(Y_t) \cdot Corr(Y_t, X) - sd(Y_c) \cdot Corr(Y_c, X)) / sd(Y_t - Y_c)$, where Y indicates privacy choices and X indicates consumer characteristics. Regions inside the black rectangles correspond to consumer types induced by the instrumental incentives.

column 2 of Panel (b)); the difference between the shared and true data is marginally significant at the 0.06 level. This selection pattern is consistent with the prediction offered by the classical economic model. Here, the classical pattern persists because intrinsic preferences are mostly homogeneous among different purchase-intent types.

In contrast, Panel (a) shows that for income sharing, the combination of intrinsic and instrumental preferences does not cause a significant selection. As is shown in Figure 4, the intrinsic and instrumental preferences for income sharing have opposite heterogeneity patterns: Wealthier participants have higher intrinsic but lower instrumental preferences. These two patterns counteract each other, rendering the selection along the income dimension insignificant.

Table 3: t-Test for Equal Means (H_1: E[D | shared] - E[D | true] $\neq$ 0)

(a) Income

	Control	Treatment
t-statistic	-0.969	1.053
p-value	0.333	0.293

(b) Purchase Intent

	Control	Treatment
t-statistic	0.190	1.847
p-value	0.849	0.065

Note: Control = Intrinsic Utility; Treatment = Intrinsic + Instrumental Utility. Shared data are constructed based on consumers' sharing decisions in the second stage; true data refers to all data collected from the first stage.

5 The Structural Model

The structural model is indispensable for my analysis for two reasons. First, it allows me to account for the endogenous nature of instrumental preferences by taking consumers' ability to anticipate the instrumental outcome (i.e., how revealing private information determines their economic payoff) as the utility primitive. Later, by fixing the degree of sophistication while allowing their instrumental preference to change with data usages, I can characterize their privacy preferences in counterfactual regimes. Second, I can estimate preference heterogeneity more efficiently by using a Bayesian structural model. The Bayesian shrinkage prior serves as a regularizer, allowing me to incorporate many covariates in the heterogeneity function without over-fitting the model. Compared with a reduced-form setup, a structural model allows me to specify parameter values that should be the same across sharing decisions, which further improves model efficiency.

5.1 Setup

Consumer i is endowed with a vector of personal data $D_i = [d_{i1}, d_{i2}, \ldots d_{iK}]$; d_{i1} is income, and d_{i2} is purchase intent. His sharing decision is characterized by a vector with equal length S_i: Each entry is an indicator of whether the associated personal variable is shared. For example, $S_i = [0, 0, 1]$ means i shares d_{i3} but not d_{i1} or d_{i2}. Sharing decision S_i brings an intrinsic privacy cost, a type-induced payoff from sharing (if the consumer is in the instrumental treatment), baseline compensation, and a random utility shock:

$$U(S_i; C_i, D_i) = \sum_k - \underbrace{c_k(X) \cdot s_{ik}}_{\text{intrinsic preference}} + 1_{instr} \cdot 1_{k \in \{1,2\}} \cdot \underbrace{\beta \cdot p_i \cdot w_k \cdot \widehat{E}[d_{ik} | S_i, D_i]}_{\text{type-induced payoff}} + \underbrace{\beta \cdot p_i \cdot s_{ik}}_{\text{util from compensation}} + \epsilon_{ik}. \tag{2}$$

$C_i = [c_1, c_2, \ldots c_K]$ is the intrinsic preference for privacy; each c_k can be expanded to a function of observables X. 1_{instr} is the instrumental-treatment indicator. $1_{k \in \{1,2\}}$ selects the data-sharing decisions that are subject to the influence of instrumental incentives. β is the marginal utility of monetary rewards. p_i is the value of gift card multiplied by 1%. w_k is the consumer's expected increase in the percentage winning probability for an adjacent, higher type; this is their *first-order belief*. Meanwhile, $\widehat{E}[.]$ is their *higher-order belief*: the consumer's expectation of the firm's expectation about his type. The baseline compensation is proportional to the amount of data shared, represented by $p_i \cdot s_{ik}$. Lastly, ϵ_{ik} is the random utility shock associated with choice S; $\epsilon_{i1}, \epsilon_{i2}, \ldots \epsilon_{iK} \overset{iid}{\sim} TIEV$.

Belief about a consumer's type depends on not only the contents of shared data, but also potentially the *sharing decision* itself: $\widehat{E}[d_{ik} | s_{ik} = 1, D_i] = d_{ik}$, $\widehat{E}[d_{ik} | s_{ik} = 0, D_i] = \widetilde{d}_k(p_i)$. I let

$\widetilde{d_k}(p_i) = \delta_{k0} + \delta_{k1} \cdot p_i$ to allow for different levels of rationality.[11] If both the firm and consumers are rational, they will expect that consumers who withhold data are more likely to consist of low types as the instrumental incentives increase, reflected by a positive δ_{k1}. If agents have naive beliefs instead, δ_{k1} is zero.

Together, the belief parameters $\{w_k, \delta_{k0}, \delta_{k1}\}$ represent the extent to which consumers understand the actual instrumental payoff, which reflects their degree of sophistication. Separating them from the actual payoff scheme allows the model to capture how the instrumental preference responds to firms' data analytic strategies when the latter is endogenized. In addition, consumers' belief about firm practices when making privacy choices itself is of interest (Stutzman et al. 2013, Ben-Shahar & Chilton 2016, Athey et al. 2017). These parameters are directly estimated in the model.

Estimating heterogeneity in intrinsic and instrumental preferences is key to understanding how consumers self-select into sharing. I characterize heterogeneity by allowing preference parameters to be functions of observables X, including demographics, time entering the experiment, time spent on each question, browser used, and device specifications. In models that allow for heterogeneity in intrinsic preferences, $c_k(X) = c_{k0} + c_{kx} \cdot X$. $\delta_{k0}(X)$, $\delta_{k1}(X)$, and $\beta(X)$ are similarly specified, except that variables in δ_k's do not include income or purchase intent so that the model can remain identified. There is also a "built-in" heterogeneity in instrumental preference, coming from the fact that instrumental incentives vary with consumer types.

Apart from privacy preferences, psychological factors can also affect data sharing choices. First is the default frame. The literature has proposed different mechanisms underlying the stickiness to default, which implies different ways that the default frame and utility parameters interact (Bernheim et al. 2015, Goswami & Urminsky 2016, Goldin & Reck 2018). To be agnostic about the mechanism, I estimate models separately for each default frame. The estimated parameters represent behavioral preferences under each frame, the relevant objects for analyzing firm-side implications of privacy choices. Section 6 focuses on the opt-in regime given the current regulatory focus; a comparison between behaviors in the two regimes can be found in Section H.1. The model also includes a behavioral response term, $m \cdot (p_i \geq 0) \cdot s_i$, to account for a combination of the mere-incentive effect and potential anchoring effects at the start of the survey. The estimation result and interpretation for this term are in Section H.2.

With the specification above, the log-likelihood can be written as the sum of log logit probabilities:[12]

$$LL = \sum_{i=1}^{N} \sum_{k=1}^{K} s_{ik} \cdot (\Delta u_{ik}) - \ln(\exp(\Delta u_{ik}) + 1), \tag{3}$$

[11] A fully parameterized $\widetilde{d_k}(p_i)$ will require strong assumptions on consumer expectation. These additional assumptions include consumer belief about privacy choices among other consumers and how they are correlated with consumer types, as well as their belief about the firm's reasoning. Imposing these assumptions does not make sense, given the premise of this paper is that even a sophisticated firm may not have the correct belief about non-sharing consumers.

[12] Expression (3) follows directly from the likelihood function: $L = \prod_{i=1}^{N} \prod_{k=1}^{K} \frac{(\Delta u_{ik})^{s_{ik}} \cdot (1)^{1-s_{ik}}}{\exp(\Delta u_{ik})+1}$.

where Δu_{ik} is the difference in mean utilities between sharing and not sharing data k, experienced by consumer i (heterogeneity functions are omitted for the clarity of exposition):

$$\Delta u_{ik} = -\underbrace{c_k}_{\text{intrinsic preference}} - 1_{i,inst} \cdot 1_{k\in\{1,2\}} \cdot \underbrace{\beta \cdot p_i \cdot w_k \cdot \left[\delta_{k0} + \delta_{k1} \cdot p_i - d_{ik}\right]}_{\text{instrumental preference}} + \underbrace{\beta \cdot p_i}_{\text{util from compensation}} + m \cdot (p_i \geq 0). \tag{4}$$

5.2 Identification

Coefficients to be estimated include c_k, $w_k, \delta_{k0}, \delta_{k1}$ for $k \in \{1,2\}$, β, and m. Parameters in c_k are identified as the utility intercept of the participants who enter the intrinsic treatment; since treatment is randomly assigned, these coefficients are the intrinsic preferences shared by all participants. Belief parameters are identified from the instrumental treatment. w_k is identified from how *different* types react differently to instrumental incentives. δ_{k0} and δ_{k1} are identified from responses to the instrumental incentives that are *common* across types. In particular, the identification of δ_{k1} comes from the interaction between the instrumental treatment and the amount of compensation. Parameter β is identified through the variation in gift-card values. Given that $\beta \cdot p_i$ is linear, and that multiple gift-card values exist across treatments, m is identified from the different responses to zero and non-zero incentives.

The key parameters in this model consist of the following: intrinsic preference, c_k; first-order belief about the instrumental consequence, w_k; and the sensitivity to income, β. Identification of these primitives allows me to construct consumers' privacy choices under different counterfactual scenarios. In particular, the first-order belief is the component that generates the adverse selection pattern created by instrumental incentives. To see this, note that w_k scales the type-dependent payoff when a consumer chooses to share his data $\widehat{E}[d_{ik}|s_{ik} = 1, D_i]$. In comparison, the higher-order belief $\widehat{E}[d_{ik}|s_{ik} = 0]$ does not affect the selection pattern, given that it is not a function of the consumer's private information. Other parameters in the model are auxiliary: They provide flexibility and absorb confounding factors that may otherwise affect the key parameter estimates. For example, δ_{k0} and δ_{k1} may reflect not only consumers' higher-order belief, but also risk preferences that are common across types. Belief related to further data sharing with other parties, if it exists, will also enter the higher-order belief term.

5.3 Estimation

I estimate the model under a Bayesian framework. I place a horseshoe prior for heterogeneity parameters (Carvalho et al. 2009) and a flat prior for the rest. Horseshoe is a form of continuous shrinkage prior; it regularizes parameters in the heterogeneity functions to avoid model over-fitting.[13] Compared with other shrinkage priors such as Bayesian Lasso, Horseshoe yields estimates closest to results from the Bayes Optimal Classifier. I leave intercepts of the heterogeneity functions unregularized to get unbiased estimates for the function mean.

I place non-negativity constraints on the sensitivity to compensation β, and bound constraints on δ such that they do not exceed the actual distribution support of consumer types. I leave the sign of $c_k(X)$ unrestricted, accounting for the possibility that consumers have a "warm glow" in sharing insensitive data to improve research.

6 Results

Table 4 compares estimation results from models with different heterogeneity specifications. To compare model performance, I calculate the expected log predictive density (ELPD) using the Watanabe-Akaike information criterion (WAIC) approximation; a higher number indicates a better out-of-sample fit. Preference estimates are very different between the model without heterogeneity (Model 1) and the models that allow for heterogeneity in intrinsic preferences (Models 2 to 4). The latter exhibit better fits, as is indicated by higher ELPD values. Meanwhile, allowing for heterogeneity in belief or sensitivity to income does not improve out-of-sample fit: Estimation results are similar across Models 2, 3, and 4, and the ELPD of Model 2 is the highest. Model 2 thus constitutes the basis for the main analysis.

6.1 Intrinsic Preferences are Highly Heterogeneous

Figure 5 shows the predicted WTA heterogeneity distribution associated with intrinsic preferences (calculated as $\frac{c_k(X)}{\beta}$), separately for each personal variable requested. Table 5 summarizes the statistics corresponding to each distribution, and Table E.1 shows credible intervals associated with these estimates. Consumers' WTA are highly heterogeneous. The mean intrinsic preferences for sharing different personal variables range from \$0.14 for gender to \$2.37 for information about their children (in the follow-up survey question, many participants describe the request for information about their children as "irrelevant" and "improper"). In comparison, the 97.5% quantiles are more than twice as large as the mean valuations. The upper-tail values are the prices that firms need to surpass to guarantee a representative dataset. For example, a data collector needs to pay \$3.82 per customer for 97.5% of them to share their income data, and \$5.08 per customer to get 97.5% of purchase-intent data. As a robustness check, Appendix F includes WTA estimates from Model 4, which also allows for heterogeneity in compensation sensitivity. The WTA distribution is similar to the main result produced by Model 2.

[13]Due to regularization, the estimated heterogeneity will be smaller than the heterogeneity displayed in raw data. This is a necessary trade-off to avoid model over-fitting.

Table 4: Intrinsic and Instrumental Preference for Privacy: Estimation Results Comparison

	Model	1. No Heterogeneity		2. Heterogeneous c		3. Heterogeneous c & δ		4. Heterogeneous c & β	
		mean	95% CI	mean	95% CI	mean	95% CI	mean	95% CI
intrinsic	c_{income}	0.57	[0.43, 0.70]	0.91	[0.59, 1.32]	0.93	[0.58, 1.51]	0.93	[0.60, 1.39]
	c_{intent}	0.55	[0.41, 0.70]	0.83	[0.42, 1.32]	0.84	[0.38, 1.38]	0.87	[0.41, 1.44]
	c_{gender}	0.02	[-0.12, 0.15]	0.19	[-0.16, 0.66]	0.24	[-0.16, 0.95]	0.20	[-0.20, 0.75]
	c_{age}	0.06	[-0.09, 0.20]	0.26	[-0.09, 0.73]	0.29	[-0.16, 0.91]	0.28	[-0.09, 0.82]
	$c_{education}$	0.37	[0.23, 0.51]	0.62	[0.33, 1.05]	0.65	[0.29, 1.29]	0.65	[0.29, 1.24]
	$c_{relationship}$	0.20	[0.06, 0.33]	0.50	[0.12, 1.01]	0.55	[0.11, 1.23]	0.50	[0.16, 1.04]
	$c_{children}$	0.74	[0.61 , 0.88]	1.11	[0.79, 1.46]	1.09	[0.71, 1.51]	1.10	[0.75, 1.55]
	c_{zip}	0.29	[0.16, 0.43]	0.56	[0.23, 1.07]	0.60	[0.18, 1.22]	0.61	[0.19, 1.13]
	c_{race}	0.29	[0.16, 0.42]	0.60	[0.29, 1.10]	0.65	[0.26, 1.26]	0.65	[0.30, 1.33]
instrumental	w_{income}	2.00	[0.15, 3.87]	2.12	[0.11, 3.99]	2.02	[0.14, 3.92]	1.90	[0.04, 3.88]
	w_{intent}	2.63	[1.07, 3.88]	1.94	[0.38, 3.76]	1.97	[0.29, 3.77]	1.90	[0.35, 3.70]
	$\widetilde{\delta}_{income,0}$	0.05	[-0.19, 0.29]	0.05	[-0.19, 0.28]	0.05	[-0.19, 0.28]	0.05	[-0.19, 0.29]
	$\widetilde{\delta}_{income,1}$	0.05	[-0.19, 0.29]	0.04	[-0.19, 0.28]	0.05	[-0.19, 0.29]	0.04	[-0.19, 0.28]
	$\widetilde{\delta}_{intent,0}$	0.08	[-0.35, 0.39]	0.06	[-0.35, 0.38]	0.07	[-0.36, 0.38]	0.06	[-0.34, 0.39]
	$\widetilde{\delta}_{intent,1}$	-0.05	[-0.36, 0.31]	-0.05	[-0.36, 0.32]	-0.05	[-0.37, 0.31]	-0.04	[-0.34, 0.32]
sensitivity to compensation	β	0.13	[0.07, 0.21]	0.15	[0.07, 0.24]	0.15	[0.06, 0.24]	0.15	[0.07, 0.25]
log posterior		-8015	[-8022,-8010]	-7476	[-7540,-7407]	-7433	[-7501, -7352]	-7525	[-7588, 7450]
ELPD$_{WAIC}$		6384	[6358 , 6410]	6460	[6431 , 6489]	6365	[6337, 6394]	6455	[6427, 6484]

Note: Variables are normalized using the Gelman method before estimation. Wherever heterogeneity is allowed, the table displays estimates on the intercept term only. The same seed is used for estimating different models. The model directly estimates $\widetilde{\delta_{ik}} \equiv \beta \cdot w_k \cdot \delta_{ik}$ instead of δ_{ik} for numerical stability. The distribution of δ_{ik} is later backed out from posterior draws.

Table 5: Posterior Predicted Distribution of WTA in Intrinsic Preference

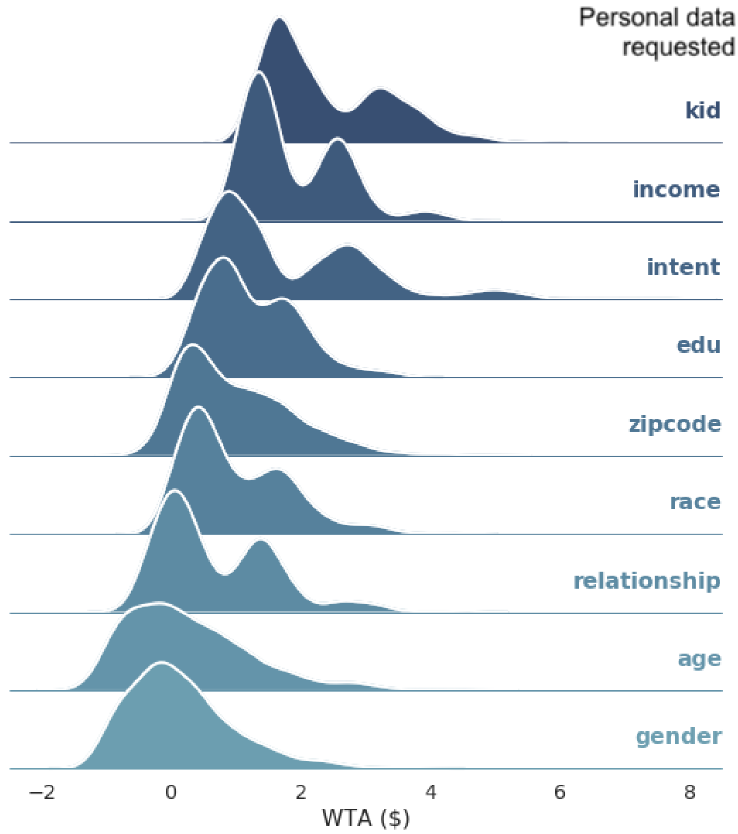
	mean	median	2.5%	97.5%
children	2.367	2.069	1.220	4.311
income	1.870	1.546	0.944	3.823
intent	1.825	1.352	0.398	5.078
education	1.228	1.051	0.228	2.845
zipcode	0.985	0.800	-0.157	2.916
race	0.980	0.737	-0.066	2.945
relationship	0.687	0.390	-0.448	2.894
age	0.260	0.084	-1.064	2.718
gender	0.142	0.006	-1.043	2.187

Note: A wider spread of distribution indicates higher preference heterogeneity.

Consumers with high intrinsic preferences tend to be male, less wealthy, and non-white. In particular, Appendix G shows the bimodal pattern in Figure 5 is mainly driven by the heterogeneity across racial groups. Younger consumers are less willing to disclose their educational background, but have otherwise similar intrinsic preferences compared with their senior counterparts.

Are these privacy-preferences high or low? One way to gauge the magnitude of intrinsic preferences is by calculating the WTA for a profile, which is essentially a bundle of different

Figure 5: Posterior Predicted Density of WTA in Intrinsic Preference



data. A simple calculation shows that the WTA for sharing the whole demographic profile has a mean of \$10.34 and a 97.5% quantile of \$29.72.[14] For more sensitive data such as browsing and location histories, the WTAs are possibly higher. Another way to gauge the magnitude of privacy preferences is by comparing them with the firm's willingness to pay for these data. This comparison is presented in Section 7.2, where I calculate the firm's valuation of personal data under different data collection strategies.

6.2 Consumer Belief on Instrumental Payoff are First-Order Consistent

As utility primitive, the first-order belief scales the magnitude of instrumental preference relative to the actual payoffs for a given information environment. In the experiment, a consumer whose type is one tier above can increase his probability of winning by 2 percentage points if he discloses his type to the firm. Mapped to the model, it means consumers' first-order beliefs are accurate if w equals 2. Column 2 of Table 4 shows that consumers' beliefs about w_{income} and w_{intent} are correct

[14]By simply summing up WTAs across individual personal variables, this calculation ignores potential substitution or complementarity across personal data.

on average. This result implies the magnitude of instrumental preferences will match the actual payoff even when the latter becomes endogenous.

Consumers' higher-order beliefs about the payoff of withholding data are much noisier, reflected by wide credible intervals for δ. This pattern makes sense, given that consumers need to conjecture the firm's and other consumers' reasoning as well as other consumers' privacy preferences when forming this belief.

Overall, the belief estimates represent the level of consumer sophistication in making privacy choices *when fully informed*, as is required by GDPR and other similar regulations. My estimates suggest that with a transparent information environment, consumers are able to engage in strategic reasoning when making data-sharing decisions, and their beliefs are accurate to the first order. In alternative policy regimes where firms can obfuscate information of their data usage, consumers' beliefs may be further away from actual firm practices.

6.3 Selection Depends on the Magnitudes of Intrinsic and Instrumental

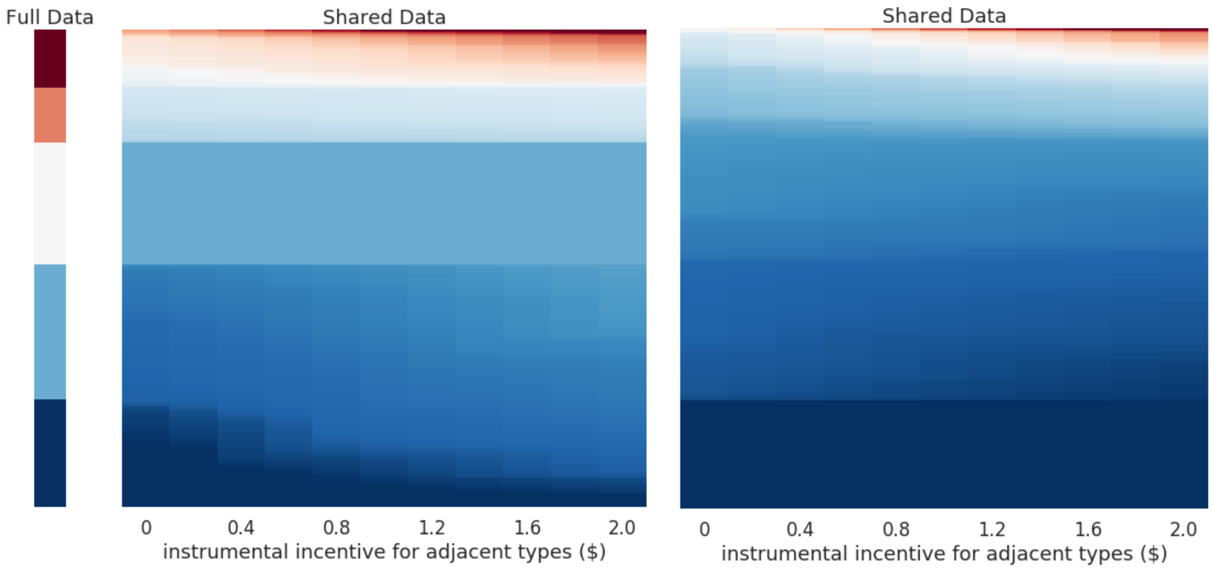
How do the relative magnitudes of the two preference components change the selection in shared data? How does the selection then affect the firm's view on its consumers? To answer these questions, I simulate consumers' sharing decisions under a range of instrumental payoffs, effectively changing the magnitudes of instrumental preferences while holding fixed their intrinsic preferences. Then I examine how different data imputation methods deployed by the firm leads to different inferences on consumer characteristics, and how these inferences compare with the true characteristics distribution. The first method imputes missing data using the median of observed data, corresponding to a view that consumers care about privacy only intrinsically and people who share and who withhold their data have similar characteristics. The second method imputes missing data using the minimum of observed data, corresponding to a view that privacy concerns are purely instrumental. This exercise is performed in the case of income sharing.

Figure 6 compares the distributions of *full* and *shared data* across instrumental incentives and imputation methods. The mean instrumental incentive among consumers ranges from 0 to 2 dollars; the latter is chosen to match the mean intrinsic preference for income. As instrumental incentive increases, the composition of consumers sharing their data tilts towards high-income cohorts, indicated by the expansion of red and the shrinkage of blue regions from left to right in both Panel (a) and (b). However, the firm always overestimates the proportion of low-income consumers, often in (a) and excessively so in (b). This bias is driven by the fact that high-income consumers have higher intrinsic preferences for privacy and are more willing to refrain from sharing. This selection bias is not fully offset by instrumental preference even when the mean instrumental and intrinsic preferences match. In Panel (a), the imputation value is median income among shared data, which is still lower than the population-level average. In Panel (b), the incorrect view that "low types are more willing to hide" exacerbates the bias in the firm's view about consumers.

Figure 6: Full vs. Shared Data across Ranges of Instrumental Incentive

(a) Impute Missing Data with Median

(b) Impute Missing Data with Minimum



In sum, taking a monolithic view about consumers' privacy preferences can result in misleading inferences about consumers and subsequent managerial decisions. Instead, firms need to either learn the joint distribution of their consumers' privacy preferences, or adopt data analytic strategies that are robust to the underlying selection pattern.

7 Counterfactuals

The co-presence of intrinsic and instrumental preferences changes how firms and researchers should conduct inference on non-sharing consumers. Section 7.1 shows that by incorporating privacy choices into the model, firms can still learn about the average type of non-sharing consumers, but may not drastically improve targeting accuracy when the non-sharing consumers consist of various types. The co-presence of these two preferences also changes the marginal value of data collection. As we will see in Section 7.2, the scale of intrinsic preferences often renders a high cost to collect data. As a result, buying additional data from consumers is profitable only when the data buying strategy leverages information externality. Through counterfactual simulation, Section 7.2 illustrates where the information externality kicks in and how the dual privacy preferences change the data collection calculus.

I demonstrate the inference and data collection strategies in the context of price targeting (e.g., Dubé & Misra 2021). Participants' conjoint responses in the first stage of the experiment allow me to calculate their *actual* willingness-to-pay for the product attributes featured in the conjoint. For a given product configuration, I can then construct the demand and the *actual* profit function that the focal firm faces, as well as the *actual* optimal prices that the firm should charge. Given a data buying strategy, the privacy preference estimates allow me to simulate data that consumers share with the firm, and the optimal prices that the firm *infers* based on shared data. I can then evaluate data analytic and collection strategies by comparing firm profits under each strategy.

The specifics of the simulation go as follows. I take a choice scenario featured in the conjoint survey to serve as the market environment (Scenario 3) and the product that the firm sells (Option C; see Figure 7). Based on market statistics, I assume the marginal cost of a smartwatch is \$50.[15] To construct data available to the firm, I simulate 300 privacy choice draws given each compensation level for data, then construct a shared dataset separately for each draw. If a consumer decides not to share variable k, the value of k is left empty. Firm data also contain a privacy choice indicator for each data-sharing decision, which equals 1 when the consumer chooses not to share k, and 0 otherwise. I assume the firm imputes missing variables using mean values among the shared data and takes competitors' prices as given when doing price optimization.[16]

Figure 7: Screenshot of the Conjoint Task and Focal Product Used for Price Optimization

If you want to buy a smartwatch and these are the available options, which one will you choose? (Scenario 3/7)

Product	A	B	C	D
Fitness Tracking	Activity	Activity + heart rate	Activity + heart rate	Activity + heart rate
Voice Control	Yes	Yes	Yes	No
Mobile Payment	No	No	Yes	Yes
Encryption	TLS 1.0	SSL 3.0	SSL 3.0	TLS 1.0
Login Option	Pin, pattern	Pin, pattern	Pin, pattern	Pin, pattern, face
Price	\$299	\$149	\$199	\$249

○ Product A ○ Product B ○ Product C ○ Product D

○ None of the above

Note: Highlights are added to illustrate the focal product used for the counterfactual. They were not present in the actual experiment.

[15]This amount is the average of the estimated production cost of Apple Watch (\$83.70) and Fitbit Flex (\$17.36). See https://www.forbes.com/sites/aarontilley/2015/04/30/the-apple-watch-only-costs-83-70-to-make/#6e981e8d2f08, and https://electronics360.globalspec.com/article/3128/teardown-fitbit-flex.

[16]If a consumer chooses not to share the choice task responses, the pricing model's outcome variable is missing. I assume the firm imputes the missing outcome using observed conjoint choices from consumers who are demographically similar in the shared data. A real-world analog is the look-alike model commonly adopted in the industry.

7.1 What Do Privacy Choices Reveal?

Sometimes firms can observe privacy choices made by consumers. Examples include the "prefer not to say" response in surveys, and the "do not track" header generated by users who decline to leave their digital traces to third-party advertisers.[17] What do these privacy choices reveal? We have seen from Section 6.3 that a simplistic "nothing to hide" argument can be misleading. Meanwhile, early draft forms of CCPA prohibit firms from setting different prices to consumers based on their privacy choices alone. The question remains whether such pricing practice benefits firms and how it impacts consumers.

In what follows, I examine what firms can learn by adding privacy choices to their model. To this end, I compare the performances of two pricing models, taking the data shared by consumers as given. In the *no-indicator* model, the firm sets prices based on the content of shared data but not their privacy choices. In the *with-indicator* model, the firm sets prices based on both the content of shared data and consumers' privacy choices. The metrics for evaluating pricing performances are

$$\Delta\pi_{no-indicator} = \pi\left(P_{d_0}(d_0)\right) - \pi\left(P_d(d)\right),\ \Delta\pi_{with-indicator} = \pi\left(P_{d_0}(d_0)\right) - \pi\left(P_{d+h}(d+h)\right),$$

where d_0 is the full consumer data, d is the data shared to the firm, and h is the privacy-choice indicator.[18] $P_x(y)$ is the firm's price targeting model trained using data x and deployed using data y. Here, x and y are always the same, but we will relax this condition in Section 7.2. Given the actual demand, profit π is determined by the firm's pricing strategy, which in turn depends on the data and the model.

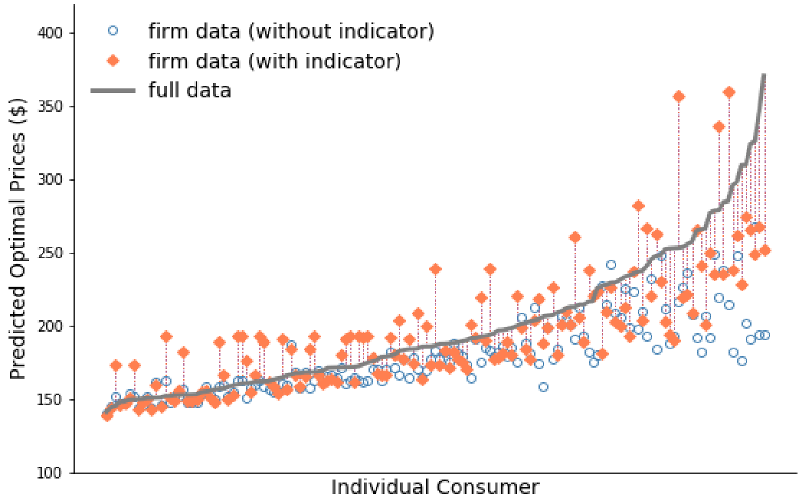
Figure 8 compares the predicted optimal prices using firm data and the actual optimal prices based on full data. Adding privacy choices to the model debiases optimal price prediction. The mean of actual optimal prices is \$199.05. The mean price received by consumers under the *with-indicator* model is \$194.26; in comparison, the mean price under the *no-indicator* model is \$179.22. One implication is that by comparing predictions from models with and without the privacy choice indicator, the firm can learn about the direction and magnitude of selection in the shared data.

Interestingly, Figure 8 also shows that predictions offered by the *with-indicator* model are not always more accurate. The *with-indicator* model surpasses the alternative model when predicting prices for consumers who have high valuations for the product, but performs worse at the other end of the spectrum. Table 6 compares the profit losses from using each pricing model. Adding privacy choices improves overall targeting performance, but only to a small extent (column 1). The performance gain comes mainly from a better calibration of prices offered to privacy-sensitive consumers (column 3). In comparison, the prediction accuracy for consumers who already share

[17] https://www.eff.org/issues/do-not-track.

[18] Here I take d as actual data shared in the experiment's second stage, instead of simulating counterfactual datasets. The latter is only necessary when evaluating data buying strategies and would add to the sampling error.

Figure 8: Inferred Optimal Prices with or without Privacy Choice Indicator



lots of data can suffer (column 2), because privacy choices add little explanatory power for the price sensitivity of these consumers.

Table 6: Profit Loss When Using Firm Data ($/1000 Consumers)

Model	Consumer Subset					
	All consumers		Share all data		Share no data	
no-indicator	2,441	[917, 5,113]	2,348	[926, 5,721]	2,492	[1,229, 4,186]
with-indicator	2,384	[957, 5,229]	2,405	[877 , 5,193]	2,419	[1,138, 3,854]

Note: This table reports posterior mean estimates, with 95% credible intervals in brackets. Profit loss is calculated as $\Delta\pi = \pi\left(P_{d_0}(d_0)\right) - \pi\left(P_d(d)\right)$; a lower number indicates a better performance.

Why does adding privacy choices debias the model while not necessarily improve individual predictions? Intuitively, privacy choices capture systematic differences in price sensitivity *between* consumers who share and who withhold their data, but not the heterogeneity in price sensitivity *within* the withholding consumers. The former reflects the impact of selection, while the latter is more useful for targeting. With greater heterogeneity in intrinsic preferences, consumers who decline to share data are more likely to have heterogeneous price sensitivities, and privacy choices will be less useful for improving targeting.

Should regulations allow price discrimination based on privacy choices? Since adding these choice variables improves aggregate-level inference and informs data buying at no additional costs, firms will want to allow prices to vary with privacy choices when allowed. The impact of such practice on consumers depends on the direction of selection bias in the shared data. In this example, we see that consumers on average receive higher prices, thus consumer surplus decrease. However, this pattern may not hold for other applications.

7.2 How to Buy Data from Consumers?

To evaluate data buying strategies, I need to first calculate the value of consumer data shared with the firm under different levels of compensation. Estimating the value of consumer data is challenging. Conceptually, it depends on what other data are available and is model- and application-specific. Computationally, estimating the value of data requires simulating many datasets and estimating a separate model for each data to smooth out the noises in data sharing choices. With a single dataset, the noises inherent in discrete privacy choices would render the value estimate imprecise and unusable. The simulation requirement brings a heavy computational burden when the researcher searches over different compensation levels for a desirable data buying strategy. The computational problem is further exacerbated when the researcher seeks to incorporate the endogenous instrumental privacy preference into the data-sharing simulation.

Below, I provide one way to calculate the value of data to the firm given the model and application domain. Let d be the data available to the firm under a specific level of compensation, and d_0 the full consumer data. The value of current data compared to full data is calculated as the profit difference when using respective datasets:

$$\Delta\pi_{total} = \pi\left(P_{d_0}(d_0)\right) - \pi\left(P_d(d)\right),$$

where a smaller difference indicates that d is more valuable. In a policy regime where consent is needed for data processing, a firm normally only observes the shared data d. However, it may still learn about the value of full data via a data broker or intermediary that observes d_0.

I then decompose $\Delta\pi_{total}$ into two parts. The first part indicates the value of data for *estimating the model*. The second part indicates its value for *profiling consumers*: that is, gathering individual consumer profiles to deploy the model, taking model parameters as fixed:

$$\Delta\pi_{total} = \Delta\pi_{model} + \Delta\pi_{profile};$$

$$\Delta\pi_{model} = \pi\left(P_{d_0}(d)\right) - \pi\left(P_d(d)\right);\ \Delta\pi_{profile} = \pi\left(P_{d_0}(d_0)\right) - \pi\left(P_{d_0}(d)\right). \tag{5}$$

Simulating d at each compensation level requires the intrinsic and instrumental preference estimates. In particular, the instrumental preference should be calculated by combining the endogenous instrumental incentive and the consumer belief estimation results. In the pricing context, the instrumental incentive is the expected payoff difference when a consumer shares versus withholds his data. Previous estimation results show that consumers are first-order rational when making data sharing decisions. However, there is no sufficient evidence that they also conduct higher-level reasoning. With this evidence, I assume consumers have first-order rational beliefs. To further simplify the analysis, I focus on the case where the firm has trained its pricing model using representative data from other customers. In other words, I only calculate the value of shared data for profiling when consumers have both intrinsic and instrumental preferences. Taking the pricing model as given, consumer i expects to receive the following prices when he withholds or shares data k:

$$E[P_i|s_{ik}=0]=\overline{P}_{i'};\text{ and } E[P_i|s_{ik}=1,d_{ik}]=\overline{P}_{i',\forall d_{i'k}=d_{ik}}.$$

Here, i' denotes all other consumers in the market; $\overline{P}_{i'}$ is the mean price for all other consumers, and $\overline{P}_{i',\forall d_{i'k}=d_{ik}}$ is the average price for all other consumers with the same attribute d_{ik}. Given that consumer i can always choose the outside option when the price is too high, his instrumental preference is the difference in log sums:

$$E[\Delta U]=\frac{1}{\beta_i}\left[\log(1+\exp(v_i-\beta_i\overline{P}_{i'}))-\log(1+\exp(v_i-\beta_i\overline{P}_{i',\forall d_{i'k}=d_{ik}}))\right], \tag{6}$$

where β_i is i's price sensitivity and v_i is his valuation for the product.

The value of consumer data to firms. As a starting point, I show the value of data when consumers only have intrinsic preferences. This situation may occur when consumers who receive sharing requests do not experience the direct economic impact from firms' data analysis. For example, Nielson maintains a consumer panel and sells the data to other firms for analysis, but these firms' focal customers may not overlap with the panel. The first two rows of Table 7 show the value of shared data at different levels of compensation when consumers value privacy only intrinsically. Having to seek consent results in a profit loss of $1,440 per thousand customers when no compensation is given, which is 3% of the total profits that the firm could have obtained using full data. Having incomplete data in the model building stage contributes to 61.4% of the total profit loss, whereas having incomplete data for profiling contributes to the other one-third.

Table 7: Profit Loss from Using Currently Shared vs. Full Data ($/1k customers)

Privacy Preference	Role of Data	Compensation for Sharing Data ($ per variable)					
		0		1		2	
Intrinsic	Model+Profile	1,440	[657, 3,040]	1,126	[536, 2,400]	883	[382, 2,193]
	Profile	556	[379, 716]	465	[317, 636]	380	[218, 592]
Intrinsic+Instrumental	Profile	862	[799, 892]	857	[779, 892]	852	[748, 890]

Note: This table reports posterior mean estimates, with 95% credible intervals in brackets. Total profit loss is calculated as $\Delta\pi_{total}=\pi(P_{d_0}(d_0))-\pi_d(d)$; profit loss associated with profiling is $\Delta\pi_{profile}=\pi(P_{d_0}(d_0))-\pi(P_{d_0}(d))$. A smaller difference indicates the current data is more valuable.

In comparison, the last row of Table 7 shows that when consumers have instrumental preference, the loss from not obtaining the full data is around twice as large as when they only have intrinsic concerns. The larger magnitude is driven by a more severe sample bias in the shared data. However, compensation for data sharing is less effective in overcoming instrumental incentives, as is indicated by smaller value differences among shared data across compensation levels. The reason is that the instrumental preference is large in magnitude for this setting, which ranges from \$20 to \$50 for each data-sharing decision.

To analyze the case where consumers have instrumental preferences when shared data is also used for modeling, one needs to solve the full equilibrium, because the pricing model and the data shared now depend on each other. This task is computationally daunting, as each iteration to search for equilibrium requires simulating multiple dataset draws and computing the pricing model for each draw. Based on the results from the intrinsic-only case in Table 7, I conjecture that when consumers have both intrinsic and instrumental preferences, the economic loss of having incomplete data for modeling is much larger than its impact solely on profiling.

Evaluating data buying strategies. Although consumer data has value, it may not be high enough to justify the price paid to overcome consumers' privacy preferences. To see if a data buying strategy is worth adopting, I compare the firm's WTP for buying data and consumers' WTA for sharing data. If the firm's WTP is lower than most consumers' WTA, matching the price for data to consumers' WTA will lead to a loss in profits; thus, no additional data buying should take place.

Given existing data d, the firm's WTP for obtaining d_0 is calculated as the profit difference divided by the unit difference between the two datasets:

$$WTP_{firm} = \frac{\Delta\pi}{N \cdot \overline{K}.}$$

$\Delta\pi$ can be either $\Delta\pi_{total}$ or one of its subparts from (5), depending on the data buying strategy considered. N is the number of consumers from whom the firm wants to collect data. $\overline{K}$ is the average number of variables withheld per consumer in dataset d. The resulting WTP is the break-even price that the firm is willing to offer per consumer and personal variable.[19]

To start, consider the *mass-buying* strategy, where the firm buys additional data from all its customers and uses the complete data for both modeling and profiling. Consider the case with only intrinsic preferences. On average, a consumer withholds 5.31 variables without compensation. With mass buying, WTP_{firm} = \$1.440/5.31 = \$0.27. In comparison, consumers' mean WTA ranges from \$0.14 to \$2.37. This gap between firm and consumer valuations indicates that mass buying is not a viable strategy. Although we do not know the exact firm WTP when consumers have both types of privacy preferences, Table 7 shows that consumer WTA grows much faster than firm WTP when instrumental preference is present, making mass-buying even less attractive.

[19]This metric represents the average value rather than the marginal value of data. However, tracing out the marginal value along different compensation levels is computationally demanding, for the reasons described at the beginning of Section 7.2. Assuming the marginal value of data decreases with the volume of data already available, the average value calculated above will serve as a lower bound for the marginal value evaluated at d.

Given the magnitude of consumers' privacy preferences, the firm will need to leverage information externality to design a more efficient data buying strategy (Acemoglu et al. 2019, Choi et al. 2019, Bergemann et al. 2020). Recall that consumer data are valuable for both modeling and profiling. Data imposes information externality when used for modeling, but not when used for profiling. Intuitively, when data are used to learn a systematic mapping between consumer characteristics and optimal prices, data coming from one consumer also improves the inferred optimal prices for other consumers. On the other hand, knowing the characteristics of a consumer does not tell the firm about the characteristics of others.

Knowing this, the firm can instead randomly sample a subset of consumers, and buy complete profiles from them to train a high-quality model. Is this strategy profitable? Prioritizing model building is sensible, since two-thirds value of obtaining full data comes from model improvement. Meanwhile, the model may not perform well when the dataset is small. The sampling strategy is worthwhile only when the subset yields approximately the value of the full data, without being too large such that the cost of data buying becomes too high. In other words, a profitable sampling percentage r should satisfy the following inequality:

$$(1.44 - 0.556)/r \geq 2.37 \times 5.31,$$

where the left-hand-side is the value of consumer profile for modeling, and the right-hand-side is the minimum that the firm needs to pay to each consumer to acquire their complete profile. Solving this inequality yields $r \leq 7.02\%$. In reality, the sampling percentage required to achieve an adequate model performance will increase with model's complexity and the underlying heterogeneity that the model intends to capture.

As regulations increasingly require firms to seek consumer consent before collecting data, firms face a quantity-representativeness trade-off when evaluating data buying strategies. Existing literature on optimal information acquisition often abstracts away from the fact that sample selection in shared data can change its value. Through this stylized analysis, I highlight the importance of considering both the quantity and representativeness of data in information acquisition design. With mass-buying, the firm seeks to collect data from more consumers, but since it cannot afford to offer a price that overcomes the preferences of most privacy-sensitive consumers, the collected data can be very biased. Sampling allows the firm to acquire high-quality, representative data but in smaller batches. A smaller but more representative data can yield better insights in many settings, especially when the selection pattern among consumers is ex ante unclear due to the heterogeneous intrinsic preferences.

8 Conclusion

Privacy choices are motivated by both intrinsic preference—a taste for privacy, and instrumental preference—the expected change in payoffs from disclosing one's private information relevant to the specific market environment. While the intrinsic preference is a utility primitive, the instrumental preference is endogenous to how the firm uses consumer data to generate targeting outcomes. Separating these two preference components can help us understand how consumers self-select into sharing data, and how this selection pattern reacts to changes in the firm's data utilization strategies. Ultimately, understanding the selection in voluntarily shared data is crucial for obtaining valid insights when collecting and analyzing consumer data.

By separating intrinsic and instrumental motives using experimental variation, I establish the following findings. Consumers' WTA corresponding to intrinsic preferences are heterogeneous and skewed to the right: The mean valuation for sharing a demographic profile is \$10, while the 97.5% quantile is \$30. When information on data usage is transparently delivered, consumer beliefs about the instrumental consequences are correct to the first order. The direction and magnitude of selection in shared data are determined by the heterogeneity and correlation of the two preference components. Firms and researchers can adopt several strategies to account for the impact of privacy-induced selection when making inferences and decisions. They can run an experiment to measure the joint preference distribution among consumers, and use this information to understand the types of consumers who withhold data. Alternatively, they can adopt strategies that are agnostic about the preference distribution. Ex ante, they can allocate resources to buying a more representative dataset rather than simply increasing its volume where information externality exists. Ex post, incorporating privacy choices into models can effectively debias the inference and prediction results.

Measuring intrinsic and instrumental preferences separately is useful for welfare analysis. Existing empirical work that evaluates privacy regulations mostly focuses on their impacts on firms instead of consumers because the former is easier to measure. In this paper, I provide an empirical framework to measure consumers' privacy preferences, and show that separating the two components can help us understand how privacy preferences respond to policy changes. I also show that intrinsic and instrumental preferences have distinct welfare implications: While the former implies a pure loss of consumer surplus caused by data collection, the latter often leads to surplus transfer among consumers and firms in the market. In doing so, my analysis complements existing work that analyzes policy impacts and provides a building block for a more comprehensive welfare analysis. Nevertheless, a formal consumer welfare analysis needs to go beyond the findings established in this paper. Quantifying changes in consumer welfare under new privacy regulations provides an exciting avenue for future research.

References

Acemoglu, D., Makhdoumi, A., Malekian, A. & Ozdaglar, A. (2019), 'Too much data: Prices and inefficiencies in data markets', *NBER Working Paper No. 26296* .

Acquisti, A., John, L. K. & Loewenstein, G. (2012), 'The impact of relative standards on the propensity to disclose', *Journal of Marketing Research* **49**(2), 160–174.

Acquisti, A., John, L. K. & Loewenstein, G. (2013), 'What is privacy worth?', *Journal of Legal Studies* **42**(2), 249–274.

Adjerid, I., Acquisti, A. & Loewenstein, G. (2019), 'Choice architecture, framing, and cascaded privacy choices', *Management Science* **65**(5), 2267–2290.

Adjerid, I., Acquisti, A., Telang, R., Padman, R. & Adler-Milstein, J. (2015), 'The impact of privacy regulation and technology incentives: The case of health information exchanges', *Management Science* **62**(4), 1042–1063.

Al-Shahi, R., Vousden, C. & Warlow, C. (2005), 'Bias from requiring explicit consent from all participants in observational research: prospective, population based study', *The BMJ* **331**(7522), 942.

Aridor, G., Che, Y.-K., Nelson, W. & Salz, T. (2020), 'The economic consequences of data privacy regulation: Empirical evidence from gdpr', *Available at SSRN 3522845* .

Athey, S., Catalini, C. & Tucker, C. (2017), 'The digital privacy paradox: Small money, small costs, small talk', *NBER Working Paper No. 23488* .

Batikas, M., Bechtold, S., Kretschmer, T. & Peukert, C. (2020), 'European privacy law and global markets for data', *Available at SSRN 3560392* .

Becker, G. S. (1980), 'Privacy and malfeasance: A comment', *Journal of Legal Studies* **9**(4), 823–826.

Ben-Shahar, O. & Chilton, A. (2016), 'Simplification of privacy disclosures: An experimental test', *The Journal of Legal Studies* **45**(S2), S41–S67.

Bergemann, D., Bonatti, A. & Gan, T. (2020), 'The economics of social data', *Cowles Foundation Working Paper* .

Bernheim, B. D., Fradkin, A. & Popov, I. (2015), 'The welfare economics of default options in 401 (k) plans', *American Economic Review* **105**(9), 2798–2837.

Bloustein, E. J. (1964), 'Privacy as an aspect of human dignity: An answer to dean prosser', *NYUL rev.* **39**, 962.

Boas, T. C., Christenson, D. P. & Glick, D. M. (2018), 'Recruiting large online samples in the united states and india: Facebook, mechanical turk, and qualtrics', *Political Science Research and Methods* pp. 1–19.

Burtch, G., Ghose, A. & Wattal, S. (2015), 'The hidden cost of accommodating crowdfunder privacy preferences: A randomized field experiment', *Management Science* **61**(5), 949–962.

Calo, M. R. (2011), 'The boundaries of privacy harm', *Indiana Law Journal* **86**(3), 8.

Carvalho, C. M., Polson, N. G. & Scott, J. G. (2009), Handling sparsity via the horseshoe, *in* 'Artificial Intelligence and Statistics', pp. 73–80.

Choi, J. P., Jeon, D.-S. & Kim, B.-C. (2019), 'Privacy and personal data collection with information externalities', *Journal of Public Economics* **173**, 113–124.

Cooper, J. C. (2017), 'Separation anxiety', *Va. JL & Tech.* **21**, 1.

Cowgill, B., Dell'Acqua, F., Deng, S., Hsu, D., Verma, N. & Chaintreau, A. (2020), 'Biased programmers? or biased data? a field experiment in operationalizing ai ethics', *Available at SSRN 3615404* .

DellaVigna, S. (2009), 'Psychology and economics: Evidence from the field', *Journal of Economic Literature* **47**(2), 315–72.

Dubé, J.-P. & Misra, S. (2021), 'Personalized pricing and consumer welfare', *Available at SSRN 2992257* .

Egelman, S., Tsai, J., Cranor, L. F. & Acquisti, A. (2009), Timing is everything? the effects of timing and placement of online privacy indicators, *in* 'Proceedings of the SIGCHI Conference on Human Factors in Computing Systems', ACM, pp. 319–328.

Farrell, J. (2012), 'Can privacy be just another good', *Journal on Telecommunications and High Technology Law* **10**, 251.

Gavison, R. (1980), 'Privacy and the limits of law', *The Yale Law Journal* **89**(3), 421–471.

Gerstein, R. S. (1978), 'Intimacy and privacy', *Ethics* **89**(1), 76–81.

Goldberg, S., Johnson, G. & Shriver, S. (2021), 'Regulating privacy online: An economic evaluation of the gdpr', *Available at SSRN 3421731* .

Goldfarb, A. & Tucker, C. (2011), 'Online display advertising: Targeting and obtrusiveness', *Marketing Science* **30**(3), 389–404.

Goldfarb, A. & Tucker, C. (2012*a*), 'Privacy and innovation', *Innovation policy and the economy* **12**(1), 65–90.

Goldfarb, A. & Tucker, C. (2012*b*), 'Shifts in privacy concerns', *American Economic Review* **102**(3), 349–53.

Goldin, J. & Reck, D. (2018), Optimal defaults with normative ambiguity, *in* 'AEA Papers and Proceedings', Vol. 108, pp. 98–102.

Goswami, I. & Urminsky, O. (2016), 'When should the ask be a nudge? the effect of default amounts on charitable donations', *Journal of Marketing Research* **53**(5), 829–846.

Harper, J. & Singleton, S. (2001), 'With a grain of salt: What consumer privacy surveys don't tell us', *Available at SSRN 299930* .

Heen, M., Lieberman, J. D. & Miethe, T. D. (2014), 'A comparison of different online sampling approaches for generating national samples', *Center for Crime and Justice Policy* **1**, 1–8.

Hsee, C. K. (1996), 'The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives', *Organizational Behavior and Human Decision Processes* **67**(3), 247–257.

Hsee, C. K. & Zhang, J. (2010), 'General evaluability theory', *Perspectives on Psychological Science* **5**(4), 343–355.

Jin, G. Z. & Stivers, A. (2017), 'Protecting consumers in privacy and data security: A perspective of information economics'.

Jin, Y. & Vasserman, S. (2018), 'Buying data from consumers: The impact of monitoring programs in u.s. auto insurance', *Working Paper* .

John, L. K., Acquisti, A. & Loewenstein, G. (2010), 'Strangers on a plane: Context-dependent willingness to divulge sensitive information', *Journal of Consumer Research* **37**(5), 858–873.

Johnson, E. J., Bellman, S. & Lohse, G. L. (2002), 'Defaults, framing and privacy: Why opting in-opting out', *Marketing Letters* **13**(1), 5–15.

Johnson, G. A., Shriver, S. & Du, S. (2020), 'Consumer privacy choice in online advertising: Who opts out and at what cost to industry?', *Marketing Science* **39**(1), 33–51.

Johnson, G., Shriver, S. & Goldberg, S. (2021), 'Privacy and market concentration: Intended and unintended consequences of the GDPR', *Available at SSRN 3477686* .

Kahneman, D. (1979), 'Prospect theory: An analysis of decisions under risk', *Econometrica* **47**, 278.

Ke, T. T. & Sudhir, K. (2020), 'Privacy rights and data security: Gdpr and personal data driven markets', *Available at SSRN 3643979* .

Kummer, M. & Schulte, P. (2019), 'When private information settles the bill: Money and privacy in google's market for smartphone applications', *Management Science* .

Lee, Y. (2019), 'Revealed privacy preferences: Are privacy choices rational?', *Working Paper* .

Marotta, V., Abhishek, V. & Acquisti, A. (2019), 'Online tracking and publishers' revenues: An empirical analysis'.

Martin, K. & Nissenbaum, H. (2016), 'Measuring privacy: An empirical test using context to expose confounding variables', *Columbia Science and Technology Law Review* **18**, 176.

Miller, A. R. & Tucker, C. (2017), 'Privacy protection, personalized medicine, and genetic testing', *Management Science* pp. 1–21.

Palmeira, M. M. & Srivastava, J. (2013), 'Free offer≠ cheap product: A selective accessibility account on the valuation of free offers', *Journal of Consumer Research* **40**(4), 644–656.

Parent, W. A. (1983), 'Privacy, morality, and the law', *Philosophy & Public Affairs* pp. 269–288.

Posner, R. A. (1981), 'The economics of privacy', *American Economic Review* **71**(2), 405–409.

Posner, R. A. (2008), 'Privacy, surveillance, and law', *The University of Chicago Law Review* **75**(1), 245–260.

Prince, J. & Wallsten, S. (2020), 'How much is privacy worth around the world and across platforms?', *Available at SSRN* .

Rainie, L. & Duggan, M. (2015), 'Privacy and information sharing', *Available at: http://www.pewinternet.org/2016/01/14/2016/Privacy-and-Information-Sharing* .

Shampanier, K., Mazar, N. & Ariely, D. (2007), 'Zero as a special price: The true value of free products', *Marketing Science* **26**(6), 742–757.

Soleymanian, M., Weinberg, C. & Zhu, T. (2019), 'Sensor data, privacy, and behavioral tracking: Does usage-based auto insurance benefit drivers?', *Marketing Science* **38**(1).

Spiekermann, S., Grossklags, J. & Berendt, B. (2001), E-privacy in 2nd generation e-commerce: Privacy preferences versus actual behavior, *in* 'Proceedings of the 3rd ACM conference on Electronic Commerce', ACM, pp. 38–47.

Stigler, G. J. (1980), 'An introduction to privacy in economics and politics', *Journal of Legal Studies* **9**(4), 623–644.

Stutzman, F., Gross, R. & Acquisti, A. (2013), 'Silent listeners: The evolution of privacy and disclosure on facebook', *Journal of Privacy and Confidentiality* **4**(2).

Tang, H. (2019), The value of privacy: Evidence from online borrowers, Technical report, HEC Paris Working Paper.

Thaler, R. (1980), 'Toward a positive theory of consumer choice', *Journal of Economic Behavior & Organization* **1**(1), 39–60.

Tucker, C. E. (2014), 'Social networks, personalized advertising, and privacy controls', *Journal of marketing research* **51**(5), 546–562.

Urminsky, O. & Kivetz, R. (2011), 'Scope insensitivity and the 'mere token' effect', *Journal of Marketing Research* **48**(2), 282–295.

Varian, H., Wallenberg, F. & Woroch, G. (2005), 'The demographics of the do-not-call list', *IEEE Security and Privacy* **3**(1), 34–39.

Warren, S. D. & Brandeis, L. D. (1890), 'Right to privacy', *Harvard Law Review* **4**, 193.

Westin, A. F. & Ruebhausen, O. M. (1967), *Privacy and freedom*, Vol. 1, Atheneum New York.

Wright, J. (1999), 'Minority groups, autonomy, and self-determination', *Oxford Journal of Legal Studies* **19**(4), 605–629.

A Consumers' Selection into Sharing Depends on Joint Distribution of Privacy Preferences

Define the means and covariances of preference components: $E[c_i] = \mu_c, Var[c_i] = \sigma_c^2; E[-T(d_i)] = \mu_t, Var[-T(d_i)] = \sigma_t^2; Corr(c_i, -T(d_i)) = \rho$. Note that $\Delta T(d_i) = T(s = 0) - T(d_i)$, where $T(s = 0)$ does not vary across consumers. Therefore, $Var[\Delta T(d_i)] = \sigma_t^2$ and $Corr(c_i, \Delta T(d_i) = \rho$. σ_c^2 and σ_t^2 respectively represent the heterogeneity of the intrinsic and instrumental preference components. Denote the total preference for privacy as g_i. Then,

$$Corr(g_i, \Delta T(d_i)) = Corr(c_i + \Delta T(d_i), \Delta T(d_i)) = \frac{Cov(c_i+\Delta T(d_i), \Delta T(d_i))}{\sqrt{Var[c_i+\Delta T(d_i)] \cdot Var[\Delta T(d_i)]}} = \frac{\rho\sigma_c+\sigma_t}{\sqrt{\sigma_c^2+\sigma_t^2+2\rho\sigma_c\sigma_t}}. \quad \text{(A.1)}$$

$Corr(g_i, \Delta T(d_i))$ captures the degree to which privacy decisions can be explained by the instrumental preference $\Delta T(d_i)$. Because a one-to-one mapping exists between instrumental preference and a consumer's type (conditional on a fixed offer to non-disclosing consumers), $Corr(g_i, \Delta T(d_i))$ is a direct assessment of the information value of non-sharing decisions for inferring consumer types. The following conclusions hold:

1. $Corr(g_i, \Delta T(d_i)) > 0$ iff $\rho + \frac{\sigma_t}{\sigma_c} > 0$: Self-selection goes in the same direction as predicted by a model with pure instrumental preference if either the correlation between the two components is positive, or the heterogeneity in intrinsic preference is not large enough to overcome the negative correlation.

2. $Corr(g_i, \Delta T(d_i))$ increases with $\frac{\sigma_t}{\sigma_c}$, and strictly increases with $\frac{\sigma_t}{\sigma_c}$ if $|\rho| < 1$: Privacy choice is more indicative of a consumer's type d_i when the heterogeneity in intrinsic preference is smaller.

The conclusions above hold regardless of the level of $T(s = 0)$. In particular, consumers need not have rational expectations, such that their beliefs about $T(s = 0)$ are consistent with the actual transfer that the firm gives to consumers who withhold their data. By the same token, firms need not have correct inferences about consumers who choose not to share data. In other words, the conclusions above are robust to scenarios where firms actively experiment or where information is inadequate for consumers or firms to form rational beliefs. They also remain valid when compensation for data sharing is present.

Note that σ_c and σ_c specifically refer to the heterogeneity across the relevant "type". For example, if the firm aims to target consumers based on their income cohorts, it is the heterogeneity of intrinsic and instrumental preferences across different income cohorts that matter.

B Displayed Compensation Schedules across Treatments

Figure B.1: Displayed Compensation Schedule: Intrinsic Treatment

(a) Main Screen

You will have the chance to win a $20 gift card if you choose to share your responses with Odde, our corporate partner. Odde is a high-end smart device manufacturer; it hopes to use the survey data to inform product development. To encourage participants to share their feedback, **it decides to increase the probability of winning for participants who share more information** (see details).

You can choose what information to share with Odde by selecting or deselecting the boxes below. Note that only the questions that you previously chose to give an informative answer to (i.e. not stating "prefer not to say") can potentially be shared. *Any information that you choose not to share with Odde will not be obtained by the company.*

☐ Choice task responses	☐ Ethnicity	☐ Age
☐ Income	☐ Marital status	☐ Gender
☐ Zipcode	☐ Education	☐ Kids at home

(b) Details Screen

A participant's winning probability is calculated by the following formula:

Probability of winning = number of boxes checked × 1%

For example, if you decide to share your responses to 5 questions that you previously gave, your probability of winning will be 5%.

Figure B.2: Displayed Compensation Schedule: Instrumental Treatment

(a) Main Screen

You will get the chance to win another $50 reward if you choose to share your responses with Odde, our corporate partner. Odde is a high-end smart device manufacturer; it hopes to use the survey data to inform product development.

Your probability of getting the reward will increase with the amount of information that you share. Meanwhile, Odde is designing a new smartwatch geared towards tech-savvy, high-income consumers, and wants to get more feedback from this group of people. As a result, it chooses to assign higher winning probabilities to participants who fit into this profile. In particular, **if it infers you to be wealthy or likely to purchase a smartwatch in the near future, the probability of you winning the reward will increase substantially** (see details).

You can choose what information to share with Odde by selecting or deselecting the boxes below. Note that only the questions that you previously chose to answer (i.e. not stating "prefer not to say") can potentially be shared and therefore be displayed. *Any information that you choose not to share with Odde will not be obtained by the company, and therefore will not be used for determining the winning probability.*

- ☐ Choice task responses
- ☐ Education
- ☐ Ethnicity
- ☐ Marital status
- ☐ Gender
- ☐ Kids at home
- ☐ Income
- ☐ Age

(b) Details Screen

Your winning probability is determined both by the baseline probability and by the adjustment terms. The baseline winning probability is calculated as follows:

Baseline probability of winning = Number of boxes checked × 1%

This baseline probability is then adjusted to encourage response sharing from the customer group that Odde intends to serve, as shown in the following chart:

Income	< $50,000	$50,000 – $75,000	> $75,000
Adjustment	**-2%**	**Unchanged**	**+2%**
Plan to purchase a smartwatch in the next 3 months	Somewhat or extremely unlikely	Neither likely nor unlikely	Somewhat or extremely likely
Adjustment	**-2%**	**Unchanged**	**+2%**

For example, if you have checked 5 boxes, then your baseline winning probability will be 5%. In addition, if the information you share indicates that your annual income is between $75,000 and $100,000, but you are unlikely to buy a smartwatch in the short run, then your final probability of winning will be 5 + 2 - 2% = 5%. The final winning probability never goes below zero.

Any information that you choose not to share with Odde will not be accessed by the company, and therefore will not be used to adjust your winning probability. Meanwhile, *Odde might still be able to use the information that you choose to share (e.g. zipcode, age, education) to infer your income level and your willingness to purchase.*

C Replicating the Experiment in the Field

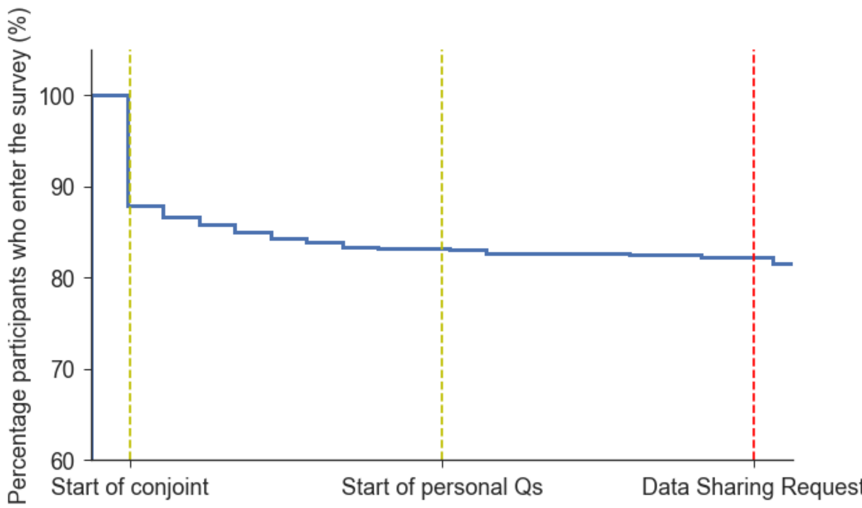
Firms and researchers can replicate my experiment to measure consumers' privacy preferences in the field. To measure the selection in shared data, having a "ground truth" dataset is necessary. In my experiment, this is achieved by having a first stage. When such design is infeasible, the ground truth data can be obtained by having a treatment group in which consumers receive enough compensation so that everyone chooses to share. Alternatively, distribution-level statistics about the relevant type (price sensitivity, risk type, etc.) may be available from a market research company or a government agency (e.g., the census bureau). Next is including treatments to induce exogenous variations of instrumental preferences. For example, if a firm intends to use consumer data for designing customized coupons, it can vary the depth of coupons across treatment arms and inform consumers about the change.

One challenge brought by the field setting is that instrumental preference is hard to be eliminated. This is because consumers' beliefs about the consequences of revealing their personal information typically anchor on the firm's routine practices of using the data. Suppose a supermarket chain asks its customers for data and promises not to use their data for business purposes. Without additional legal guarantees, such a promise will not have commitment power: Users may still expect the supermarket owner to use these data to customize coupons and promotions.

Fortunately, as long as variation in actual instrumental payoffs exists, we can still separate the two preference components by leveraging an additional assumption on consumer belief. In particular, assume the degree to which consumers internalize the instrumental payoff is stable. Consumers' privacy choices across treatments allow us to back out changes in instrumental preferences and compare them with changes in actual instrumental payoffs. By comparing the two, we can estimate how consumers internalize the actual instrumental consequences when forming privacy preferences. With the additional assumption, we can then calculate privacy preferences and data sharing choices in a hypothetical scenario where the instrumental preference is zero, thus backing out the intrinsic preference among consumers.

D Attrition

Figure D.1: Percentage of Participants Remained Throughout the Survey



E Credible Intervals for Intrinsic Preference Estimates (WTA)

Table E.1: Posterior Estimates of Mean and Standard Deviation of the Intrinsic WTA

(a) WTA Mean

	mean	95% CI
income	1.870	[1.012, 3.518]
intent	1.825	[0.981, 3.534]
gender	0.142	[-0.285, 0.709]
age	0.260	[-0.172, 0.805]
education	1.228	[0.619, 2.337]
relationship	0.687	[0.249, 1.454]
children	2.367	[1.337, 4.523]
zipcode	0.985	[0.450, 1.992]
race	0.980	[0.437, 2.008]

(b) WTA Standard Deviation

	mean	95% CI
income	0.906	[0.379, 1.840]
intent	1.337	[0.702, 2.615]
gender	0.929	[0.438, 1.965]
age	1.078	[0.536, 2.173]
education	0.805	[0.330, 1.602]
relationship	0.998	[0.477, 1.973]
children	1.001	[0.465, 1.990]
zipcode	0.982	[0.455, 1.953]
race	0.906	[0.406, 1.801]

F Intrinsic WTA Estimates with Heterogeneous Sensitivity to Income

As a robustness check, I calculate consumers' WTA distribution corresponding to Model 4, which allows consumers to have heterogeneous preferences in both the intrinsic value for privacy and monetary compensation. The estimated sensitivity to income is similar across consumers. The median sensitivity is 0.16; for consumers at the bottom 2.5% quantile, $\beta = 0.13$, while for the top 2.5% quantile, $\beta = 0.18$. Table F.1 reports the posterior distribution of intrinsic WTA from Model 4. Compared to the main estimates, the WTA estimates from Model 4 are slightly more heterogeneous for high-value variables and less so for low-value ones. That said, the two sets of estimates are similar both qualitatively and quantitatively.

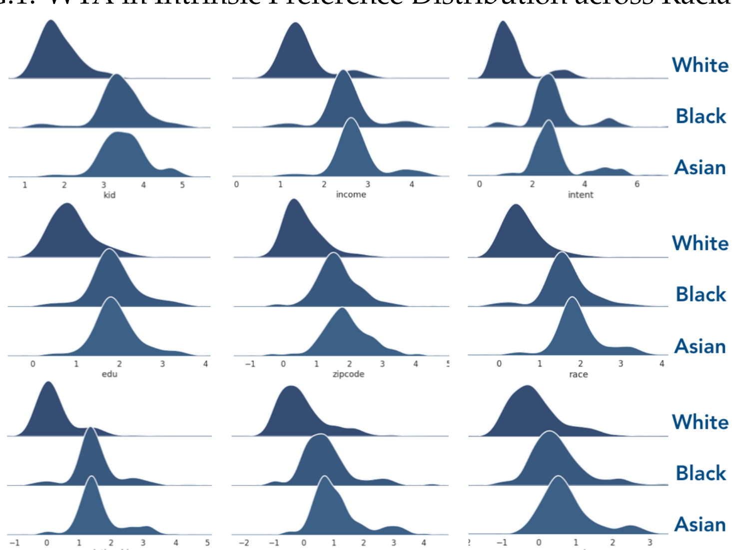
Table F.1: Posterior Distribution of WTA in Intrinsic Preference (with Heterogeneous Sensitivity to Income)

	mean	median	2.5%	97.5%
kid	2.300	2.032	1.059	4.590
income	1.823	1.547	0.801	3.964
intent	1.788	1.275	0.342	5.232
edu	1.220	1.021	0.225	3.061
zipcode	0.980	0.747	-0.115	3.039
race	0.976	0.745	-0.059	2.978
relationship	0.709	0.405	-0.361	2.913
age	0.286	0.076	-0.944	2.723
gender	0.164	-0.011	-0.910	2.187

Note: Numbers in this table refer to statistics associated with the estimated WTA distribution among consumers; these are measures of preference heterogeneity.

G Intrinsic Preference Distribution: Bimodal Pattern Decomposition

Figure G.1: WTA in Intrinsic Preference Distribution across Racial Groups

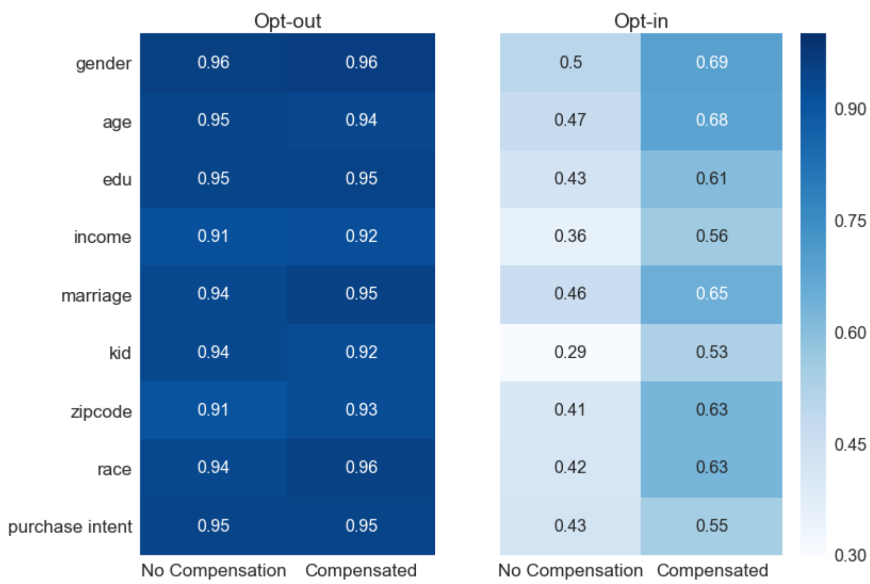


H Psychological Factors

H.1 The Default Frame

Figure H.1 visualizes the data-sharing frequency in different default regimes. Under the opt-out regime, almost everyone shares everything, regardless of the amount and format of compensation. The lack of choice variation in the opt-out regime does not per se imply a weaker preference for privacy or economic incentives; it simply means the impact of a "share-all" frame is strong enough to dominate other components in utility.

Figure H.1: Frequency of Data Sharing under Different Policy Regimes



Interaction between the default regime and privacy preferences. The literature has widely acknowledged *that* the default frame influences choices (Kahneman 1979, Thaler 1980, Johnson et al. 2002). However, little consensus exists on *how* or *how much* default affects choices. To flexibly characterize how default influences privacy choices, I estimate separate models for each default frame. Table H.1 displays the estimated privacy preferences under opt-in and opt-out regimes. In the comparison below, I acknowledge the scaling differences across the models, and normalize parameters to the same (dollar) scale when needed. The scaling does not affect the sign of parameters nor the sensitivity ranking across categories of data within the same model. The comparison of belief parameters w and δ are not affected by the scaling either.[20]

To compare intrinsic-preference parameters across models, Figure H.2 displays the willingness to pay (WTP) of intrinsic preferences, which are heavily influenced by the default frame. The negative WTPs imply that once data are obtained by companies, consumers will not take back

[20]To see this, note that if the instrumental utility is $w \cdot \beta \cdot \Delta E[d]$ in the utility space, then its dollar value is simply $w \cdot \Delta E[d]$.

Table H.1: Privacy Preferences across Default Frames

	Default Frame	Opt-In		Opt-Out	
		mean	95% CI	mean	95% CI
intrinsic	c_{income}	0.906	[0.588, 1.323]	-1.903	[-2.705, -1.134]
	c_{intent}	0.826	[0.419, 1.322]	-2.704	[-3.653, -2.127]
	c_{gender}	0.189	[-0.162, 0.664]	-2.988	[-3.956, -2.184]
	c_{age}	0.262	[-0.088, 0.733]	-2.429	[-3.127, -1.729]
	$c_{education}$	0.624	[0.329, 1.051]	-2.739	[-3.301, -2.161]
	$c_{relationship}$	0.497	[0.124, 1.010]	-2.734	[-3.331, -2.105]
	$c_{children}$	1.109	[0.790, 1.461]	-2.143	[-2.692, -1.380]
	c_{zip}	0.560	[0.227, 1.066]	-2.093	[-3.448, -1.328]
	c_{race}	0.604	[0.285, 1.104]	-2.660	[-3.518, -1.805]
instrumental	w_{income}	2.118	[0.108, 3.989]	1.994	[0.136, 3.893]
	w_{intent}	1.942	[0.383, 3.762]	1.995	[0.109, 3.909]
	$\widetilde{\delta}_{income,0}$	0.047	[-0.186, 0.282]	0.054	[-0.183, 0.280]
	$\widetilde{\delta}_{income,1}$	0.037	[-0.192, 0.284]	0.052	[-0.185, 0.286]
	$\widetilde{\delta}_{intent,0}$	0.059	[-0.352, 0.379]	-0.121	[-0.391, 0.350]
	$\widetilde{\delta}_{intent,1}$	-0.049	[-0.362, 0.324]	-0.129	[-0.384, 0.281]
sensitivity to compensation	β	0.146	[0.070, 0.2359]	0.046	[0.001, 0.141]
log posterior		-7476	[-7540, -7407]	-2075	[-2166, -1981]

Note: The models are estimated separately for each default frame. Variables are normalized using the Gelman method before estimation. Both models allow for heterogeneity in intrinsic preferences.

their control over personal data, unless they are incentivized by the amount indicated by the WTP. In my data, the gap between median WTA and median WTP amounts to $69.18 (income) to $88.06 (gender). In comparison, previous literature estimates dollar values of default in 401(k) plan enrollment decisions that range from $37–$54 (Bernheim et al. 2015) to $1,200 (DellaVigna 2009). However, the WTP estimates are very noisy, due to the fact that the estimated sensitivity to compensation in the opt-out regime is close to zero (see Table H.2 for credible interval estimates).

Interestingly, Table H.1 shows that the default frame does not heavily influence consumer beliefs about the instrumental payoff. The differential impacts of default suggest that while subjective evaluations are more susceptible to the influence of the default condition, objective evaluations—beliefs about the instrumental payoff—are less so. In view of this fact, distinguishing between the intrinsic and instrumental preferences also reveals how default (and potentially other psychological factors) influences different privacy motives differently.

The managerial implication is immediate. With a regulation that mandates opt-out consent, firms can still collect most customer data even if consumers are fully informed when making privacy choices. However, once the firm moves to an opt-in regime, it will incur substantial losses in the amount of available data. The default paradigm is also useful for thinking about the real

Figure H.2: Posterior Predicted Density of WTP in Intrinsic Preference

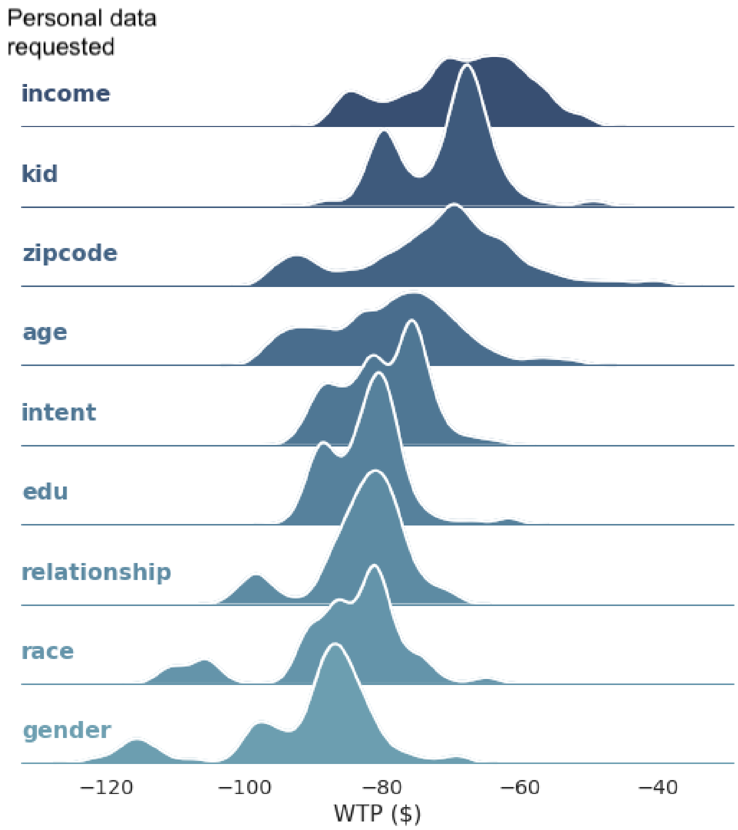


Table H.2: Posterior Estimates of Parameters Associated with Intrinsic WTP Distribution

(a) WTP Mean

	mean	95% CI
income	-66.59	[-621.55, -6.92]
intent	-78.87	[-733.52, -8.29]
gender	-89.84	[-866.19, -9.28]
age	-76.57	[-722.37, -8.03]
education	-81.10	[-767.11, -8.41]
relationship	-82.10	[-773.98, -8.63]
children	-70.15	[-634.81, -7.34]
zipcode	-70.52	[-653.87, -7.46]
race	-86.69	[-834.71, -8.97]

(b) WTP Standard Deviation

	mean	95% CI
income	70.23	[2.41, 479.59]
intent	35.52	[0.96, 249.39]
gender	57.36	[1.92, 360.06]
age	76.40	[2.70, 522.52]
education	34.59	[1.22, 248.68]
relationship	43.54	[1.22, 286.02]
children	29.53	[1.37, 185.01]
zipcode	57.63	[3.04, 405.01]
race	28.37	[1.30, 213.90]

impact of data-portability rights.[21] Taking the incumbent as the default choice, consumers are less likely to opt out of incumbent tracking and transfer data to its competitors, unless the expected utility gain from switching is substantially large.

[21] GDPR Article 20 and CCPA Title 1.81.5, Section 1798.100 (d).

H.2 Other Psychological Factors

The model includes a behavioral response term $m \cdot (p_i \geq 0) \cdot s_i$, to account for a combination of a mere-incentive effect and potential anchoring effects at the start of the survey. Behavioral response to the mere presence of incentives is well documented in the psychology literature (Shampanier et al. 2007, Urminsky & Kivetz 2011, Palmeira & Srivastava 2013), which can be explained by the theory that people are insensitive to scopes when evaluating options separately (Hsee 1996, Hsee & Zhang 2010). In treatment groups that distribute positive amounts of compensation, participants are told at the beginning that they can enter a gift-card lottery upon finishing the survey. This information may inadvertently create an additional anchoring effect, making all participants in these groups more inclined to share their data in order to get the anticipated gift-card rewards. The parameter m captures the combination of these two forces. Under the second mechanism, the additional anchoring effect will be stronger for participants in the opt-in group (because an opt-out condition per se also has a substitutive anchoring effect); this possibility is accounted for by having separate m's for different default conditions.

In the opt-in frame, the point estimate for m is 0.76, with the 95% credible interval being $[0.65, 0.87]$. In the opt-out frame, the point estimate is 0.07, with the credible interval being $[-0.17, 0.30]$. The strong effect asymmetry and the fact that the effect is almost non-existent in the opt-out condition suggest anchoring is more likely to be the main driver of this effect.