Valuing Intrinsic and Instrumental Preferences for Privacy

Tesary Lin∗

August 30, 2019

[Job Market Paper. Click here for the latest version]

Abstract

In this paper, I propose a framework for understanding why and to what extent people value their privacy. In particular, I distinguish between two motives for protecting privacy: the intrinsic motive, that is, a “taste” for privacy; and the instrumental motive, which reflects the expected economic loss from revealing one’s “type” specific to the transactional environment. Distinguishing between the two preference components not only improves the measurement of privacy preferences across contexts, but also plays a crucial role in developing inferences based on data voluntarily shared by consumers. Combining a two-stage experiment and a structural model, I measure the dollar value of revealed preference corresponding to each motive, and examine how these two motives codetermine the composition of consumers choosing to protect their personal data. The compositional differences between consumers who withhold and who share their data strongly influence the quality of firms’ inference on consumers and their subsequent managerial decisions. Counterfactual analysis investigates strategies firms can adopt to improve their inference: Ex ante, firms can allocate resources to collect personal data where their marginal value is the highest. Ex post, a consumer’s data-sharing decision per se contains information that reflects how consumers self-select into data sharing, and improves aggregate-level managerial decisions. Firms can leverage this information instead of imposing arbitrary assumptions on consumers not in their dataset.

Keywords: privacy, revealed preference, value of data, experiment, pricing

∗The University of Chicago Booth School of Business; tesarylin@chicagobooth.edu. I am grateful to Pradeep Chintagunta, Sanjog Misra, Brad Shapiro and Oleg Urminsky for their advice and support. Thanks go to Michael Dinerstein, Jean-Pierre Dubé, Avi Goldfarb, Austan Goolsbee, Sylvia Hristakeva, Garrett Johnson, Sarah Moshary, Eric Posner, Anita Rao, Robbie Sanders, Abigail Sussman, Dana Turjeman, 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 advent of privacy regulations across the globe, companies can no longer collect consumers’ personal data without their explicit consent.1 From the General Data Protection Regulation (GDPR) in the EU to the California Consumer Privacy Act (CCPA) in the US, new privacy laws mandate firms to deliver transparent information and seek explicit consent before data processing, while endowing consumers with more rights and choices. The GDPR, for example, gives data controllers (e.g., firms) specific guidelines on delivering “clear and plain” information about data processing, and requires them to seek “affirmative” opt-in consent.2 It also allows data subjects (e.g., consumers) to rectify, delete, or transfer their data to another firm at any time.3

Under this new regulatory regime, consumers’ preferences for privacy, namely, their desire to control and safeguard their personal data (Westin & Ruebhausen 1967), play a central role in determining what data firms are able to collect. For instance, the GDPR requires websites to seek permission before tracking consumers. In compliance with this requirement, European sites lost 10% of their recorded traffic and revenue as a result of increased consumer vigilance (Goldberg et al. 2019). Old business practices trying to bypass consumer awareness often backfire: for example, Google was charged a $56.8 million fine for obfuscating information and pre-ticking the sharing option when seeking consent for data processing.4 Instead, firms need to understand consumers’ preferences for privacy, in order to know when consumers share their personal data and how consumers’ privacy choices influence their inferences and decisions based on consumer data.

Despite the pressing need to understand privacy preferences, existing research often finds that they are hard to be reliably measured. One of the main puzzles is that preferences for privacy vary across contexts: consumers’ privacy choices vary with the reasons their data are collected, entities that can access their data (Rainie & Duggan 2015, Martin & Nissenbaum 2016), and information that changes their perceived usage of the data (John et al. 2010, Athey et al. 2017). Consequently, privacy preferences measured in one scenario are hard to generalize to other contexts (Acquisti et al. 2016).

I propose that we can better understand consumers’ privacy choices across contexts and their impacts on how firms collect and analyze personal data, if we distinguish between two different motives for protecting privacy. Privacy preferences can emerge because privacy itself is valued as an intrinsic right (Warren & Brandeis 1890); it can also arise because of its instrumental value, namely, the economic payoff of preventing their private “type” from being revealed through data (Stigler 1980, Posner 1981). Consumers hold both types of privacy preferences. Intrinsically, most people find it “creepy” to have smart thermostats tracking their activities at home, even when

---

1 “Personal data” refers to any combination of data that can be used to identify a natural person. See GDPR Article 4(1): https://gdpr-info.eu/art-4-gdpr/.
4 https://digiday.com/media/google-fined-57m-french-regulator-breaching-gdpr/
their behaviors are not objectionable (Rainie & Duggan 2015). Motivated by an instrumental perspective, reckless drivers can be less willing than safe drivers to install telematics offered by insurance companies that monitor their driving habits (Soleymanian et al. 2019, Jin & Vasserman 2018). Because a person’s relevant type varies across market environments, the instrumental preference also varies across economic contexts. By contrast, the intrinsic preference for privacy does not depend on a person’s type.

**Research objective.** This paper empirically distinguishes between intrinsic and instrumental preferences for privacy, due to the following reasons. First, separately measuring these two preference components allows us to understand what types of consumers tend not to share their personal data, which has important regulatory and managerial implications. The “if you’ve got nothing to hide, you’ve got nothing to fear” doctrine is only valid when people harbor purely instrumental preferences for privacy. On the other hand, assuming consumers value privacy only because of intrinsic motives may lead to the misleading conclusion that people who value privacy more are no different from the rest of the population. Accounting for the coexistence of intrinsic and instrumental preferences is essential for making correct inferences about consumers while avoiding erroneous attributions based on privacy decisions.

Second, understanding the composition of privacy preferences is crucial for analyzing the information value of personal data voluntarily shared by consumers, and for designing compensation schemes to encourage data sharing. The interplay between intrinsic and instrumental preferences determines not only how many people share their data, but also who chooses to share. Understanding how different instrumental incentives change the composition of consumers who refrain from tracking is crucial not only for developing valid inference based on voluntarily shared data, but also for designing effective incentive schemes that can collect personal data with higher information value.

Lastly, measuring the relative magnitudes of each preference component can help us understand the extent to which privacy preferences vary systematically across economic contexts. The answer to this question will contribute to the discussion about relative merits of the omnibus (a single law that regulates privacy across industries) and the sectoral approach (different laws for different sectors) to privacy regulation. The omnibus approach is justified by the notion that privacy is a universal and intrinsic right, whereas the sectoral approach works better if privacy preferences vary across economic scenarios. Understanding when the intrinsic or instrumental motive dominates in determining privacy choices is key to assessing the degree of this systematic, utility-induced context dependence.

**Research methodology and preview of findings.** To characterize how intrinsic and instrumental preferences shape privacy choices, I first construct a model of consumer disclosure with two-dimensional preferences. This model illustrates how the interplay between the two preference components determines the type of consumers who share their data. In particular, it shows that the
heterogeneity of and the correlation between intrinsic and instrumental preferences codetermine to what extent the “nothing to hide” argument can be refuted.

I then design an experiment that solicits revealed preferences for privacy by requesting consumers to share sensitive data in a market research survey. A central element of the experiment is its 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 for consumers who choose not to share their personal data, and is the key to characterizing instrumental preference. The experiment then generates three layers of variation needed to identify the model: (a) the intensity of instrumental incentives allows me to separate the two preference components; (b) the amount of compensation enables me to calculate the dollar values of privacy preferences; and (c) the default-choice condition permits comparison of privacy choices in different policy regimes. The experiment also contains a conjoint survey, which allows me to calculate the information value of personal data for price optimization in the counterfactual analysis.

Data and structural estimation results reveal the following findings. First, personal data are not valued equally by consumers: without instrumental incentives, the mean willingness to accept (WTA) to share personal data ranges from $0.14 to $2.37, with substantial heterogeneity across both consumers and categories of data. Second, privacy choices vary with economic context. Third, self-selection into data sharing is not always adverse, deviating from the “nothing to hide” doctrine and predictions offered by canonical economic models. Lastly, privacy choices are more responsive to incentives in the opt-in (not sharing by default) than in the opt-out (sharing by default) regime.

In the counterfactual analysis, I calculate the information value of voluntarily shared personal data in the context of pricing, then examine what firms can do to improve this information value and increase subsequent profits. Two findings follow: First, firms can improve the efficiency of their data-acquisition strategy by sub-sampling consumers when personal data are used for model estimation. Second, apart from the contents of personal data, privacy choices per se contain information that reflects the nature of sample bias induced by heterogeneous privacy choices and improves firms’ aggregate-level inference on consumers.

**Contribution to the literature.** First and foremost, my paper formalizes and extends Becker’s (1980) and Farrell’s (2012) notion of privacy preferences. Both Becker and Farrell distinguish the valuation for privacy as a “final good” and the valuation for privacy induced by the economic consequences of data sharing. My framework is also related to Wathieu & Friedman’s (2009) categorization of direct vs. indirect privacy concerns. Compared to these papers, my paper characterizes how the coexistence of intrinsic and instrumental preferences determine the selection pattern in data shared by consumers, and how inference based on these data should be adjusted accordingly.
Second, my paper separately measures intrinsic and instrumental preferences in terms of dollar values, and provides a replicable method for firms and researchers to measure consumers’ heterogeneous privacy preferences across contexts. It builds on existing work that measures the revealed preference for privacy as a whole, including Goldfarb & Tucker (2012b), Athey et al. (2017), and Kummer & Schulte (2019). It also contributes to the literature of measuring dollar values of privacy preferences (Hui 2007, Acquisti et al. 2013, Benndorf & Normann 2018) by focusing on a broader set of consumers and different categories of personal data.

My paper also contributes to the literature on context-dependent privacy preferences by highlighting the economic and strategic reasons that generate context effects. 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. Moreover, it explains why privacy choices respond to information treatments that change how consumers perceive the economic consequences associated with privacy decisions (John et al. 2010, Miller & Tucker 2017, Athey et al. 2017).

Last but not least, by discussing how consumers’ privacy choices affect firms’ inferences and resultant profits, my paper adds to the research on how privacy preferences influence firms’ managerial outcomes, including effectiveness of advertising (Goldfarb & Tucker 2011, Tucker 2014, Rafieian & Yoganarasimhan 2018), funds raised (Burtch et al. 2015), and profits (Johnson 2013, Johnson et al. forthcoming). Compared to existing studies, mine emphasizes the impact of consumers’ self-selection into data sharing on the quality of firms’ data-driven decisions.

The paper proceeds as follows. Section 2 introduces the conceptual framework, clarifies definitions, and illustrates the implications of the dual-privacy-preference framework. Section 3 describes the experiment design, followed by Section 4, which provides descriptive evidence of intrinsic and instrumental preferences. Sections 5 and 6 present the structural model and estimation results. Section 7 describes the counterfactual analysis, and Section 8 concludes.

2 The Conceptual Framework

This section uses an example to clarify the distinction between intrinsic and instrumental preferences for privacy. It describes how the instrumental motive is endogenously derived from the economic context, how consumers self-select into sharing or protecting their personal data, and how this selection pattern differs from predictions generated by models that assume monolithic privacy preference.

---

5This angle is related to the contextual-integrity literature in sociology (Nissenbaum 2004, Martin & Nissenbaum 2016), which states that privacy rights should be evaluated based on what data are collected and how they are used. Whereas this strand of literature emphasizes the normative aspects of data collection practices, my paper focuses on a positive analysis, that is, whether and to what extent instrumental preference explain privacy decisions.
2.1 Setup

Consider a firm that sells a product or service to many consumers. Consumers have different types, which calls for customized offers; denote consumer i's type as $d_i$. The firm requests personal data from consumers in order to know their types. At a later stage, the firm gives customized transfer $T(.)$ to consumer $i$, which maximizes the firm’s expected profits conditional on the firm’s understanding of $i$’s type. For example, $T(.)$ can be price discount and $d$ is price sensitivity; or $T(.)$ can be the annual limit in an insurance contract and $d$ is risk type. To encourage data sharing, the firm may incentivize consumers using compensation, denoted as $P$. Examples of compensation include perks offered to consumers who sign up for a loyalty program, or gift cards for sharing email.

Consumer $i$ owns personal data that can reveal their type. Without loss of generality, assume the shared data perfectly reveal their type. Therefore, a one-to-one correspondence exists between the content of personal data and a consumer’s type. We can always construct consumer types such that the transfer $T(d)$ is monotonic in $d$. For example, suppose that $d$ is age and that the middle-aged group has the lowest price sensitivity, followed by the older and then the youngest. Then we can label the middle-aged group as $d = 1$, the older group as $d = 2$, and the youngest group as $d = 3$. Without loss of generality, I define consumer types such that $T(.)$ is increasing in $d$, and refer to consumers with higher $d$ (who can obtain higher economic payoffs upon revealing their type) as the high type.

Consumers decide whether to share their data with the firm. $s_i \in \{0, 1\}$ indicates whether $i$ decides to share $d_i$: $s_i = 1$ means the piece of data is shared. For consumers who share no information, the firm forms beliefs on their types and chooses the amount of transfer accordingly: $T(s = 0) = T(F_d(d|s = 0))$, where $F_d(d|s = 0)$ is the distribution of consumer type conditional on the consumer choosing not to share his data. For consumers who share, the transfer can condition on the content of the data, written as $T(d_i)$.

2.2 Privacy Preferences

A consumer has an intrinsic motive for privacy $c_i$, which is a taste for protecting his data regardless of the economic consequences induced from revealing his type. He also has an instrumental motive for privacy, namely, the (expected) economic gain from not revealing his type: $\Delta T(d_i) \equiv T(F_d(d|s = 0)) - T(d_i)$. For example, suppose $T(.)$ is price discount and $d$ is age. If older consumers have higher price sensitivity, the firm will choose to give them higher discounts upon learning their

---

5One way to understand the one-to-one correspondence assumption between data content and consumer type is the following. Suppose the data do not perfectly reveal the exact level at which the consumer values the product, but indicates a range for the consumer’s valuation; in this case, we define the range of valuation as his type. In cases where two income levels correspond to the same level of product valuation, we can code the two income levels as having the same value. Another way to understand the model is that some firms only care about the predictive performance of the pricing model but not consumers’ price sensitivity per se; in this case, a direct mapping occurs from $d_i$ to $P(d_i)$, and the intermediary “type” is unnecessary.

6
age. Anticipating this outcome, the instrumental preference for older consumers to protect their personal data will be low.

The key distinction between intrinsic and instrumental preferences is whether they are generated from the consequences of the consumer’s type being learned.\footnote{Under this taxonomy, costs that may at first glance appear instrumental still belong to the “intrinsic” bin (e.g., concern for identity theft, because this concern is common for all consumers).} The intrinsic preference is a utility primitive, which persists regardless of the market environment and the consumer’s “type” relevant to this market. By contrast, the instrumental preference is derived from the economic environment; thus, it changes with the payoff function $T(.)$ as well as his type in this particular market $d_i$. The intrinsic preference can also be correlated with a consumer’s type. However, when the same consumer is placed in a different economic environment, his instrumental preference for privacy changes accordingly, but his intrinsic preference does not.\footnote{This statement does not contradict the fact that intrinsic preference can be shifted by psychological factors. Rather, it distinguishes psychological and economic shifters, which have different impacts on the expression of the two preference components.}

Instrumental preference and the utility from compensation are also distinct constructs, even though both are economic payoffs. The instrumental motive is closely related to the value of private information. It is a function of the hidden type that the firm cares about, that is, information about the consumer that can shift the level of optimal transfer between the firm and consumers. On the other hand, compensation does not necessarily hinge on a consumer’s type; it is more properly viewed as the price for personal data.

### 2.3 Who Chooses Not to Share Personal Data?

A consumer shares data iff the privacy cost is offset by the compensation that the firm provides:

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

(1)

Firms often want to figure out the characteristics of consumers choosing not to share their data, in order to optimize the transfer to consumers that maximizes profits. A model that assumes privacy preferences to be purely instrumental will generate the following prediction: only low types choose to withhold their data in equilibrium, because these are the consumers who incur a larger loss upon sharing data (Grossman & Hart 1980, Milgrom 1981, Jovanovic 1982). Such reasoning is the underpinning of the “nothing to hide” statement. Alternatively, a theory that assumes privacy preference to be pure intrinsic may fail to capture the nuance of consumers’ self-selection into sharing.

The dual-preference framework paints a more nuanced picture of how consumers self-select into sharing personal data. The intrinsic preferences for privacy are likely to be heterogeneous among consumers. This heterogeneity changes the firm’s inference $F_d(d|s = 0)$ because nondis-
The closure no longer signals low-type customers unambiguously.\footnote{The hidden heterogeneity that the firm cannot observe is what shapes inference.} The degree to which privacy choice reveals information about a consumer’s type depends on both the relative heterogeneity of the intrinsic preference and its correlation with the instrumental preference. This is formally characterized by the proposition below (see proof in Appendix A):

**Proposition 1.** Denote the standard deviation of intrinsic and instrumental preferences respectively as \( \sigma_c \) and \( \sigma_t \), and their correlation coefficient as \( \rho \). The following conclusions hold:

(A) In data shared to the firm, sample selection goes in the same direction as predicted by a model with pure instrumental preference iff \( \rho + \frac{\sigma_t}{\sigma_c} > 0 \).

(B) Privacy choice is more indicative of a consumer’s type \( d_i \) when \( \frac{\sigma_t}{\sigma_c} \) is higher.

To illustrate this proposition, suppose older consumers (who would have obtained better discounts upon sharing their age) care more about privacy intrinsically, and that the intrinsic preference is highly heterogeneous compared to the instrumental. Then the intrinsic preference will play a dominant role in privacy decisions: on average, consumers who choose not to share their data are more senior and should receive more generous discounts. This pattern forms a stark contrast to the case with a pure instrumental preference for privacy.

This framework can account for scenarios where data sharing increases consumers’ expected utility by improving product quality, for example, when the shared data serve as inputs to a recommendation engine. If the firm knows how much each consumer values the quality improvement, this improvement is a form of compensation. Otherwise, the heterogeneous preference for quality improvement plays the same role the heterogeneous intrinsic preference does in the firm’s inference problem.

In sum, the dual presence of intrinsic and instrumental privacy preferences has two main implications. First, although the intrinsic preference is a utility primitive, the instrumental preference is endogenously determined by the market environment. This fact explains why preferences for privacy vary across the contexts of data used, who gets access to the data, and what data are requested. Second, when the intrinsic preference for privacy is heterogeneous, privacy choice no longer unambiguously signals a specific type of customer. The more the heterogeneous intrinsic preference is compared to the instrumental, the less we can assume a consumer’s type based on his privacy decisions. Accounting for this fact is essential for analyses based on voluntarily contributed personal data.

## 3 The Experiment

My experiment serves three main purposes. First, it generates variation crucial for separately identifying intrinsic and instrumental utility parameters. Second, it provides a setup that sequentially
records consumers’ private information and privacy choices, which enables the characterization of preference heterogeneity. Lastly, it creates choice environments that match features of the new privacy regulatory regime that are yet to be realized in the US. In this section, I first introduce the empirical challenges and explain how my experiment addresses them, and then describe key features of the new policy regime, which the experiment intends to examine. This is followed by the introduction of experiment design and implementation.

3.1 Empirical Challenges and Solutions

To further understand how consumers self-select into different privacy choices, one needs to separately measure the intrinsic and instrumental preferences. However, doing so using observational data is difficult. First, the economic incentive is usually fixed in observational settings, making it implausible to separate instrumental preferences from the intrinsic.\(^\text{10}\) 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 may keep using Gmail even after learning that Google analyzes all their email texts, due to either their low preferences for privacy or their high valuation of Gmail service.\(^\text{11}\) The last and the most difficult challenge is that both consumers’ privacy choices and their private types need to be observed to identify my model, yet privacy choices are precisely the decision concerning whether to reveal these private types. If this challenge is not accounted for, the collected data will exhibit self-selection as long as variation in privacy decisions exists.

To circumvent these problems, I design an experiment that includes three main features. First, instrumental incentives are turned on or off across different treatment conditions. 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 novel two-stage design, where the first stage collects participants’ private information, and the second stage solicits revealed preferences for privacy. The details of the experiment design are described in Section 3.3.

\(^{10}\)One way to identify instrumental preferences without experimental variation is to apply a rational expectation assumption, and backing out the perceived instrumental utility from a firm’s actual practices with the data. However, the existing literature often finds that consumers often have incorrect perceptions of how their data are used (Athey et al. 2017, Rainie 2016). By using an experiment, I can avoid imposing this strong rationality assumption.

\(^{11}\)It is possible to address this confound by finding exogenous shifters that change only the cost of protecting privacy or only the quality of products. In fact, several excellent studies utilize this type of shifters to identify privacy preferences (Varian et al. 2005, Goh et al. 2015, Johnson et al. forthcoming, Miller & Tucker 2017). However, a good exogenous shifter is sometimes hard to find, and it cannot solve the other challenges in my empirical analysis.
3.2 Examination of the New Policy Regime

Specifying relevant choice environments is particularly important for the measurement of privacy preferences given its context-dependent nature. Below, I describe key features of new privacy regulations and explain how these features are matched to my design of choice environments.

The new policy regime features three main elements: transparent information, consumer rights, and affirmative consent. The first two elements are similar across continents. For example, the CCPA and the GDPR have many overlapping provisions, including the right to know, the right to deletion, the right of access, and the right of data portability.\textsuperscript{12} Specific to the right to know, the GDPR requires information related to personal data processing to be easily accessible and conveyed using “clear and plain language.” It also requires data processors to articulate the purpose of data processing, potential risks and safeguards, as well as consumers’ rights.\textsuperscript{13} Similarly, the CCPA endows California consumers with the right to know what data are collected and the corresponding purposes. To match these elements, my experiment explains clearly the usage and flow of the data, and explicitly notifies participants about their options related to data sharing.

The last element, affirmative consent, varies across regulatory frameworks in terms of the default choice. While EU laws (GDPR and ePrivacy Regulation) implement opt-in consent, practices in the US are mixed. Although many existing laws still adopt opt-out consent, opt-in has been used when regulating more sensitive data.\textsuperscript{14} Whether opt-in or opt-out should become the new regulation standard has been one of the central debates in recent policy discussions.\textsuperscript{15} Regardless of the default regime, requests effectively operate in an opt-in condition for data that are not generated by default, such as survey responses, tests, and membership sign-ups. The literature has widely acknowledged the fact that default frame influences choices (Kahneman 1979, Thaler 1980, Johnson et al. 2002, Acquisti et al. 2013). However, little consensus exists on how or how much default affects choices. In view of the policy and academic relevance of studying the default effect, my experiment includes both opt-in and opt-out conditions. The empirical analysis focuses mainly on the opt-in regime given its prevalence, but will compare privacy choices in different consent regimes.

3.3 Experiment Design

The experiment uses a survey as an instrument but solicits revealed preference instead of stated attitude. This is achieved by including personal questions with varying degrees of sensitivity: a

\textsuperscript{12}https://www.bakerlaw.com/webfiles/Privacy/2018/Articles/CCPA-GDPR-Chart.pdf
\textsuperscript{13}https://gdpr-info.eu/recitals/no-39/
\textsuperscript{14}Examples include Illinois’s Biometric Information Privacy Act, and the Fair Credit Reporting Act on data used for employment purposes or that contains medical information.
\textsuperscript{15}Recent privacy bills that promote an opt-in standard include the Information Transparency and Personal Data Control Act and the Consent Act. The divergence of opinions can be seen, for example, by comparing testimonies given by Jon Leibowitz and by Victoria Espinel in one of the Senate Hearings. See https://www.commerce.senate.gov/public/index.cfm/2019/2/policy-principles-for-a-federal-data-privacy-framework-in-the-united-states.
participant's decision to share the response to a particular question indicates his level of privacy cost associated with this piece of data. This technique has previously been deployed by Acquisti et al. (2012) and Goldfarb & Tucker (2012b). 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 not only incentive compatible, but also more relevant than attitudes 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 their intent to purchase a digital device in the near future. These 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 to answer are thus allowed not to respond rather than being forced to fabricate a response. Appendix B.1 shows examples of the conjoint and demographic questions.

Stage one serves two purposes: The primary purpose is to record consumers' private information that will shift instrumental preferences in the subsequent stage. Second, the conjoint questions provide inputs for calculating the information value of data to firms, which is central to the counterfactual analysis. The conjoint questions also disguise the real purpose of the survey so that participants are not primed to consider privacy.

Stage two solicits privacy choices. After finishing the survey, participants are directed to a new screen. Here, they are requested to share survey responses with a third-party business partner, which is a smartwatch manufacturer who wants to use the data to inform its product-design decision. Participants can choose whether to share each demographic variable and the purchase intent variable separately via check boxes. Data sharing is encouraged by compensation in the form of a gift-card lottery. Crucially, participants are not told about the data-sharing step until they answer all questions in stage one; once consumers reach the second stage, the “return” button is disabled, preventing them from deliberately changing previous responses to facilitate sharing. These two features, together with the presence of aforementioned "prefer not to say" options, are included to ensure responses in the first stage are truthful.

Stage two is also where all treatments take place. Figure 1 displays the three layers of treatments: the incentive scheme, the amount of compensation, and the sharing default. These treatments are orthogonal to each other. The first layer varies the incentive scheme across treatment groups:

---

16 Only first-stage responses that are informative (willingness to pay, responses other than “prefer not to say”) are allowed to be shared in stage two.

17 Allowing for and modeling response fabrication as a measure of privacy preferences is possible (Lwin & Williams 2003). However, doing so will eliminate the privacy-money tradeoff that consumers face.

18 There is one exception: by design, participants who receive zero compensation do not receive any instrumental incentives.
Note: The three layers of treatments are orthogonal to each other. In each layer, treatments are assigned with equal probability.

- Treatment 1 (pure compensation): the amount of compensation increases proportionally to the amount of data shared and is common across all participants. In particular, sharing one additional piece of data is associated with a 1% level increase in the probability of winning the gift card.

- Treatment 2 (compensation + instrumental incentive): a baseline level of compensation exists that takes the same form as in Treatment 1. The amount of compensation is subsequently adjusted based on whether the company perceives the participant to be a potential customer based on the data it obtains. Participants who are more likely to be their customers receive higher compensation than the baseline, whereas unlikely customers receive a lower 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.

The incentive scheme is transparent to participants within each treatment. Appendix B.2 shows the information displayed to the participants. In sum, the instrumental preferences for privacy 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. Privacy choices in Treatment 1 identify intrinsic privacy preferences alone: the stated purpose of data collection is intentionally neutral; moreover, the company is unknown to all participants and therefore has negligible instrumental consequences for them. By contrast, choices in Treatment 2 are motivated by both intrinsic and instrumental preferences. Therefore, the differential responses between Treatments 1 and 2 identify instrumental preferences for privacy.
The other treatments are designed as follows. The second treatment layer changes the value of the gift card across participants, creating variations for measuring the dollar values of privacy preferences. The third layer varies default choice, which is set to either sharing all data (opt-out) or sharing none (opt-in). Within each layer, treatments are assigned with equal probability.

To measure if participants understand and trust the validity of incentive treatments, a set of follow-up questions are prompted after participants make the data-sharing choices. These questions include the 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.

Two features of this experiment intentionally diverge from current business practices. First is its use of gift cards instead of differential prices for inducing instrumental incentives. Although inducing instrumental incentives via pricing is more natural, it is infeasible in my experiment. Given that the featured company is previously unknown to participants (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.

Second, the choice environment also differs from common business practices, where firms try to obfuscate information via obscure privacy policies. As previously mentioned, one goal of this experiment is to emulate the choice environment required by the new regulatory regime, which features transparent information provision. Moreover, minimizing information friction allows me to measure privacy preferences cleanly, rather than confounding consumers’ unawareness with their lack of care for privacy.

Some aspects of the experiment may bring confounds to the preference measurement. First is the use of a lottery instead of sure rewards for compensation. If participants are not risk neutral, their perceived gain from the gift-card lottery will differ from the objective expected value of the lottery. In particular, if participants are risk averse, the estimated dollar value of privacy preferences will be an upper bound of the true values, whereas the opposite holds if participants are risk seeking. In the follow-up survey, 35% of the participants prefer the lottery, while the rest prefer the sure reward: although inconclusive, this pattern suggests that risk aversion is not a dominating feature. Another potential concern is that participants may share less in the instrumental-incentive treatment if they dislike an “unfair” compensation scheme. Fairness is not among the main concerns listed by the participants when they indicate reasons for (not) sharing personal information.

3.4 Replicating the Experiment across Contexts

Firms and researchers can replicate the experiment to measure consumers’ privacy preferences in the scenario they are interested in. In particular, they can randomly sample a group of consumers
and give them compensation high enough so that everyone chooses to share data; this treatment group will serve the same purpose as the first stage of my experiment; the rest of the treatment groups can then adopt a design similar to my second stage. By comparing the “full data” collected in the high compensation group and the voluntarily shared data in other groups, firms and researchers will be able to learn about the nature of privacy preferences and the nature of selection in data in their context of interest.

4 Data and Descriptive Evidence

In what follows, I describe the data source and sample characteristics, and then present model-free patterns of intrinsic and instrumental preferences. I start by describing privacy choices in the opt-in regime. Data show how consumers purposefully share some data while protecting others, how the economic context changes the composition of consumers that share data, and how this compositional shift changes the quality of data shared. Then I compare privacy choices in different default regimes. Compared to the opt-in condition, privacy choices in the opt-out condition provide less variation that reveals preferences. I conclude by discussing its implications for subsequent analysis.

4.1 Data Source and Cleaning

Participants of the experiment come from Qualtrics Panels. Existing work finds that the Qualtrics panel is more demographically and politically representative than alternative online panels (Heen et al. 2014, Boas et al. 2018). To further reduce possible discrepancies, stratified sampling is applied so that the demographics of participants entering the survey resemble the distribution given by the 2018 US census. 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 a first stage is still necessary.

A total of 4,142 participants enter the survey; 3,406 of them proceed to the data-sharing-request stage. Most attrition occurs during the phase when participants answer conjoint questions. For people who leave the survey upon seeing the request, I code their choices as sharing nothing, regardless of the default condition. 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 at the lowest decile. The cleaned data include 2,583 participants, comparable to other large-scale experiments that study consumers’ utility from digital consumption, such as Brynjolfsson et al. (2019) and Allcott et al. (2019).

For respondents using the same IP address, I keep the first response when the finishing time of the first respondent does not overlap with the starting time of the second respondent. If these times overlap, I discard both responses.
4.2 Sample Characteristics

Attrition and sample cleaning may 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 of the 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 main 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 to the population, participants who finish the survey tend to be female, less educated, and have lower income. The sample differences should be taken into account when interpreting the preference-measurement results.

Table 1: Demographics of Experiment Participants (Cleaned Sample)

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


Note: For discrete variables, values in the survey are collapsed 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, mean values are reported with standard deviation in parenthesis.

Apart from demographic data, another important variable is purchase intent—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 assigned
different scores indicating the probability of buying. 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.

To examine ex-post covariate balance, I plot out the standardized mean differences across treatments for each variable. Appendix C shows the covariate balance plots and assessment methodology. Both the instrumental incentive and the default treatments are well balanced. Participants in the zero-incentive group differ from the incentivized group in when they start and finish the survey. These variables are controlled for during estimation.

### 4.3 Intrinsic Preferences

Table 2 shows how the frequency of sharing varies by the category of data in Treatment 1 (pure compensation), where privacy choices reflect pure intrinsic preference. Preferences for protecting income data and information about their children are higher than the others. Gender information is the least sensitive: participants are indifferent between whether this information is revealed, as is shown by the 50% frequency of sharing without compensation. Overall, the table shows that different data are valued differently by consumers and that participants make attentive trade-offs in the experiment.

<table>
<thead>
<tr>
<th>Category of Data</th>
<th>Compensation = 0</th>
<th>Compensation &gt; 0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gender</td>
<td>Age</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>0.70</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Note: “Relationship” corresponds to their responses about marital status. “Kids” corresponds to responses about the number of children they have.

### 4.4 Instrumental Preferences

The extent to which instrumental incentives influence privacy choices is theoretically ambiguous. In Treatment 2, participants benefit more if they are perceived as wealthy or intend to buy digital products in the short term (hereafter, “high types”). The presence of instrumental incentives makes high-type participants more willing to disclose their personal data in Treatment 2 than in Treatment 1. However, the influence of instrumental preferences on privacy choices can be small when the intrinsic preferences are more heterogeneous, especially when the two preference components are negatively correlated.
Figure 2 shows both the influence of instrumental incentives and how this influence is mitigated by intrinsic motives. Panel (a) plots the proportion of participants who choose to share their income data across income cohorts for each incentive treatment. Participants in the highest-income cohorts are more likely to share in Treatment 2, whereas the differences in other cohorts are not significant. Interestingly, wealthier participants have stronger intrinsic preferences for privacy than their low-income counterparts, which is opposite to the direction that instrumental preferences indicate. This negative correlation explains why the impact of instrumental preferences is mitigated. Panel (b) shows the same plots for the sharing decision of purchase intents. Here, instrumental incentives have a more significant influence on privacy choices. Taking the messages together, we see that instrumental incentives make low-type consumers more willing to protect their privacy than in the case without, but the strength of this influence and the final sorting pattern depends on the nature of intrinsic motives.

One may be concerned that the systematic variation between consumer type and privacy choices across treatments is driven by variables correlated with consumer types, rather than the types per se. To address this concern, I replicate Figure 2 using the propensity forest (Wager & Athey 2018), which allows me to estimate a non-parametric treatment effect while controlling for a rich set of covariates (see Appendix D for implementation details and results). The estimated differences caused by instrumental incentives are qualitatively the same as in Figure 2.

To further examine how and to what extent instrumental incentive has a significant impact on the distribution of data shared, I compare the mean income and purchase intent between shared data and the true mean demographics, separately for each treatment group. Table 3 displays the t-test statistics for this comparison. In the instrumental treatment, the shared data describe a sample with higher income and higher purchase intent than the true distribution, with the difference in purchase intent being more significant. In comparison, the mean income and mean purchase intent are not significantly different from the true population average in the non-instrumental incentive treatment. To sum up, instrumental incentives change the representativeness of voluntarily shared data across contexts, subject to the moderation from intrinsic motives.

<table>
<thead>
<tr>
<th>(a) Income</th>
<th>(b) Purchase Intent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treatment 1</td>
</tr>
<tr>
<td>Statistics</td>
<td>-0.969</td>
</tr>
<tr>
<td>p-value</td>
<td>0.333</td>
</tr>
</tbody>
</table>

Note: Treatment 1 = Intrinsic Utility; Treatment 2 = Intrinsic + Instrumental Utility.
Figure 2: Frequency of Data Sharing across Incentive Treatments

(a) Income Sharing

(b) Purchase-Intent Sharing

Note: Intrinsic Utility = Treatment 1; Intrinsic + Instrumental Utility = Treatment 2. Frequency is calculated as the proportion of participants who share their income data within each income cohort (not across). The sum of bar heights can be greater than 1.

Purchase intent 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?” A higher value indicates higher purchase intent.

4.5 The Impact of the Default Regime

Figure 3 visualizes the data-sharing frequency in different default regimes. Under the opt-out (old policy) regime, almost everyone shares everything, regardless of the amount and format of compensation. The lack of choice variation does not per se imply diminishing preferences for privacy or economic incentives; it simply means the impact of a “share-all” frame is so strong that it dominates all other components in utility. This pattern is, in fact, consistent with casual observations in the current business model: with the opt-out consent framework, most consumers continue sharing their personal data even after encountering consent-request banners. However,
as the world shifts toward the new policy regime, understanding privacy choices in the new framework becomes crucial.

Figure 3: Frequency of Data Sharing under Different Policy Regimes

5 The Structural Model

The structural model serves three main purposes. First, by backing out the preference primitives, it estimates the dollar value of privacy preferences. Recently, scholars have been proposing a “data market” with property rights of data assigned to consumers (Arrieta-Ibarra et al. 2018, Zingales & Rolnik 2017). A dollar-value measure of consumers’ privacy preferences can help us understand what the price for data in such a market will be. Second, it clarifies how instrumental incentives shift privacy choices by changing consumers’ beliefs about economic payoffs. Lastly, the utility primitive estimates allow me to later simulate privacy choices and the information value of shared data under counterfactual compensation schemes.

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 piece of data 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:

\[
U(S_i; C_i, D_i) = \sum_{k} - c_k(X) \cdot s_{ik} + 1_{instr} \cdot 1_{k \in \{1, 2\}} \cdot \beta \cdot p_i \cdot w_k \cdot \hat{E}[d_{ik}|S_i, D_i] + \beta \cdot p_i \cdot s_{ik} + \epsilon_{ik}. \tag{2}
\]

\[
C_i = [c_1(X), c_2(X), ... c_k(X)] \text{ is the intrinsic preference for privacy; each } c_k(X) \text{ is a function of observables } X \text{ (more details below). } 1_{instr} \text{ is the instrumental-treatment indicator. } 1_{k \in \{1, 2\}} \text{ selects the data-sharing decisions that are subject to the influence of instrumental incentives. } \beta \text{ is the marginal utility of monetary rewards. } p_i \text{ is the value of the gift card multiplied by } 1\%. w_k \text{ is the consumer’s expected increases in the winning probability for an adjacent, higher type. } \hat{E}[.] \text{ is “belief about belief”: the consumer’s expectation of the firm’s expectation about his type, conditional on the data shared and the sharing decision. The baseline compensation is scaled by the amount of data shared, represented by } p_i \cdot s_{ik}. \epsilon_{ik} \text{ is the random utility shock associated with choice } S; \epsilon_{i1}, \epsilon_{i2}, ... \epsilon_{iK} \sim TIEV.
\]

In the main analysis of this paper, belief about a consumer’s type only depends on the directly associated data shared: \( \hat{E}[d_{ik}|s_{ik} = 1, D_i] = \hat{d}_{ik}, \hat{E}[d_{ik}|s_{ik} = 0, D_i] = \tilde{d}_{ik}(p_i). \) I let \( \tilde{d}_{ik}(p_i) = \delta_{i0}(X) + \delta_{i1}(X)p_i \) to allow for different levels of rationality: If both the firm and consumers are rational, the conjectured type of consumers not sharing their data can change with the stake of instrumental incentives. If instead agents form naive beliefs, \( \delta_1 \) is zero. By treating consumer beliefs on type-related payoffs \( (w_k, \delta_{i0}, \delta_{i1}) \) as free parameters, the model imposes no assumption on consumer rationality. This flexibility is important, because previous literature finds that consumer beliefs often fail to match actual usage of consumer data (Athey et al. 2017, Rainie 2016), due to firms’ purposeful obfuscation (Ben-Shahar & Chilton 2016) and their evolving data-utilization practices (Stutzman et al. 2013). Estimating consumers’ beliefs under full information allows us to understand the degree of consumer sophistication separately from the information effect. Identification of different belief parameters is made possible by the ability to observe consumer types, which is a unique feature of this study.

Correctly estimating heterogeneity in intrinsic versus instrumental preferences is key to understanding how consumers self-select into sharing. I characterize heterogeneity by allowing privacy preference parameters to be functions of observables \( X \), including demographics, time entering the experiment, time spent in each question, browser used, and device specifications. In particular, \( c_k(X) = c_{k0} + c_{k1} \cdot X \); \( \delta_{i0}(X) \) and \( \delta_{i1}(X) \) follow the same specification.\(^{20}\) Note that there is also a “built-in” heterogeneity in instrumental preference, coming from the fact that instrumental incentives vary with consumer types.

\(^{20}\)When estimating heterogeneity of consumer beliefs, income and purchase intent are excluded from covariates \( X \) for identification reasons.
Psychological factors other than privacy preferences also affect choices. The most important factor is the default frame. How default affects choices finds little consensus. Importantly, different mechanisms imply different interaction effects between the default frame and utility parameters (Bernheim et al. 2015, Goswami & Urminsky 2016, Goldin & Reck 2018). The main focus of this paper is backing out behavioral preferences under each frame, which are the relevant objects for analyzing firm-side implications of privacy choices. Therefore, I choose to be agnostic about the mechanism by estimating the model separately for each default frame. 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.

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

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

where \( \Delta u_{ik} \) is the difference in mean utilities between sharing and not sharing data \( k \), experienced by consumer \( i \):

\[
\Delta u_{ik} = -c_k(X) - 1_{i,\text{inst}} \cdot 1_{k \in [1,2]} \cdot \beta \cdot p_i \cdot w_k \cdot \left[ \delta_{k0}(X) + \delta_{k1}(X) \cdot p_i - d_{ik} \right] + \beta \cdot p_i + m \cdot (p_i \geq 0),
\]

which is the empirical analog of the conceptual framework.

5.2 Identification

Coefficients to be estimated include \( c_k(X) \), \( w_k \), \( \delta_{k0}(X) \), \( \delta_{k1}(X) \) for \( k \in [1,2] \), \( \beta \), and \( m \). Parameters in \( c_k(X) \) is identified as the utility intercept of the participants who enter the intrinsic treatment; given that 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 the instrumental incentives. \( \delta_{k0}(X) \) and \( \delta_{k1}(X) \) are identified from responses to the instrumental incentives that are common across types: in particular, \( \delta_{k1} \) comes from the interaction between the instrumental treatment and the amount of compensation.\(^{21}\)

Parameter \( \beta \) is identified through the variation in gift-card values. Given that \( \beta \cdot p_i \) is linear, and that there are multiple gift-card values, \( m \) is identified from the different responses to zero and non-zero incentives.

\(^{21}\)To allow for identification, the set of variables in \( \delta_{k0}(X) \) and \( \delta_{k1}(X) \) do not include income and purchase intents; for \( c_i(X) \) these two variables are included.
5.3 Estimation

I estimate the model under a Bayesian framework. Flat priors are placed for major parameters, while horseshoe priors are used for the heterogeneity parameters $c_kx$ and $\delta_kx$ (Carvalho et al. 2009). A horseshoe is similar to a Bayes lasso prior (Park & Casella 2008) in assuming sparseness. Their difference is that the former applies less shrinkage to non-zero parameters and is more robust and adaptive (Carvalho et al. 2010). This prior can be viewed as a regularizer from a frequentist perspective. Intercepts $c_{k0}, \delta_{k0}$ are left unregularized to obtain unbiased mean estimates on functions $c_k(X)$ and $\delta_k(X)$. I also place non-negativity constraints on the sensitivity to compensation $\beta$, and bound constraints on $\delta$ such that they do not exceed the actual distribution support of consumer types. In addition, the model directly estimates $\tilde{\delta}_{ik} \equiv \beta \cdot w_k \cdot \delta_{ik}$ instead of $\delta_{ik}$ for numerical stability.\(^{22}\) The distribution of $\delta_{ik}$ can be easily backed out from posterior draws.

6 Estimation Results

6.1 Model Comparison

Table 4 compares estimation results from models with different heterogeneity specifications: Model 1 assumes no heterogeneity; Model 2 allows for heterogeneity in intrinsic preferences $c_k(X)$; Model 3 allows for heterogeneity both in intrinsic preferences $c_k(X)$ and in consumer beliefs $\delta_k(X)$. To compare model performance, I calculate the expected log predictive density (elpd) using the Watanabe-Akaike information criterion (WAIC); a higher number indicates a better out-of-sample fit (Watanabe 2010). The elpd statistics are displayed in the last row of Table 4. Preference estimates are very different between the model without heterogeneity (Model 1) and the models that allow for heterogeneity in intrinsic preferences (Models 2 and 3): the latter exhibit better fits, as is demonstrated by higher elpd values. On the other hand, allowing for heterogeneity in belief does not bring better out-of-sample fit: estimation results are similar across Models 2 and 3, and the elpd of Model 2 is higher. Model 2 constitutes the basis for subsequent analysis.

6.2 Intrinsic Preferences

The willingness to accept (WTA) to give up one’s privacy due to intrinsic motives is calculated as $c_k(X) \overline{\beta}$.\(^{23}\) Figure 4 shows the predicted distribution of heterogeneous WTA for different data, and Table 5 summarizes the statistics corresponding to each distribution (see Table E.1 for credible intervals associated with these estimates). The mean intrinsic preference for sharing different pieces of data vary from $0.14$ (gender) to $1.87$ for income and $2.37$ for information about

\(^{22}\)That is, I estimate $p_i \cdot (\beta \cdot w_k \cdot d_{ik} - \tilde{\delta}_{k0} - \tilde{\delta}_{k1} \cdot p_i)$ instead of $\beta \cdot p_i \cdot w_k \cdot (d_{ik} - \delta_{k0} - \delta_{k1} \cdot p_i)$ as the instrumental-preference component.

\(^{23}\)Factors used for scaling $p_i$ are multiplied back to get the correct dollar measure.
Table 4: Intrinsic and Instrumental Preference for Privacy: Estimation Results Comparison

<table>
<thead>
<tr>
<th></th>
<th>Model 1. No Heterogeneity</th>
<th>2. Heterogeneous c</th>
<th>3. Heterogeneous c and δ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>95% CI</td>
<td>mean</td>
</tr>
<tr>
<td><strong>intrinsic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c_{\text{income}}$</td>
<td>0.568 [0.431, 0.703]</td>
<td>0.906 [0.588, 1.323]</td>
<td>0.933 [0.579, 1.510]</td>
</tr>
<tr>
<td>$c_{\text{intent}}$</td>
<td>0.550 [0.407, 0.699]</td>
<td>0.826 [0.419, 1.322]</td>
<td>0.844 [0.383, 1.377]</td>
</tr>
<tr>
<td>$c_{\text{gender}}$</td>
<td>0.015 [-0.121, 0.146]</td>
<td>0.189 [-0.162, 0.664]</td>
<td>0.241 [-0.162, 0.947]</td>
</tr>
<tr>
<td>$c_{\text{age}}$</td>
<td>0.057 [-0.090, 0.202]</td>
<td>0.262 [-0.088, 0.733]</td>
<td>0.287 [-0.159, 0.906]</td>
</tr>
<tr>
<td>$c_{\text{edu}}$</td>
<td>0.368 [0.234, 0.514]</td>
<td>0.624 [0.329, 1.051]</td>
<td>0.647 [0.289, 1.258]</td>
</tr>
<tr>
<td>$c_{\text{relationship}}$</td>
<td>0.198 [0.061, 0.331]</td>
<td>0.497 [0.124, 1.010]</td>
<td>0.549 [0.112, 1.232]</td>
</tr>
<tr>
<td>$c_{\text{kid}}$</td>
<td>0.740 [0.606, 0.875]</td>
<td>1.109 [0.790, 1.461]</td>
<td>1.090 [0.710, 1.506]</td>
</tr>
<tr>
<td>$c_{\text{zip}}$</td>
<td>0.294 [0.160, 0.430]</td>
<td>0.560 [0.227, 1.066]</td>
<td>0.603 [0.178, 1.215]</td>
</tr>
<tr>
<td>$c_{\text{race}}$</td>
<td>0.292 [0.157, 0.423]</td>
<td>0.604 [0.285, 1.104]</td>
<td>0.652 [0.262, 1.264]</td>
</tr>
<tr>
<td><strong>instrumental</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$w_{\text{income}}$</td>
<td>2.000 [1.48, 3.867]</td>
<td>2.118 [1.018, 3.989]</td>
<td>2.020 [1.140, 3.918]</td>
</tr>
<tr>
<td>$w_{\text{intent}}$</td>
<td>2.630 [1.065, 3.883]</td>
<td>1.942 [0.383, 3.762]</td>
<td>1.973 [0.286, 3.768]</td>
</tr>
<tr>
<td>$\tilde{\delta}_{\text{income},0}$</td>
<td>0.051 [-0.187, 0.288]</td>
<td>0.047 [-0.186, 0.282]</td>
<td>0.049 [-0.187, 0.284]</td>
</tr>
<tr>
<td>$\tilde{\delta}_{\text{income},1}$</td>
<td>0.052 [-0.187, 0.286]</td>
<td>0.037 [-0.192, 0.284]</td>
<td>0.050 [-0.188, 0.291]</td>
</tr>
<tr>
<td>$\tilde{\delta}_{\text{intent},0}$</td>
<td>0.084 [-0.345, 0.385]</td>
<td>0.059 [-0.352, 0.379]</td>
<td>0.066 [-0.355, 0.379]</td>
</tr>
<tr>
<td>$\tilde{\delta}_{\text{intent},1}$</td>
<td>-0.054 [-0.355, 0.309]</td>
<td>-0.049 [-0.362, 0.324]</td>
<td>-0.046 [-0.367, 0.3065]</td>
</tr>
<tr>
<td><strong>sensitivity to compensation</strong></td>
<td>$\beta$</td>
<td>0.130 [0.065, 0.208]</td>
<td>0.146 [0.070, 0.2359]</td>
</tr>
<tr>
<td>log posterior</td>
<td>-8015 [-8022, -8010]</td>
<td>-7476 [-7540, -7407]</td>
<td>-7433.06 [-7501, -7352]</td>
</tr>
<tr>
<td>elpdWAIC</td>
<td>6384 [6358, 6410]</td>
<td>6460 [6431, 6489]</td>
<td>6365 [6337, 6394]</td>
</tr>
</tbody>
</table>

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.

their children (in the follow-up survey, many participants describe the request for information about their children as “irrelevant” and “improper”). The heterogeneity of privacy preferences is substantial compared to their mean levels. The upper tails are worth more attention from a data-acquisition perspective. 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.

Are these privacy-preference estimates big or small? To interpret their magnitude, I compare these numbers with findings in the previous literature, as well as the market price for personal data when consumers are not involved. Two other studies have estimated the dollar value of revealed preference for privacy. In Hui (2007), consumers’ willingness to answer one additional sensitive question in a market research survey amounts to 2.78 Singapore dollars (2.04 USD).24 Acquisti et al. (2013) estimate consumers’ WTA of attaching their names to transactions associated with a gift card to be $1.04. Considering the variation caused by different economic contexts and categories of data requested, the magnitudes of WTA in these studies are similar to the mean preference in my data.

---

24According to their estimation, one additional Singapore dollar is associated with 0.39 util while answering one more sensitive question decreases utility by 0.14. The WTA estimate (0.39/0.14) should be interpreted as the WTA for the “marginal” question that the consumer chooses to answer.
On the other hand, consumers’ valuations for privacy are quite high compared to current market prices, where consumers remain uninformed. Castelluccia et al. (2014) find the average price for cookie impression is $0.69 CPM in the US. Miller & Skiera (2017) calculate that the lifetime value of a cookie in the EU has a mean of €1.43 and a median of €0.02. Supposing these prices are
the same as the prices for individual demographic data and that consumers are fully informed, €0.02 ($0.02) will not be high enough to convince half of the consumers to share their personal data, and €1.43 ($1.61) will still leave out the most privacy-concerned consumers from the data being collected.

A third way of gauging the magnitude of intrinsic preferences is to consider how the WTAs for individual pieces of data translate to the WTA for a profile that is essentially a bundle of different data. For example, if cookies used to identify online users are associated with different demographic tags examined above, the WTA for sharing the whole demographic profile will have a mean of $10.34 and a 97.5% quantile of $29.72. For categories of personal data that are highly granular and intimate, such as browsing and location histories, the WTAs are possibly even higher.

6.3 Instrumental Preferences

To visualize how the presence of instrumental motives changes the preferences for privacy, Figure 5 displays the distribution of individual WTAs with or without instrumental incentives across different consumer types. The instrumental incentives systematically shift the privacy preferences of the high types (the right side of the x-axis) downwards, and the low types upwards.

Consumers’ beliefs about type-related payoffs play a key role in the formation of instrumental preferences. One particular interest is the extent to which consumers’ beliefs correspond to the actual data usage. In the experiment, a consumer whose type is one tier above receives an additional 2% probability of winning the gift card if his type is disclosed to the firm; in the model, this means the true \( w \) values are 2. Column 2 of Table 4 shows that consumers’ beliefs about \( w_{\text{income}} \) and \( w_{\text{intent}} \) are accurate on average. Their beliefs about the payoff of being anonymous are much noisier, as is reflected by the wide credible intervals for \( \delta \). This pattern is consistent with the fact that guessing the payoff from withholding data requires higher-level thinking (consumers need to form beliefs about how the firm will perceive them when they withhold data).

Overall, the belief estimates provide a first glance at the level of consumer sophistication in making privacy choices when fully informed. Extrapolation of these belief estimates should be done with caution: in the field, the legalese in privacy policies often creates more opaque and confusing information environments than the one featured in my experiment. My estimates suggest that consumers make strategic reasoning when making privacy decisions, but this reasoning may or may not be based on precise beliefs about firm practices.

The impact of instrumental incentives on privacy preferences will persist as long as the first-order belief (\( w \)) is accurate. In other words, as long as consumers have relatively accurate beliefs regarding how the economic payoff is related to their private types, they will respond to the instrumental incentives by adjusting privacy choices. Rational expectation is not required for the pattern to be sustained.
6.4 The Impact of the Default Regime

To flexibly characterize how default influences privacy choices, I estimate separate models for each default frame. Table 6 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 \( \delta \) are not affected by the scaling either, since these parameters directly apply to the sensitivity to compensation parameter \( \beta \).\(^{25}\)

To compare intrinsic-preference parameters across models, Figure 6 displays the willingness to pay (WTP) of intrinsic preferences, which are heavily influenced by default. The negative WTPs

\(^{25}\)To see this point, 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 6: Privacy Preferences across Default Frames

<table>
<thead>
<tr>
<th>Default Frame</th>
<th>Opt-In</th>
<th>Opt-Out</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>95% CI</td>
</tr>
<tr>
<td>intrinsic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ε_income</td>
<td>0.906</td>
<td>[0.588, 1.323]</td>
</tr>
<tr>
<td>ε_intent</td>
<td>0.826</td>
<td>[0.419, 1.322]</td>
</tr>
<tr>
<td>ε_gender</td>
<td>0.189</td>
<td>[-0.162, 0.664]</td>
</tr>
<tr>
<td>ε_age</td>
<td>0.262</td>
<td>[-0.088, 0.733]</td>
</tr>
<tr>
<td>ε_edu</td>
<td>0.624</td>
<td>[0.329, 1.051]</td>
</tr>
<tr>
<td>ε_relationship</td>
<td>0.497</td>
<td>[0.124, 1.010]</td>
</tr>
<tr>
<td>ε_kid</td>
<td>1.109</td>
<td>[0.790, 1.461]</td>
</tr>
<tr>
<td>ε_zip</td>
<td>0.560</td>
<td>[0.227, 1.066]</td>
</tr>
<tr>
<td>ε_race</td>
<td>0.604</td>
<td>[0.285, 1.104]</td>
</tr>
<tr>
<td>instrumental</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w_income</td>
<td>2.118</td>
<td>[0.108, 3.989]</td>
</tr>
<tr>
<td>w_intent</td>
<td>1.942</td>
<td>[0.383, 3.762]</td>
</tr>
<tr>
<td>δ_income,0</td>
<td>0.047</td>
<td>[-0.186, 0.282]</td>
</tr>
<tr>
<td>δ_income,1</td>
<td>0.037</td>
<td>[-0.192, 0.284]</td>
</tr>
<tr>
<td>δ_intent,0</td>
<td>0.059</td>
<td>[-0.352, 0.379]</td>
</tr>
<tr>
<td>δ_intent,1</td>
<td>-0.049</td>
<td>[-0.362, 0.324]</td>
</tr>
</tbody>
</table>

sensitivity to compensation β  

| 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.

imply that once data are obtained by companies, consumers will not take back 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 E.2 in Appendix E for credible interval estimates).

Interestingly, Table 6 shows that consumer beliefs about instrumental payoff are not heavily influenced by the default frame. The differential impacts of default suggest that while subjective evaluations are more susceptible to the influence of the default condition, objective evaluations—beliefs on 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) affects 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 data collected. The default paradigm is also useful for thinking about the real impact of data-portability rights.\textsuperscript{26} 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.

Although understanding the mechanism of the default effect is not the main focus of my paper, it is important for measuring consumer welfare and evaluating privacy regulations. Appendix F.1 discusses how patterns in the data can be used to distinguish between different default mechanisms.

6.5 Summary

The estimation results show that the instrumental preferences for privacy imply a systematic context dependence in privacy choices. This contextual effect is a part of consumer utility and should be accounted for in welfare analysis. The extent to which instrumental preferences determine privacy decisions depends on the heterogeneity of intrinsic preferences. The estimation results

\textsuperscript{26}GDPR Article 20 and CCPA Title 1.81.5, Section 1798.100 (d).
show a substantial amount of heterogeneity in intrinsic preferences for privacy across consumers and categories of data.

The rich heterogeneity in intrinsic motives moderates the power of instrumental motives in explaining consumers’ privacy choices. A simplistic “nothing to hide” argument will lead to misleading inferences and firm decisions. For example, Cloudflare’s decision to give Tor users more CAPTCHA challenges ignited great controversy and eventually led to the retraction of this discriminating practice.27 On the other hand, viewing consumers who withhold their personal information as the same as consumers who willingly share personal data is also implausible.

The heterogeneity in privacy preferences also has profound implications for firms’ data-collection strategy. When firms’ compensation for data is common across categories of data and consumers, the amount of compensation required to collect the majority of customer data is determined by the right-tail value of privacy preferences. In my data, these tail values are often more than twice as large as the mean value. Whether data collected from privacy-insensitive consumers suffice for inference depends on the representativeness of these consumers.

7 Counterfactuals

In this section, I examine how consumers’ privacy choices influence firms’ inferences and subsequent profits in the new regulatory framework, and what firms can do to improve the quality of their inference. I investigate these questions in the context of price targeting. The application context is chosen to reflect the specific role of personal data: from a firm’s perspective, personal data differ from regular data in their ability to reveal a consumer’s private type (“profiling”) and thereby facilitate targeting. Moreover, pricing is also an area that has witnessed substantial efficiency improvement due to the recent influx of personal data. A good example is Dubé & Misra (2019), who use state-of-the-art machine learning techniques combined with big data to optimize individual-level prices. The following analysis focuses on two angles:

Ex ante: What is a firm’s WTP to compensate for consumers’ preference for privacy, and how does this WTP change with the function of personal data? How should the firm allocate its resources for data collection?

Ex post: Can the choice of whether to share personal data itself reveal information about the type of consumers who withhold their data? How much will this information improve a firm’s inference and subsequent decisions?

7.1 Setup: Price Targeting

The focal firm is the third-party company featured in the second stage of the experiment. I take a choice scenario featured in the first-stage conjoint survey to serve as the market environment (Task 3) and the product that the firm sells (Option C); they are displayed in Figure 7. Consumers’ valuation of product features and price sensitivity come directly from the conjoint survey.

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

Data available to the firm are evaluated under the consent-seeking regimes (opt-in and opt-out). Firm data are constructed in the following manner: First, I simulate 300 privacy choice draws under each counterfactual choice environment; if not specified, these are sharing choices when no compensation is given. Then I construct firm data based on these privacy choices: If a consumer decides not to share data $k$, the value of variable $k$ is left empty. Firm data also contain a “not sharing k” indicator, which equals 1 when the consumer chooses not to share $k$, and 0 otherwise. The firm’s performances in the consent-seeking regimes are compared to a benchmark scenario, where firms can amass full data without the need to request consent: this dataset contains all information given by consumers in the first stage.

To calculate optimal prices, I apply two assumptions: the marginal cost of a smartwatch is $50, the firm imputes missing variables using mean values among the data available, and takes competitors’ prices as given when doing price optimization.29 To evaluate the profitability of

---

28 This amount is the average of the estimated production cost for Apple Watch ($83.70) and the cost of 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.

29 If a consumer chooses not to share the choice task responses, the outcome variable in the pricing model is missing. In this case, I assume the firm imputes the missing outcome using observed conjoint choices from consumers who are demographically similar to this observation. In practice, this assumption amounts to leaving the whole observation out of the pricing model and predicting the missing outcome afterwards.
pricing schemes based on different datasets, I estimate consumer demand based on the full data and view this demand as the ground truth.

The main object of interest is the difference in (true) firm profits between consent-seeking and benchmark scenarios due to the difference in personal data used for pricing. In particular, I distinguish the information value of personal data for model building and for customer profiling by comparing the following two constructs:

\[
\Delta \pi_{\text{total}} = \pi(P_{d_0}(d_0)) - \pi(P_d(d)); \quad \Delta \pi_{\text{profile}} = \pi(P_{d_0}(d_0)) - \pi(P_{d_0}(d)).
\] (4)

Here, \(d\) and \(d_0\) denote data available to the firm in the consent-seeking regime and the full data, respectively. \(P_{d_0}(d)\) is the firm’s pricing model trained using \(d_0\) and takes \(d\) as input. \(\pi\) is the true profit function when the firm sets prices according to \(P(.)\). While \(\Delta \pi_{\text{profile}}\) captures the impact of a limited data set solely on profiling, \(\Delta \pi_{\text{total}} - \Delta \pi_{\text{profile}}\) measures its effect on model estimation. \(^{30}\)

The reason for decomposing the information value is two-fold. First, an information externality exists among consumer data through the model-estimation phase, but not necessarily the profiling phase. As we see below, this information externality points to strategies that a firm can use to economize on data acquisition. Second, consumers’ instrumental preferences are activated by profiling, that is, when their data are used to set prices that they receive; however, this preference may not be as strong when their data are used only for model estimation. This aspect is useful for incentive design when instrumental incentives become the main obstacle for getting representative data. Both aspects are discussed extensively in Section 7.2.

Some caveats are worth noting. Previous literature has documented that demographic variables offer less information than other personal data do when used for predicting consumer choices (e.g., see Rossi et al. 1996). The information value shown in the following exercise should be taken as a lower bound for other types of personal data, such as browsing and purchase histories. The information value may also differ in cases where personal data are used for other purposes, such as targeted advertising and product recommendation.

### 7.2 Firm’s WTP to Overcome Consumers’ Privacy Concern

By comparing the firm’s WTP with consumers’ WTA, I examine when it is worthwhile for the firm to compensate consumers for additional personal data under the new policy regime. I start by comparing cases in which consumers’ instrumental incentives are absent versus present, focusing on the opt-in regime. Within each case, I decompose the information value of personal data by their function and calculate the corresponding firm WTP, and then discuss how firms can design better data-acquisition strategies based on the findings. Lastly, I examine the information value of

\(^{30}\)It is more natural to measure the information value on model building as \(\pi(P_{d_0}(d)) - \pi(P_d(d))\) instead of \(\pi(P_{d_0}(d_0)) - \pi(P_d(d_0))\), given that data at the profiling phase are often harder to obtain than are the data for modeling.
data shared in the opt-out regime, and discuss the managerial and policy implications of different consent regimes.

A firm’s WTP to overcome privacy concerns is a metric that reflects the marginal value of additional personal data, given the data already shared with the firm. It is calculated as

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

Here, $\Delta \pi$ is the profit difference defined in (4), $N$ is the number of consumers corresponding to the profit number, and $\bar{K}$ is the average number of variables withheld per consumer when they receive no compensation for sharing data.

**Data shared as motivated by intrinsic preferences.** In some cases, consumers who receive requests for data provision do not directly experience the economic impacts from the firm’s data analysis. For example, Nielson and ComScore maintain a panel of consumers and provide the data to other firms for analysis, but these firms’ focal customers may not overlap with the panel (although they have similar demographics). Alternatively, a wedding vendor has one-off transactions with most of its customers, and those who already use its service will not expect direct economic consequences from sharing their data. In situations like these, consumers have little or no instrumental incentives.

Figure 8 illustrates how consumers’ data-sharing choices affect inferred optimal prices when these choices are motivated by pure intrinsic motives. Panel (a) shows that when firm data are incomplete both in the model estimation and profiling phases, the inferred prices are less accurate. Panel (b) shows having full data in the model-estimation phase reduces pricing mistakes, although it does not fully resolve the mistakes on the high end. Having to seek consent in the opt-in regime results in a profit loss of $1,439.91 per thousand customers when no compensation is given; this amount is 3% of the total profits that could have been obtained using full data. Inaccuracy in profiling accounts for 38.6% of the total profit loss. Panel (a) of Table 7 further describes the posterior mean and credible intervals of the profit losses at different levels of compensation.

Based on the profit numbers, we can calculate the firm’s WTP to buy additional customer data. On average, a consumer withholds 5.31 pieces of data without compensation. Suppose first that the firm can only pay the same price to different categories of data and all consumers. In this case, $WTP_{firm} = \frac{1,439.91}{5.31} = 0.27$. This price is lower than the mean consumer WTA for most personal data in my study, and certainly insufficient to persuade all consumers to share their data. In fact, a price of $1 per piece of data only persuades each consumer to share an additional 0.78 pieces of data on average and increases profit by $314.21 per thousand customers.

However, the firm can do better by allocating resources to collect data at the model-estimation phase. The key is to realize that information externalities abound in the estimation stage. That is, information from the estimation sample is incorporated into the model, and the outcome of
model deployment is applied to all customers. To leverage this information externality, firms can approximate the information value from full data by subsampling, and obtain more representative data by giving each consumer a higher amount of compensation. Assuming the model can be well calibrated by sampling 1% of the consumers, \( WTP_{firm} = (1.43991 - 0.5558) \times 100/5.31 = 16.65 \), four times as much as the 97.5% quantile of consumer WTA for the most precious data.

In reality, the performance of the model will increase with the size of the estimation sample. For subsampling to be a valid strategy to improve the efficiency of data collection, a critical condition is that estimation data has decreasing returns to scale. This condition is empirically supported by recent literature, such as Bajari et al. (2018). The proper sampling percentage will depend on the complexity of the model and how much heterogeneity the model intends to capture.

Data shared as motivated by intrinsic and instrumental preferences. In cases where a firm solicits data from its own consumers and applies its model to them, instrumental preferences will be present. In the context of price targeting, consumers’ instrumental motives are derived from price differences that they expect to receive when sharing versus withholding their data.\(^{31}\)

Importantly, the sharing decision depends on consumers’ beliefs about a firm’s pricing practice. A perfectly rational consumer will have beliefs consistent with the firm’s actual pricing practice conditional on available data in equilibrium. However, in the domain of privacy, consumers may only have limited information and cognitive capacity. As a result, their beliefs may not be perfectly consistent with the firm’s actual practice, especially when the firm’s pricing technology evolves at a fast pace. Below, I describe the case where consumers have “approximately rational” beliefs, and calculate their privacy decisions accordingly. The quantitative results are subject to the influence of belief assumption, and should be interpreted with caution.

To further simplify the analysis, I focus on the case in which the firm has previously collected full data from another set of consumers and trained its pricing model accordingly. That is, I

\(^{31}\)The instrumental incentive induced by my experiment does not apply, because it features a different economic context.
calculate the value of shared data for profiling only.\textsuperscript{32} Taking the pricing model as given, consumer $i$ expects to receive different prices when he withholds or shares data $k$:

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

Here, $i'$ denotes all other consumers in the market, $\overline{P}_i'$ is the mean price for all other consumers, and $\overline{P}_{i', Yd_{i,s}=d_{ik}}$ 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:

$$E[\Delta U] = \frac{1}{\beta_i'} \left[ \log(1 + \exp(v_i - \beta_i'\overline{P}_i')) - \log(1 + \exp(v_i - \beta_i'\overline{P}_{i', Yd_{i,s}=d_{ik}})) \right].$$ \hspace{1cm} (5) 

Here, $\beta_i$ is $i$'s price sensitivity and $v_i$ is his valuation for the product. With true rational expectation, the expected price difference between sharing and withholding data $k$ will also depend on other data-sharing decisions. Therefore, Equation (5) is a first-order approximation of true rational expectation.

Figure 9 shows the inferred optimal prices when consumers harbor both intrinsic and instrumental preferences when sharing data used for profiling. Compared with Panel (b) of Figure 8, the inferred prices here exhibit severe downward bias, especially for consumers who should be charged high prices. This pattern reflects the fact that consumers only share data that describe them as price sensitive. Despite the sample-selection bias, the resulted profit loss is moderate when the firm can manage to obtain full data for model estimation, as shown in Panel (b) of Table 7. Table 7 also indicates that compensation for data sharing is less effective in overcoming instrumental incentives, because the expected price differences in this scenario range from $20 to $50 for each data-sharing decision.

Taken together, these patterns reinforce the message that firms can be better off by allocating resources to collect data for estimation. Given the loss of inference quality in the presence of instrumental preference, firms may also gain from using a separate consumer panel for estimation as a way to “insulate” the estimation sample from instrumental concerns.

**Data shared in the opt-out regime.** In contrast to the opt-in regime, the amount of personal data that the firm can obtain in the opt-out regime is almost the same as in the benchmark scenario: on average, each consumer holds back 0.59 pieces of data without compensation. The associated profit loss is $104.91 per thousand customers (95% credible interval [$42.29, $237.74]). This amount is less than 10\% of the profit loss in the opt-in regime, or 0.3\% of the total profits that can be obtained with full data. In other words, in the opt-out regime, the firm can obtain almost full information.

\textsuperscript{32}To analyze the scenario where consumers have instrumental preferences when sharing estimation data, one needs to solve the equilibrium, because in this scenario the pricing model and the data shared depend on each other. This analysis is part of my future work.
Table 7: Profit Loss ($/1,000 customers) with Different Prices for Data (Opt-In Regime)

(a) Privacy choices motivated by intrinsic reasons

<table>
<thead>
<tr>
<th>Effect scope</th>
<th>Price per piece of data ($)</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>1,439.91</td>
<td>1,125.7</td>
<td>[536.34, 2,400.43]</td>
<td>882.89</td>
</tr>
<tr>
<td>Scoring</td>
<td>555.80</td>
<td>464.65</td>
<td>[316.91, 635.37]</td>
<td>379.77</td>
</tr>
</tbody>
</table>

(b) Privacy choices motivated by intrinsic + instrumental reasons

<table>
<thead>
<tr>
<th>Effect scope</th>
<th>Price per piece of data ($)</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scoring</td>
<td>861.64</td>
<td>856.94</td>
<td>[778.95, 891.08]</td>
<td>852.49</td>
</tr>
</tbody>
</table>

Note: This table reports posterior mean estimates, with 95% credible intervals in brackets. Total profit loss is calculated as $\Delta \pi_{\text{total}} = \pi(P_d(d_0)) - \pi(d)$; profit loss associated with profiling is $\Delta \pi_{\text{profile}} = \pi(P_{d_0}(d_0)) - \pi(P_{d_0}(d))$.

value from consumers’ data without giving any additional compensation, even when consumers are fully informed when granting consent.

7.3 The Information Value of Privacy Choices

At the stage when available consumer data are a given, what can firms do to improve the quality of their inference? Incorporating consumers’ privacy choices (i.e., decisions of whether to share personal data per se, apart from the contents of data) into the model may help. Intuitively, privacy choices can capture systematic differences in types *between* consumers who share and who withhold their data. However, their value may be limited, in that they cannot reflect the heterogeneity in
types *within* the withholding consumers. The extent to which privacy choices can improve firms’ inference remains an empirical question.

On the other hand, the California Consumer Privacy Act (CCPA) gives ambiguous requirements regarding whether firms can base their pricing and product provision on consumers’ privacy choices. Section 1798.125 (a) states (italics added by author),

A business shall not discriminate against a consumer because the consumer exercised any of the consumer’s rights under this title, including, but not limited to, by:

....

(B) Charging different prices or rates for goods or services, including through the use of discounts or other benefits or imposing penalties.

(C) Providing a different level or quality of goods or services to the consumer.

But it then goes on to say,

*Nothing in this subdivision prohibits a business from charging a consumer a different price or rate, or from providing a different level or quality of goods or services to the consumer, if that difference is reasonably related to the value provided to the consumer by the consumer’s data.*

This ambiguity has created considerable uncertainty in companies’ preparation in response to the coming regulation.

In what follows, I examine (a) how alternative interpretations of the CCPA anti-discrimination clause will affect a firm’s ability to draw inference from consumer data, and how this effect translates into the difference in firm profits; and (b) how consumers who share more or less data are affected differently in terms of the prices that they receive. I evaluate the inferred optimal prices under two models, reflecting different interpretations of CCPA. In the first model (CCPA-strict), the firm sets prices based on the content of data provided by consumers, but not their privacy choices. In the second model (CCPA-lenient), the firm sets prices based on *both* the content of available data and consumers’ privacy choices. Given that the information loss is larger in the opt-in regime, I evaluate the models within this policy scenario. Since the goal of this exercise is to compare model performance, I take actual privacy choices from the experiment to construct firm data, instead of simulating counterfactual choices and datasets which would add unnecessary noise to the results. The metrics for evaluating pricing performances are

\[ \Delta \pi_{\text{ccpa strict}} = \pi \left( P_{d_0} (d_0) \right) - \pi \left( P_{d} (d) \right), \quad \Delta \pi_{\text{ccpa lenient}} = \pi \left( P_{d_0} (d_0) \right) - \pi \left( P_{d+c} (d + c) \right), \]

where \( c \) refers to the privacy-choice indicator.

**Uniform pricing.** Examining the information value of privacy choices for uniform pricing serves as a useful starting point. Table 8 compares the inferred optimal price and associated profit

---

[^33]: https://leginfo.legislature.ca.gov/faces/billCompareClient.xhtml?bill_id=201720180AB375
[^34]: Note that CCPA does not prevent firms from price discrimination based on the content of shared data.
loss under each model. The model without privacy choices underestimates the optimal price, because consumers who care more about privacy are less price sensitive. Adding privacy choices brings the inferred optimal price close to the actual optimum, and cuts down the profit loss by 50%.

Table 8: Inferred Optimal Flat Prices under Different Models

<table>
<thead>
<tr>
<th>Data &amp; Model</th>
<th>Firm (CCPA strict)</th>
<th>Firm (CCPA lenient)</th>
<th>Full</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inferred Optimal Price ($)</td>
<td>174.87 [164.99, 186.16]</td>
<td>184.80 [172.13, 201.12]</td>
<td>185.80 [176.0, 196.07]</td>
</tr>
<tr>
<td>Profit Loss ($/1000 customer)</td>
<td>353.46 [1.03, 1310.15]</td>
<td>158.28 [0.16, 760.39]</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: This table reports posterior mean estimates, with 95% credible intervals in brackets. “Actual” refers to the benchmark scenario in which the firm sees all personal data. Profit loss here refers to $\pi(P_d(d_0)) - \pi(P(d))$, where $d$ includes privacy choices in the lenient interpretation and excludes them in the strict interpretation.

Customized pricing. Figure 10 compares the individual-level optimal prices predicted by these two models. Prices predicted by the CCPA-lenient model are less biased: the mean price that consumers receive under this model is $194.26, close to the mean price $199.05 when the firm has all data; in comparison, the mean prices under the CCPA-strict model is $179.22. However, predictions generated by the CCPA-lenient model are not necessarily more accurate at first glance. The CCPA-lenient model surpasses the alternative model when predicting prices for consumers who have high valuations for the product, but performs worse at the opposite end of the spectrum.

Table 9 compares the profit losses from using each pricing model, and in particular, the profit differences when applying the predicted prices to specific consumer subsets, defined by their privacy choices. Adding privacy choices to the model improves the predicted prices for
consumers who share no personal data. On the other hand, the prediction accuracy for consumers who already share lots of data can suffer, because privacy choices add little additional explanatory power for the preferences of these consumers. This pattern suggests that in the CCPA-lenient domain, the firm may be able to obtain better inference by taking the ensemble of the two models.

Table 9: Profit Loss When Using Firm Data ($/1000 Consumers)

<table>
<thead>
<tr>
<th>Consumer Subset</th>
<th>Model</th>
<th>All consumers</th>
<th>Share all data</th>
<th>Share no data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CCPA strict</td>
<td>2,440.48 [916.84, 5,113.17]</td>
<td>2,347.53 [926.25, 5,720.58]</td>
<td>2,492.27 [1,229.00, 4,185.65]</td>
</tr>
<tr>
<td></td>
<td>CCPA lenient</td>
<td>2,384.15 [956.78, 5,229.10]</td>
<td>2,405.26 [877.12, 5,192.52]</td>
<td>2,418.52 [1,137.89, 3,853.76]</td>
</tr>
</tbody>
</table>

Note: This table reports posterior mean estimates, with 95% credible intervals in brackets.

To further understand how the CCPA-lenient model affects consumers who make different privacy choices, Figure 11 separately displays prices for consumers who withhold at most one piece of data and prices for consumers who withhold most of their data. For consumers who already share a lot, prices set in the two models are not very different. For consumers who choose not to share most personal data, the actual optimal prices are, on average, higher than those for privacy-insensitive consumers (as can be seen by comparing the gray lines across the two panels). The privacy-choice indicators pick up this information, resulting in a rise in inferred optimal prices. The fact that privacy choices only convey aggregate-level information about consumer types is reflected by the fact that the new inferred prices are shifted up by almost the same amounts when compared to the original prices inferred.

Figure 11: Inferred Optimal Prices with or without Privacy Choices (by Consumer Subsets)

Taken together, these results paint a nuanced picture of the information value of privacy choices. Incorporating consumers’ privacy choices into a firm’s decision model can improve the model’s aggregate-level inference by adjusting for sample selection. This adjustment is agnostic
about the direction of selection bias. In fact, it can be used as a tool for assessing the direction of bias caused by sample selection. On the other hand, the information value is limited when the goal is to improve individual-level pricing. The reason is that the privacy-choice indicators only contain information about the average type of consumers choosing not to share data.

7.4 Summary

The counterfactual studies suggest that firms can improve their inference on consumers by subsampling when collecting personal data used for estimation, and by using privacy choices to learn about self-selection bias when estimating models. Both strategies avoid the use of arbitrary assumptions on consumers’ privacy motives when making inferences. The quantitative results are subject to the influences of assumptions on firm behavior and the application context. However, the strategies developed based on the qualitative findings are generic to inference problems: they can be applied not only when personal data are collected for pricing, but also when they are used for other managerial decisions such as targeted advertising and customized product recommendation. They can also be applied when personal data are requested to conduct general-interest research with a goal of inference.35

8 Conclusion

Privacy choices are motivated both by intrinsic preference—a taste for privacy, and by instrumental preference—the utility change from disclosing one’s type relevant to the specific market environment. While the instrumental preference for privacy is induced by economic contexts, the intrinsic preference is a utility primitive. The interplay of these two preference components has important implications for measuring privacy preferences across contexts, improving inferences using personal data shared by consumers, and evaluating the effectiveness of data-collection incentives.

By separating intrinsic and instrumental motives using experimental variation, I discover the following findings. First, the WTA corresponding to intrinsic preferences has a mean ranging from $0.14 to $2.37; however, this WTA is highly heterogeneous and skewed to the right. Second, instrumental preferences shift privacy choices across economic contexts. Third, consumers self-select into sharing data based on their privacy preferences; the direction and magnitude of this self-selection are determined by the heterogeneity and correlation of the two preference components. Counterfactual analysis shows that firms and researchers can take the following steps to improve their inferences based on consumers’ personal data: Ex ante, they can strategically allocate more resources to collect high-quality data where the information gain is largest. Ex post, incorporating

35See https://socialscience.one/blog/first-grants-announced-independent-research-social-media%E2%80%99s-impact-democracy.
privacy choices into models can allow them to learn about the impact of consumers’ self-selection on inference and improve aggregate-level decision-making.

My paper sheds light on how consumers’ privacy choices affect firms’ profits by changing the quality of their inferences. Recent literature (Miller & Tucker 2009, Goldfarb & Tucker 2012a, Adjerid et al. 2015, Jia et al. 2018) has been focusing on how privacy choices affect innovation, which is another important aspect. Policy analysis needs to account for both inference and innovation angles when estimating the impact of privacy regulations on the industry. A more thorough welfare analysis also needs to account for the externalities in consumers’ privacy choices. Although my paper does not estimate social welfare, it reveals how the information externality may vary with firms’ strategic uses of data. Given that the externality affects consumers mainly through instrumental consequences, the equilibrium impact of the externality will depend on the relative magnitudes of the two privacy-preference components.

The experiment is designed to exclude confounds for measuring privacy preferences while being as close to a natural market environment as possible. Future studies can be carried out to study privacy choices in a broader range of market contexts. For instance, researchers can compare information environments that feature transparent versus obscure information about the purpose of data collection; this design will be useful to study consumers’ belief and its role in moderating privacy choices. Researchers can also collect privacy choices across multiple economic contexts and compare privacy choices of similar consumers.

Future analysis will further explore the implications of the coexistence of intrinsic and instrumental preferences. One direction is to investigate the substitution and complementarity in data-sharing decisions, which is likely present given that one variable often contains information about another. Another direction is to develop better models to extract information from consumers’ data-sharing decisions. A third extension is to explore optimal compensation for data procurement that incorporates instrumental incentives.
References


274.

54(2), 442–92.


and technology incentives: The case of health information exchanges’, Management Science 62(4), 1042–
1063.


pp. 38–42.


of Legal Studies 45(S2), S41–S67.

1678.

Economics 120(4), 1260–1278.

American Economic Review 105(9), 2798–2837.

Berkheim, B. D. & Rangel, A. (2010), Behavioural public economics, in ‘Behavioural and Experimental

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.

Brynjolfsson, E., Collis, A. & Eggers, F. (2019), ‘Using massive online choice experiments to measure changes


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

97(2), 465–480.

Castelluccia, C., Olejnik, L. & Minh-Dung, T. (2014), Selling off privacy at auction, in ‘Network and Dis-
tributed System Security Symposium (NDSS)’.

97(3), 745–788.

47(2), 315–72.


Miller, K. & Skiera, B. (2017), ‘Economic damage of cookie lifetime restrictions’, *Available at SSRN 3046254*.


A Proof for Proposition 1

First, define the notation for the means and covariances of preference components: $E[c_i] = \mu_c$, $Var[c_i] = \sigma_c^2$, $E[-T(d_i)] = \mu_t$, $Var[-T(d_i)] = \sigma_t^2$; $\text{Cov}(c_i, -T(d_i)) = \rho$. Note that $\Delta T(d_i) = T(F_d(d|s = 0)) - T(d_i)$, where $T(F_d(d|s = 0))$ does not vary across consumers. Therefore, $Var[\Delta T(d_i)] = \sigma_t^2$ and $\text{Corr}(c_i, \Delta T(d_i)) = \rho$.

Denote the total preference for privacy as $g_i$. Then,

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

$\text{Corr}(g_i, \Delta T(d_i))$ 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 transfer to non-disclosing consumers $T(F_d(d|s = 0))$), $\text{Corr}(g_i, \Delta T(d_i))$ is a direct assessment of the information value of non-sharing decisions for inferring consumer types. The following observations hold:

1. $\text{Corr}(g_i, \Delta T(d_i)) > 0$ if and only if $\rho + \frac{\mu_t}{\sigma_c} > 0$.
2. $\text{Corr}(g_i, \Delta T(d_i))$ increases with $\frac{\mu_t}{\sigma_c}$, and strictly increases with $\frac{\mu_t}{\sigma_c}$ if $|\rho| < 1$.
3. $\text{Corr}(g_i, \Delta T(d_i))$ increases with $\rho$ if $\sigma_c + \rho \sigma_t > 0$, and decreases with $\rho$ if $\sigma_c + \rho \sigma_t < 0$.

Observation 3 reveals a more nuanced relationship between the explainability of instrumental preference and the correlation between the two preference components. In particular, if $\sigma_t > \sigma_c$, a regime $\rho \in [-1, -\frac{1}{\sigma_c}]$ exists where an increase in $\rho$ leads to a decrease in $\text{Corr}(g_i, \Delta T(d_i))$. The reason is that when $\rho$ is close to -1, the variation in instrumental preference dominates intrinsic preference ($\sigma_t > \sigma_c$), leading to a perfect correlation between total preference $g_i$ and instrumental preference $\Delta T(d_i)$. Once $\rho$ deviates away from -1, this relationship is loosened. Note that when $\sigma_t < \sigma_c$, $\text{Corr}(g_i, \Delta T(d_i))$ always increases with $\rho$.

The proof will still go through regardless of the level of $T(F_d(d|s = 0))$. In particular, consumers need not have rational expectations, such that their beliefs about $T(F_d(d|s = 0))$ are consistent with the actual transfer that the firm gives to consumers who withhold their data. In other words, the conclusions above are robust to the mismatch between consumer beliefs and firm practices, and can account for scenarios where firms actively experiment or where information is inadequate for consumers to form rational beliefs. The proof also remains valid when compensation for data sharing is present.
B Screenshots of the Survey

B.1 Conjoint and Demographic Questions

Figure B.1: Example Questions in the First Stage

(a) Conjoint Question

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

<table>
<thead>
<tr>
<th>Product</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitness Tracking</td>
<td>Activity</td>
<td>Activity + heart rate</td>
<td>Activity + heart rate</td>
<td>Activity + heart rate</td>
</tr>
<tr>
<td>Voice Control</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Mobile Payment</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Encryption</td>
<td>TLS 1.0</td>
<td>SSL 3.0</td>
<td>SSL 3.0</td>
<td>TLS 1.0</td>
</tr>
<tr>
<td>Login Option</td>
<td>Pin, pattern</td>
<td>Pin, pattern</td>
<td>Pin, pattern</td>
<td>Pin, pattern, face</td>
</tr>
<tr>
<td>Price</td>
<td>$299</td>
<td>$149</td>
<td>$199</td>
<td>$249</td>
</tr>
</tbody>
</table>

- Product A
- Product B
- Product C
- Product D
- None of the above

(b) Demographic Question

What is your total household income for the last calendar year, before taxes?

- Less than $25,000
- $25,000 to $50,000
- $50,000 to $75,000
- $75,000 to $100,000
- $100,000 to $200,000
- $200,000 or more
- Prefer not to say
B.2 Compensation Schedules across Treatments

Figure B.2: Displayed Compensation Schedule: Intrinsic Treatment

(a) Main Screen

You will have the chance to win a $50 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
- Kids at home
- Age
- Gender
- Marital status
- Income
- Education
- Ethnicity

(b) Details Screen

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

\[
\text{Probability of winning} = \text{number of boxes checked} \times 1\%
\]

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.3: 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:

\[
\text{Baseline probability of winning} = \frac{\text{Number of boxes checked}}{100} \times 1\%
\]

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:

<table>
<thead>
<tr>
<th>Income</th>
<th>&lt; $50,000</th>
<th>$50,000 - $75,000</th>
<th>&gt; $75,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjustment</td>
<td>-2%</td>
<td>Unchanged</td>
<td>+2%</td>
</tr>
<tr>
<td>Plan to purchase a smartwatch in the next 3 months</td>
<td>Somewhat or extremely unlikely</td>
<td>Neither likely nor unlikely</td>
<td>Somewhat or extremely likely</td>
</tr>
<tr>
<td>Adjustment</td>
<td>-2%</td>
<td>Unchanged</td>
<td>+2%</td>
</tr>
</tbody>
</table>

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 Covariate Balance Assessment

To examine if treatment randomization satisfies ex-post covariate balance, I decompose all discrete variables into binary values and plot out the standardized mean differences across treatments for each variable. Stuart et al. (2013) propose 0.1 as a conservative threshold for assessing covariate balance. Figure C.1 shows that for both the instrumental-incentive treatment and the default-condition treatment, all standardized mean differences are within this threshold. For the compensation-amount treatment, at most 3 variables out of 121 (the number of binary variables after discretization) pass the threshold, which correspond to the start time and the end time of survey responses. The difference comes from the fact that data collection finishes at different times in zero-incentive versus incentivized treatments, and that they have different lengths. Other covariance-balance-test metrics, including the exact multinomial tests (for discrete) or t-tests (for continuous), yield similar results.

Figure C.1: Mean Difference between Standardized Variables across Treatment Groups
To the extent that the systematic variation shown in Figure 2 may be driven by variables correlated with consumer types, a model that controls for additional variables serves as a robustness check. To this end, I use a propensity forest (Wager & Athey 2018) to estimate how privacy choices respond to the instrumental incentives. The causal forest model can efficiently estimate causal parameters without imposing linearity restriction, allowing me to account for potential influences from over 120 variables and obtain estimates robust to specification errors. I choose the propensity forest over the two-sample forest because the former can account for endogeneity. Although the treatment itself is exogenous, the instrumental preference is identified as the interaction effect between treatment and consumer types, and types may be correlated with other covariates, which is the motivation for using this model.

I build two propensity forests: one to estimate causal effects of instrumental incentives on income-sharing decisions, and the other on purchase-intent sharing. Covariates include the following variables: demographics, purchase intent, other treatment indicators, survey quality indicators, survey starting time, browser used, and device specifications. These variables are further screened by a test of treatment overlap; only variables that pass the test enter the estimation model. The final data for estimation include 125 explanatory variables.

Figure D.1 plots the distribution of the heterogeneous treatment effect of instrumental incentives on privacy choices. Panel (a) compares the effect across income groups. Among the lower end of the income spectrum, the effects on adjacent income groups are harder to distinguish from each other. Meanwhile, the distribution corresponding to the lowest-income cohort (“< $25,000”) is stochastically dominated by others. The distribution separation is also clearer when we focus on income groups not adjacent to each other (e.g., “< $25,000” and “$100,000 – $200,000”). In Panel (b), the separation pattern is much stronger: distributions of sharing decisions for lower purchase intent cohorts are located more to the left, indicating a lower level of instrumental utility. Future versions of this analysis will add confidence intervals associated with each distribution, which can be calculated via the infinitesimal jackknife (Wager et al. 2014).

As a cautionary note, the distances between distributions do not necessarily reflect the magnitude of instrumental preference. The reason is that the coefficients are estimated as the marginal effect on shifting privacy choices, which is affected by the baseline probability of sharing and the magnitude of unexplained variation. To directly compare intrinsic and instrumental preferences, having a cardinal measure is essential.

---

36 These indicators include whether a participant chooses the same response throughout the conjoint survey, whether a participant gives nonsensical answers to open-ended questions, and time spent on the survey.

37 These specifications include the device type, operating system, and screen size.

38 The overlap condition requires that for the covariates of interest, 0 < Pr(T = 1|X_i = x) < 1. This is one of the key assumptions (besides unconfoundedness and SUTVA) that need to be satisfied for the treatment effect to be estimated. Even with treatment randomization, the overlap condition may still be violated with high-dimensional covariates or rare covariate values; therefore, an overlap test is necessary.
Figure D.1: Heterogeneous Treatment Effect of Instrumental Incentive on Privacy Choices

(a) Across Income Groups

(b) Across Purchase Intent Groups

E  Credible Intervals for Privacy Preference Dollar Value Estimates

Table E.1: Posterior Estimates of Parameters Associated with WTA Distribution

<table>
<thead>
<tr>
<th>Parameter</th>
<th>(a) WTA Mean</th>
<th>(b) WTA Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>95% CI</td>
</tr>
<tr>
<td>income</td>
<td>1.870</td>
<td>[1.012, 3.518]</td>
</tr>
<tr>
<td>intent</td>
<td>1.825</td>
<td>[0.981, 3.534]</td>
</tr>
<tr>
<td>gender</td>
<td>0.142</td>
<td>[-0.285, 0.709]</td>
</tr>
<tr>
<td>age</td>
<td>0.260</td>
<td>[-0.172, 0.805]</td>
</tr>
<tr>
<td>edu</td>
<td>1.228</td>
<td>[0.619, 2.337]</td>
</tr>
<tr>
<td>relationship</td>
<td>0.687</td>
<td>[0.249, 1.454]</td>
</tr>
<tr>
<td>kid</td>
<td>2.367</td>
<td>[1.337, 4.523]</td>
</tr>
<tr>
<td>zipcode</td>
<td>0.985</td>
<td>[0.450, 1.992]</td>
</tr>
<tr>
<td>race</td>
<td>0.980</td>
<td>[0.437, 2.008]</td>
</tr>
</tbody>
</table>
Table E.2: Posterior Estimates of Parameters Associated with WTP Distribution

(a) WTP Mean

<table>
<thead>
<tr>
<th>Parameter</th>
<th>mean</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>income</td>
<td>-66.59</td>
<td>[-621.55, -6.92]</td>
</tr>
<tr>
<td>intent</td>
<td>-78.87</td>
<td>[-733.52, -8.29]</td>
</tr>
<tr>
<td>gender</td>
<td>-89.84</td>
<td>[-866.19, -9.28]</td>
</tr>
<tr>
<td>age</td>
<td>-76.57</td>
<td>[-722.37, -8.03]</td>
</tr>
<tr>
<td>edu</td>
<td>-81.10</td>
<td>[-767.11, -8.41]</td>
</tr>
<tr>
<td>relationship</td>
<td>-82.10</td>
<td>[-773.98, -8.63]</td>
</tr>
<tr>
<td>kid</td>
<td>-70.15</td>
<td>[-634.81, -7.34]</td>
</tr>
<tr>
<td>zipcode</td>
<td>-70.52</td>
<td>[-653.87, -7.46]</td>
</tr>
<tr>
<td>race</td>
<td>-86.69</td>
<td>[-834.71, -8.97]</td>
</tr>
</tbody>
</table>

(b) WTP Standard Deviation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>mean</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>income</td>
<td>70.23</td>
<td>[2.41, 479.59]</td>
</tr>
<tr>
<td>intent</td>
<td>35.52</td>
<td>[0.96, 249.39]</td>
</tr>
<tr>
<td>gender</td>
<td>57.36</td>
<td>[1.92, 360.06]</td>
</tr>
<tr>
<td>age</td>
<td>76.40</td>
<td>[2.70, 522.52]</td>
</tr>
<tr>
<td>edu</td>
<td>34.59</td>
<td>[1.22, 248.68]</td>
</tr>
<tr>
<td>relationship</td>
<td>43.54</td>
<td>[1.22, 286.02]</td>
</tr>
<tr>
<td>kid</td>
<td>29.53</td>
<td>[1.37, 185.01]</td>
</tr>
<tr>
<td>zipcode</td>
<td>57.63</td>
<td>[3.04, 405.01]</td>
</tr>
<tr>
<td>race</td>
<td>28.37</td>
<td>[1.30, 213.90]</td>
</tr>
</tbody>
</table>

F Psychological Factors

F.1 The Default Frame

When does the mechanism matter? When the research goal is to estimate consumer welfare, it becomes necessary to figure out to what extent the friction created by the default frame is part of the welfare preference (experienced utility) as opposed to behavioral preference (decision utility). Different mechanisms imply different ways that the welfare preference should be calculated. For example, if consumers stick to default due to inattention, the default effect is separate from welfare utility. On the other hand, if default shifts choices via implicit endorsement, that is, by providing information that changes consumers’ evaluation of different options, it should be part of welfare utility (see Bernheim & Rangel (2010), Goldin & Reck (2018) for a review).

How does default matter? Previous literature has proposed different explanations for how the default condition shifts choices. Within my experiment context, three mechanisms are the most likely (broadly categorized):

- **Inattention/cost of gathering information.** Both of them can be either rational or irrational: rational means a decision maker endogenously determines whether it is worthwhile to pay attention or gather additional information, depending on the stakes of the decision. Irrational means no such ex-ante trade-off is involved (Karlan et al. 2016).

- **Anchoring (including endorsement effect).** It means that the default option serves as an attraction point. This can be caused by the fact that its advantages are more salient (Goswami &

---

39 Other mechanisms do not apply to my setting. For example, choice procrastination is unlikely to be the driving factor because participants are not given a long time frame for making choices. Mechanical switching cost is also unlikely, since it only takes one click for participants to change their decision.
Urminsky 2016), or because it is viewed as an implicit recommendation (Madrian & Shea 2001).

- **Loss aversion (also called reference dependence).** It means that loss looms larger than gains (Kahneman 1979, Thaler 1980).

Below, I discuss how patterns in my data can help distinguish between different mechanisms:

**Sign switch between WTA and WTP is inconsistent with pure loss aversion.** Loss aversion merely scales gains and losses, which is sign preserving. The coexistence of positive WTA and negative WTP implies that other mechanisms must be at work.

**Changes in sensitivity ranking across data suggest anchoring effect.** Neither inattention (information gathering cost) nor loss aversion can generate a switch of rankings across different pieces of data. Under inattention, the decision maker either sticks to default, or makes attentive choices that rank different options in the same way across frames. Loss aversion implies privacy costs for sharing different data are scaled by the same factor. Under anchoring, however, values that are more certain to the decision maker are less susceptible to the influence of defaults, which can generate changes in relative sensitivity among data. The fact that beliefs about the instrumental payoff are less influenced by the default frame is also consistent with this explanation.

**Greater sensitivity to compensation in the opt-in frame is consistent with rational inattention but not loss aversion.** The impact of rational inattention on sensitivity to economic payoffs can be best illustrated by Figure F.1. Since the utility from sharing data increases with the amount of compensation, the welfare utility is upward sloping (the dashed line). The behavioral utility in the opt-in frame is always closer to zero than the welfare utility (the blue line below), while behavioral utility in the opt-out frame exhibits the opposite pattern (the red line above). Under rational inattention, the impact of the default frame diminishes as the stake increases, generating a steeper utility response in the opt-in frame and a flatter response in the opt-out frame. By contrast, under loss-aversion theory, not sharing means “gaining” privacy and “losing” money in the opt-out frame; therefore, decision makers should be more sensitive to monetary payoff when the default is opt-out.

To directly compare sensitivity to economic payoffs across frames, data from different default conditions should be contained in the same model to avoid mechanical differences caused by scaling. Table F.1 displays the estimation result corresponding to $\beta$ from the pooled regression. It shows that participants in the opt-in frame are indeed more sensitive to compensation than in the opt-out frame.

To conclude, my data suggest that inattention/information-acquisition costs and anchoring are the main drivers of the default effect in my experiment. The fact that anchoring is one of the drivers implies the impact of default is likely to be asymmetric: previous literature shows that the welfare utility is often closer to the behavioral utility in the opt-in frame (Madrian & Shea 2001,
Figure F.1: Sensitivity to Compensation with Rational Inattention

![Diagram showing sensitivity to compensation with rational inattention.](image)

Table F.1: Sensitivity to Economic Payoff across Default Frames (Pooled Estimation)

<table>
<thead>
<tr>
<th>Default Frame</th>
<th>Opt-In</th>
<th>Opt-Out</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>95% CI</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.138</td>
<td>[0.059, 0.231]</td>
</tr>
</tbody>
</table>

Tannenbaum & Ditto 2012). Given this normative ambiguity, I refrain from trying to back out the welfare utility given the current data constraints.

F.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.