Choice Architecture, Privacy Valuations, and Selection Bias in Consumer Data

Tesary Lin* Avner Strulov-Shlain†

July 12, 2024

Abstract

How does choice architecture used during data collection influence the quality of collected data in terms of volume (how many people share) and representativeness (who shares data)? To answer this question, we run a large-scale choice experiment to elicit consumers' valuation for their Facebook data while randomizing two common choice frames: default and price anchor. An opt-out default decreases valuations by 22% compared to opt-in, while a \$0–50 price anchor decreases valuations by 37% compared to a \$50–100 anchor. Moreover, some consumer segments are influenced by frames more while having lower average privacy valuations. As a result, conventional frame optimization practices that aim to maximize data volume can exacerbate bias and lower data quality. We demonstrate the magnitude of this volume-bias trade-off in our data and provide a framework to inform optimal choice architecture design.

Keywords: privacy, choice architecture, market for data, selection bias, experiment

*Boston University Questrom School of Business; tesary@bu.edu

†University of Chicago Booth School of Business; avner.strulov-shlain@chicagobooth.edu.

The authors thank Guy Aridor, Dan Bartels, Josh Dean, Berkeley Dietvorst, Dante Donati, Kwabena Donkor, Andrey Fradkin, Sam Goldberg, Ali Goli, Avi Goldfarb, Michael Grubb, Tanjim Hossain, Yufeng Huang, Alex Imas, Jihye Jeon, Garrett Johnson, Yucheng Liang, Nina Mazar, Sanjog Misra, Ilya Morozov, Olivia Natan, Omid Rafieian, Heather Sarsons, Ananya Sen, Tim Simcoe, K. Sudhir, Oleg Urminsky, Giorgos Zervas, and Jinglong Zhao for their helpful comments. We thank Christy Kang, Paulina Koenig, and Kaushal Addanki for their excellent research assistance. This research is funded by the Becker Friedman Institute at the University of Chicago and the Willard Graham Research Fund at Chicago Booth. It was approved by the Institutional Review Boards at the University of Chicago (IRB21-1376) and Boston University (IRB-6239X).

1 Introduction

Companies often deploy some form of "choice architecture" (Thaler & Sunstein 2008) when collecting consumer data, which are choice environments designed to nudge consumers towards sharing more data, all else equal. [1] For instance, after the GDPR took effect, websites have been using a combination of default settings, salient options, and obstructions to nudge their users toward sharing cookie identifiers (Matte et al. 2020, Nouwens et al. 2020). Several major consent management platforms, such as OneTrust and Usercentrics, offer products and resources to help websites design user interfaces that will maximize opt-ins.[2]

Existing choice architecture optimization practices emphasize maximizing the volume of data collection, often neglecting a second dimension of data quality—the representativeness of the data collected. Biased input data often leads to biased insights and decision-making (Cochran et al. 1954). For instance, Cao et al. (2021) show that gender bias in Product Hunt's product votes leads to severe bias in the predicted appeal of new products, causing female-focused products to experience 45% less growth than male-focused counterparts. As another example, Bradley et al. (2021) show that using large but unrepresentative samples leads to overestimates of COVID-19 vaccine uptake by 14-17 percentage points compared to using a small representative sample. Bias in shared consumer data can often ensue when different consumer groups have different valuations of their data (Lin 2022). However, little attention is paid to how choice architecture influences the representativeness of data shared when different groups of consumers respond to choice architecture to varying degrees.

In this paper, we ask the following questions: How do choice frames influence consumers' privacy valuations, and what is the heterogeneity of the choice frame effects? How do the choice frames change the composition of consumers sharing their data, beyond their influence on the quantity of data shared? How should firms choose their choice architecture to optimize data quality? The economic returns of data to firms depend on both its quantity and its representativeness. Therefore, to assess how choice architecture affects the quality of data collected and the efficiency of data collection, we must account for its effect on both the *volume* and the composition of shared data, henceforth its *bias*.

To answer these questions, we developed a simple conceptual framework and ran an experiment to illustrate the key mechanisms that affect the volume-bias trade-off. Our conceptual framework connects a firm's choice of frames and price for data with their impact on the volume and bias of data collected and provides sufficient statistics to aid decision making. Empirically, we recruited 5,028 Facebook users, and elicited their willingness-to-accept (WTA) to sell their data while randomizing the choice frames they faced. Within participants, we asked their valuation

[1]We use "choice architecture", "choice frames", and "frames" interchangeably throughout the paper.

[2]https://www.onetrust.com/blog/onetrust-launches-consent-rate-optimization-to-maximize-opt-ins/; https://usercentrics.com/resources/opt-in-optimization/

for sharing the following variables with researchers and advertisers: *about me* (their information on the profile page), *posts*, *likes*, *friends and followers*, and *responses to our survey*. For each variable, we elicited incentive-compatible WTA using a multiple price list (Kahneman et al. 1990, Andersen et al. 2006), followed by a free-text entry. Across participants, we randomized the choice default and the price anchor. Both frames possibly affect choices without restricting them. *Default* varies between opt-in, opt-out, and active choice. *Price anchor* is the range of prices in an initial multiple price list, which is either \$0-\$50 (low) or \$50-\$100 (high). We also collect consumer characteristics to explore heterogeneity in their privacy valuations and responses to choice frames. These variables include demographics, social media and Internet usage, and their belief about what data are already available to Facebook and the public.

Consumers' valuations for data are substantially different across both individuals and data types. Across individuals, valuations for the same data range from \$0 all the way up to infinity, with 20% having their valuations above \$100, the maximal value that we offer in the multiple price list. When we top-code these values at \$100, the mean valuation across data is \$54.2. The difference in mean WTA between the most valuable (friends and followers) and the second most valuable data (posts) is \$9.1, while the difference between the most valuable and least valuable data (survey responses) is \$29.1.

We also find a significant influence of choice frames on consumers' valuations. Consumers decrease their valuations by 37.4% in the *low* versus *high* price anchor condition. Compared to an *opt-out* default, *active choice* increases the valuation of data by 11.8%, and *opt-in* increases the valuation by 22.1%. On average, the difference in valuation is \$16.2 lower due to the *low* price anchor, and is \$2.4 and \$5.2 higher in the *active* and *opt-in* defaults.

To explore how choice frames affect the composition of shared data, we estimate causal forest models (Athey et al. 2019) to see what consumer attributes correlate with the heterogeneity in data values and frame effects. The causal forest models allow us to nonparametrically and efficiently characterize the joint distribution of privacy valuations and frame effects. We find that consumers' valuations of their data and their responses to choice frames are negatively correlated overall. Such a negative correlation is stronger across specific consumer segments. Consumers who value their data *less* across frames are overall younger, poorer, less educated, and more likely to click on ads while on Facebook. Interestingly, these attributes also predict *larger* frame effects.

The negative correlation between privacy valuations and choice frame effects creates a potential trade-off between volume-maximizing and bias-minimizing objectives during data collection. Under a volume-maximizing frame, consumer groups who already value their data less now give up their data even more willingly due to the frame. In this case, the collected data may oversample these groups even more, while possibly undersampling other consumer groups if the firm also sets a lower price for data due to the supply expansion. Instead, there is often a different frame that will minimize bias at the expense of collecting less data. The strength of this trade-off will affect the preferred choice architecture.

We perform counterfactual analysis to explore the volume-bias trade-off within our conceptual framework. We start by comparing the quality of data collected under the volume-maximizing (*vol-max*) frame and the bias-minimizing frame (*bias-min*). We find that different frames indeed prioritize volume and bias differently, confirming the empirical existence of the trade-off. In particular, there are substantial relative differences in bias between the vol-max frame and the bias-min frames. In fact, sometimes the vol-max frame exacerbates the bias even more than a no-optimization benchmark where a firm randomly selects among possible frames.

We then cast the problem in a decision-theoretic framework to derive and estimate three key statistics that determine frame choice: the elasticity of volume to price; the elasticity of bias to price; and the ratio between the relative changes in bias and volume when switching from the volume-maximizing frame to the bias-minimizing one. Our estimates show that a firm will prefer to move away from the volume-maximizing frame if it values a 10% bias reduction more than a 0.36% increase in volume. The bias-volume trade-off is often more pronounced when the sample data is small, while the two frames tend to coincide as the sample data approaches 100%. This finding suggests that firms may face varying degrees of data quality trade-off depending on their stage in consumer data gathering. The trade-off shifts in favor of bias mitigation once we allow the firm to personalize its frame assignment. With personalized frames, the ratio of bias reduction to volume reduction nearly doubles, indicating that a volume-maximizing frame is less likely to be optimal.[3]

Our experiment is not intended to replicate the current data-sharing environments that consumers face. Rather, it is a practical demonstration of how frame effect heterogeneity impacts the quality of shared data. We intend our experiment to be a proof of concept that shows qualitatively how choice architecture can influence data quality, rather than a quantification of specific frame effects. By uncovering the components of the volume-bias trade-off when companies optimize their frames, our experiment informs the design of choice architecture broadly, beyond just the choice of default and anchoring frames.

We show how a well-designed choice architecture can balance reach and distributional goals. This point applies well beyond the data market settings: Whenever distributional goals are present in government programs (Linos 2018, Misra 2023) or marketing campaigns (Dubé & Misra 2023), understanding the frame effect heterogeneity allows us to design a choice architecture that balances reach and distribution. Our framework is directly applicable to these settings as well.

Our paper contributes to the existing literature on several fronts. Our main contribution lies in showing how the heterogeneous effects of choice architecture create selection issues in the data market. Building on the literature that identifies consumers' privacy valuations using a revealed preference approach (Goldfarb & Tucker 2012, Athey et al. 2017, Kummer & Schulte 2019), we extend empirical studies quantifying the value of data to consumers, an area that is still

[3]This result is driven by the fact that in our setting, the frame that encourages data sharing the most is generally the same (low anchor + opt-out) across consumer subgroups, while personalization brings a more substantial gain in bias reduction.

nascent (Acquisti et al. 2013, Lin 2022, Tang 2019, Collis et al. 2020). One reason for the scarcity of papers that quantify privacy values is that privacy preferences are context-specific and hard to measure (Martin & Nissenbaum 2016). Lin (2022) highlights consumers' economic reasoning in different data usage scenarios (the *instrumental preference*) as a contributor to the context effect. In comparison, we focus on the behavioral contributor to the context effect—the choice architecture, and explore how it affects the composition of consumers who share data.

A wealth of literature has examined how choice architecture affects privacy choices (Johnson et al. 2002, Acquisti et al. 2012, 2013, Adjerid et al. 2019, Kormylo & Adjerid 2021, Mrkva et al. 2021, D'Assergio et al. 2022, Tomaino et al. 2022). Our innovation lies in linking the heterogeneity of frame effects with consumers' self-selection into sharing data and the quality of sample data. This connection bridges two previously disjointed areas: behavioral biases in privacy valuations and data market efficiency (Arrieta-Ibarra et al. 2018, Acemoglu et al. 2022, Bergemann et al. 2022, Ichihashi 2021, Markovich & Yehezkel 2021). We also augment the empirical literature on behavioral industrial organization that quantitatively takes consumers' biases and fallibility into firms' decision-making processes (Peltzman 1981, Rao & Wang 2017, Strulov-Shlain 2023, Miller et al. 2022).

Our project also contributes to the literature on the value of consumer data to firms (Rossi et al. 1996, Trusov et al. 2016, Miller & Skiera 2017, Bajari et al. 2019, Aridor et al. 2021, Rafieian & Yoganarasimhan 2021, Sun et al. 2021, Wernerfelt et al. 2022, Lei et al. 2023, Peukert et al. forthcoming) by emphasizing the critical role of data composition in evaluating its value and determining collection strategies. Different papers theorize about the shape of the data value function (Jones & Tonetti 2020, Iansiti 2021, Farboodi & Veldkamp 2023) but have not reached a consensus. We posit that the value depends on both volume and bias, and provide a conceptual framework along with empirical evidence to analyze how data collectors' decisions impact these factors within a decision framework.

Lastly, our work is connected to recent studies on how biased input data leads to biased algorithms (Cao et al. 2021), biased estimates of public opinions and behaviors (Bradley et al. 2021), degraded performance in business analytics (Lin 2022, Neumann et al. 2022, Tucker 2023, Lee et al. 2023, 2024) and other market outcomes (Johnson et al. 2020). While the sources of input data bias vary, individual differences in privacy concerns is often one of them. Our work explores this angle further by examining how firms' choice architectures can either exacerbate or mitigate this bias.

The rest of the paper is organized as follows. Section 2 describes existing marketplaces for data and the prevalence of choice architecture in these marketplaces. Section 3 uses a conceptual model to illustrate why accounting for the joint heterogeneity of privacy valuations and frame effects on the consumer side is crucial for understanding how frames affect sample bias. Section 4 illustrates the design of our experiment. Section 5 describes our data and reduced-form evidence, followed by heterogeneity analysis for privacy valuations and choice frame effects. Section 6 introduces our counterfactuals to exemplify and quantify the bias-volume trade-off that firms face using a decision-theoretic framework, and Section 7 concludes.

2 Choice Architecture in Consumer Data Markets

2.1 Markets for Consumer Data

We use "data markets" to refer to the broad settings where firms offer compensation to consumers in exchange for their consent to share and use data. Examples of data shared include consumers' surveys, their history with a company (e.g., via membership), web cookies, device identifiers, access to their data from other providers (e.g., bank accounts), and other personal data. Similarly, compensation for data sharing also takes many forms, including product provision and upgrade, cash, or membership benefits. The practice of offering compensation in exchange for consumer data has existed at least since American Airlines launched the first modern loyalty program in 1981.[4] Compensation can take the form of explicit or implicit prices. Examples of *implicit* compensation include personalized search results, feeds, ads, and recommendations. *Explicit* prices can be loyalty points, discounts, or incentives for responding to surveys or being in a consumer research panel. Our focus excludes B2B data marketplaces, which involve data trades between intermediaries and end user firms, as our goal is to understand how choice architecture affects consumers' data sharing decisions and data quality.

Presenting consumers with compensation versus no-tracking alternatives became popular after the implementation of the General Data Protection Regulation (GDPR), Apple's App Tracking Transparency Framework (ATT), and other similar privacy regulations and changes that require companies to get consumers explicit consent for sharing data with third parties. In its most basic form, companies present customers with two options: being tracked with free or better products, or keeping their data with the expense of forgoing the product or service. More recently, some online publishers and e-commerce websites in Europe started setting up a "pay or tracking" interface, where consumers can choose between using the product for free while sharing data for personalized advertising, paying for using the product without tracking, or forgoing the product altogether. The most high-profile example is Meta, which charges €12.99 for using their platform without sharing data (see Figure A.1). Similar practices have been adopted by top publishers in Germany, France, Italy, and Austria (Mueller-Tribbensee et al. 2024). According to a recent large-scale web crawl, 317 out of 431 top publisher sites in Germany have adopted the pay-or-be-tracked interface as of 2023 (Morel et al. 2023).

[4]Source: https://loyaltyrewardco.com/the-true-history-of-loyalty-programs/

Centralized data marketplaces are also emerging following recent privacy regulations.[5] These "data exchanges" allow consumers to share various forms of data, such as website visits and interactions, social media data, and transaction histories to companies they select, in exchange for compensations that are either a lump sum, a monthly payment, or a fraction of the platform's revenue from selling their data as it occurs.

2.2 The Prevalence of Choice Architecture for Collecting Consumer Data

Choice architecture is the organization of "the context in which people make decisions" (Thaler & Sunstein (2008) page 3). It includes "nudges" when benign, and "sludges" or dark patterns when malicious, and is commonly used by companies to maximize consumers' consent for data collection. For example, Utz et al. (2019) show that after GDPR took effect, 57.4% of EU websites used choice architecture to increase data collection; Nouwens et al. (2020) focus on UK websites and documented the presence of choice architecture in over 80% of the sampled sites. Similarly, mobile apps on iOS platforms often use a customized banner, known as a pre-prompt, to encourage users into sharing their device identifiers and other personal data after Apple's implementation of its ATT Framework.

Although default is the most commonly used choice architecture (Utz et al. 2019, Nouwens et al. 2020), price anchors are present in settings where an explicit price is given. For instance, Datacy shows consumers an estimate of their data's worth along the "sell your data" button (see Figure A.2).

Companies typically use a combination of multiple elements to maximize consent rate. Many consent management platforms, including OneTrust and Usercentrics, offer companies with services to test and optimize the choice architecture used in consent banners. With the ability to experiment, modern companies can deploy various choice architectures "more easily, frequently, and at a much larger scale than traditional brick-and-mortar retailers" (FTC 2022). The ability to identify effective choice architectures via experimentation means that companies can readily substitute across different designs, especially when existing manipulative practices are prohibited by law (Hils et al. 2020). This evolving nature means that empirical exercises aimed at measuring the effects of specific design practices can quickly get out-of-date. Therefore, when interpreting our experiment, we focus less on specific frame effects but on the impact of choice architecture optimization practices in general.

Perhaps surprisingly, discussions and efforts on optimizing choice architecture for consent often focus on maximizing the consent rate–the volume of the data that companies collect, while ignoring the representativeness of the consumers that opt in. As one of the leading consent optimization platforms, OneTrust promotes consent rate as the key objective for evaluating their

[5] Examples include *Datacy*, *Datacoup*, *Rewarded Interest*, *TIKI*, and *Tartle*. See: https://datacy.com, https://datacoup.com, https://rewardedinterest.com, https://mytiki.com, https://tartle.co/sellers

banner testing and optimization services.Similarly, efforts to optimize pre-prompts for iOS devices often feature phrases like "increase your ATT opt-in rates" and "how to maximize opt-ins".[6]

Regulators increasingly focus on companies' choice architecture practices, especially in the domain of online data collection. For example, the UK's Information Commissioner's Office and the Competition and Markets Authority both expressed concern that companies' choice architecture deployment can potentially generate consumer harm.[7] In the US, the FTC also issued a report highlighting the use of "design elements that trick customers into sharing data" as an area that needs increasing regulatory focus,[8] and has brought cases to deter these practices.[9]

We focus on the impact of choice architecture on sample data bias because such bias can have far-reaching negative consequences not only for organizations that seek to use this data but also for consumers and the broader society. For example, Blattner & Nelson (2021) show that insufficient credit history data from under-served consumers leads to higher noises in their credit scores, which in turn leads to lenders less willing to extend loans to these consumers. As a more recent example, Bradley et al. (2021) show that using large but biased panels to gauge the take-up rate of COVID-19 vaccines can lead to persistent estimation biases, which can lead to miscalibration in public health assessments and subsequent disease prevention recommendations.

3 Conceptual Framework

In this section, we lay out the framework that underpins our experimental design and subsequent analysis. The experiment elicits consumers' willingness-to-accept for data sharing, and we explain how we interpret it here as data supply; in Section 6.2 we close the model by incorporating the firm's demand for data. We start by distinguishing two types of data valuation. The first is a hypothetical frame-neutral valuation v_0, which is the would-be valuation if all choice frames were absent. This valuation reflects consumers' best guess about the true value of their privacy. We note that neither the researcher nor the consumer observe this true value. The observed valuation is what we call the expressed valuation $\widetilde{v}$, which is a function of the frame-neutral valuation but is influenced by choice architecture, θ:

$$\widetilde{v} = f(v_0; \theta). \tag{1}$$

In other words, the choice architecture creates a gap between the expressed valuation and the consumers' prior judgment of the value of their data.

[6]See: https://www.appsflyer.com/blog/tips-strategy/apps-boost-att-opt-in/, https://www.inmobi.com/blog/how-to-increase-your-att-opt-in-rates

[7]*Harmful design in digital markets: how online choice architecture practices can undermine consumer choice and control over personal information*. ICO and CMA joint position paper, 2023.

[8]https://www.ftc.gov/business-guidance/blog/2022/09/ftc-issues-illuminating-report-digital-dark-patterns

[9]https://www.ftc.gov/news-events/news/press-releases/2017/02/vizio-pay-22-million-ftc-state-new-jersey-settle-charges-it-collected-viewing-histories-11-million

We situate the two valuations in the context of a data market, where firms directly interact with consumers to get their consent for sharing data. In such a market, consumers form the supply for data based on their privacy valuations; the firm has a demand for data, offers a price for buying data, and can choose the choice frame to influence the supply from consumers. Although $\widetilde{v}$ may be larger or smaller than v_0, the firm often chooses a frame that delivers the lowest $\widetilde{v}$ to maximize its gain per dollar offered. In essence, the firm-chosen frame pushes the consumers' data supply curve toward the right (see Figure 1a). As a result, the equilibrium quantity of data traded is larger and the equilibrium price is lower compared to the frame-neutral equilibrium.

So far, we have shown how choice frames can induce behavioral distortions and change data collection by treating data as a standard commodity. However, consumer data is not a standard commodity. One unique feature of data is that its value to firms depends not only on its volume but also on its representativeness. When evaluating the impact of choice architecture, the firm should care about not just its impact on the volume of data shared, but also which consumers are more likely to share and how that affects the bias in the data sample.

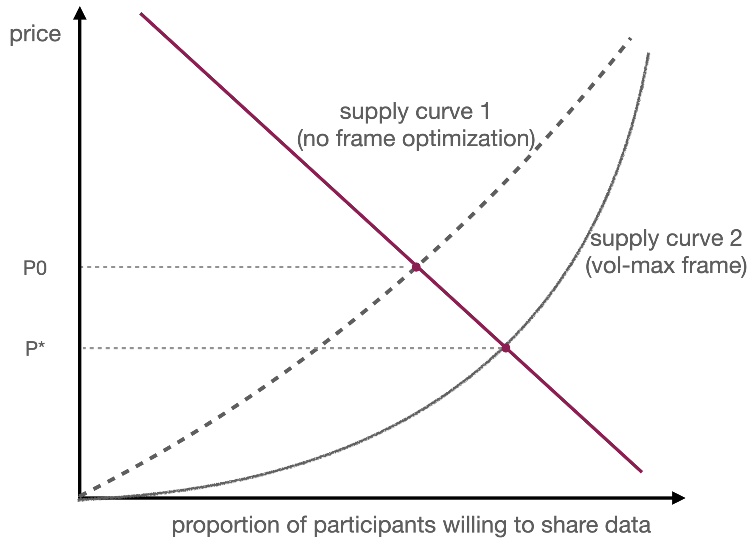
How does the consideration of sample bias affect our evaluation of frame effects? Consider the example in Figure 1b. Absent the choice frame effects, low-income consumers (in light yellow) value their privacy less compared with wealthier ones on average; however, their valuations are also more susceptible to the influence of choice architecture. Without the influence of choice architecture, for any given price the firm would have under-sampled wealthier consumers while over-sampling poorer ones. With a choice frame that pushes the data supply downward, such selection bias becomes more severe because poorer consumers adjust their valuations downward even further compared to the wealthier ones. A trade-off emerges: The volume-maximizing choice frame would have helped the firm in a regular commodity market, but can end up harming the firm by inducing more bias in the collected data.

In our illustrative example above, deploying a volume-maximizing frame ends up exacerbating bias in the sample data due to the negative correlation between consumers' privacy valuations and their frame responses. However, this need not always happen. For instance, suppose all consumers who share their data in the baseline frame belong to the low-income segments. In this case, deploying the volume-maximizing frame will never exacerbate the bias, and may even alleviate the bias by including more consumers in the sample. In other words, the volume-maximizing frame can also have a bias-mitigation effect through supply expansion.

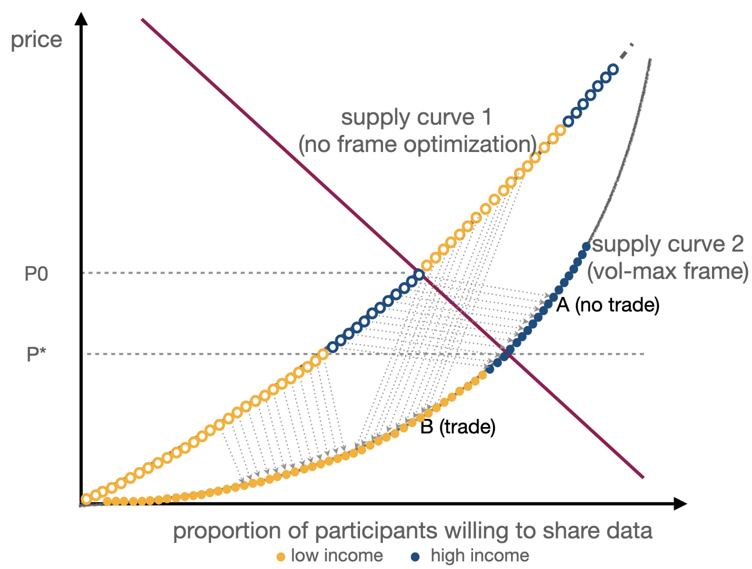
To examine the potential bias-volume trade-off in practice, we use an experiment to randomize choice frames between participants, while collecting consumer characteristics to explore the heterogeneity in their privacy valuations and frame effects. This approach allows us to construct an empirical analog of Figure 1b, and unpack how the trade-off present in frame choices depends on demand and supply conditions in a data market.

Figure 1: Impacts of Choice Frames on Data Quality: An Illustration

(a) Example 1



(b) Example 2

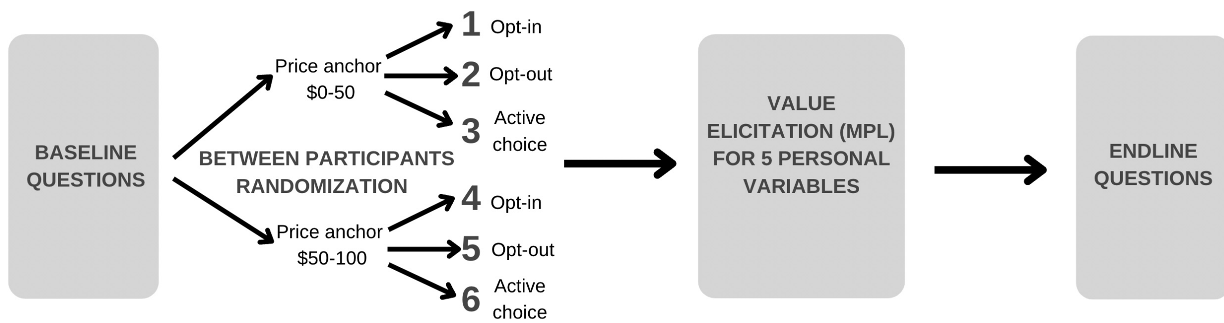


Notes: The top panel shows the differences in volume and price. The bottom panel further illustrates potential effects on composition, with each color representing a different consumer type. The yellow types are more affected by the volume maximizing frame and thus are more over-represented compared to curve 1.

4 Experiment

The main component of our experiment is a multiple price list (MPL) that elicits consumers' valuation of their Facebook data in an incentive-compatible fashion. To test the effects of choice frames on reported valuation, we randomize the default choice and the range of price lists between participants when implementing the MPL. We also use baseline and endline questions to measure consumer characteristics, their internet and social media usage, and their beliefs about which of their data are already available. Figure 2 summarizes the flow of our experiment. Below, we start with the participant recruiting procedure, then introduce the value elicitation components, the choice frame treatments, and end with the design of survey questions that measure consumer characteristics.

Figure 2: Experiment Overview



4.1 Participant Sources

We recruit our participants from two sources: Facebook Ads and Prolific. Using these sources confers two advantages. First, both sources allow us to screen participants based on the availability of their Facebook accounts.[10] Including only participants who have an active Facebook account ensures that their data sharing decision reflects only their privacy preferences and not the availability of their data. Second, including participants from both sources facilitates external validity analysis. Existing studies that measure privacy preference often use survey panels or student populations, who may have lower valuations for privacy or respond to choice frames differently compared with the population. Having both the Facebook and survey panel (Prolific) participants allows us to examine this possibility. We restrict our participants to English speakers living in the US, with an age between 18 and 64.

To minimize selection into the experiment, we do not disclose the specific research topic in the Facebook recruiting ad (see Figure B.1) or the study description on the Prolific recruitment site. A user who clicks on the survey invitation link from the recruiting ad or Prolific task is directed to our study introduction page. The introduction explains that we are university researchers who want to understand the public's social media usage and perceptions,[11] then ask for their consent to enter the study.

[10] When recruiting the Facebook participants, we restrict our ad placement to positions only viewable by logged-in Facebook users, which excludes Facebook Audience Network and Instagram. For the Prolific participants, we use the internal selection tool provided by the platform to enlist only Facebook users.

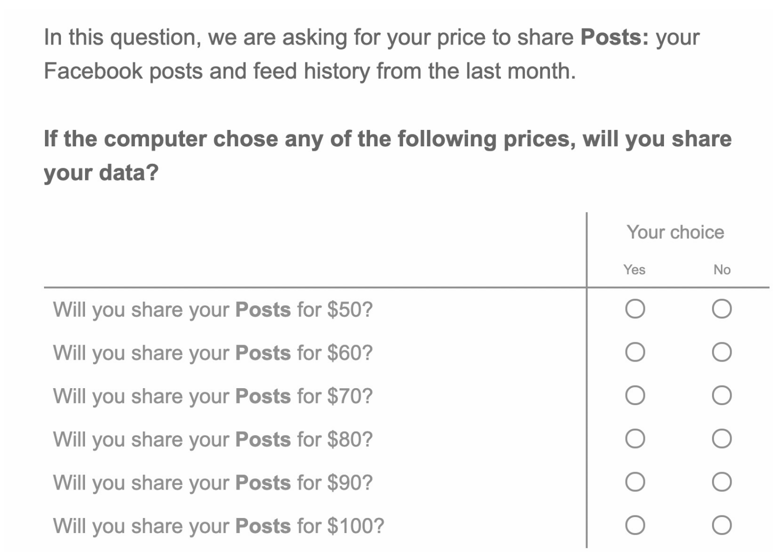
[11] This information includes their data sharing behavior and privacy attitudes on social media, but to minimize selection bias, we did not explicitly mention that on the introduction page.

4.2 The Multiple Price List and Choice Frame Treatments

We measure participants' valuation of their Facebook data using a multiple price list (MPL). MPL resembles Becker-DeGroot-Marschack (Becker et al. 1964) in its use of a lottery to ensure incentive compatibility (Andersen et al. 2006). Its advantage is simplicity: Instead of a second-price auction as in BDM, MPL uses simple take-it-or-leave-it offers repeated at different price points. With a simpler value elicitation format, it is easier for participants to understand the procedure and why telling the truth is optimal.

Standard MPLs only give us an interval that includes a consumer's valuation of their data (see Figure 3). To get more granular numbers, we follow each MPL with an open-text question asking participants their exact valuation of the data. If a participant chooses not to share data in all the MPL questions displayed, we also give them the option to indicate "I do not want to share my data at any price" in the open-text prompt (see Figure 4). Suppose a consumer's response is chosen by the lottery. In this case, the data-money exchange occurs if the number in their free-text entry is lower than or equal to the offer price randomly generated by the computer. This mechanism guarantees that reporting the true value is a dominant strategy.

Figure 3: The Multiple Choice List Procedure: Step 1

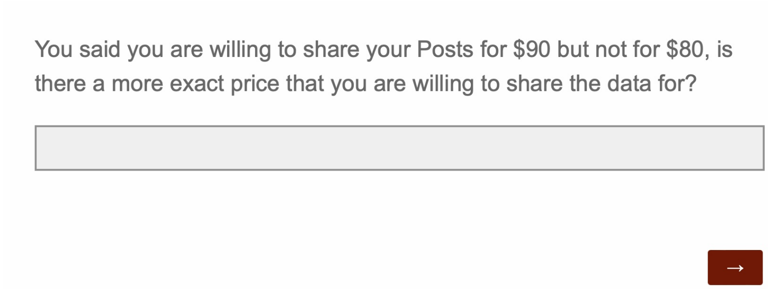


Notes: This screenshot shows what a participant sees in the first step of our MPL procedure in the *high price range + active choice* condition. A similar screen is shown in other treatment conditions with the following changes. In the *opt-in* conditions, all the "yes" options are pre-selected; in the *opt-out* conditions, the "no" options are pre-selected; in the *low price range* conditions, the prices listed range from $0 to $50, instead of $50 to $100.

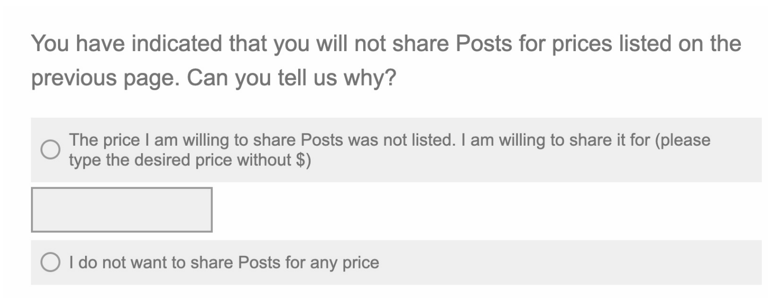
By using the above MPL procedure, we use an explicit price instead of products (with an implicit subjective value) for WTA elicitation. This approach allows us to trace out the data supply curves across choice architecture conditions. Estimating the supply curves for data is crucial: It allows us to characterize the effect of choice architecture in dollar terms, and to examine how the

Figure 4: The Multiple Choice List Procedure: Step 2

(a) Scenario 1: participant says "yes" to some prices but not all



(b) Scenario 2: participant says "no" to all prices listed



Notes: The screenshots above show two examples of what a participant sees in the second step of our MPL procedure. In the first scenario, a participant agrees to share their posts data for $90 or above but not for $80 or below. In the second scenario, a participant chooses not to share their *posts* data for all the prices listed in the first step. In this case, we give them the option to indicate that their valuation for the *posts* data is infinity.

volume-bias trade-off depends on the price for data. Backing out the supply curve would be more difficult if we had used implicit prices (such as personalized services) as we do not observe how much value the consumer places on a particular product.

After answering the baseline questions, participants see a message asking for their valuation to share data with advertisers. An example of the MPL procedure follows, showing them how their choices and the random price generated by the computer co-determine whether the data exchange will occur. The exact explanation is in Appendix B.1.

Participants then go through a practice question, where we ask them to imagine selling a gift card worth $14.5 by responding to a multiple price list. If a participant gives a value different from $14.5, we display an error prompt, asking them to think again and showing how a truthful response is optimal (see Figure B.2). Since the price anchor treatments do not span the entire range, we also train participants to report values outside their allotted treatment and to get familiar with the exact elicitation interface. For example, a participant in the low price range treatment (seeing an MPL with a price range of $0-$50) will be asked to report the value of selling a hypothetical $60 gift card. To do so, they need to choose "No" for all prices in the MPL; they should then type in $60 in the free-text question. Similarly, the high price range participants need to agree to share with all prices greater or equal to $50 and then report $40 as the free-text response.

After the practice round, we show participants the actual MPL questions to get their valuation for different types of data. We inform participants that our offer price is randomly drawn between \$0-\$95 and does not depend on their response. Each participant receives five rounds of MPLs in random order, one for each personal variable. The following list shows the five variables and the definition we show to participants:

- *"About me" page: your Facebook information page;*
- *Posts: your Facebook posts and feed history;*
- *Likes: the posts and pages you liked on Facebook;*
- *Friends and followers: the people you befriended and followed, and people who followed you;*
- *Survey answers: the answers you gave earlier on your browsing behavior and demographics.*

We independently vary the following choice frames across participants (see Figure B.3 for different treatment interface variations). The first frame is the **default** choice in the multiple price list. A pre-selected "yes" for questions "Will you share your posts for \$X" is an *opt-out* default, while a pre-selected "no" is *opt-in*. For *active choice*, neither option is pre-selected; participants will have to click on one of them to answer the questions and proceed to the next screen. The second frame is the **price range** offered in MPL: the *low price* condition has prices between \$0 and \$50, while the *high price* condition ranges from \$50 to \$100. Importantly, our price anchor treatment only moves the range of prices that participants can select during the MPL step, but not the actual price draw distribution, nor does it restrict the value participants can report in the open-text response. We communicated this information to the participants before showing them any multiple price lists. Participants receive the same combination of choice frames throughout all 5 data sharing questions.

Apart from the choice frames, we also randomize the time range of behavioral data (*posts* and *likes*) between participants, which varies between *one month*, *one year*, and *since joining Facebook*. In doing so, we vary the value of data in a direction known to us. The goal is to see if consumers respond to the scope of data requested when valuing their personal data.

We inform participants that one of their responses will be chosen to implement the data-price exchange with a 2% probability. All participants receive a flat participation fee immediately after finishing the survey: We ensure they receive the participation payment in time to build trust that we will also pay them if they are selected by the lottery and send us their data. Participants who were chosen by the lottery and had valuations lower than our offer price, received a step-by-step guide to download the Facebook variable. They receive our offer price within 24 hours after sending their data to us.

4.3 Baseline and Endline Questions

The survey includes questions that capture consumer characteristics, their internet and social media usage, as well as their beliefs about personal data availability. In general, we put questions related to data sharing in the endline survey, so that participants are not primed to consider privacy before the MPL questions; otherwise, we include the questions in the baseline survey.

The baseline questions measure participants' social media consumption behavior and their demographics. We ask participants about their time spent on Facebook and online, when they started using Facebook, and their engagements with merchants and ads on Facebook. For demographics, we record their age, gender, ethnicity, income, and education.

The endline questions include measures of information-seeking behavior and participants' beliefs about data availability. To measure information seeking, we include the following question: "*Did you look up additional information when answering the questions that ask how much you value your Facebook data?*" If they answer "yes", we ask what kind of information they looked up. To measure their belief about data already available to various parties in the market, we ask the following questions: (a) *Which of your personal information on Facebook do you think is available to the public?* (b) *What information do you think advertisers on Facebook already know about you?*

4.4 Discussion

Using explicit prices for data. Our use of explicit prices can result in different levels of consumer privacy valuation compared to those expressed when implicit prices are used. Tomaino et al. (2022) show that companies may be able to collect more data by using services with equal values rather than using posted prices, all else equal. However, the main takeaway from our experiment should remain valid despite this level change. What we show is that the trade-off companies face when designing frames depends on the joint distribution of privacy valuations and frame effects. We do not claim that the precise numbers, whether the privacy valuations or the frame-effect distributions, will remain the same in other field settings. In fact, privacy valuations in one (field) setting do not map well to another setting, known as context dependence (Martin & Nissenbaum 2016). Given this fact, we view the data from our experiment as a case study to demonstrate the mechanism rather than a quantification exercise.

Generalizability of choice architecture effects. Our experiment examines the effects of default and price anchor, two implementations among all possible frames that firms can use to nudge consumers into sharing data. Our goal here is not to design a "policy evaluation" experiment that mimics the current data-sharing environment: Companies often substitute across frames, especially when their tried-and-true tactics are banned (Willis 2020), making it hard to pick the frames guaranteed to be common in all current and future regulatory regimes. An attempt to include all possible frames companies use will also be prohibitively costly. Instead, by providing an example that shows the importance of the joint distribution of consumers' privacy preferences and their frame responses, we argue that the findings are generalizable to other settings. While the effect sizes themselves might be different, these overlooked interactions should be taken into account. In our setting, the key supply-side factors that affect the trade-off are the joint distribution between consumers' privacy preferences and their frame responses. Thus, any combination of frames that shifts privacy valuations in a heterogeneous manner will fulfill our purpose.

Mappings to intermediary vs. end-user firm settings. Our experiment design aligns closely with a data intermediary setting, as we ask consumers if they are willing to share data with advertisers in addition to us. That said, the effect of choice architecture on privacy decisions and shared data quality is independent of whether it is the intermediary or an end-user firm that asks for data. Both parties can set up design interfaces that nudge consumers one way or another, and both can use explicit or implicit prices as a means of compensation. We also believe that both the intermediary and the end-user should care about getting data representative of the consumer population that they are interested in learning about. In short, although our experiment is closer to an intermediary setting, our framework and key takeaways do not hinge on who is buying data.

Incentive compatibility of valuations above price range. In the experiment, we informed participants that the randomized price drawn from the computer was between \$0 and \$95. If a consumer's actual valuation is within this range, she has the incentive to report truthfully. If her valuation is above \$95, reporting any value above \$95 is optimal for her. In essence, it means that the reported values above \$95 should be considered as "partially truthful": they truthfully reveal the fact that the underlying valuation is greater than \$95, but are otherwise a stated preference.

Debates abound over whether stated preferences for privacy are truthful (Spiekermann et al. 2001, Singleton & Harper 2002); Prince & Wallsten (2020) argue that the gap between stated and revealed preferences diminishes once the context is controlled for. Nevertheless, we adopt various strategies in the data analysis to emphasize the valuations within the incentive-compatible range. Our reduced-form analysis focus on log valuation as our preferred specification. In addition, we allow truncation at different finite points above \$95 in Table D.3 and the results are robust to different truncation point specifications.

5 Data and Model Evidence

We recruited a total of 5,028 participants: 2,010 from Facebook during February 11-27, and 3,018 from Prolific during March 7-10, both in 2022. Table 1 provides summary statistics of our participants. Compared with the US representative demographics,[12] ours includes more females and are overall better educated; the distributions of age, income, and ethnic majority are similar

[12]https://www.census.gov/quickfacts/fact/table/US/PST045221; https://www.census.gov/library/visualizations/2022/comm/aging-nation-median-age.html.

to the national average. Compared to those recruited from Prolific, the Facebook ad participants include more females, are older, wealthier, better educated, more likely to be minorities, and spend more time on Facebook; they also click on ads and shop on Facebook more often. Thus, having participants from both sources allows us to cover a wider demographic range, which allows us to demonstrate how privacy valuations and frame effects differ across the demographic spectrum. Tables C.2 and C.3 show that our participants are balanced across the six treatment conditions.[13]

Table 1: Summary Statistics of Participant Characteristics

	Overall		**Facebook**		**Prolific**		**U.S. Census**
	Mean	SD	Mean	SD	Mean	SD	Mean
Number of participants							
N	5028		2010		3018		
Race (percentage)							
White	0.8	0.4	0.73	0.44	0.85	0.35	0.76
Black	0.07	0.25	0.08	0.27	0.06	0.23	0.14
Asian	0.12	0.32	0.16	0.37	0.08	0.28	0.06
Other	0.05	0.05	0.06	0.06	0.04	0.04	0.05
Gender (percentage)							
Female	0.59	0.49	0.76	0.43	0.48	0.5	0.51
Median age							
Median age	39.5	12.77	39.5	13.23	39.5	12.22	38.8
Median household income ($)							
Median household income	62500	48675	62500	51629	62500	45224	64994
Education (percentage)							
High school graduate or higher	0.99	0.1	0.99	0.09	0.99	0.11	0.89
Bachelor's degree or higher	0.64	0.48	0.73	0.44	0.58	0.49	0.33
Facebook Questions							
Average time spent on FB (h)	1.41	1.38	1.97	1.37	1.03	1.25	
FB membership duration (y)	5.76	0.88	5.75	0.87	5.76	0.89	
Average time spent on internet (h)	3.52	1.84	3.5	1.82	3.54	1.86	
Active user (percentage)	0.31	0.46	0.47	0.5	0.21	0.41	
Purchase from FB or Instagram (times/mo)	0.41	0.78	0.63	0.95	0.27	0.6	
FB or Instagram ad click (times/mo)	1.66	1.81	2.47	1.92	1.12	1.51	

Notes: "Overall" means the entire sample in our study, and "Facebook" and "Prolific" break down the sample by the recruitment source. The last column provides statistics from the 2020 US Census for comparison.

The second step in the WTA elicitation is unrestricted; thus participants can be inconsistent. For example, a participant may say they are willing to sell their posts for $40 but not for $30 in the MPL, then ask for $56 on the next screen. Appendix Figure C.1 shows responses from all participants who completed our study, with their free-text valuations on the Y-axis and the implied

[13] Here, we separate the covariate balance tests for the attributes measured in the baseline and endline surveys. One concern about the endline survey responses is that they may be influenced by the treatments, especially when it comes to beliefs about data usage. Table C.3 shows that the endline responses do not differ significantly across treatments and thus can be included in our heterogeneity analysis.

WTA from their MPL responses on the X-axis. We find that 93% of participants give consistent valuations throughout. Among participants who have given inconsistent answers, many deviate to a range close to where they were. More importantly, 83% of consumers who give inconsistent valuations only do so occasionally, suggesting that their inconsistency is more a sign of regret than a byproduct of inattentiveness.

Consumers' valuations for their data are heavily skewed to the right: In fact, they report a value of infinity (i.e., "I do not want to share for any price") for 18.3% of the time. We allowed for infinity reporting because earlier work posits that some consumers can be "privacy fundamentalists" who would reject any benefits from data usage in exchange for privacy (Westin 2003, Woodruff et al. 2014), and we wanted the valuation measurement to allow for this possibility. Nevertheless, most consumers in our experiment do not seem to adopt a fundamentalist attitude in that they are selective when reporting infinite values. For example, 7% of participants report infinity valuation for their survey answers compared to 31.7% for friends and followers. Only 3.7% of participants report infinity values for all items.

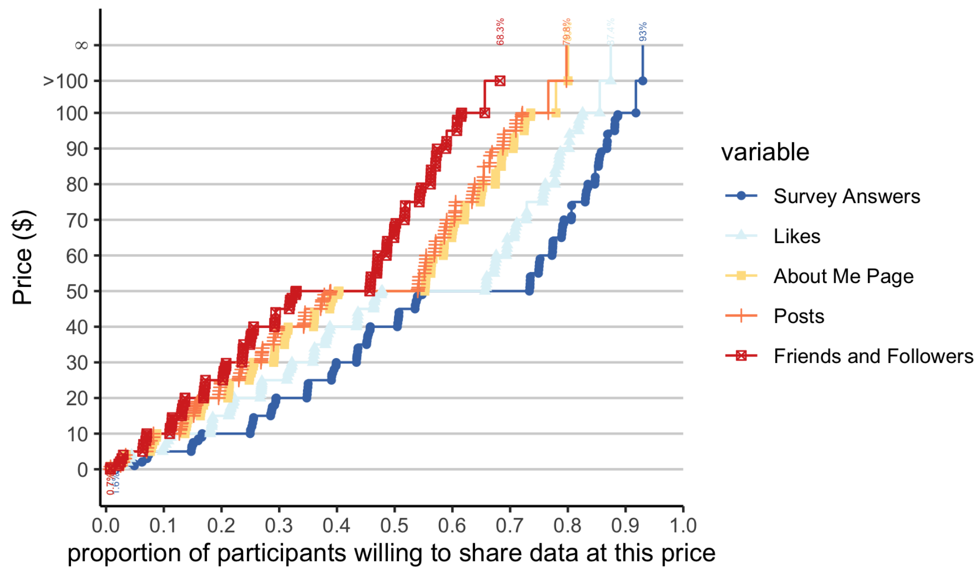
Genuine or not, the infinity values create challenges in reporting summary statistics and reduced-form analysis. We adopt several strategies to account for this challenge. In the model-free evidence, we report most results in the form of data supply curves, with the percentage of consumers with infinite valuations at the top of each curve. Since the data supply curves are essentially cumulative distribution functions of the valuation, they transparently visualize the distribution of consumer valuations and the impact of choice frames on different parts of the curve. In the reduced-form models, we use log valuation as the outcome variable in our main specification and indicate where we top-code the data. Focusing on log valuations makes the results robust to the impacts of extremely high values and our top-code choices. We note that the long-tail pattern in privacy valuations is common (e.g., see Collis et al. 2020 and Lin 2022) and is not unique to our setting.

5.1 Valuations across Consumer Data

Our first result shows that consumers have systematically different valuations for different types of personal data (Figure 5). They value *friends and followers* data the most; next comes *posts* and *about me* information; *survey answers* are valued the least. When top-coded at \$100, the mean valuation for *friends* data is \$64.6; the average valuation across all Facebook data is \$57.2, compared to an average of \$42.1 for sharing the survey responses. Given that 20% of consumers report their data valuation at above \$100, these valuations serve as the lower bounds for the actual valuations.

Are expressed valuations arbitrary, or do participants have a consistent valuation of their data? The "coherent arbitrariness" theory (Ariely et al. 2003) argues that people can have coherent differences in valuation after choosing an arbitrary starting point. One may expect such a phenomenon to be likely for private data valuations, as the pros and cons of sharing data can

Figure 5: Data Supply Curves by Data Type



Notes: The figure shows supply curves for different data types, pooling valuations across treatments. Each participant contributes a data point for each data variable. The top vertical line represents the share of participants who reported a finite WTA.

be uncertain. To see if this hypothesis holds in our setting, we leverage the fact that the order of personal variables is also randomized across consumers and examine the valuations using only the first question each consumer encounters. The differences in valuation across data persist with a similar magnitude, even when we focus on only the first valuation (see Appendix D.1). As another test of choice coherence, we compare valuation for Posts and Likes data between participants who are randomized to different data duration conditions, and find that they assign a higher value to their online history data when it has a longer duration (see Appendix Figure D.2). These patterns show that consumers' valuation for privacy is not arbitrary, but coherent both within and across people.

5.2 The Effects of Choice Architecture on Data Supply Curves

Despite the coherence, consumers' privacy valuations are prone to the influence of choice architecture. Table 2 summarizes the magnitude of average treatment effects using Tobit regressions. We use Tobit regressions to address the challenges from the non-incentivized high or infinite reported valuations. Throughout the paper, we top-code the data at $100 when estimating the effects on the level of valuations, and at $1000 when taking logs for estimating relative effects and further reducing the sensitivity to extreme values.[14] Our preferred specification is Model 4, which includes the types of personal variables as additional controls. Compared to active choice, an opt-out frame decreases the average valuation by 11.8%, while opt-in increases the valuation by

[14] In Appendix D.2, we repeat our analyses for different top-coding and specifications.

10.3%. The influence of a price anchor is more substantial. Consumer valuations for data decrease by 37.4% on average when elicited in the low-price as opposed to the high-price condition.[15]

Table 2: Average Treatment Effects and Valuation across Data: Tobit Regressions

	WTA	WTA	log(WTA)	log(WTA)
Intercept	63.884 ***	48.772 ***	4.371 ***	3.633 ***
	(0.876)	(0.901)	(0.041)	(0.042)
Price Anchor = Low	-16.112 ***	-16.165 ***	-0.372 ***	-0.374 ***
	(0.947)	(0.947)	(0.045)	(0.045)
Default = Active	2.377 *	2.376 *	0.118 *	0.118 *
	(1.139)	(1.139)	(0.054)	(0.054)
Default = Opt-in	5.178 ***	5.170 ***	0.222 ***	0.221 ***
	(1.147)	(1.147)	(0.055)	(0.055)
Likes		8.474 ***		0.433 ***
		(0.467)		(0.023)
About Me Page		18.225 ***		0.869 ***
		(0.538)		(0.027)
Posts		20.056 ***		0.944 ***
		(0.556)		(0.028)
Friends and Followers		29.070 ***		1.453 ***
		(0.648)		(0.034)
Num.Obs.	25140	25140	25140	25140

Notes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Outcomes variables are top-coded at \$100 in models using level WTA as the outcome and at \$1000 in models with log WTA as outcomes; standard errors are clustered at the participant level. The omitted category is a high price anchor with an opt-out default, and the survey answers valuations in columns 2 and 4. For log outcome models, we transform the valuation using $\log(Y + 1)$ to account for the presence of zero valuations.

Figure 6 compares the data supply curve across treatments. The default conditions shift the supply curve uniformly, with the supply curve corresponding to active choice sitting squarely between opt-in and opt-out. In contrast, the anchor price distorts different regions of the supply curve in different ways. In particular, the gap between supply curves is the largest in the middle region, due to valuations bunching towards the endpoints of the price range.[16] Paradoxically, a low-price anchor also triggers more consumers to report extremely high (greater than \$100) and infinite values. In the low-price condition, the average percentage of infinity values across variables is 21.2%, compared to 15.5% in the high-price condition. As a result, the supply curves from the two treatments intersect around the \$95 price point, demonstrating the non-monotonicity of the price anchor effect.[17]

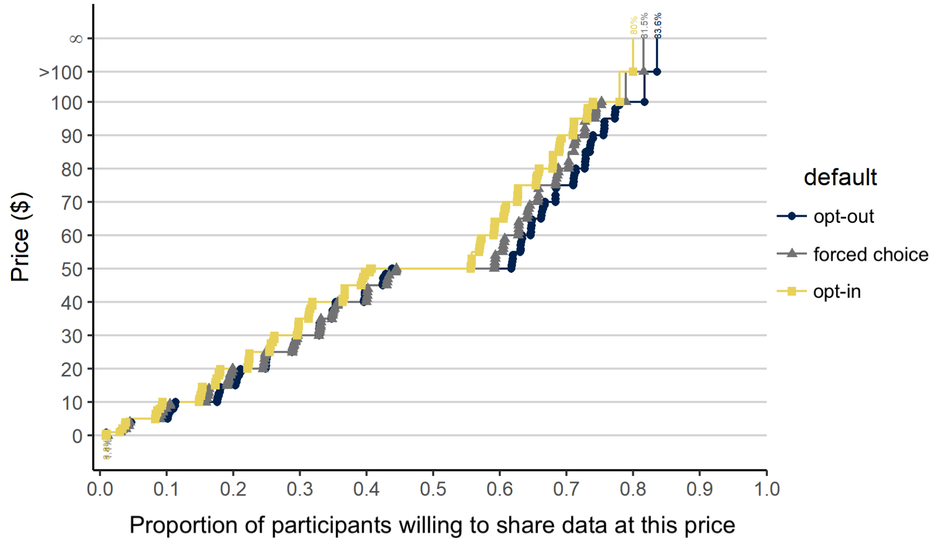
[15] For log outcome models, we transform the outcome using $\log(Y+1)$ instead of $\log(Y)$ to account for the presence of zero valuations. In our setting, Y is often reasonably large, thus we directly read off the coefficients as percentage changes in the outcome induced by the treatments.

[16] Although bunching is prevalent, it is not persistent within individual. In fact, only 3% of our participants choose their valuations at one of the price range endpoints persistently across five personal variables.

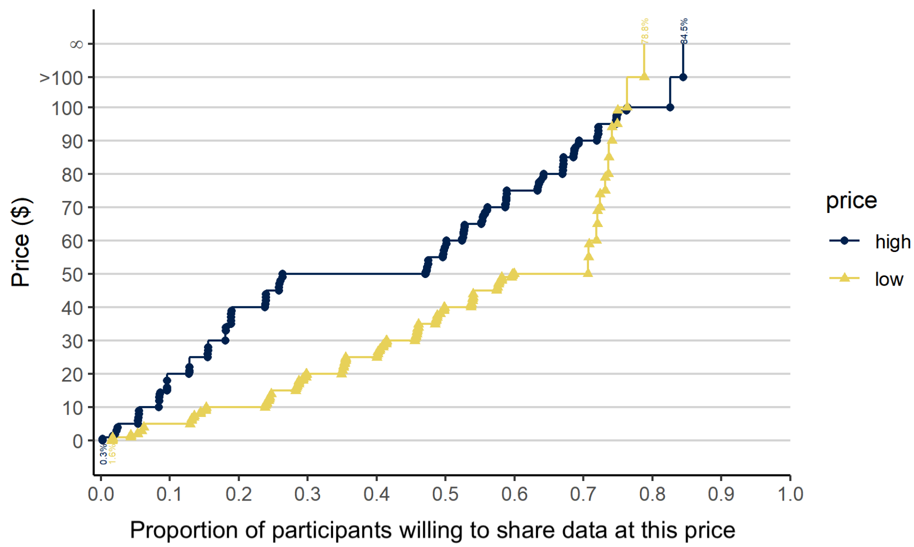
[17] Appendix Table D.2 includes further controls and variables of interest. Consumers who ask for more than the token's value in the practice round also report higher WTA for their data. In addition, consumers who believe specific data are already available to Facebook and to the public value their data less when it comes to sharing data with advertisers. Interacting these beliefs with the treatments leads to mixed and non-significant results: price range effects are stronger for those who believe data are available, while the default effects are weaker.

Figure 6: Data Supply Curves by Treatment Condition

(a) Default option



(b) Price range



Notes: The figure shows supply curves for the treatments in the experiment — participants' WTA on the vertical axis against their share on the horizontal axis. Panel (a) shows the supply curves for default treatments, and panel (b) shows the supply curves for the price anchor treatments. Each participant contributes 5 data points, one for each data variable. The top vertical line represents the share of answers of an infinite WTA.

Table D.3 shows versions of the log outcome model with different top codes for the data valuation. As the top code becomes less stringent, the coefficient representing the price anchor effect decreases in magnitude while all other coefficients increase, driven by the fact that a high anchor decreases the percentage of "infinite" values. One way to explain this finding is that price anchor creates "backlash" in a subset of consumers. Another potential explanation is that in the low price anchor treatment, consumers bother less with reporting a precise value because their data valuation is further away from the highest MPL price, and opt for infinity as a mental shortcut.

5.3 Heterogeneous Effects of Choice Architecture

In Section 3, we show that the correlation between consumers' baseline privacy valuations in the neutral benchmark condition and their responses to choice frames is the key driver that creates the tension between volume-maximizing and bias-minimizing effects across frames. Simply comparing data supply curves under different frames is insufficient for capturing such correlation, as the supply curves do not show how much the same consumer (segment) would value their data under different conditions. Instead, we use heterogeneous effect models to characterize the joint distribution of consumers' privacy valuations and the frame effects. For robustness, later we also subset the raw data to capture simpler forms of heterogeneity.

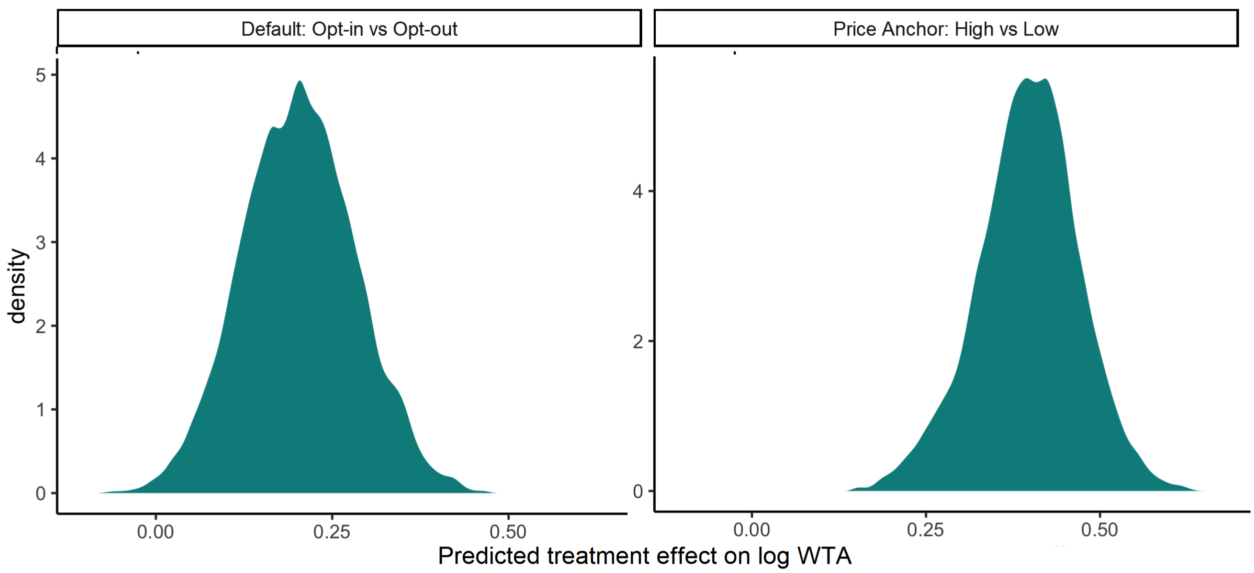
To achieve this goal, we estimate causal forest models proposed by Athey et al. (2019). These models allow us to efficiently and flexibly capture the treatment effects across consumer subgroups.[18] We adopt two specifications. Our preferred specification is a multi-arm forest, which captures potential interaction effects between default and price range treatments. The caveat of the multi-arm forest model is that it does not account for censored values (i.e., our top codes for infinite values). Our second specification is a survival forest, which takes care of censoring in a Tobit-model fashion, but only compares two treatments at a time and thus does not capture interaction effects. The heterogeneity patterns in the two models are qualitatively similar, though the results from the multi-arm forests are smaller in magnitude due to censoring. In what follows, we use results from the multi-arm forest for our analysis. We include details of model tuning and a comparison of estimation results across the two models in Appendix D.

We are interested in a consumer's baseline valuation absent frame effects, but this is unobservable. In view of this fact, we construct the benchmark privacy valuation as the average of a participant's would-be valuations across the six treatments, predicted by our model as $E[Y|X_i]$ (hereafter the *average* privacy valuation). This average valuation underpins a consumer's data-sharing choices when the firm randomly chooses among the six possible frames. It is possible that a true frame-neutral valuation lies closer to one of the frames or even outside this range, but this approach allows us to get a comparable valuation across participants without imposing strong parametric assumptions. In our main specification, we use log valuation as the outcome variable, where the valuations are top-coded at $1,000 before taking the log. The covariates in the model include demographics, Facebook and general internet usage, and consumer beliefs about what data are already available to Facebook as well as the public.

[18] In comparison, simpler models such as linear regression do not have regularization built in and can perform poorly with a large number of consumer feature covariates.

Figure 7 shows the distribution of the heterogeneous treatment effects estimates for the two frames. The average treatment effects for both frames are significantly above zero. Consistent with the raw data, the price range has a larger overall effect compared to the default. Together, the effects of opt-in and high price anchors increase WTA for almost all participants and vary in magnitude. Appendix D.3 shows the treatment effect estimates separately for each individual, along with standard errors.

Figure 7: Heterogeneous Treatment Effect Estimates: Multi-Arm Causal Forests

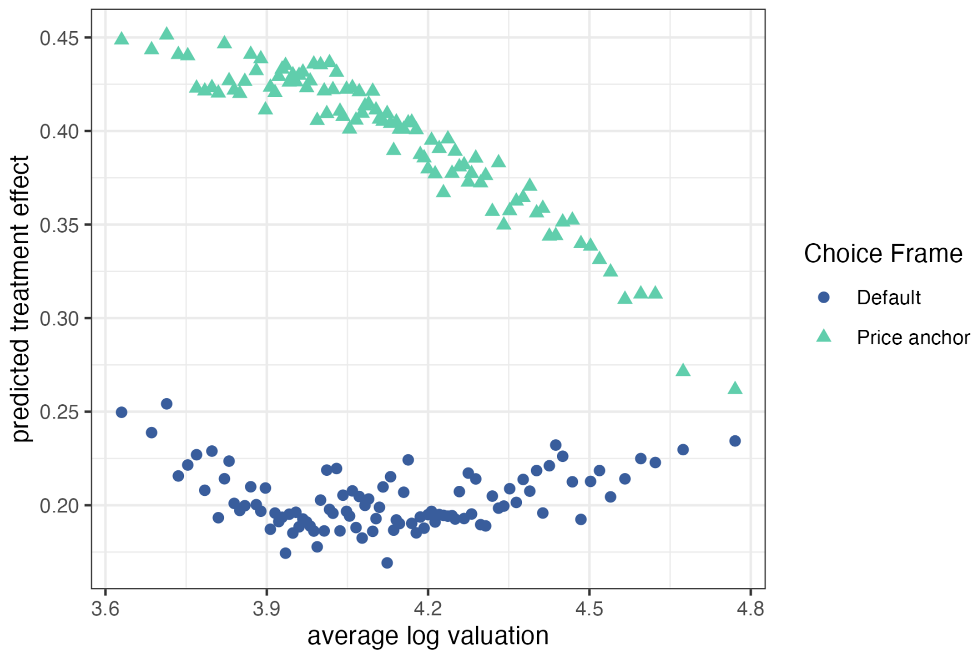


Notes: This figure shows the densities of estimated treatment effects on WTA at the individual level; a larger spread indicates more heterogeneity. Outcomes are log valuations truncated at $1000; standard errors are clustered at the subject ID level.

Figure 8 shows the correlation between frame effects and the average log valuation across participants. The effect of default does not systematically vary with valuations (correlation is -0.04). On the other hand, the effect of price anchor is negatively correlated with the average valuation: Consumers with lower privacy valuations are also more likely to be swayed by a price anchor (correlation is -0.46). This pattern is consistent with Collis et al. (2020), who find that consumers who receive an informative signal on the data value often revise their valuation upwards when their prior is lower than the signal, but rarely update their belief downwards when their prior is high. We note that there is no strong theoretical reason to expect a correlation would always exist. However, when a negative correlation does exist, it can lead to a volume-bias trade-off, as we discuss in the next section.

We want to ask if the negative correlation between average valuation and frame effects reflects systematic differences across consumer subgroups: As is shown in Section 3, it is the differences across the subgroups that give rise to the tension between volume maximization and bias mitigation

Figure 8: Correlation between Choice Frame Effects and Average Log Valuation

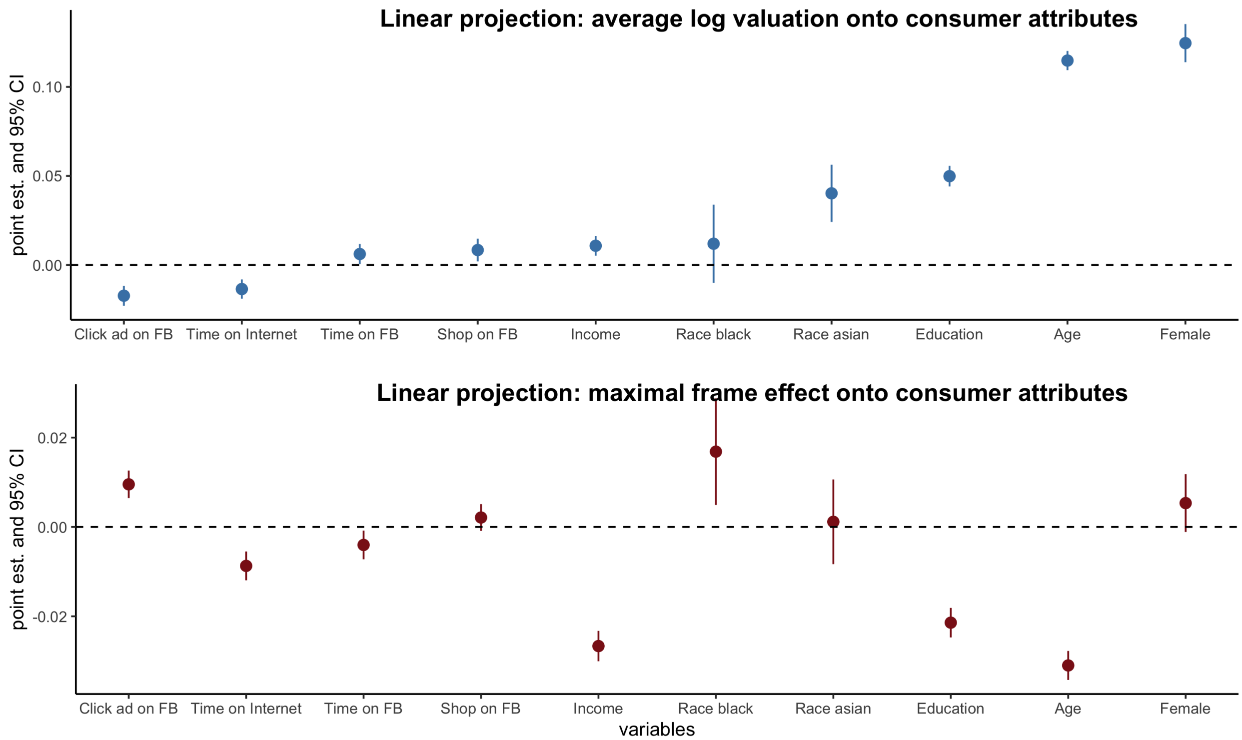


Notes: The figure shows treatment effects against average value across participants. Treatment effects are estimated using multi-arm forests, with the outcome as log valuation truncated at $1000. The average valuations are generated alongside the model as the predicted outcome marginalized across the choice frames. Each point represents 1% of participants. The circles are the estimated effects of opt-out versus opt-in, and triangles are the effects of high versus low price range.

when it comes to frame choice. To further explore this question, we project the average valuation and the heterogeneous "maximal frame effects" to the consumer attributes using linear projection, an approach advocated by the recent literature (Semenova & Chernozhukov 2021) as an efficient way to summarize heterogeneity in a conditional expectation estimate. We define the maximal frame effect as the difference between the two frames that give the highest and lowest average treatment effects. In other words, we calculate for each consumer the difference in predicted valuation between the frame that maximizes the average reported valuation (opt-in, high price range) and the frame that minimizes it (opt-out, low price range).

Figure 9 shows the linear projection estimates and 95% confidence intervals. The top panel shows the predictors of a high average valuation. Consumers who value their personal data more are older, richer, better educated, more likely to be female and Asian, spend less time on the Internet, and are less likely to click ads on Facebook. In comparison, the bottom panel shows the predictors of susceptibility to frames. Those more easily influenced by our choice frames are overall younger, poorer, less educated, and more likely to click on ads while using Facebook. Overall, these attributes predict stronger choice frame effects while also predicting lower valuations for personal data.

Figure 9: Heterogeneous Privacy Valuations and Treatment Effects by Consumer Subgroups



Notes: This figure visualizes estimates from our linear projection models, with dots indicating point estimates and vertical lines representing 95% confidence intervals. In the top panel, the outcome is a consumer's predicted log valuation averaged across the choice frame treatments. In the bottom panel, the outcome is the predicted *maximal treatment effect*, represented as the difference between the choice frames with the highest and lowest ATE. Both outcome metrics come from the multi-arm causal forest estimates, with log valuation censored at $1000 as the outcome. All covariates are standardized before entering the model.

These heterogeneity results are broadly in line with findings in the existing literature. Regarding privacy valuations, Goldfarb & Tucker (2012) find that older people and women value their privacy more, and Collis et al. (2020) show that high-income consumers and Asian communities value their Facebook data more. As to frame effects, Mrkva et al. (2021) similarly find that lower socioeconomic status predicts stronger choice architecture effects. The consistency between our findings and prior work strengthens our belief that our finding is robust.

Meanwhile, we want to emphasize that our results here, whether it is the privacy value distribution or the correlational structure between privacy values and frame effects, may not generalize to all settings. One reason is that consumers can have different privacy preferences in different economic contexts when their data are used for different purposes (Lin 2022), which means the heterogeneity pattern of the preference is also likely to change.

6 Exploring the Volume-Bias Trade-off

Our analysis has shown that choice frames can change the composition of consumers who share data since different consumers respond to frames differently—the question is how. With different frames, consumers have different willingness to share data. Does the frame chosen by the firm exacerbate or alleviate the existing selection bias? A company is often inclined to choose an architecture that maximizes the volume of data collected. Is this the optimal frame for data collection, or does volume come at the cost of representativeness?

To answer these questions, we conduct counterfactual simulations based on the raw data, leveraging the unconditional random assignment; and based on our forest-model estimates, leveraging the counterfactual predictions in all treatments for each participant. We also provide a framework to interpret the findings. Our goal is to examine the volume and bias in data collected under different choice frames. In particular, we compare the performance of two frames: the *volume-maximizing* (hereafter *vol-max*) frame, which maximizes the supply of data at each price point; and the *bias-minimizing* (hereafter *bias-min*) frame, which minimizes the level of average bias at each price point. Our results show that a volume-maximizing choice frame can have opposite impacts on the bias in sample data through two distinct mechanisms. We also quantify the trade-off between the volume and representativeness objectives at the time of frame optimization, and show that the trade-off can depend on whether the firm is able to personalize its frame assignment.

6.1 When Does Bias in Data Degrade Its Value?

We start by clarifying what kind of data bias the firm cares about and when such bias decreases the value of consumer data. The value of data to the firm comes from its information value in learning and prediction. Put formally, there is a series of outcomes Y_i with distribution $f(Y|X)$ that the firm wants to learn about using consumer data. Examples of Y include consumers' willingness to pay for certain products, interest in certain political topics or product categories, and price sensitivity. The firm observes X (e.g., demographics) from all consumers at the point of data collection, but can only see Y for consumers who agree to share their data at price P. A biased sample means the sample distribution $f(X|\tilde{v}(\theta, X) < P)$ does not equal the target population distribution $f(X)$.

Note that the target population can be the firm's desired customer database and need not represent the general population. In other words, all the analysis of bias in the rest of the paper can be equally applied to any known ideal distribution $f(X)$, even if that distribution is not representative.

A biased sample can compromise the statistical accuracy of data-driven insights and degrade the value of shared consumer data in the following scenarios:

1. The firm values learning the **average** outcome $V = E[Y] = E[f(Y|X) \cdot f(X)]$; the sample data gives $f(Y|X) \cdot f(X|\tilde{v}(\theta, X) < P) = f(Y|\tilde{v}(\theta, X) < P)$, and the firm does not know either $f(X)$ or $f(X|\tilde{v}(\theta, X) < P)$. In this scenario, the firm is unable to recover $E[Y]$ by reweighting the sample data to match the target population distribution. This scenario is common in survey research, where researchers can see the attributes of participants in their survey but often lack information about the attributes distribution in their population of interests (e.g., Williams et al. 2024).

   In situations where the firm observes attributes in both the sample and population distribution, they may use reweighting to alleviate the sample bias. However, reweighting is not a panacea as it decreases the effective sample size (ESS). The ESS decreases more as the variance of the weights increases (Stantcheva 2022), which happens when the sample data is more biased compared to the target population distribution.

2. The firm cares about the full, **heterogeneous** distribution of the outcome $f(Y|X)$ and is able to observe X from the sample data. That allows them to learn $f(Y|X)$; however, some consumers $X = x_1$ are underrepresented in the dataset, which means that the accuracy of estimates for $f(Y|X = x_1)$ is low. This is typical in settings where the firm (or other agencies) needs to know the heterogeneity of outcomes to design targeted campaigns and interventions, as reflected by data challenges facing clinical trials (Ma et al. 2021), credit score estimates (Blattner & Nelson 2021), among others.

In theory, one can argue against representativeness as a metric for data quality, since outcomes may not always vary systematically with observable characteristics. However, as data is used to learn about many different outcomes, this possibility becomes rare in practice. This is especially true when the data buying firm is an intermediary (e.g., Meta, Google, Experian) that later resells the raw or derived data to different end-user firms.[19] Since different end-user firms have different customer bases and different analytic objectives, the data-buying firm has to consider all these use cases. Even in cases where the data-buying and end-user firms are identical, the firm still needs to form an expectation over potential future uses of data, hence the phrase “data is a strategic asset”. In situations like these, the objective of minimizing biases in the learned outcome vector Y boils down to the objective of minimizing biases in observed attributes X in the sample data, while the firm remains agnostic about the specific data production in each use case. Motivated by this observation, we focus on the representativeness of observables X in the sample as a key data quality metric. We note that our focus on representative data based on observable characteristics has always been the norm in sampling research (Cochran et al. 1954).

[19]Selling derived data can mean selling impressions from certain consumer segments, where the advertisers specify what segments they want to reach based on demographics or inferred consumer interests.

6.2 A Decision-Theoretic Model for Optimizing Choice Architecture

Next, we outline the firm's frame and price optimization problem, showing how a firm can maximize its profits—the information value from data minus data acquisition costs—by balancing data volume and bias. This framework allows us to formally evaluate the tradeoffs between these two factors. In particular, we identify sufficient statistics that can be estimated from the data to inform optimization.

A firm aims to maximize the information value of data minus the price paid to acquire it. There are K groups of consumers, each with potentially different preferences and behaviors. The firm can control two variables: the price offered to each group and the frame each group sees. It selects a frame for each group from a finite set of possible frames $\mathcal{T} = \{\theta^1, ..., \theta^I\}$, and a price for each group from $\mathbb{R}^+$. Essentially, the firm chooses an assignment of frames and prices $\Theta = (\theta_1, ..., \theta_K)$ and $\vec{P} = (p_1, ..., p_K)$.

We denote $\vec{N} = (n_1, ..., n_K)$ as the vector representing the number of consumers who share their data in each group, and $N = \sum_k n_k$ as the total number of people sharing (i.e., the volume). For each group k, data sharing depends on both the price and the choice architecture they face. In particular, let $\widetilde{v}_k = f_k(v_{0,k}; \theta_k)$ be the observed privacy valuation for group k. Then $n_k(\theta_k, p_k) = \int [f_k(v_{0,k}; \theta_k) \leq p_k]\, dv \equiv F_k(p_k)$. We define $B(\vec{N})$ as the bias in the data. Examples of bias measures include the sum of normalized distance between sample and population means (e.g., the t-statistic), as well as distribution distance measures (e.g., the Kolmogorov–Smirnov statistic, Chi-squared distance, Kullback–Leibler divergence, among others).

The value of data for the firm is represented by $V(N, B)$, a function of both data volume (N) and bias (B). We assume V is twice differentiable, with positive but decreasing marginal returns in volume ($\frac{\partial V}{\partial N} > 0$, $\frac{\partial^2 V}{\partial^2 N} \leq 0$) and decreasing in bias ($\frac{\partial V}{\partial B} < 0$). In Appendix E.3, we provide three examples of data value functions from existing literature in economics, marketing, and statistics, and show that they all satisfy these properties.

The firm's profit function is the value of data minus the acquisition cost:

$$\Pi = V(N, B) - \sum_k p_k \cdot n_k.$$

The firm chooses frames and prices to influence the volume and bias, as an intermediate step to maximize profits.

We now turn to solving the problem. We start by focusing on **the case of a uniform frame and price** offered to all customers. In certain settings, the firm lacks the ability to offer either customized prices or frames across consumers. In this case, the profit simplifies to a function of a single price p and frame θ, assigned to the entire consumer base:

$$\Pi(\theta, p) = V\left(N(\theta, p), B\left(\vec{N}(\theta, p)\right)\right) - pN.$$

The optimal price given a frame. Under our assumptions about the shape of V, we can find the optimal price for a given θ that satisfies the following condition (see proof in Appendix E)[20]:

$$p_\theta = \left(\eta_N^V \epsilon_p^N + \eta_B^V \epsilon_p^B\right) \bar{V}. \tag{2}$$

Here, η_N^V is the elasticity of value to volume, η_B^V is the elasticity of value to bias, and $\bar{V}$ is the average value per customer. Similarly, ϵ_p^N and ϵ_p^B are the elasticities of volume and bias with respect to price, respectively. Typically, ϵ_p^N is positive, as higher prices encourage more data sharing. The sign of ϵ_p^B is ambiguous, but we expect it to be negative when the price is high relative to consumers' average valuation, since higher prices in this region result in better coverage and hence reduced bias. Equation 2 is an implicit condition because V, N, and B are all functions of p. If price effects on bias reduction and volume increase are stronger, the optimal price will be higher.

The optimal choice frame. So far, we have derived the optimal price for a given choice architecture. To find the globally optimal price and frame assignment, the firm needs to compare profits across different frames $\forall \theta \in \mathcal{T}$ and choose the price-frame pair that maximizes profits. Since V increases with volume and decreases with bias, two candidate frames are θ_{maxN} (maximizing volume) and θ_{minB} (minimizing bias).[21]

Globally, we ask which frame is better, each with its own optimal price:

$$V\left(N_{maxN}, B_{maxN}\right) - p_{maxN} N_{maxN} \lesseqgtr V\left(N_{minB}, B_{minB}\right) - p_{minB} N_{minB}.$$

Given that at status quo firms use the volume-maximizing frame to collect data, we establish a sufficient condition for preferring the bias-minimizing frame. In particular, we compare profits under the volume-maximizing and the bias-minimizing frames, both with the vol-max optimal price p_N^*. If profits under the bias-minimizing frame are higher even with the volume-maximizing frame's optimal price, the bias-minimizing frame-price pair will be globally better.

$$\begin{aligned} V\left(N(p_N^*, \theta_{maxN}), B(p_N^*, \theta_{maxN})\right) - V\left(N(p_N^*, \theta_{minB}), B(p_N^*, \theta_{minB})\right) \\ - p_N^* \left(N(p_N^*, \theta_{maxN}) - N(p_N^*, \theta_{minB})\right) < 0. \end{aligned}$$

As derived in Appendix E, this condition can be approximated by:

$$\frac{\eta_N^V}{|\eta_B^V|}(1 - \epsilon_p^N) + \epsilon_p^B < \left|\gamma_N^B\right|, \tag{3}$$

[20] N and B are both conditional on θ, but we omit θ for brevity.

[21] Depending on the specific value function, a choice architecture that balances volume and bias can be better than either the volume-maximizing or bias-minimizing frames, but we choose these two frames to represent the bounds of the firm's preferences.

where $\gamma_N^B \equiv \frac{\%\Delta B}{\%\Delta N}$ is the arc elasticity of bias to volume, representing the bias reduction relative to volume loss when switching from a volume-maximizing to a bias-minimizing frame. Condition 3 implies that the greater γ_N^B is, the more efficient the bias-minimizing frame becomes in increasing data value, and the more likely that the bias-minimizing frame is preferable. On the other hand, if data value is more sensitive to volume than bias, the volume-maximizing frame is more likely to be better.

Although the formulas above are based on the case of uniform frame and price, we can extend the same framework to settings where the firm has to set a uniform price but can personalize the choice architecture.[22] Since frames are discrete, Condition 2 still holds, only that $\vec{N}$ is evaluated at assignment $\Theta = (\theta_1, ..., \theta_K)$ with $\theta_k \in \mathcal{T}$ for each $k \in \{1, ..., K\}$ instead of at a uniform frame $(\theta', ..., \theta')$ for some $\theta' \in \mathcal{T}$. The firm can then calculate expected profits for each frame assignment policy Θ given the optimal price for each frame, and again choose the best among those pairs.

The derivation above shows that to decide which frame is optimal while allowing for price adjustments, the firm needs to know how consumers respond to prices and frames (the elasticities of volume and bias to price and the relative changes in bias to volume due to frame changes), as well as its data value function V and associated elasticities: the elasticities of volume and bias to price. In what follows, we use the data from our experiment to estimate the data volume and bias non-parametrically as a function of prices and frames, and calculate the different statistics above. We do not impose any specific value function V here, since our experiment cannot speak to the value for data collectors.

6.3 Counterfactual Simulation: The Bias and Volume of Data across Frames

We construct data supply curves under each choice frame using data from our experiment: Consumers with data valuations lower than the offer price are those who share their data.[23] Knowing who will share data for a given price and frame, we can construct the *sample data* under different price points and frames. We calculate sample data for prices ranging from \$25 and \$90, in \$1 increments.[24] For each sample data, we calculate its bias and volume:

- *Bias* is represented as the standardized difference in attributes between consumers who share data and all experiment participants. For continuous attributes like income, this is the standardized difference in means; for discrete attributes, this is the difference in

[22] If the firm is able to personalize price, then the problem becomes trivial as it can use personalized prices and the volume-maximizing frame to maximize volume and representativeness. However, existing regulations, such as CCPA, prohibit price discrimination based on data-sharing decisions, thus we believe that this scenario is empirically less likely.

[23] When constructing the supply curve, we averaged each consumer's valuations across 5 personal variables, as we do not focus on the distinction across personal variables in this counterfactual.

[24] We choose the range of price grid such that the lowest price still guarantees that a nonzero percentage of consumers will share data in the average frame. If the percentage of consumers sharing the data is zero, we will not be able to calculate the bias metric. The highest price is more flexible, though due to the long-tail nature of consumer valuation, increasing prices in \$1 increments at the higher end does not move the amount of data shared as much, thus we stop at \$90.

attribute percentages. We first focus on biases in individual attributes to illustrate the different mechanisms, then average over the standardized mean difference to form an aggregate bias measure.

- *Volume* is represented as the percentage of consumers captured in the sample data. Although alternative metrics for volume exist, using a percentage metric better characterizes how the volume-maximizing frame changes the bias in data across different supply curve regions.

As a first step, we compare the performance of a vol-max frame with a non-optimized benchmark where the firm chooses one of the available frames at random (see Appendix Figure D.7), an empirical analog of Figure 1. To construct valuations and choices under the no-optimization benchmark, we average the valuations across the frames in our experiment. Under the no-optimization benchmark, at each price level, low and high income consumers have similar inclinations to share data, represented by the red (low income) and blue (high income) consumers evenly distributed along the supply curve. When it comes to the vol-max frame, however, more low-income consumers are willing to share data than wealthier ones at any given price point, indicated by the accumulation of red dots towards the right and blues dots towards the left. In short, with the negative correlation between privacy valuations and frame effects, the vol-max frame could potentially exacerbate bias even more than when the firm does not perform frame optimization at all.

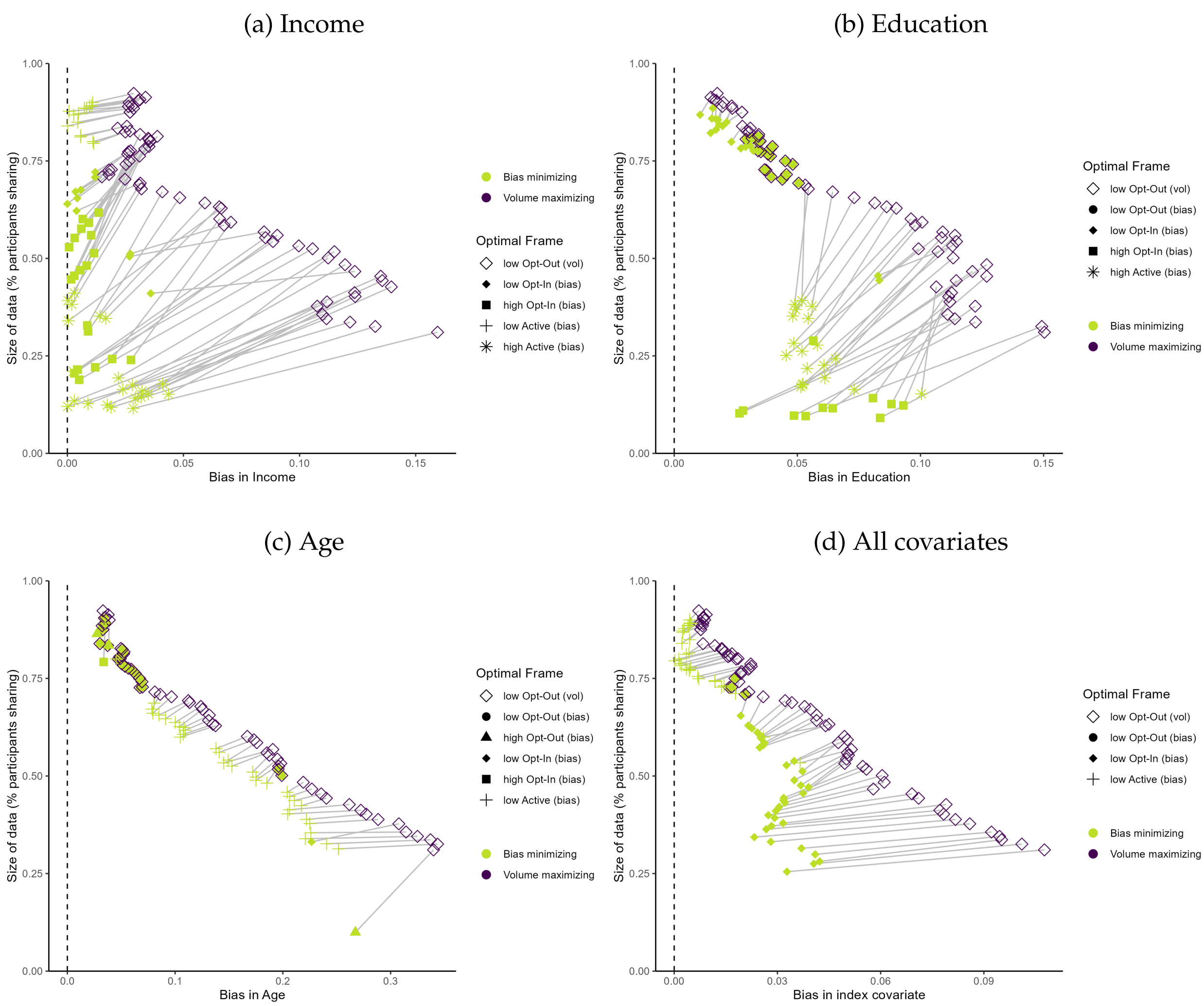
With the optimization framework in mind, we now visually examine the supply and trade-offs of volume and bias relative to the frame and price choices, then move to estimate the key elasticities. We illustrate the bias-variance trade-off for consumer subgroups with a negative correlation between privacy valuation and frame response: *income*, *age*, and *education*, as identified by our forest model estimates. Focusing on covariates in separation is a common approach of assessing similarity between samples (e.g., in covariate balance tables like the one we have in Table C.2). We then expand the trade-off comparison by including all demographic covariates to create a single bias index.[25] Constructing a uni-dimensional bias index allows us to examine a simple personalization approach below.[26]

Figure 10 shows the biases and volumes of sample data collected under the vol-max and bias-min frames separately for the four different variables. In this figure, each point represents the sample data collected under a specific choice architecture and price. The dark purple color represents the vol-max frame, which is consistently the low price anchor with an opt-out default. The light green points correspond to the bias-min frames, which can be different across price points. The horizontal axis indicates data bias, and the vertical axis represents volume. The gray lines connect pairs of samples collected at the same price but under different frames. For many prices and covariates, the vol-max and bias-min frames do not coincide. The vol-max frame can substantially

[25]To create a simple and transparent index, we standardize each covariate and then take an average. The interpretation is that more is better—higher in income, education, age, being female, and engaging more digitally online.

[26]In Appendix D.7, we take an alternative approach, where we calculate the standardized bias for each covariate separately, and then average across these biases as our bias measure.

Figure 10: Data Quality under Vol-Max and Bias-Min Frames



Notes: Each point represents a sample dataset collected under a choice frame × price combination. The prices for data vary from \$25 to \$90 in \$1 increments; points connected by a gray line are collected under the same price but different frames. All covariates are standardized to have mean of 0 and standard deviation 1.

exacerbate bias in sample data compared to the bias-min frame, especially at lower price points: Intuitively, when consumer segments with initial low data valuations are more influenced by the frame, they become much more willing to share data, leading to over-representation in the sample.

On the other hand, the quality of data collected under these two frames converges as the price increases. Differences in bias and volume between the frames become smaller. In fact, at higher prices, the vol-max frame is more likely to coincide with the bias-min frame.

Finally, the slopes of the gray lines vary substantively. These slopes represent the ratio of bias reduction to volume reduction from switching between the vol-max and bias-min frames, which is the arc elasticity γ_N^B. The flatter the line, the higher the reduction in bias compared to volume loss.

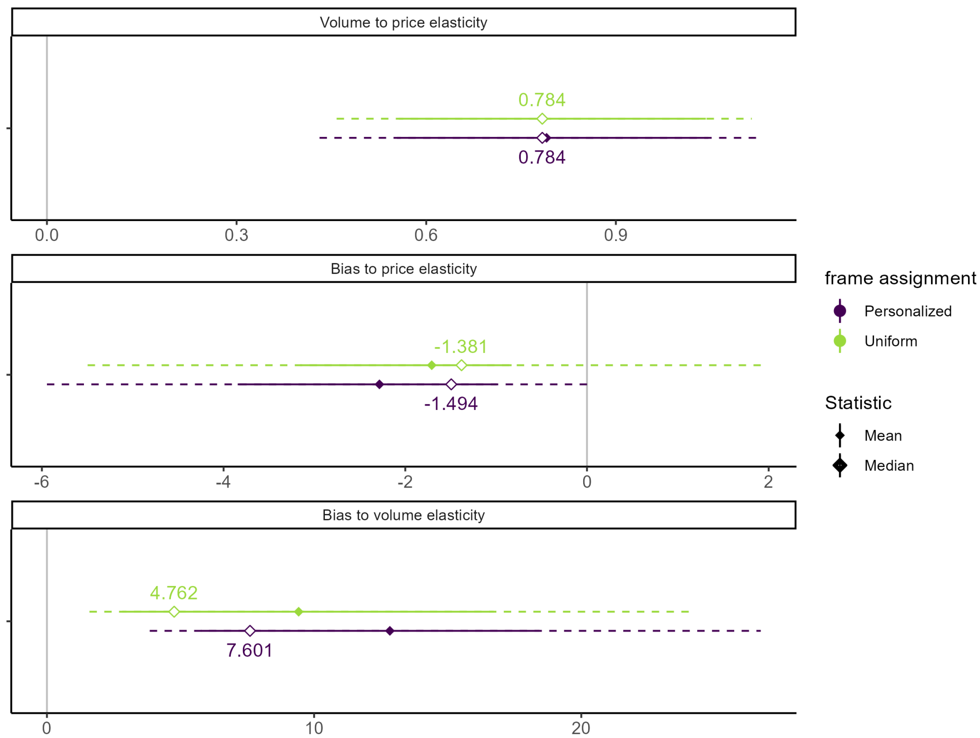
As shown in Section 6.2, the frame choice depends on the responsiveness of volume and bias to prices and frames, and on the sensitivity of data value to volume and bias. We now quantify the key elasticities and the trade-offs in switching between vol-max and bias-min frames for the same price. We focus on the covariate index for a disciplined comparison (see Figure 10d). The flat slopes of the gray lines imply that in our setting, the firm can sacrifice relatively little volume to achieve substantial gains in bias reduction: Conditional on the same price, moving from the vol-max to the bias-min frames can substantially increase sample representativeness while reducing volume only slightly. Mapping it to our framework, it means the ratio γ_N^B is high, making the sufficient condition for preferring a bias-minimizing frame 3 more likely to hold.

The gap between vol-max and bias-min frames is large when the sample size is small, and generally decreases as the sample size grows. This pattern suggests that the trade-off between the two objectives is more pronounced during the initial stage of gathering consumer data, typical for new entrants and smaller firms. In comparison, larger and more established firms are more likely to find that volume-maximizing frames effectively reduce bias.

In accordance with Condition 3, we estimate sufficient statistics to quantify how the volume-bias trade-offs determines the frame choice. First, we calculate the relative bias and volume reductions as the firm moves from the vol-max to the bias-min frame, and take their ratio to get γ_N^B. Here, we define bias by comparing the average of standardized consumer covariates between the sample and the full data. The bias reduction is zero if the two frames coincide and one if all bias is eliminated; volume reduction is defined similarly. We then calculate two price elasticities: The volume-to-price elasticity, ϵ_p^N, is the ratio of the percentage increase in volume to the percentage increase in price under the vol-max frame; the bias-to-price elasticity, ϵ_p^B, is the ratio of the percentage change in bias to the percentage increase in price under the vol-max frame. We calculate the elasticity conditional on the firm using the vol-max frame as this is the status-quo choice among practitioners. Calculating the elasticity for the bias-min frame is less natural, because the bias-min frame often changes as the price increases, violating the "fixed frame" idea in our elasticity definition.

The trade-off and elasticities can depend on whether the firm can customize the frame for each consumer or is constrained to uniform frame assignment. For example, the firm may know customers' demographic attributes before asking them to share their private behavioral data and can customize the frames accordingly, which is a feature that many consent optimization products offer. In view of this possibility, we calculate these trade-off metrics separately for uniform and personalized frame assignments. To demonstrate the potential of personalization while adhering to a model-free visualization scheme, we divide participants into above and below median covariate index levels and search for the best among all 6×6 potential frame assignments. We provide the summary statistics of the trade-off metrics in Figure 11.

Figure 11: Summary Statistics of Bias-Variance Trade-off Metrics



Notes: The figure shows summary statistics of the three different elasticities estimated at $1 increments. Solid lines represent the 25th-75th percentile range, and dashed lines are 10th-90th percentile ranges of the respective distributions. The numbers in the figure represent the median of the distributions.

With a uniform frame, the median volume-to-price elasticity is 0.78 and the distribution is tight and positive. In contrast, the bias-to-price elasticity is more variable across prices and is negative for the majority of prices, with a median value of -1.38. Finally, moving from the vol-max frame to the bias-min frame results in a large and variable ratio of bias reduction to volume reduction, with a median of 4.76. These numbers quantify the flat slopes in Figure 10d, showing that a smaller volume reduction is needed for a larger bias reduction gain. Based on condition 3, these values imply that if the elasticity of value to volume is less than roughly 28 times the elasticity of value to bias, it is better to use the bias-minimizing frame.

The trade-off metrics change further in favor of bias minimization when the firm is able to customize frames across consumers. Personalized frame assignment results in a higher ratio of bias reduction to volume reduction, further favoring bias-minimizing frames. As Figure 11 shows, the median ratio of bias reduction to volume reduction increases by 60% in our data. The smaller volume loss may come from the fact that the frame most effective in decreasing consumers' valuations is mostly the same across consumers (opt-out and low-anchor) in our setting. On the other hand, the greater bias reduction gain is likely robust, as personalization gives more leverage for balancing consumer composition by encouraging one segment to share and discouraging another. Our personalization algorithm is rudimentary, but the fact that even this simple procedure has a large impact proves the potential of more sophisticated personalization techniques to provide greater benefits.

Although we compare only vol-max and bias-min frames for illustrative purposes, other frames that balance the volume and representativeness goals may be preferable in certain situations. As an example, Appendix Figure D.8 shows the bias (for income) and volume of sample data under all possible frame assignments for a specific price, with the frontier of undominated choice architectures highlighted in color. In this example, another frame on the frontier that does not maximize volume nor minimize bias could potentially result in higher data value gain. Appendix Figure D.9 shows the frontiers for several prices. As these figures show, personalization allows the firm to push the frontier towards the left, making it easier to find a frame that reduces bias without much sacrifice to volume.

7 Conclusion

Choice architecture can substantially distort consumers' privacy valuations and thus the supply of consumer data. The choice frames we examine—default and price anchor—shift the average consumer valuation for their Facebook data by 22% and 37%, respectively. Moreover, for some consumer segments, the susceptibility to frame influence is negatively correlated with their valuation for data across frames. Younger, lower-income, and less educated consumers tend to respond more strongly to changes in the frame; they also value their personal data less on average.

The fact that different consumers respond to frames differently implies that choice architecture can substantially change the composition of data that firms collect. We argue that conventional practices of choosing a frame that maximizes data volume while ignoring its impact on data bias can be suboptimal, and derive metrics that help companies assess when this is true. We also show how alternative frame designs can allow us to achieve substantial gains in bias reduction without sacrificing much volume, especially when frames can be personalized. We provide an empirical example to show the mechanisms under which choice architecture can create tension between the two data quality objectives.

In business settings, companies can implement these lessons to improve their data collection practice. The key to optimizing the frame for data collection lies in knowing the joint heterogeneity of privacy preferences and choice frame responses, which allows the firm to estimate the key statistics needed for optimization. One potential tool is a sequential optimization strategy, where the firm uses contextual bandit to simultaneously learn about this joint heterogeneity and choose the optimal frame given its current knowledge. Using a multi-armed bandit to search for the optimal design is an existing feature in experiment-as-a-service platforms.[27] Previous work (Qiang

[27]See: https://vwo.com/blog/multi-armed-bandit-algorithm/, and https://www.optimizely.com/optimization-glossary/multi-armed-bandit/.

& Bayati 2016, Schwartz et al. 2017, Misra et al. 2019) shows that enriching MAB with a model can substantially improve learning efficiency by reducing the dimension of the parameter space to learn about. We believe that such a framework has great potential to improve the framing choices and data collection efficiency over existing practices.

Our study is related to classic adverse selection and moral hazard problems, but distinct in a few key dimensions. In our setting, selection is exemplified by our findings of different valuations across user groups. As such, groups with lower privacy valuations will share more data. If these consumers are worth less to data collectors, then it becomes a classic adverse selection problem. In the paper, we argue that regardless of which group of users are initially more valuable, privacy valuation differences result in sample imbalance, and the under-represented group becomes more valuable as a result. This is not exactly classical adverse selection, in the sense that consumers' value to the firm is in part "induced" by them being less willing to share data rather than the other way around. Nevertheless, they are conceptually related and could generate similar patterns in some cases. Moral hazard in our setting means that consumers could change their behavior to resemble higher types in order to be offered higher prices, in the event of price discrimination. This concern is less relevant in our particular setting as we only consider uniform prices. However, we acknowledge that moral hazard may be present if personalized pricing emerges in markets for data, or in settings where data collection is tied to downstream price discrimination in product markets.

Our results contribute to the broad discussions about the efficiency of the data market. Several unique features of consumer data can lead to market failures, such as externalities (Bergemann et al. 2022), incomplete information (Jin 2018), and non-rivalry (Jones & Tonetti 2020). Here, we show another feature that contributes to market inefficiency: A data market that allows consumers to make their own choices can often create bias in data collected and decrease the returns to data. In the context of choice architecture evaluation, we show when companies' efforts to maximize data collection can exacerbate or alleviate the bias in different market conditions. However, a data collector attuned to reducing bias can leverage choice architecture as a tool rather than a hindrance. We believe the bias aspect is worth emphasizing, as it represents a novel channel in data markets that affects its functioning.

References

Abis, S. & Veldkamp, L. (2024), 'The changing economics of knowledge production', *The Review of Financial Studies* **37**(1), 89–118.

Acemoglu, D., Makhdoumi, A., Malekian, A. & Ozdaglar, A. (2022), 'Too much data: Prices and inefficiencies in data markets', *American Economic Journal: Microeconomics* **14**(4), 218–256.

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

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

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

Andersen, S., Harrison, G. W., Lau, M. I. & Rutström, E. E. (2006), 'Elicitation using multiple price list formats', *Experimental Economics* **9**(4), 383–405.

Aridor, G., Che, Y.-K. & Salz, T. (2021), The effect of privacy regulation on the data industry: Empirical evidence from gdpr, *in* 'Proceedings of the 22nd ACM Conference on Economics and Computation', pp. 93–94.

Ariely, D., Loewenstein, G. & Prelec, D. (2003), "'coherent arbitrariness": Stable demand curves without stable preferences', *Quarterly Journal of Economics* **118**(1), 73–106.

Arrieta-Ibarra, I., Goff, L., Jiménez-Hernández, D., Lanier, J. & Weyl, E. G. (2018), Should we treat data as labor? moving beyond 'free', *in* 'AEA Papers and Proceedings', Vol. 108, pp. 38–42.

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

Athey, S., Tibshirani, J. & Wager, S. (2019), 'Generalized random forests', *The Annals of Statistics* **47**(2), 1148–1178.

Athey, S. & Wager, S. (2019), 'Estimating treatment effects with causal forests: An application', *Observational studies* **5**(2), 37–51.

Bajari, P., Chernozhukov, V., Hortaçsu, A. & Suzuki, J. (2019), The impact of big data on firm performance: An empirical investigation, *in* 'AEA Papers and Proceedings', Vol. 109, pp. 33–37.

Becker, G. M., DeGroot, M. H. & Marschak, J. (1964), 'Measuring utility by a single-response sequential method', *Behavioral science* **9**(3), 226–232.

Bergemann, D., Bonatti, A. & Gan, T. (2022), 'The economics of social data', *The RAND Journal of Economics* **53**(2), 263–296.

Blattner, L. & Nelson, S. (2021), 'How costly is noise? data and disparities in consumer credit', *arXiv preprint arXiv:2105.07554* .

Bradley, V. C., Kuriwaki, S., Isakov, M., Sejdinovic, D., Meng, X.-L. & Flaxman, S. (2021), 'Unrepresentative big surveys significantly overestimated us vaccine uptake', *Nature* **600**(7890), 695–700.

Cao, R., Koning, R. M. & Nanda, R. (2021), 'Biased sampling of early users and the direction of startup innovation', *NBER Working Paper No. 28882* .

Cochran, W. G., Mosteller, F. & Tukey, J. W. (1954), 'Principles of sampling', *Journal of the American Statistical Association* **49**(265), 13–35.

Collis, A., Moehring, A. & Sen, A. (2020), 'Economic value of data: Quantification using online experiments'.

D'Assergio, C., Manchanda, P., Montaguti, E. & Valentini, S. (2022), 'The race for data: Gaming or being gamed by the system?', *Available at SSRN 4250389* .

Dubé, J.-P. & Misra, S. (2023), 'Personalized pricing and consumer welfare', *J. Polit. Econ.* **131**(1), 131–189.

Farboodi, M. & Veldkamp, L. (2023), 'Data and markets', *Annual Review of Economics* **15**, 23–40.

FTC (2022), Bringing dark patterns to light, Staff report, FTC.
URL: *https://www.ftc.gov/reports/bringing-dark-patterns-light*

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

Hils, M., Woods, D. W. & Böhme, R. (2020), Measuring the emergence of consent management on the web, *in* 'Proceedings of the ACM Internet Measurement Conference', pp. 317–332.

Iansiti, M. (2021), 'The value of data and its impact on competition', *Harvard Business School NOM Unit Working Paper* **22-002**.

Ichihashi, S. (2020), Dynamic privacy choices, *in* 'Proceedings of the 21st ACM Conference on Economics and Computation', pp. 539–540.

Ichihashi, S. (2021), 'The economics of data externalities', *Journal of Economic Theory* **196**, 105316.

Jin, G. Z. (2018), 'Artificial intelligence and consumer privacy', *The Economics of Artificial Intelligence: An Agenda; Agrawal, A., Gans, J., Goldfarb, A., Eds* pp. 439–462.

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

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

Jones, C. I. & Tonetti, C. (2020), 'Nonrivalry and the economics of data', *American Economic Review* **110**(9), 2819–58.

Kahneman, D., Knetsch, J. L. & Thaler, R. H. (1990), 'Experimental tests of the endowment effect and the coase theorem', *Journal of political Economy* **98**(6), 1325–1348.

Kormylo, C. & Adjerid, I. (2021), 'Reconsidering privacy choices: The impact of defaults, reversibility, and repetition', *Available at SSRN 3846620* .

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

Lee, J. Y., Yang, J. & Anderson, E. (2023), 'Using grocery data for credit decisions', *Wharton Customer Analytics* .

Lee, J. Y., Yang, J. & Anderson, E. (2024), Who benefits from alternative data for credit scoring? evidence from peru.

Lei, X., Chen, Y. & Sen, A. (2023), 'The value of external data for digital platforms: Evidence from a field experiment on search suggestions', *Available at SSRN* .

Liang, A. & Madsen, E. (2020), Data and incentives, *in* 'Proceedings of the 21st ACM Conference on Economics and Computation', pp. 41–42.

Lin, T. (2022), 'Valuing intrinsic and instrumental preferences for privacy', *Marketing Science* .

Linos, E. (2018), 'More than public service: A field experiment on job advertisements and diversity in the police', *Journal of Public Administration Research and Theory* .

Ma, M. A., Gutiérrez, D. E., Frausto, J. M. & Al-Delaimy, W. K. (2021), Minority representation in clinical trials in the united states: trends over the past 25 years, *in* 'Mayo Clinic Proceedings', Vol. 96, Elsevier, pp. 264–266.

Markovich, S. & Yehezkel, Y. (2021), "'for the public benefit": who should control our data?', *Available at SSRN 3945108* .

Martin, K. & Nissenbaum, H. (2016), 'Measuring privacy: an empirical test using context to expose confounding variables', *Colum. Sci. & Tech. L. Rev.* **18**, 176.

Matte, C., Bielova, N. & Santos, C. (2020), Do cookie banners respect my choice?: Measuring legal compliance of banners from iab europe's transparency and consent framework, *in* '2020 IEEE Symposium on Security and Privacy (SP)', IEEE, pp. 791–809.

Miklós-Thal, J., Goldfarb, A., Haviv, A. & Tucker, C. (2024), 'Frontiers: Digital hermits', *Marketing Science* .

Miller, K., Sahni, N. S. & Strulov-Shlain, A. (2022), 'Sophisticated consumers with inertia: Long-term implications from a large-scale field experiment', *SSRN Electron. J.* .

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

Misra, K., Schwartz, E. M. & Abernethy, J. (2019), 'Dynamic online pricing with incomplete information using multiarmed bandit experiments', *Marketing Science* **38**(2), 226–252.

Misra, S. (2023), 'Algorithmic nudges', *mimeo* .

Morel, V., Santos, C., Fredholm, V. & Thunberg, A. (2023), Legitimate interest is the new consent-large-scale measurement and legal compliance of iab europe tcf paywalls, *in* 'Proceedings of the 22nd Workshop on Privacy in the Electronic Society', pp. 153–158.

Mrkva, K., Posner, N. A., Reeck, C. & Johnson, E. J. (2021), 'Do nudges reduce disparities? choice architecture compensates for low consumer knowledge', *J. Mark.* **85**(4), 67–84.

Mueller-Tribbensee, T., Miller, K. M. & Skiera, B. (2024), 'Paying for privacy: Pay-or-tracking walls', *arXiv preprint arXiv:2403.03610* .

Neumann, N., Tucker, C. E., Kaplan, L., Mislove, A. & Sapiezynski, P. (2022), 'Data deserts and black box bias: The impact of socio-economic status on consumer profiling'.

Neyman, J. (1934), 'On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection', *Journal of the Royal Statistical Society Series A: Statistics in Society* **97**(4), 558–606.

Nie, X. & Wager, S. (2021), 'Quasi-oracle estimation of heterogeneous treatment effects', *Biometrika* **108**(2), 299–319.

Nouwens, M., Liccardi, I., Veale, M., Karger, D. & Kagal, L. (2020), Dark patterns after the gdpr: Scraping consent pop-ups and demonstrating their influence, *in* 'Proceedings of the 2020 CHI conference on human factors in computing systems', pp. 1–13.

Peltzman, S. (1981), 'The effects of ftc advertising regulation', *The Journal of Law and Economics* **24**(3), 403–448.

Peukert, C., Sen, A. & Claussen, J. (forthcoming), 'The editor and the algorithm: Recommendation technology in online news', *Management science* .

Prince, J. T. & Wallsten, S. (2020), 'How much is privacy worth around the world and across platforms?', *Journal of Economics & Management Strategy* .

Qiang, S. & Bayati, M. (2016), 'Dynamic pricing with demand covariates', *arXiv preprint arXiv:1604.07463* .

Rafieian, O. & Yoganarasimhan, H. (2021), 'Targeting and privacy in mobile advertising', *Marketing Science* **40**(2), 193–218.

Rao, A. & Wang, E. (2017), 'Demand for "healthy" products: False claims and ftc regulation', *Journal of Marketing Research* **54**(6), 968–989.

Rossi, P. E., McCulloch, R. E. & Allenby, G. M. (1996), 'The value of purchase history data in target marketing', *Marketing Science* **15**(4), 321–340.

Schwartz, E. M., Bradlow, E. T. & Fader, P. S. (2017), 'Customer acquisition via display advertising using multi-armed bandit experiments', *Marketing Science* **36**(4), 500–522.

Semenova, V. & Chernozhukov, V. (2021), 'Debiased machine learning of conditional average treatment effects and other causal functions', *The Econometrics Journal* **24**(2), 264–289.

Singleton, S. M. & Harper, J. (2002), 'With a grain of salt: What consumer privacy surveys don't tell us', *Available at SSRN 299930* .

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

Stantcheva, S. (2022), 'How to run surveys: A guide to creating your own identifying variation and revealing the invisible', *Annual Review of Economics* **15**.

Strulov-Shlain, A. (2023), 'More than a penny's worth: Left-digit bias and firm pricing', *The Review of Economic Studies* .

Sun, T., Yuan, Z., Li, C., Zhang, K. & Xu, J. (2021), 'The value of personal data in internet commerce: A high-stake field experiment on data regulation policy', *Available at SSRN 3962157* .

Tang, H. (2019), 'The value of privacy: Evidence from online borrowers'.

Thaler, R. H. & Sunstein, C. R. (2008), *Nudge: Improving decisions about health, wealth, and happiness*, Penguin.

Tomaino, G., Wertenbroch, K. & Walters, D. J. (2022), 'Express: Intransitivity of consumer preferences for privacy', *Journal of Marketing Research* .

Trusov, M., Ma, L. & Jamal, Z. (2016), 'Crumbs of the cookie: User profiling in customer-base analysis and behavioral targeting', *Marketing Science* **35**(3), 405–426.

Tucker, C. (2023), Algorithmic exclusion: The fragility of algorithms to sparse and missing data, *in* 'Brookings', Center on Regulation and Markets at Brookings, pp. 1–26.

Utz, C., Degeling, M., Fahl, S., Schaub, F. & Holz, T. (2019), (un) informed consent: Studying gdpr consent notices in the field, *in* 'Proceedings of the 2019 acm sigsac conference on computer and communications security', pp. 973–990.

Veldkamp, L. & Chung, C. (2024), 'Data and the aggregate economy', *Journal of Economic Literature* **62**(2), 458–484.

Wernerfelt, N., Tuchman, A., Shapiro, B. & Moakler, R. (2022), 'Estimating the value of offsite data to advertisers on meta', *Available at SSRN 4176208* .

Westin, A. F. (2003), 'Social and political dimensions of privacy', *Journal of social issues* **59**(2), 431–453.

Williams, A., Snoke, J., Bowen, C. & Barrientos, A. (2024), 'Disclosing economists' privacy perspectives: A survey of american economic association members' views on differential privacy and the usability of noise-infused data', *Harvard Data Science Review* .

Willis, L. E. (2020), 'Deception by design', *Harv. JL & Tech.* **34**, 115.

Woodruff, A., Pihur, V., Consolvo, S., Brandimarte, L. & Acquisti, A. (2014), Would a privacy fundamentalist sell their dna for $ 1000... if nothing bad happened as a result? the westin categories, behavioral intentions, and consequences, *in* '10th Symposium On Usable Privacy and Security (SOUPS 2014)', pp. 1–18.

A Institutional Background Appendix

Figure A.1: Meta's Pay-or-Tracking Offer in Europe

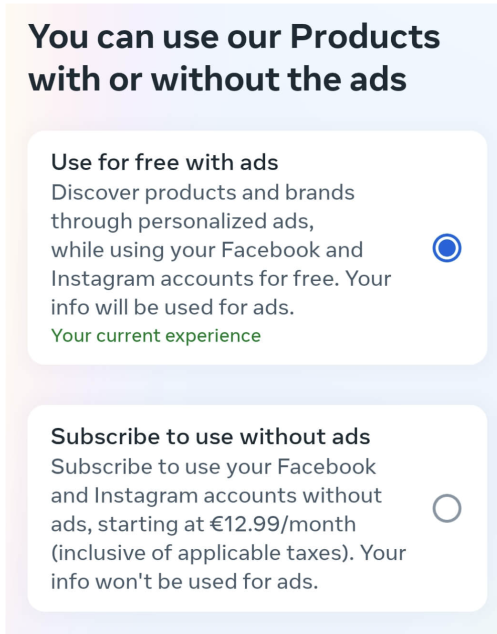
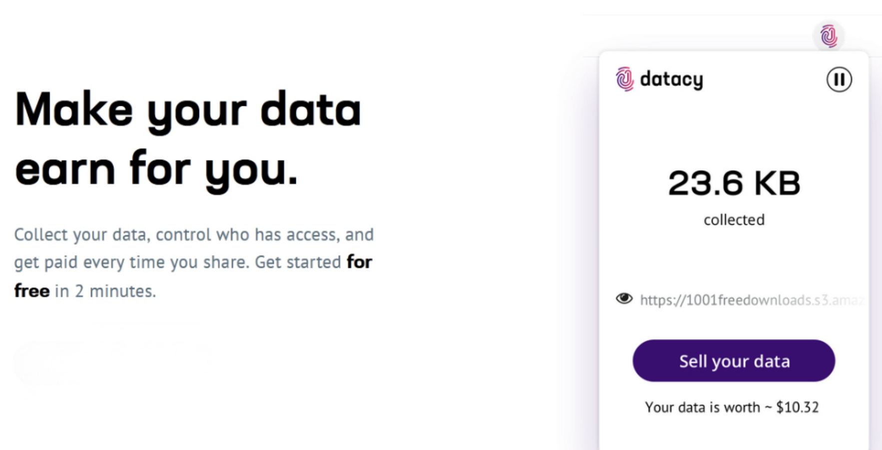


Figure A.2: Datacy's Price Anchor When Buying Consumer Data



B Experiment Appendix

Figure B.1: Example Recruiting Ad

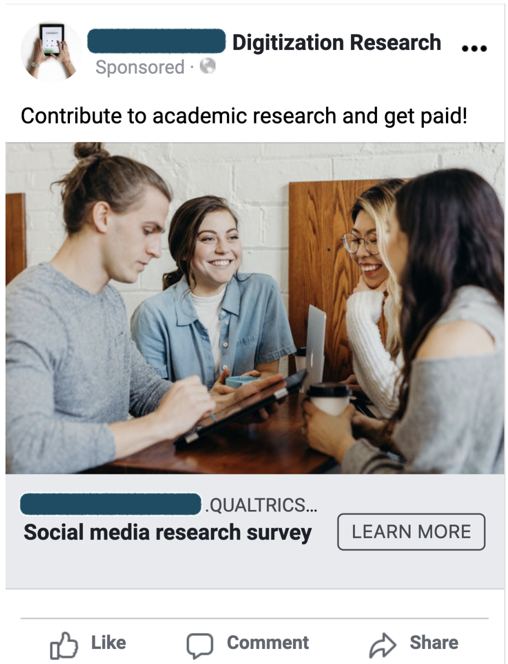
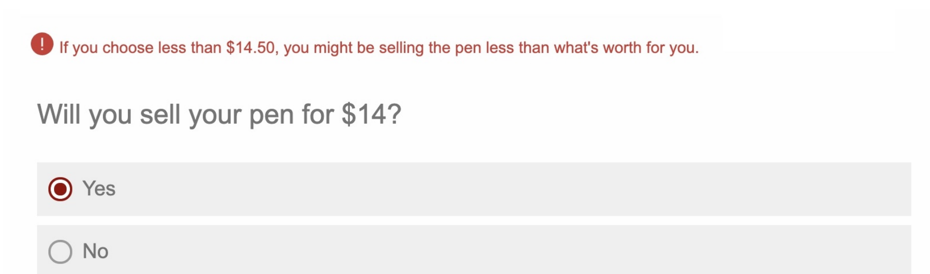


Figure B.2: Information Prompts in the MPL Practice Question

(a) Accepting an offer price too low



(b) Rejecting an offer price too high

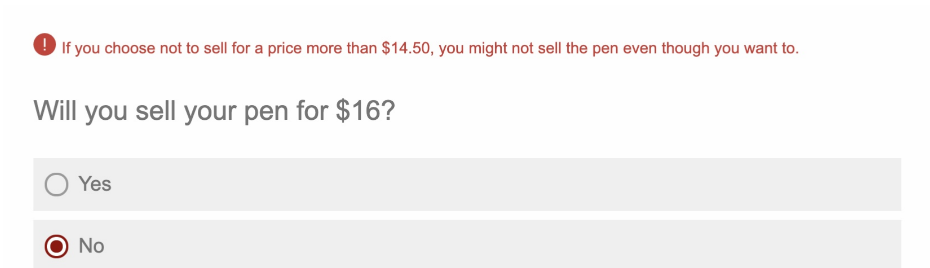


Figure B.3: Treatment Choice Architecture Variations

(a) High price anchor, opt-in default

In this question, we are asking for your price to share **Posts:** your Facebook posts and feed history from the last month.

If the computer chose any of the following prices, will you share your data?

	Your choice: Yes	Your choice: No
Will you share your **Posts** for $50?	○	◉
Will you share your **Posts** for $60?	○	◉
Will you share your **Posts** for $70?	○	◉
Will you share your **Posts** for $80?	○	◉
Will you share your **Posts** for $90?	○	◉
Will you share your **Posts** for $100?	○	◉

(b) Low price anchor, opt-out default

In this question, we are asking for your price to share **Posts:** your Facebook posts and feed history from the last month.

If the computer chose any of the following prices, will you share your data?

	Your choice: Yes	Your choice: No
Will you share your **Posts** for $0?	○	◉
Will you share your **Posts** for $10?	○	◉
Will you share your **Posts** for $20?	○	◉
Will you share your **Posts** for $30?	○	◉
Will you share your **Posts** for $40?	○	◉
Will you share your **Posts** for $50?	○	◉

(c) High price anchor, opt-out default

In this question, we are asking for your price to share **Posts:** your Facebook posts and feed history from the last month.

If the computer chose any of the following prices, will you share your data?

	Your choice: Yes	Your choice: No
Will you share your **Posts** for $50?	◉	○
Will you share your **Posts** for $60?	◉	○
Will you share your **Posts** for $70?	◉	○
Will you share your **Posts** for $80?	◉	○
Will you share your **Posts** for $90?	◉	○
Will you share your **Posts** for $100?	◉	○

(d) Low price anchor, opt-out default

In this question, we are asking for your price to share **Posts:** your Facebook posts and feed history from the last month.

If the computer chose any of the following prices, will you share your data?

	Your choice: Yes	Your choice: No
Will you share your **Posts** for $0?	◉	○
Will you share your **Posts** for $10?	◉	○
Will you share your **Posts** for $20?	◉	○
Will you share your **Posts** for $30?	◉	○
Will you share your **Posts** for $40?	◉	○
Will you share your **Posts** for $50?	◉	○

(e) High price anchor, active choice

In this question, we are asking for your price to share **Posts:** your Facebook posts and feed history from the last month.

If the computer chose any of the following prices, will you share your data?

	Your choice: Yes	Your choice: No
Will you share your **Posts** for $50?	○	○
Will you share your **Posts** for $60?	○	○
Will you share your **Posts** for $70?	○	○
Will you share your **Posts** for $80?	○	○
Will you share your **Posts** for $90?	○	○
Will you share your **Posts** for $100?	○	○

(f) Low price anchor, active choice

In this question, we are asking for your price to share **Posts:** your Facebook posts and feed history from the last month.

If the computer chose any of the following prices, will you share your data?

	Your choice: Yes	Your choice: No
Will you share your **Posts** for $0?	○	○
Will you share your **Posts** for $10?	○	○
Will you share your **Posts** for $20?	○	○
Will you share your **Posts** for $30?	○	○
Will you share your **Posts** for $40?	○	○
Will you share your **Posts** for $50?	○	○

B.1 MPL Instructions Excerpt

The following is the text explaining the MPL procedure to participants. It is followed by practice questions.

> Your answers to the survey questions, and other information, can help us understand browsing behavior better. It can also help companies and advertisers provide more products that they think you like, and show you fewer products that you are less likely to buy. **Would you be willing to share more data with us and advertisers?**
>
> **If you will, we will pay you a fair price.** When you started the survey, the computer already randomly selected if you will be asked to provide data at the end of the survey, and also randomly chose a price we will pay for it. If you are selected to participate, your payment and data shared with advertisers will base on what you choose. So you should answer carefully!
>
> For example: suppose you choose to share some data for a price of $Y. If the computer chose a price lower than $Y, you will not be asked to share the data and will not be paid. If the computer chose a price larger than $Y, you will be asked to download a copy of your data from Facebook and send them to us; you will then get the price the computer has chosen.

C Sample Description Appendix

Table C.1: Attrition throughout the Study

	Facebook	**Prolific**
Exposed to the ad	158453	-
Click on the ad	10135	-
Consent to participate	2348	3119
Complete all questions	2008	3018

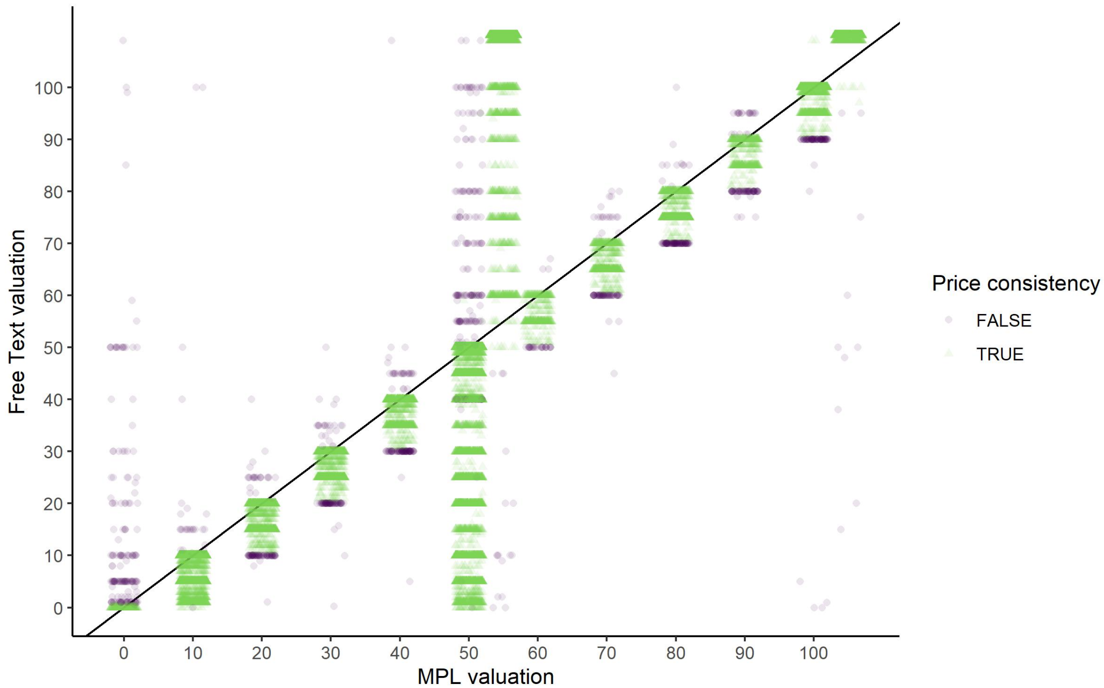
Table C.2: Covariate Balance across Treatments: Baseline Variables

	Treatment						T-test stat
	low: opt-out	low: active	low: opt-in	high: opt-out	high: active	high: opt-in	p-value
Number of participants							
n	786	824	817	832	815	831	0.321
Race (percentage)							
White	0.838	0.8	0.819	0.802	0.801	0.812	0.457
Black	0.055	0.073	0.059	0.073	0.071	0.059	0.773
Asian	0.106	0.112	0.105	0.108	0.118	0.125	0.42
Other	0.034	0.045	0.049	0.056	0.052	0.047	0.012
Gender (percentage)							
Female	0.592	0.586	0.595	0.594	0.574	0.592	0.191
Age							
Age	39.398	37.885	38.359	38.864	39.124	38.526	0.146
Income ($)							
Income	66724	67433	72080	70354	73156	66347	0.962
Education (percentage)							
College education and above	0.644	0.623	0.647	0.637	0.659	0.643	0.87
Some college	0.19	0.184	0.177	0.198	0.179	0.182	
Less than college	0.167	0.193	0.175	0.165	0.162	0.176	
Facebook and Internet usage							
Avg time spent on FB (h)	-0.051	0.061	0.01	-0.046	0.039	-0.027	0.755
FB membership duration (y)	0.037	0.028	0.095	0.019	0.023	0.014	0.712
Avg time spent on internet (h)	0.009	0.054	-0.013	-0.021	0.002	-0.016	0.872
Active user (percentage)	0.298	0.301	0.296	0.321	0.326	0.289	0.236
Purchase from FB or Instagram (times/mo)	-0.023	-0.032	0.003	-0.012	-0.02	-0.065	0.365
FB or Instagram ad click (times/mo)	-0.007	-0.024	0.01	-0.024	-0.002	0.042	0.019

Table C.3: Covariate Balance across Treatments: Endline Variables

	Treatment						T-test stat
	low: opt-out	low: active	low: opt-in	high: opt-out	high: active	high: opt-in	p-value
Number of participants							
n	786	824	817	832	815	831	0.321
Information available to the public							
Info from the about page	0.692	0.704	0.689	0.702	0.714	0.745	0.457
Posts	0.375	0.396	0.381	0.375	0.404	0.397	0.773
Photos	0.427	0.428	0.414	0.401	0.415	0.438	0.42
Lists of likes	0.309	0.317	0.321	0.312	0.324	0.337	0.012
Friends and followers	0.469	0.527	0.504	0.493	0.482	0.51	0.962
Don't know	0.162	0.126	0.151	0.16	0.144	0.144	0.191
Other information	0.085	0.104	0.075	0.079	0.106	0.066	0.146
Information available to Facebook advertisers							
Name and email	0.725	0.718	0.704	0.69	0.733	0.729	0.755
Info from the about page	0.822	0.816	0.815	0.831	0.847	0.836	0.712
Posts	0.531	0.535	0.526	0.507	0.531	0.521	0.872
Photos	0.468	0.45	0.454	0.425	0.456	0.452	0.236
Lists of likes	0.711	0.708	0.71	0.69	0.739	0.71	0.365
Friends and followers	0.585	0.631	0.624	0.588	0.632	0.611	0.019
Don't know	0.109	0.103	0.115	0.105	0.109	0.105	0.353
Other information	0.038	0.038	0.037	0.035	0.022	0.022	0.484
Participants looking up information							
Looked up information when aswering questions	0.019	0.023	0.021	0.023	0.028	0.031	0.891
The type of information participants look up							
How advertisers use data for targetting	0.008	0.007	0.004	0.01	0.011	0.011	0.659
What Facebook shares with advertisers	0.006	0.01	0.011	0.013	0.017	0.011	0.429
How much each data is worth	0.003	0.004	0.006	0.006	0.01	0.01	0.182
Other information	0.01	0.012	0.006	0.008	0.007	0.017	0.974
Frequency of encountering cookie banners/day	1.523	1.657	1.575	1.561	1.609	1.552	0.15

Figure C.1: Consistency between Final Price and MPL Choice



D Empirical Results Appendix

D.1 Alternative Data Supply Curves

Figure D.1: Data Supply Curves by Data Type (First-Round Answer Only)

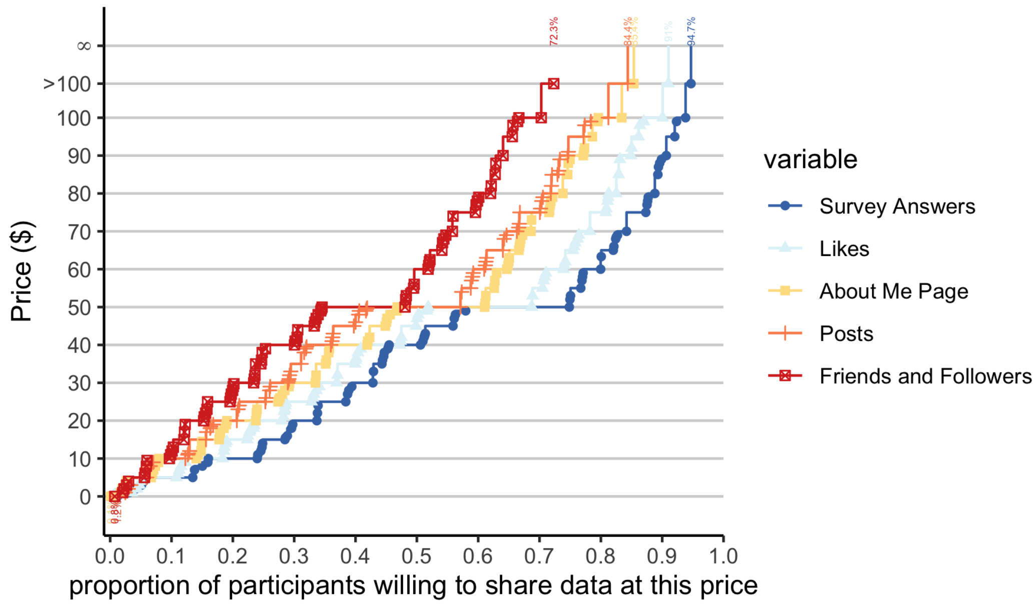
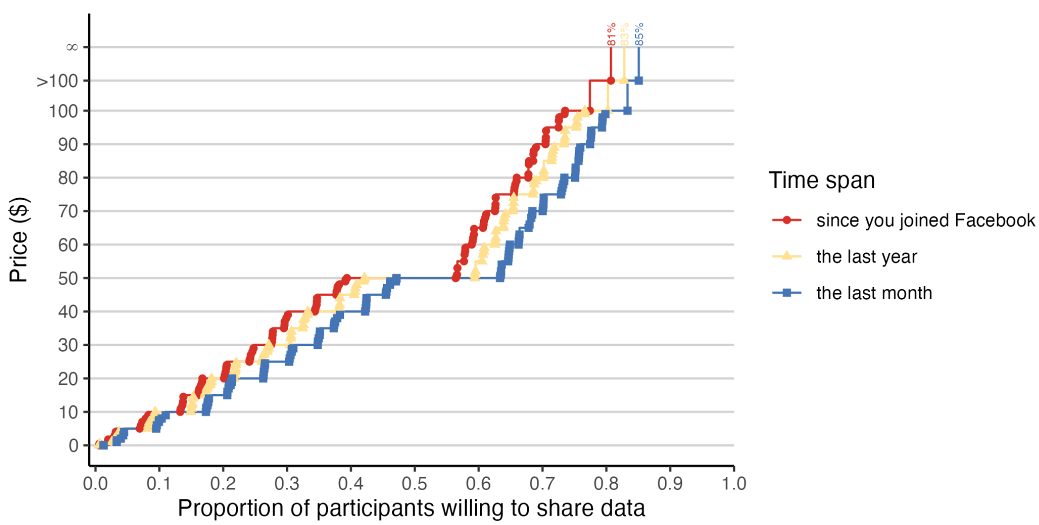


Figure D.2: Data Supply Curves by Data Duration



D.2 Alternative Model Specifications

Table D.1: Treatment Effects with OLS Specification

	WTA	log(WTA)	WTA	log(WTA)
Intercept	60.438 ***	3.884 ***	48.375 ***	3.556 ***
	(0.723)	(0.022)	(0.760)	(0.024)
Price Anchor = Low	-16.522 ***	-0.536 ***	-16.522 ***	-0.536 ***
	(0.762)	(0.023)	(0.762)	(0.023)
Default = Active	1.763 +	0.042	1.763 +	0.042
	(0.927)	(0.029)	(0.927)	(0.029)
Default = Opt-in	4.185 ***	0.105 ***	4.185 ***	0.105 ***
	(0.929)	(0.029)	(0.929)	(0.029)
Likes			6.964 ***	0.226 ***
			(0.381)	(0.012)
About Me Page			14.833 ***	0.409 ***
			(0.421)	(0.013)
Posts			16.283 ***	0.446 ***
			(0.429)	(0.013)
Friends and Followers			22.236 ***	0.562 ***
			(0.467)	(0.014)
Num. Obs.	25140	25140	25140	25140
R2	0.060	0.071	0.111	0.108

Notes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The outcomes are top-coded at $100; standard errors are clustered at the participant level.

Table D.2: Treatment Effects with Additional Controls: Tobit Regressions

	WTA	WTA	WTA	log(WTA)	log(WTA)	log(WTA)	log(WTA)
Intercept	63.884 ***	58.839 ***	71.972 ***	4.371 ***	4.166 ***	4.779 ***	4.704 ***
	(0.876)	(0.802)	(1.048)	(0.041)	(0.039)	(0.051)	(0.065)
Price Anchor = Low	-16.112 ***	-4.917 ***	-16.155 ***	-0.372 ***	0.087 +	-0.374 ***	-0.278 ***
	(0.947)	(0.991)	(0.938)	(0.045)	(0.048)	(0.045)	(0.075)
Default = Active	2.377 *	1.930 +	2.607 *	0.118 *	0.100 *	0.130 *	0.143
	(1.139)	(1.029)	(1.131)	(0.054)	(0.051)	(0.054)	(0.091)
Default = Opt-in	5.178 ***	2.978 **	5.315 ***	0.222 ***	0.131 **	0.228 ***	0.296 **
	(1.147)	(1.026)	(1.135)	(0.055)	(0.051)	(0.054)	(0.091)
Practice round deviation		0.695 ***			0.028 ***		
		(0.021)			(0.001)		
Believes data is available (BDA)			-12.304 ***			-0.621 ***	-0.508 ***
			(0.861)			(0.043)	(0.075)
BDA x Price Anchor = Low							-0.144 +
							(0.085)
BDA x Default = Active							-0.022
							(0.104)
BDA x Default = Opt-in							-0.102
							(0.102)
Num.Obs.	25140	25140	25140	25140	25140	25140	25140

Notes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Outcome variables are top-coded at $100 in models using level WTA as the outcome and at $1000 in models with log WTA outcomes. "Practice round deviation" is the difference in reported valuation from a benchmark objective value in the pre-treatment training. "Believes data is available (BDA)" is an end-survey answer to whether the participant thinks a valued data type is available to advertisers or not. Standard errors are clustered at the participant level.

Table D.3: Average Treatment Effects and Data Valuation across Topcodes: Tobit Regressions

Top-code	DV: Free-Text valuation					
	100	250	385	500	750	1000
Intercept	3.553 ***	3.597 ***	3.607 ***	3.614 ***	3.625 ***	3.633 ***
	(0.024)	(0.033)	(0.035)	(0.037)	(0.040)	(0.042)
Price Anchor = Low	-0.540 ***	-0.462 ***	-0.434 ***	-0.417 ***	-0.392 ***	-0.374 ***
	(0.023)	(0.034)	(0.037)	(0.040)	(0.043)	(0.045)
Default = Active	0.042	0.084 *	0.096 *	0.102 *	0.111 *	0.118 *
	(0.029)	(0.042)	(0.045)	(0.048)	(0.051)	(0.054)
Default = Opt-in	0.105 ***	0.169 ***	0.186 ***	0.196 ***	0.210 ***	0.221 ***
	(0.029)	(0.042)	(0.046)	(0.048)	(0.052)	(0.055)
Likes	0.229 ***	0.338 ***	0.368 ***	0.387 ***	0.413 ***	0.433 ***
	(0.012)	(0.017)	(0.019)	(0.020)	(0.022)	(0.023)
About Me Page	0.412 ***	0.654 ***	0.720 ***	0.761 ***	0.822 ***	0.869 ***
	(0.013)	(0.020)	(0.022)	(0.023)	(0.026)	(0.027)
Posts	0.449 ***	0.716 ***	0.789 ***	0.833 ***	0.895 ***	0.944 ***
	(0.013)	(0.021)	(0.023)	(0.024)	(0.027)	(0.028)
Friends and Followers	0.564 ***	1.043 ***	1.169 ***	1.248 ***	1.365 ***	1.453 ***
	(0.014)	(0.024)	(0.027)	(0.029)	(0.032)	(0.034)
Num.Obs.	25140	25140	25140	25140	25140	25140

Notes: + $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The outcomes are log valuations with top-codes indicated in the first row; standard errors are clustered at the participant level.

D.3 Causal Forest Models Tuning and Evaluation

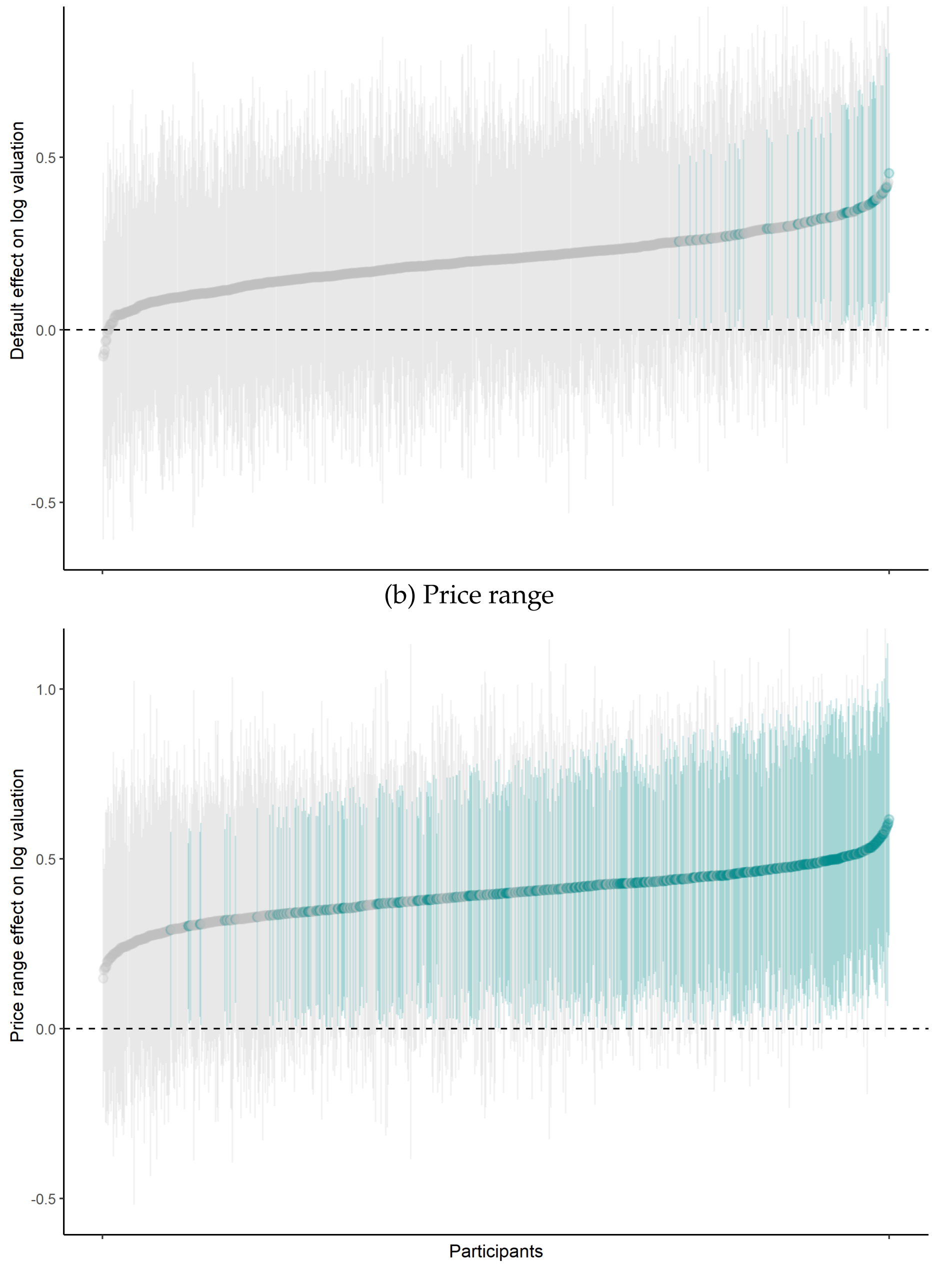
For our heterogeneity estimation, we use the ‘grf’ package to implement a variety of causal forest models. We cluster the standard error at the participant level, since privacy valuations from the same individual may be correlated. In practice, clustering means that the model estimation will sample all five valuations from an individual at a time to estimate its heterogeneity. As such, the model will only pick up valuation heterogeneity that is common across the five variables, and may underestimate the general degree of heterogeneity as a result (Athey & Wager 2019). Another implication of the clustering specification is that our forest model will not recover the bunching pattern in the raw data, because bunching is rarely universal across the data valuations for a given individual.

Ex ante, we suspect that poor choices of hyperparameters may lead to poor out-of-sample fit and degraded performance of the model. To guard against this possibility, we search for the best model across three parameter dimensions: number of trees, minimal node sizes, and maximum imbalance of a split. We evaluate the model fit using the R-loss (Nie & Wager 2021), which ensures that model fit is evaluated based on the treatment effect estimation rather than its performance in predicting the outcome. We find that the choice of hyperparameters does not substantially change the qualitative heterogeneity patterns, though the choices of ‘min.node.size’, ‘honest.fraction‘ and ‘honest.prune.leaves’ can substantially affect the R-loss. Our final model has ‘min.node.size’ = 3, ‘num.trees = 3500’, and the rest of the hyperparameters are consistent with the package default value.

D.4 Individual-Level Treatment Effect Estimates

Figure D.3: Individual-Level Treatment Effect Estimates

(a) Default option



(b) Price range

Notes: In the figures above, dots represent HTE point estimates and vertical lines represent 95% confidence intervals; significant estimates are shaded in green. Note that The standard errors for heterogeneous treatment effects are generally larger than for the average treatment effect. In addition, clustering standard error is known to mask heterogeneity patterns substantially when combined with causal forests (Athey & Wager 2019). Therefore, it is natural for default to have a significant treatment effect in Table 2 while having insignificant effects here.

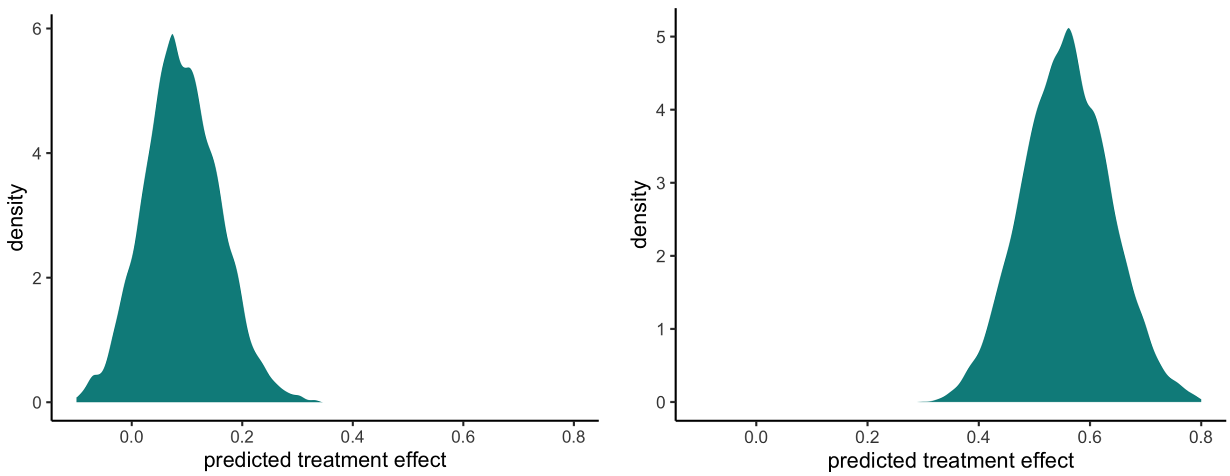
D.5 Comparison between Multi-Arm and Survival Forests

Since the multi-arm forest does not account for the fact that the latent censored values are greater than the censoring point, the magnitudes in its estimates are potentially biased compared to the survival forest. Figure D.4 shows that the estimated default effect distribution from the multi-arm causal forest is slightly biased upwards while the price anchor effect distribution is slightly biased downwards. Nevertheless, the magnitudes in the two models are similar, and the average treatment effects from both models are in line with estimates from our Tobit model.

Figure D.4: Heterogeneous Treatment Effect Estimates and Standard Errors: Survival Forests

(a) Default option

(b) Price range

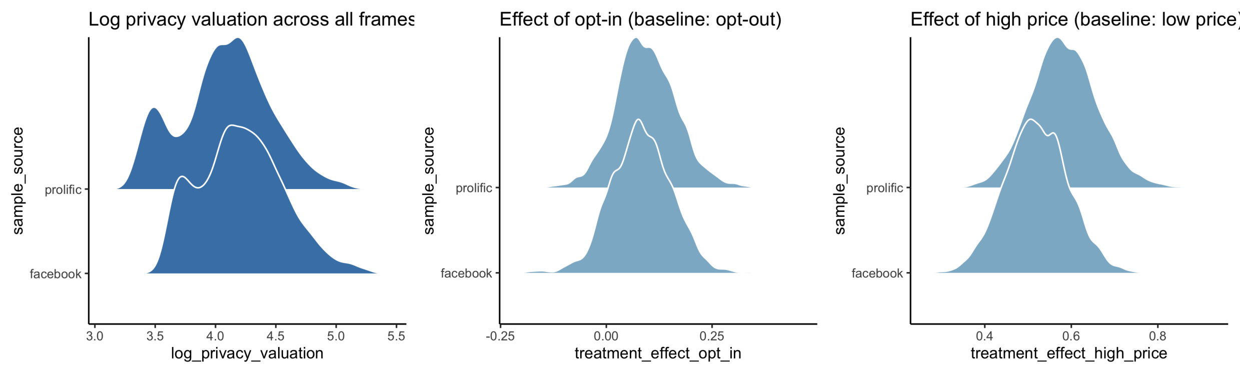


Notes: Outcomes: log valuations truncated at $1000; standard errors are clustered at the subject ID level. ATE estimates from the causal survival forest: ATE_{default} = 0.09 (se = 0.03); ATE_{default} = 0.56 (se = 0.03).

D.6 Treatment Effects by Sample Source

We separately compare the average log valuation and choice frame effect sizes from our two participant sources, which is shown in Figure D.5. The Facebook participants have a higher privacy valuation overall and is less influenced by frames. Although we do not know if the Facebook participants are more representative of the total population than the Prolific panel, our results suggest that studies using multiple participant sources have merit in assessing the robustness of the effects they aim to study.

Figure D.5: Average Log Valuation and Choice Frame Effect Distributions by Participant Source



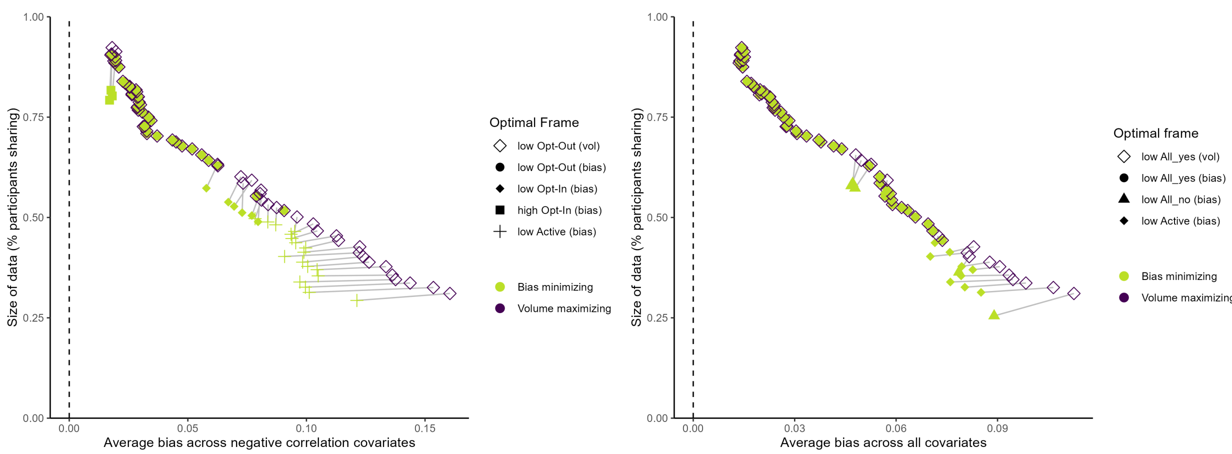
Notes: In these figures, the treatment effects are estimated using survival forests, with the outcome as log valuation truncated at $1000.

D.7 Frame Assignments and the Bias-Volume Trade-off

We provide an alternative approach of bias calculation for multiple covariates. In Figure 10d we reduce the dimensionality of multiple covariates to a single covariate index, and then proceed as if this is the covariate to achieve balance for. Here, we instead calculate the samples under different uniform frame and price combinations, find the bias for each standardized covariate, and then use the average bias as our measure. The results are shown in Figure D.6. Consistent with our prediction in Section 3, the gaps between the bias-min and vol-max frames are larger when we focus on only the covariates that exhibit negative correlation (see Panel a), and smaller when we include all covariates (Panel b).

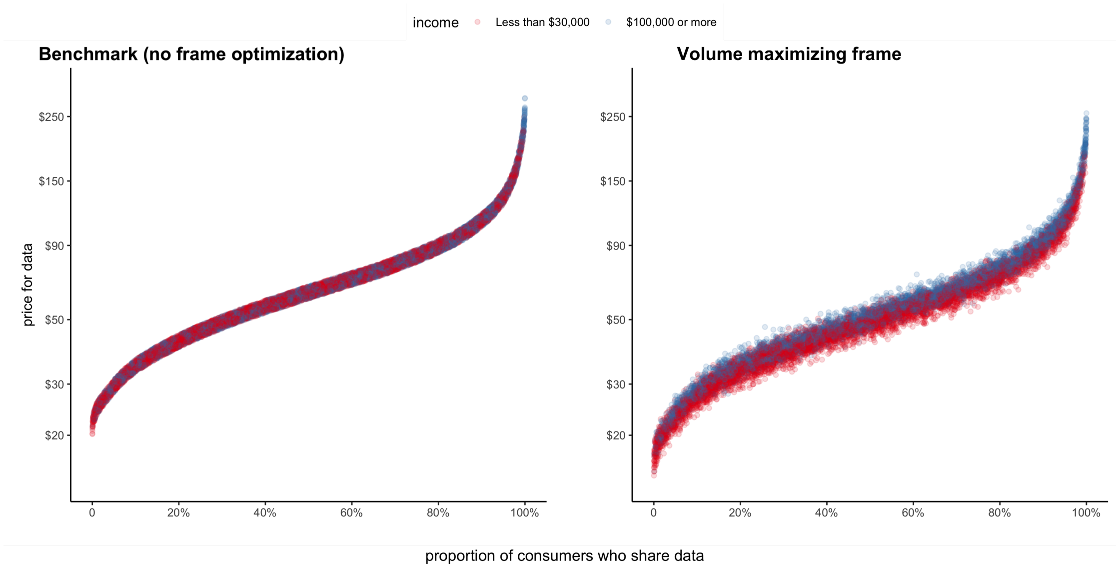
Figure D.6: Data Quality under Vol-Max and Bias-Min Frames for Individual Variables

(a) Negative correlation covariates (b) All covariates



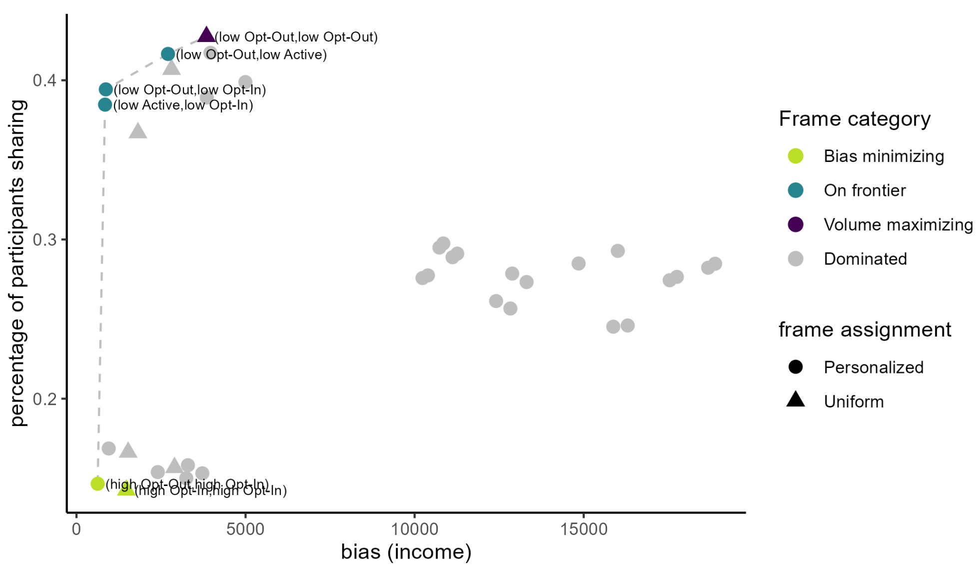
Notes: Each point represents a sample dataset collected under a choice frame × price combination. The prices for data vary from \$25 to \$90 in \$1 increments; points connected by a gray line are collected under the same price but different frames.

Figure D.7: Data Quality Comparison: No-Optimization Benchmark versus Vol-Max Frame



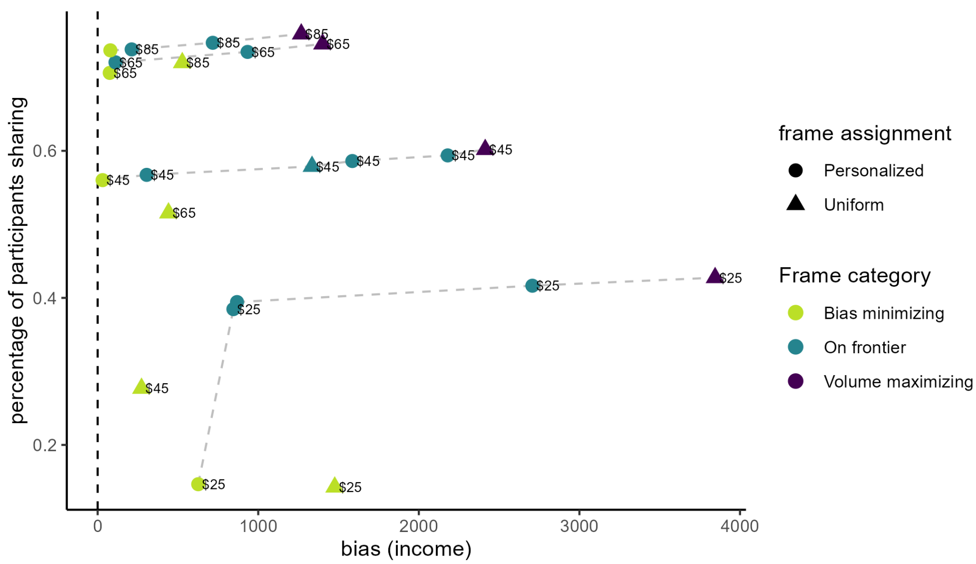
Notes: Each point here represents one consumer. In both figures, the X-axis reflects the percentile of valuations that a participant occupies in the "no-optimization benchmark" position. The left panel includes a 15% jitter to facilitate the exposition of consumer composition, while the right panel does not include jittering. Thus, the "fuzziness" of the curve on the right panel reflects the fact that the vol-max frame shifts the valuation ranking among some participants.

Figure D.8: The Volume-Bias Frontier (p = $25)



Notes: The figure shows the volume and bias of average income in the sample for all possible frame assignments with an offer price of $25. Consumers are divided into two groups: above- and below-median income. Each dot represents sample data under a different frame assignment. Triangles are uniform frames (both types see the same frame), and circles are cases in which each type is exposed to a different frame. The dark triangle is the volume-maximizing frame, and the light triangle and circle are the bias-minimizing uniform and personalized frame assignments, respectively. Other dark circles, connected by a line, represent the frames on the bias-volume frontier. The rest are dominated assignments: there are some frames that deliver higher volume and lower bias.

Figure D.9: The Volume-Bias Frontier (Multiple Price Points)



Notes: The figure shows the frontier frame assignments increasing volume and reducing bias of average income in the sample for four different prices ($25, $45, $65, and $85). Each dot represents sample data under a different frame assignment. Triangles are uniform frames (both types see the same frame), and circles are cases in which each type is exposed to a different frame. The Dark triangle is the volume maximizing frame, and the light triangle and circle are the bias minimizing uniform and personalized frame assignments, respectively. Other dark circles, connected by a line, represent the frontier of bias reducing and volume increasing frame assignment.

E Framework Proofs Appendix

E.1 Optimal Price Derivation

Here, we derive the results shown in Section 6.2, starting with the profit function and finding the optimal price:

$$\Pi(\theta, p) = V\left(N\left(\theta, p\right), B\left(\vec{N}\left(\theta, p\right)\right)\right) - pN.$$

To find the optimal price that maximizes profit, we take the first order condition (sufficient given the concavity assumptions about V):

$$\frac{\partial \Pi}{\partial p} = \frac{\partial V}{\partial N}\frac{dN}{dp} + \frac{\partial V}{\partial B}\frac{dB}{dp} - N = 0.$$

We use ϵ_p^N and ϵ_p^B to represent the elasticities of N and B with respect to price, respectively. We can rewrite the above as

$$\frac{\partial V}{\partial N}\frac{N}{p_\theta}\epsilon_p^N + \frac{\partial V}{\partial B}\frac{B}{p_\theta}\epsilon_p^B - N = 0.$$

Rearranging the equation above gives us an implicit equation for the optimal price (given a specific frame θ):

$$\begin{aligned} p_\theta &= \frac{\partial V}{\partial N}\epsilon_p^N + \frac{\partial V}{\partial B}\frac{B}{N}\epsilon_p^B \\ &= \left(\eta_N^V \epsilon_p^N + \eta_B^V \epsilon_p^B\right)\bar{V}, \end{aligned}$$

where η_N^V is the elasticity of value to volume, η_B^V is the elasticity of value to bias, and $\bar{V} = V/N$ is the average value per customer. Similarly, ϵ_p^N and ϵ_p^B are the elasticities of N and B with respect to price, respectively.

E.2 Sufficient Condition Approximation

We start with the sufficient condition for the volume-maximizing frame to be worse than the bias-minimizing frame, fixing the price as the optimal one for the volume-maximizing frame:

$$\mathrm{V}\big(N(p_N^*, \theta_{maxN}), B(p_N^*, \theta_{maxN})\big) - V\left(N(p_N^*, \theta_{minB}), B(p_N^*, \theta_{minB})\right) - p_N^*\left(N(p_N^*, \theta_{maxN}) - N(p_N^*, \theta_{minB})\right) < 0.$$

First-order approximation gives

$$\frac{\partial V\left(p_N^*, \theta_{maxN}\right)}{\partial N}\left(N(p_N^*, \theta_{maxN}) - N(p_N^*, \theta_{minB})\right) + \frac{\partial V\left(p_N^*, \theta_{maxN}\right)}{\partial B}\left(B(p_N^*, \theta_{maxN}) - B(p_N^*, \theta_{minB})\right)$$
$$- p_N^*\left(N(p_N^*, \theta_{maxN}) - N(p_N^*, \theta_{minB})\right) < 0,$$

which is a sufficient but not necessary condition for the bias-minimizing frame to yield higher profit than the volume-maximizing frame. A simpler way to express the inequality is

$$\left(\frac{\partial V_{maxN}}{\partial N} - p_N^*\right)\Delta N + \frac{\partial V_{maxN}}{\partial B}\Delta B < 0,$$

or equivalently,

$$\left(\eta_N^V - \frac{p_N^*}{\bar{V}}\right)\frac{\Delta N}{N} + \eta_B^V\frac{\Delta B}{B} < 0.$$

Here, $\frac{\Delta N}{N}$ is the relative change in N due to moving from the volume maximizing frame to the bias minimizing frame, estimated at the volume maximizing optimal price; similarly for $\frac{\Delta B}{B}$. Using the optimal price formula from above to substitute $\frac{p_N^*}{\bar{V}}$, we get

$$\left(\eta_N^V(1 - \epsilon_p^N) - \eta_B^V\epsilon_p^B\right)\frac{\Delta N}{N} + \eta_B^V\frac{\Delta B}{B} < 0.$$

Since η_B^V is always negative and $\frac{\Delta N}{N}$ is negative when switching from the vol-max to other frame-price pairs, dividing the inequality above with $|\eta_B^V|$ and $\frac{\Delta N}{N}$ gives the following:

$$\frac{\eta_N^V}{|\eta_B^V|}(1 - \epsilon_p^N) + \epsilon_p^B < \frac{\%\Delta B}{\%\Delta N} \equiv \gamma_N^B,$$

where γ_N^B is the arc elasticity of bias to volume, representing the ratio between the percent change in bias and the percent change in volume due to switching the frame from volume maximizing to bias minimizing.

E.3 Firm's Choice Architecture Optimization: Example Data Value Functions

In this section, we provide examples of data value as a function of both its volume and representativeness, drawing inspiration from the existing literature.

Efforts to model data value functions in the existing literature take one of three approaches. The first one explicitly characterizes data as an input to a firm's production function, common in macroeconomics (Jones & Tonetti 2020) and finance literature (Farboodi & Veldkamp 2023, Abis & Veldkamp 2024, Veldkamp & Chung 2024). Existing papers using this approach do not explicitly acknowledge the value of having representative data. Thus, we adapt the production function to explicitly acknowledge the value of having less bias for knowledge production:

$$V(N,B) = A \cdot N^{\eta_1} B^{\eta_2},$$

where $A > 0, 0 < \eta_1 < 1,$ and $\eta_2 < 0$. Most work in this literature acknowledges data (volume) has decreasing returns to scale, thus $0 < \eta_1 < 1$ is uncontroversial.

The second approach models data value either as improving the precision of the firm's belief under a Bayesian framework (Ichihashi 2020, Liang & Madsen 2020), or as decreasing the mean squared error of a key parameter that matters for a firm's decision-making (Acemoglu et al. 2022, Miklós-Thal et al. 2024). For example, imagine the firm has a prior on the optimal action P (e.g., pricing policy, advertising frequency) to maximize its profits. The prior has mean u_0 and precision $1/\sigma_0^2$. Data collected by the firm serve as an informative signal to update its posterior. Suppose the ground truth is u^* (unobservable to the firm). We can characterize the signal extracted from the data as having a mean $u^* + B$ and precision N/σ^2. Further, assume both the prior and the signal are normally distributed. Then the firm's posterior takes the following form:

$$\text{Posterior mean} = \frac{\frac{u_0}{\sigma_0^2} + \frac{u^*+B}{\sigma^2/N}}{\frac{1}{\sigma_0^2} + \frac{1}{\sigma^2/N}};$$

$$\text{Posterior precision} = \frac{1}{\sigma_0^2} + \frac{1}{\sigma^2/N}.$$

With a diffuse prior ($\sigma_0 = \infty$), the mean and precision above further reduce to $u^* + B$ and N/σ^2, respectively, consistent with a frequentist formulation of estimator mean and variance. The mean squared error of the firm's estimate is thus $B^2 + \frac{\sigma^2}{N}$. Theory work often assumes that the firm's profit loss is proportional to the mean squared error of this key estimate. Thus we can write the data value function as

$$V(N,B) = -A \cdot (B^2 + \frac{\sigma^2}{N}),$$

where $A > 0$ is a scaling constant.

Lastly, the literature on sampling research (see Cochran et al. 1954 for a review) explicitly characterizes the value of having representative sample data to obtain accurate insights while acknowledging the value of adequate sample size in reducing variance; work on sample recruitment reflects the focus on balancing representativeness and sample size goals to minimize the mean squared error in estimates. In this last example, we introduce the framework that originated from the seminal work by Neyman (1934), where the firm is able to construct an unbiased estimator via sample reweighting even when facing a biased sample.[28] In particular, suppose the firm intends to learn the average preference of its consumers for a new product, denoted as $\overline{w}$. The unbiased

[28]In contrast, in our second example we assume the firm has no means to correct for sample bias when estimating the parameter; the first example is agnostic about this possibility.

estimator via sample reweighting is

$$\widetilde{w} = \sum_{j=1}^{J} f_j \overline{w_j},$$

where $\overline{w_j} \equiv \frac{\sum_i =1^{N_j} w_i}{N_j}$ is the mean preference among consumers in group j, N_j is the number of group j consumers in the sample, and f_j is the proportion of group j consumers in the target population. The bias of this estimator is zero,[29] thus the mean squared error equals the variance of the estimator:

$$MSE(\widetilde{w}) = Var(\widetilde{w}) = \sum_{j=1}^{J} f_j^2 \frac{\sigma_j^2}{N_j}.$$

In the special case where σ_j is homogeneous across groups, Neyman (1934) shows the optimal sample recruiting scheme (subject to a fixed total sample size constraint) is one where the sample is representative, i.e.,

$$\frac{N_j}{\sum_{j=1}^{J} N_j} = f_j, \forall j.$$

More generally, for any given sampling scheme, the MSE can now be expressed as $\sum_{j=1}^{J} f_j^2 \frac{\sigma^2}{p_j N}$, where p_j is the proportion of group j consumers in the sample. Noting that $\sum_{j=1}^{J} \frac{f_j^2}{p_j} - 1$ corresponds to the chi-squared distance between the population and sample distribution, we can define $B = \sum_{j=1}^{J} \frac{f_j^2}{p_j}$ as our bias measure. In this case, the value function that minimizes MSE becomes

$$V(N, B) = -A\frac{B}{N}.$$

[29]Here, we follow the convention in optimal sampling research and ignore selection on unobservables.