

# Privacy-Seeking Behavior in the Personal Data Market

Joy Wu

Institute for Strategy, Technology and Organization,  
Ludwig-Maximilians-Universität München (LMU Munich),  
80539 Munich, BA, Germany  
joy.wu@lmu.de

**Version: November 5, 2021\***

**([Latest Version Here](#))**

Firms are looking to commercialize, trade, and monetize the personal data they collect and receive from consumers. Internet users regularly choose to disclose and share their personal data in return for goods and services. This study examines whether a data recipient’s ability to exploit data in a secondary market can motivate users’ privacy behavior. An online experiment elicited individuals’ willingness to share and reservation prices for sharing their personally-identifiable psychometric data when faced with real privacy consequences. I found that individuals’ information disclosure behaviors were misaligned with their willingness to allow data recipients to monetize their data and trade with a third party. Individuals behaved more privately—by refusing to share data or by demanding greater benefits in exchange for privacy losses—when they became more aware of a data recipient’s ability to sell their data for money. Moreover, when individuals considered allowing access to and exposing their data to many recipients, the privacy responses were weaker than the responses to just *one* recipient’s exploitation abilities.

*Key words:* data privacy; privacy valuation; psychometric data; data markets; economic experiment

---

\*This project was funded by a seed grant from Cornell University’s Institute for the Social Sciences. The study received IRB approval from Cornell University and was pre-registered at the American Economic Association’s registry for randomized controlled trials (AEARCTR-0004005). Special thanks to Aija Leiponen, David Just, Vicki Bogan, Ted O’Donoghue, and Chris Forman for their encouraging guidance. I also thank participants of the Cornell Behavioral Economics Research Group; graduate student seminars in Applied Economics and Management at Cornell; Consortium on Competitiveness and Cooperation Doctoral Conference; Innovation, Entrepreneurship, and Technology Brownbag at Cornell; the Institute for Strategy, Technology and Organization at LMU Munich; the Applied Economics and Policy Seminar at Cornell; the Technology and Innovation Management Group at ETH Zürich; the 2021 Academy of Management Conference; Munich TIME Colloquium; and 2021 DRUID Conference for helpful comments. All errors are my own.

## 1. Introduction

Digitization has revolutionized commerce in information markets, including the market for individual information. Much of the digital economy is financed and facilitated by user-generated data—often personally or uniquely identifiable—containing information related to individuals’ behaviors, intentions, and attributes. Users of various digital technologies passively and actively generate personal data in return for goods and services. Digital footprints of consumers are collected, curated, stored, and verified at very low marginal costs (Goldfarb and Tucker 2019). Previously, firms have been mostly focused on how to analyze these data and exploit them internally. More recently, firms are looking towards opportunities to externally exploit user data and trade them in digital ecosystems (Thomas and Leiponen 2016). Average people are often excluded from this secondary data market.

There are some prominent, recent examples of personal data exploitation in secondary markets. In 2018, the *Cambridge Analytical* scandal involved a personality survey hosted on the Facebook platform that was used to curate users’ psychometric data (Graham-Harrison and Cadwalladr 2018). The information contained in these data was used to predict the marginal voters in the 2016 U.S. presidential campaign, and even targeted those voters with persuasive information to induce shifts in their voting behavior.<sup>1</sup> In 2020, a new firm *Clearview AI* had harvested user-published facial images to fuel their highly successful facial recognition technology (Hill 2020). Both of these recent cases demonstrate how the exploitation of user-generated data can be commercially valuable, even motivating public discourse on regulating secondary data markets. Interestingly, both cases also involve data—while “personal”—previously disclosed or published by the individual and, therefore, exist beyond the boundary of what economists would normally consider as information in a *private* state. Whether individuals actually want privacy regimes that grant them control over who can access and use these categories of data is still an open question.

The idea that secondary market activities may incur privacy costs is not new. Varian (1996) explained that extrinsic, nuisance costs to the individual can occur as a result of secondary transactions, because there is a clear externality that arises from the misalignment of the individual and third-party interests. While the costs incurred from third-party data access can potentially motivate individuals’ privacy behavior, this paper examines an adjacent motivator: the *second party*’s ability to benefit from an external trade with third parties. My paper is the first, to my knowledge, to identify individuals’ privacy responses to a second party’s data monetization potential in external markets.

<sup>1</sup> The underlying psychometric information was based on the Five Factor model, and these data are effective at predicting and even influencing people’s behavior (Matz et al. 2017).

Using an economic experiment, I test whether increasing the salience about a second party's data exploitation abilities causes individuals to be more privacy-seeking and demand greater benefits in exchange for privacy-losses. To capture relative privacy valuations, the experiment compares users' privacy-seeking behavior in response to various information provisions about the conditions for releasing their data to a second party. This study specifically evaluates willingness-to-share *personally-identifiable psychometric data* in the form of Five Factor scores (i.e., the same data exploited by *Cambridge Analytica*) in return for monetary incentives. The primary treatment condition includes explicit information about the recipient's ability to monetize data by selling to third parties. As a qualitative benchmark for understanding the magnitude of these privacy responses, a second condition stipulates that data would be released to many recipients. Various secondary treatments are tested to examine more fully these privacy responses to data exploitation. These include provisions that exclude information about sales to third parties; include many recipients who can exploit data; and defines the "third party" as another person recruited for the study.

Throughout my results, I find consistent evidence that information signals about the second party's ability to exploit data motivates privacy-seeking behavior—both in their reluctance to participate in a data market (i.e., selecting out) and in greater minimum prices (i.e., demanding more benefits) for releasing their data. These privacy responses are also stronger than the condition where data would be released to many data recipients. There are several conclusions from these findings.

First, privacy preferences are not fully revealed in any one disclosure choice. Even after individuals disclose their information in a first-stage survey to generate their psychometric data, many exhibit privacy-seeking behavior when deciding whether to share that data with additional parties. In fact, 21% of individuals reject the maximum possible price \$2.99 and choose to *not* participate in the data market (absent any information provisions about that recipient's exploitation abilities).

Second, the individual's taste for privacy is not independent from a data recipient's ability benefit from the data transfer. Increased salience about a recipient's exploitation abilities decreases the odds of participating in the data market by 2.8 times. Variants of this main information treatment find that aversion to both the recipient's ability to make money and sale to third parties separately contribute to individuals' exploitation-aversion. Moreover, this privacy-seeking response to a recipient's exploitation signals is robust to replacing the "third party" with another person in the study.

Third, in the category of psychometric data, individuals are relatively insensitive to increased exposure to many second party data recipients. Compared to data-sharing with *thirty* recipients, their odds of exiting the data market were 1.8 times greater when treated with information about just *one* data recipient's exploitation abilities. In the treatment variant where the third party

is contained in the study, I effectively find that two recipients with salient exploitation signals motivate stronger privacy responses than exposure to 30 recipients with opaque signals

These findings contribute to our understanding of data privacy preferences in three ways. First, this study demonstrates that information disclosure behavior can be an unreliable measure of downstream data-sharing preferences. The second finding provides a new perspective on the privacy costs from secondary data market activities and third party data usage, by demonstrating evidence that individuals' privacy behavior responds to the *second party's ability to benefit* from data exploitation in external markets. Finally, the third finding offers a valuable comparison between tastes for exposure versus exploitation, and finds that an aversion to others experiencing one's personal information can be less motivating for privacy behavior in this category of commercially valuable data.

## 2. Related Literature

Traditional models of privacy support a theory of individuals acting strategically in information markets and calculating all the costs and benefits to data-sharing, suggesting that privacy regulations are inefficient for markets (Stigler 1980, Posner 1981). Following the "privacy calculus" framework from Laufer and Wolfe (1977), the privacy literature has found a wide range of concerns that enter into the cost-benefit analysis of each disclosure decision (Culnan and Armstrong 1999, Dinev and Hart 2006).

In addition, a behavioral economic perspective on privacy decision-making has emerged, which finds that disclosure choices respond to non-normative, environmental factors, including choice architecture, framing, perceptions of control, and contextual cues (John et al. 2010, Brandimarte et al. 2012, Acquisti et al. 2013, Adjerid et al. 2019). This paper supports and extends these applications of behavioral economics in measuring privacy preferences. A privacy response to a data recipient's private gain is—although intuitively reasonable—not economically obvious. A rational economic model would assume that the private benefits to the potential buyer should be independent of the individual's minimum willingness-to-share their personal data. My paper contributes to behavioral economic perspectives by finding evidence—within the domain of personal data-sharing behavior—that contradicts this classical assumption.

Empirical research has documented an association between individuals' privacy-seeking attitudes and the secondary-use of information (Culnan 1993, Angst and Agarwal 2009, Sutanto et al. 2013).<sup>2</sup> This study will further our understanding of secondary data markets beyond these studies in two

<sup>2</sup> In Culnan (1993), those less concerned (or with positive attitudes) related to secondary-use were correlated with less concern about other privacy features, including control over personal information access and the nuisance of privacy invasions. Sutanto et al. (2013) found that an information technology solution that prevents third-party data-sharing reduced perceived privacy intrusions.

ways. First, I focus on and measure privacy costs associated with the exploitation abilities of the second party (i.e., the data recipient) in the secondary market, rather than specific risks associated with third party uses. Second, the economic experiment in this study will elicit revealed, incentive-compatible behavioral responses to real privacy consequences. As a field experimental study by [Athey et al. \(2017\)](#) has demonstrated, it is challenging to use survey and stated attitudes about data privacy, as these observations may be slanted from the privacy regime individuals may actually want.

Relying on disclosure behavior is also not without its challenges. Research has argued that there is instability to valuations people place on their privacy ([Acquisti et al. 2015](#), [Adjerid et al. 2013](#)). Most of this instability is attributed to a lack of awareness and incomplete information users have about disclosure outcomes ([Acquisti and Grossklags 2005a](#)). My study finds upstream disclosure behavior can be misaligned with willingness-to-participate in downstream data markets, which support this understanding that the estimated value—either implicit or explicit—people place on their personal information are uncertain, unstable, and prone to misdirection.

To correct for individuals' lack of awareness about data-sharing consequences, policy-oriented research study the effects of notice and consent policies ([Athey et al. 2017](#), [Acquisti et al. 2016](#), [Tsai et al. 2011](#)). However, researchers are still challenged with finding what precise features or consequences of secondary data markets can actually motivate users' privacy behavior. Perhaps the closest study related to mine is [Buckman et al. \(2019\)](#), where the authors experimentally elicited privacy valuations that were treated with more salient, negative risks and consequences of personal data disclosure related to third party access. They found statistically indistinguishable changes to prices subjects were willing-to-accept in return for privacy losses. Deviating from past notice-and-choice interventions in the literature, my study finds a novel provision related to the recipient's private benefit of users' data—which I find is not independent from users' reservation prices for sharing their data. My study deepens and extends our knowledge by finding that decision-makers do reveal privacy preferences that are sensitive to the exploitation abilities of data recipients in these secondary markets, and exhibit relatively weak privacy responses to data access by and exposure to more parties.

Research has found monetary rewards to be effective in motivating information disclosure ([Preibusch 2015](#), [Xu et al. 2010](#), [Hui et al. 2007](#)). However, observed values vary widely. A price elicitation can be challenging to implement in privacy choice settings, where individuals have no prior experience with data prices. My work contributes in improving methodology for eliciting privacy valuations in three ways. First, as shown in the [Buckman et al. \(2019\)](#) experiment, a posthoc analysis discovered suggestive evidence of individuals seeking to “price out” of the market after

perceiving a greater risk of third-party data access. In order to capture this “all-or-nothing” decision, my elicitation method allows for a choice *to participate* in a data market in addition *to price* data conditional on participation. Second, my price list reflects typical amounts seen in digital platforms that trade “free goods” in return for consumer data. Finally, the coarseness of my price list is effectively a small collection of binary decisions, which removes the need for individuals to form sophisticated point estimates about the value of their data.

In summary, this paper will contribute to the landscape of privacy decision-making research in several important areas. First, I offer a behavioral perspective on how valuations of one’s personal data are non-independent from the data exploitation abilities of the data recipient. My results show the importance of this factor in privacy choices, especially on the decision to enter or exit the data market, as well as in comparison to exposing one’s data to many data recipients. Second, I demonstrate how disclosure behavior is unreliable in approximating individuals’ willingness to forgo privacy protections and control over their data in secondary markets. Similar to [Buckman et al. \(2019\)](#), I address the prior literature’s limitations by measuring privacy valuations by adopting a price elicitation methodology from experimental economics. However, I extend the basic elicitation design, allowing for “all-or-nothing” privacy choices in addition to pricing decisions about personal data (details are provided in Section 4.4). Finally, I provide empirical measures of privacy preferences over commercially-valuable psychometric data, which deviates from the more standard demographic and identifier categories of personal data in the prior literature (see a longer discussion in Section 2.1).

## 2.1. The Unique Privacy and Exploitation Concerns of Psychometric Data

This study examines individuals’ privacy responses to the market value of personally-identifiable psychometric information. From the individual side, personal (or unique) identification is a required feature of user-generated data that invokes privacy-seeking behavior ([Glasgow and Butler 2017](#)). On the other hand, psychometric data—as an appendage to personal identifiers—have the potential to be extremely valuable in commercial settings.

Prior studies have examined individuals’ disclosure behavior of identifying, tracking, or sensitive contents of personal information (e.g., email addresses, shopping behaviors, medical history) ([John et al. 2010](#), [Acquisti et al. 2013](#), [Athey et al. 2017](#), [Buckman et al. 2019](#)). Therefore, it is not immediately obvious that psychometric data is important for studying privacy concerns. On the spectrum of information secrecy, Five Factor personality information is rather neutral and, arguably, observable to others (i.e., close friends and colleagues can have good estimates of one’s personality score). In fact, universities and workplace settings easily and often elicit responses to Five Factor surveys, simply because individuals benefit from assessments of their personality. However, from the perspective of party that is interested in exploiting data internally or externally,

the value of consumer data is (1) not only a function of the ability to identify individuals and (2) not necessarily correlated with the degree of secrecy or (social) sensitivity of the content.<sup>3</sup> While the inclusion of personal identifiers is characteristic of personal data, the ability to uniquely or personally identify may be becoming less valuable. As consumer data becomes more vast and interconnected, with more sophisticated search and verification technologies, identification and identifiers may become easier—and less costly—to obtain. For example, [Acquisti and Gross \(2009\)](#) show how social security numbers can be predicted from publicly available data.

Psychometric data based on the Five Factor model has been shown in psychology to have the ability to understand, predict, and discriminate attitudes and behaviors among individuals in real-life outcomes ([Goldberg 1992](#), [McCrae and John 1992](#), [Matz et al. 2017](#)). The public-domain International Personality Item Pool (IPIP) for administering the Five Factor personality measurement has been pervasively used across the world for various research and assessment purposes ([Goldberg et al. 2006](#))—which contributes to its predictive and persuasive power over human behavior and attitudes. Similarly, [Miller and Tucker \(2018\)](#) make this important characterization about the predictive power of a person’s genetic data on future health risks. These data (unlike cookie data or email addresses) contain information about other behaviors and traits that have consequences for long-term welfare. One of the most vivid, exploitative uses of psychometric data in recent memory was, of course, conducted by *Cambridge Analytica* in targeting and persuading the behavior of marginal voters in the 2016 United States presidential campaign ([Graham-Harrison and Cadwaladr 2018](#)). For these reasons, psychometric data is a unique category of data for studying privacy preferences and their interaction with the data exploitation abilities of parties that collect these data.

### 3. Framework and Hypotheses

Classical perspectives on privacy economics theorize that data markets can sufficiently trade-off between the economic benefits of using consumer data and the privacy concerns of consumers ([Posner 1981](#), [Stigler 1980](#)). This trade-off is described in [Acquisti et al. \(2013\)](#) and then in [Buckman et al. \(2019\)](#) as an individual’s utility over wealth and privacy:  $u(w, p)$ . Suppose the consumer with  $p^+$  amounts of privacy is considering to enter a state with  $p^- < p^+$  amounts of privacy. In order for the decision-maker to agree to this change, she needs to receive at least the minimum amount in benefits (i.e., reservation price)  $r$ , where  $u(w + r, p^-) = u(w, p^+)$ .

<sup>3</sup> Consider the example of a data set linking individuals’ social security numbers with their names. In an identity theft, these social security numbers have high exploitation value. However, to a firm that seeks to forecast individual-level online shopping behaviors, the randomness of how social security numbers are generated holds little explanatory power. Personal identifiers contain information valuable for surveillance and communications targeting but hold fewer dimensions of analytical power than behavioral and attitudinal data that can be linked to those identifiers.

The individual’s ability to determine what benefits,  $r$ , they would accept in return for privacy losses requires an ability to account for all costs to one’s disclosure behavior. This is an unrealistic expectation: People often have an incomplete and limited capacity for attention when it comes to understanding the consequences of sharing personal information and the reason behind why their information is being collected (Acquisti and Grossklags 2005b). Asymmetric information is the most obvious challenge to privacy choices—hence, much of the empirical privacy research studies privacy responses under various “notice and choice” regimes.

Lack of awareness is a uniquely challenging issue for personal data markets. Consequences may result from failed-to-imagine scenarios and incorrect presumptions about data property and exclusionary rights. As an example, facial data shared by users from decades ago—when facial recognition technology was relatively unknown—are now being harvested for commercial use by *Clearview AI* (Hill 2020). Moreover, the complexity and novelty of data markets, the abstractness of information goods property, and the inalienability of personal data to the individual are all considerations that exacerbate the salience challenges to individuals’ privacy decisions (Koutroumpis et al. 2019, 2020).

In most data markets individuals cannot re-appropriate information they have shared.<sup>4</sup> Therefore, the user’s most consequential privacy choice is often an initial decision to disclose non-digital information, which then generates data. When generating personal data, individuals are often faced with tangible, immediate benefits in return. Behavioral economics research has popularized individuals’ tendencies to over-weight payoffs closer to the present time (O’Donoghue and Rabin 1999). If individuals myopically focus on immediate benefits to information-sharing while ignoring the less vivid downstream outcomes of data trading, then they do not accurately reveal their willingness-to-accept for all—and especially more opaque—privacy consequences.

Given all these obstacles towards perfectly informed choices, observing disclosure behaviors in isolation likely leads to poor measures of the individual’s true willingness-to-accept for data-sharing in *all situations* (i.e., a full transfer or loss of data “ownership”), especially for downstream markets where privacy costs are opaque at the time of information disclosure. Therefore, my first prediction is that individuals’ disclosure behavior is unreliable in revealing their privacy preferences.

**HYPOTHESIS 1.** *Individuals’ upstream information disclosure choices are misaligned with their willingness-to-share personal data in downstream markets.*

An additional challenge arises in choosing which consumer “notices” should be provisioned to support individuals’ privacy decision-making. Even if it is possible to provision the entire universe

<sup>4</sup> Even under the European General Data Protection Regulation (GDPR) provisions on individual control rights, the non-excludable nature of data makes the enforcement of erasure rights to personal data difficult and costly to both individuals and firms.



of potentially relevant privacy information to the decision-maker, this only swings the pendulum back to salience issues related to limited capacities for attention. Therefore, privacy researchers are tasked with finding the most relevant information provisions to support individuals' privacy calculus—that is, information that can meaningfully inform and motivate their privacy choices.

In this paper, I examine the effect of two potential influences on the individuals' privacy valuation: (1) exposure to more data recipients and (2) an increase in the salience of exploitation by data recipients. The representative decision-maker of this study operates in a world in which her data contains commercial value for others. What motivates her privacy behavior may be determined by both anti-surveillance and anti-exploitation preferences. First, she considers her taste for the others experiencing—or “seeing”—this data. Second, she contemplates her taste for others exploiting—or “making money” from—this data in a secondary market. As examples, the number of data recipients who can access her data increases exposure, and information about her recipient's ability to earn profits after obtaining data can be a relevant signal for her exploitation tolerance.

To organize and demonstrate how privacy behavior can change in response to exposure and exploitation, I adopt a framework for inattention from DellaVigna (2009, p. 349). First, to follow the previous notation, suppose some amount of utility the individual loses from data-sharing is  $V \leq u(w, p^+) - u(w, p^-)$ . Consider this amount of utility (loss) to be made up of two components:  $V = v(e) + o(e)$ . The first component  $v$  is this individual's taste for others' experiencing her personal information. The second component  $o$  is this individual's taste for others exploiting her information. Both components are a function of  $e \geq 0$ , representing the exposure (e.g., number of data recipients) in data-sharing.<sup>5</sup> Now, suppose that the individual is inattentive over the  $o$  component (e.g., how her recipients can monetize her data in a secondary market). Then, she perceives  $\hat{V} = v(e) + [1 - \theta(s)] o(e)$ . Here,  $\theta$  is the inattention parameter as a function of salience  $s \in [0, 1]$  of  $o$ . Assuming  $\theta'(s) < 0$ ,  $\theta(1) = 0$  is full awareness with a fully salient signal, and  $\theta(0) = 1$  is complete blindness with no salient signal (which follows psychological theory that information attention is non-decreasing in salient signals).

Seminal works theorizing the existence and origins of the intrinsic value of privacy would theorize—practically, by definition—that the value of privacy is determined by minimizing one's personal information from being exposed to others (Westin 1967). However, few studies have empirically documented a relationship between the value of privacy and increased exposure to more parties who can experience the information in one's personal data. In one study I am aware of, Schudy and Utikal (2017) experimentally examined sharing personal addresses with varying numbers of data recipients, and they found that willingness-to-share decreased with exposure to more anonymous recipients.

<sup>5</sup> Note that this can generalize to cases where there is positive utility gained from data exposure and exploitation.

In order to empirically document privacy responses to data exposure (for at least the category of psychometric data examined in this study), the experiment first measures privacy responses to exogenous manipulations of exposure  $e$ —specifically, in the number of data recipients the individual is considering sharing data with. Thus, if changes in this number influences the individual’s taste for recipients experiencing or exploiting her data, this effects her predicted portion of her privacy loss,  $\hat{V}$ , and she will reveal this in her reservation price  $r$  for sharing data.

**HYPOTHESIS 2.** *Individuals’ privacy-seeking behavior over their personally-identifiable psychometric data is increasing in exposure to more data recipients.*

Next, the main treatment of this study varies information about the data recipient’s ability to exploit personal data for profit in a secondary market, which should influence the privacy decision-maker’s taste for the recipient exploiting her data and not her taste for the recipient experiencing her data. External data exploitation is now a widely considered and profitable digital business strategy among firms. Consumers in the modern, digital economy have evolved beyond aversions to spying and intrusion from an anti-surveillance era ([Westin 1967](#)). Moreover, data-based analytics have evolved privacy issues beyond the nuisance costs of unwanted solicitation theorized in [Varian \(1996\)](#), towards more targeted advertising and even digital mass persuasion ([Matz et al. 2017](#)). Today, some of the largest and most profitable digital companies are built on personal data. New entrants into data-based businesses such as *Clearview AI* have profited from the harvesting of user-published, digital goods.

There are two conclusions to be made when information provisions about how recipients can exploit personal data—i.e., increasing salience  $s$ —are shown to impact the individual’s privacy behavior. The first is that individuals lack full awareness about the value of their data to others. The second is that individuals are averse to a data recipient’s ability to monetize personal data, and—when data exploitation in secondary markets is not salient—their data-sharing behavior understates their value of privacy. If strengthening information  $s$  increases the individual’s awareness for data exploitation *and* she is averse to this data-sharing consequence, then her taste for data exploitation is revealed in her price for data-sharing.

**HYPOTHESIS 3.** *Privacy-seeking behavior over personal data is stronger when it becomes more salient to the individual that their data recipient is able to benefit from secondary data market exploitation.*

Finally, the influence of exposure and exploitation over privacy responses can be compared. This study designs one comparison case (i.e., one and thirty data recipients) for only one particular kind of data (i.e., psychometric data). This study will use the change in privacy choices under greater

exposure as a qualitative comparison for individuals’ aversion to data exploitation in secondary markets. However, the results of this comparison may vary depending the context of the decision, the type of data in question, and the amount of exposure considered.

## 4. Online Experiment

The online experiment was designed to simulate a personal data market where users generated their psychometric data and faced real decisions to share that data with others in return for benefits. It was conducted at a U.S. institution’s business school lab (hereafter the “Lab”). The Lab maintains an Institutional Review Board (IRB) approved subject pool for online and in-person studies. This study was advertised with the title “How well do you know yourself? An economic decision study” in order to avoid priming potential participants with the idea that the study was meant to examine privacy preferences.<sup>6</sup> In fact, the words “privacy” and “security” are not used for the entirety of the study, up until the exit survey questions related to privacy attitudes.<sup>7</sup> Also advertised was a minimum payment of \$2 in Amazon gift cards for a 15 minute study, with the possibility to earn more based on the survey taker’s decisions within the study.

The data was collected across three different samples, with a combined total of 1,188 participants in Spring 2019, Fall 2019, and Summer 2020.<sup>8</sup> The primary conditions, experimental paradigm, and survey recruitment and collection methods were replicated exactly for each sample collected. A set of three secondary experimental conditions varied by sample. A summary of the survey chronology is shown in Figure 4. The details of the experimental design are described in the proceeding sections.

### 4.1. Collecting Personal Identifiers

The Lab was well positioned to conduct this study, as the personal identifiers of subjects (e.g., names, emails) were available to the researcher prior to the survey launch and then attached to survey takers’ psychometric data during the experiment.<sup>9</sup> In general, the Lab monitors and removes users who register under aliases, click-through a survey without engaging with the content, or have a history of incomplete studies. The subject pool is primarily made-up of students affiliated

<sup>6</sup> The collection and intended use of respondents’ data, the data-sharing decisions subjects would be asked to make, and the real outcomes of data-sharing were approved by the University’s Institutional Review Board (IRB) prior to any interventions.

<sup>7</sup> For example, [Adjerid et al. \(2019\)](#) showed that individuals’ had a higher propensity towards privacy outcomes when prompted to make decisions about their “privacy” settings versus their “survey” or “app” settings.

<sup>8</sup> The the four-condition, within-subject design (including, but not limited to, the method of randomization, the primary and secondary outcome elicitation methods, prices, type of data being shared) was pre-registered with the AEA RCT Registry. Prior to samples collected subsequent to Sample 1, the pre-registration was updated with the newly anticipated total number of participants.

<sup>9</sup> In order to have been in the Lab’s subject pool and participated in any of its advertised studies, individuals needed to first register with names, email addresses, and answer some pre-screening questions. Surveys were distributed through personalized email links.

**Figure 1** Example Self-Assessment Questions and Responses.

	Very Inaccurate	Moderately Inaccurate	Neither Accurate Nor Inaccurate	Moderately Accurate	Very Accurate
I am relaxed most of the time.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
I leave my belongings around.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
I have difficulty understanding abstract ideas.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I pay attention to details.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
I keep in the background.	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

with the university, and many register with accurate personal identifiers in order to receive course credit.<sup>10</sup>

Other platforms with arguably more nationally representative samples could have been used (e.g., Amazon Mechanical Turk, Prolific) for this online study. However, none offered higher quality identifiers on their participants. Eliciting identifiers within the survey could have contaminated the experimental treatments. Subjects also received personalized survey links through Qualtrics to the email they registered with at the Lab. In these emails, the content was addressed to the subject’s name, so that surveyers were reasonably aware that the data generated from their personality assessment could be linked to their identity.

#### 4.2. User-Generated Psychometric Data

This experiment collected self-disclosed responses to a 50-item Five Factor questionnaire about the respondent’s attitudes, personality, and habits. This survey was taken from the standard sample of Likert-type assessment statements in the International Personality Item Pool (IPIP), which is widely used in psychology research. A small selection of the items are shown in Figure 1. Responses to these self-assessments generated each person’s Five Factor personality scores across the traits: Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Intellect. All self-assessment statements required a response from the participant. Skipped questions prevented the survey from continuing, and a failure to complete prevented the respondent from earning the \$2 participation earnings. After completing all 50 items in the self-assessment, respondents were presented with their scores and personal identifiers (see example in Figure 2) as well as information on how to interpret their scores.<sup>11</sup>

Similar to previous work measuring privacy valuations over personal data gathered within the study, untruthful responses could have occurred in the information people disclosed. A number of considerations were included in this design to minimize this occurrence. First, this self-assessment

<sup>10</sup> While this experiment only offered monetary payment, many other studies overlapping in time with this study offered course credits.

<sup>11</sup> Each score is an integer in the range of 10 to 50. A high score in Extraversion, for example, indicates high extraversion and low introversion.

**Figure 2 Example User-Generated, Psychometric Data.**

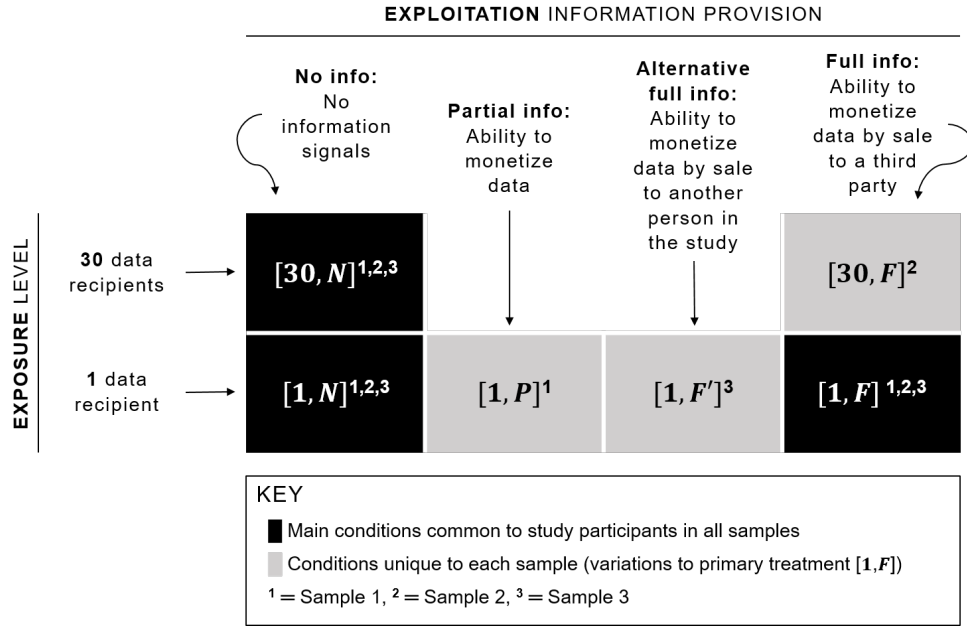
First Name	Last Name	Extraversion	Agreeableness	Conscientiousness	Emotional Stability	Intellect
<b>Jane</b>	<b>Doe</b>	<b>19</b>	<b>31</b>	<b>21</b>	<b>42</b>	<b>48</b>

occurred at the start of the study, without revealing to participants that the study intended to explicitly elicit their willingness-to-share data with other persons in the study. Subjects would have been reasonably, myopically focused on the immediate benefits of receiving a personality assessment. Second, the Lab is hosted in a business department, where MBA students regularly take Five Factor score assessments in their leadership courses. Third, endogenous motivations (e.g., “nothing to hide” or “something to hide”) of privacy behavior should be orthogonal to the experimental treatments in this study; certain scores are not more economically valuable than other scores and this indifference is made clear to subjects both prior to the self-assessment and prior to data-sharing decisions.<sup>12</sup> Fourth, it is unclear whether untruthful answers, if revealed, are more costly as a false representation of a person’s characteristics. Finally, and most compellingly, the IPIP surveys have been extremely successful in countless psychology research studies in eliciting accurate and discriminating personality measures that can predict the real behaviors and traits of individuals.

### 4.3. Data-Sharing Conditions

In the second stage of the survey, conditions were created to elicit and quantify individuals’ willingness-to-accept downstream data exposure and secondary market exploitation. There are six distinct conditions that vary the number of data recipients and the salience of a data recipient’s exploitation abilities. For the latter variation, there are either 1 or 30 data recipients. As shown in Figure 3, the six data-sharing conditions are denoted  $[1, N]$ ,  $[1, P]$ ,  $[1, F]$ ,  $[1, F']$ ,  $[30, N]$ , and  $[30, F]$ . All subjects made decisions under three conditions  $[1, N]$ ,  $[30, N]$ , and  $[1, F]$  (i.e., these conditions were replicated across all three samples). Their fourth data-sharing condition varied by sample. Each of the subjects’ four data-sharing decisions were randomly assigned to either their first, second, third, or fourth decision period.

<sup>12</sup> Prior to the self-assessment portion of the experiment, respondents are told, “Your responses for each statement will NOT determine your earnings in this study.” Prior to the data-sharing decisions, participants are not given any information that their scores determine the available prices or the exploitation abilities of data recipient(s).

**Figure 3 Six Experimental Conditions.**

A limit of four data-sharing decisions per subject was a feature of the design in order to reduce inattention and survey fatigue. On the other hand, replicating the experiment across three samples allowed one of the four conditions to be alternated out as a variation of an exploitation treatment.

Prior to any data-sharing decisions, the survey explained to the subject they would be considering “several” scenarios for sharing their data with others in the study. To make the decisions incentive-compatible, they were told that one of their data-sharing scenarios would be made real at the end of the study. One week following survey submission, the experimenter randomly selected data recipients to receive personal data from other study participants and bonus earnings from a hypothetical commercial data activity.<sup>13</sup>

#### 4.4. Eliciting Willingness-to-Share Data and Other Outcomes

Monetary rewards are not only appropriate for eliciting reservation values for sharing data, empirical privacy work has found monetary rewards to be effective in obtaining private information (Hui et al. 2007, Xu et al. 2010). For each data condition, decision-makers in the experiment considered five potential prices—\$0.01, \$0.49, \$0.99, \$1.99, \$2.99—to either accept or reject.<sup>14</sup> Following the Becker-DeGroot-Marschak (BDM) (Becker et al. 1964) incentive-compatible method for eliciting willingness-to-accept, the subjects were told one of the prices would be selected at random to be

<sup>13</sup> Subjects had a less than one percent chance of being selected as a data recipient who could earn money from others’ personal data; however, they were not made aware of this chance.

<sup>14</sup> The price range was chosen based on pilot surveys of this study, where participants were asked to share their data under a one-price choice.

the final price for sharing their data. The BDM mechanism is commonly used in behavioral and experimental economics to reveal an individual’s reservation price of a good, even for commodities without established market prices (List and Shogren 2002). For the incentive-compatibility to work, the individual should understand that the potential, final price is randomly drawn from the price range. Therefore, prior to knowing the final price, it is in the best interest of the decision-maker to reveal their true, minimum willingness-to-accept.

Implementing a BDM price elicitation is not without its challenges. First, the instructions for how to choose a minimum price given an unknown final price drawn from a random distribution can be unusual and confusing. It is easy to conflate choosing the *minimum* acceptable with *any* acceptable price. Moreover, it is important to recognize that data are not sold for explicit prices in the real world,<sup>15</sup> so it is reasonably challenging for the average decision-maker to deduce their reservation price without prior experience with market prices for data. To remedy this challenge, I use a short and coarse list of only five prices (i.e., the survey taker did not need to consider a continuum of prices), which is neither long nor granular. The survey taker could easily consider each price in sequence, and imagine whether they would “take it or leave it” if that price were the final price. It is much easier to imagine if one can accept \$2.99 in exchange for data-sharing, rather than form a point estimate of one’s minimum price. On the other hand, price-reversing behavior in a multiple price list can happen (e.g., accepting \$1.99 while rejecting \$2.99). However, I did not design survey mechanisms to prevent this, in order to observe whether survey takers were attentive to the price elicitation in post-experimental analyses.

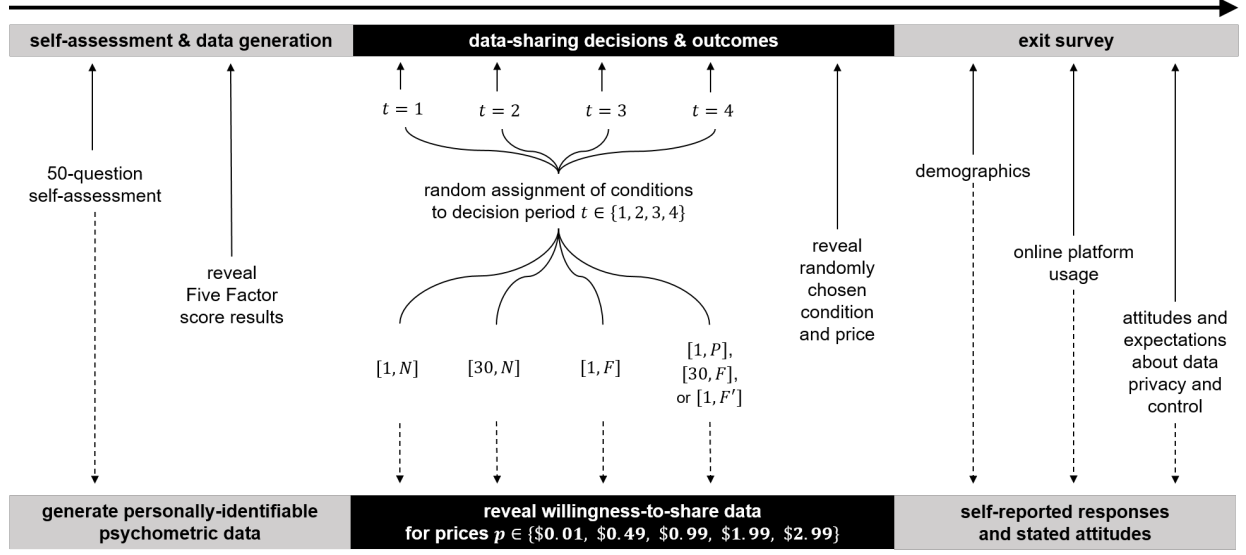
Following the data-sharing decisions and outcomes, the exit survey included voluntary questions on the individual’s demographics, social media usage, and general attitudes and expectations about their data privacy. Demographic questions included gender, age, employment, marital status, and education. Social media questions included usage and frequency of using the digital platforms Facebook, LinkedIn, and Twitter. Finally, a series of privacy and data ownership attitudes were elicited using a Likert-type assessment of statements related to privacy concerns and data usage by firms.

## 5. Measures and Estimation

Given the nature of the price elicitation, acceptable prices were chosen simultaneously with whether to accept *any* price in the available range. Selecting “I do not accept” for all prices allowed individuals to exit the experiment’s data market. My method of analyzing data market outcomes is to separately measure (1) the odds or likelihood of data market participation by individuals and

<sup>15</sup> Rather, there are implicit prices in trading data for goods and services.

Figure 4 Survey Chronology.



(2) the price demanded among those who did participate. Unlike many labor market participation questions in economics that focus on changes to the price variable, the question of selecting into data markets is more interesting and relevant for firm data strategy. Real data markets operate under an “all or nothing” and rarely under a “how much” in exchange for data. In addition, while measuring price changes are useful for inferences internal to the experiment, it is difficult to map these magnitudes to external contexts, especially where benefits for data are not monetary amounts.

While I recognize there exist behavioral mechanisms that influence data market entry or exit (i.e., the selection decision) separately from price; however, identifying these mechanisms is beyond the scope of this study. A Heckman-style selection model is not my estimation method, given the lack of a valid exclusion restriction. Therefore, a two part estimation—without a correction for selection bias—is the preferred style of inference for the results of this study. First, it estimates how the explanatory variables impacted data market participation. Whether individual  $i$  chose to participate in the data market under decision-period  $t$  is indicated by

$$Participation_{it} = \begin{cases} 1, & y_{it}^* \leq 2.99 \\ 0, & y_{it}^* > 2.99. \end{cases} \quad (1)$$

where  $y_{it}^*$  is the true, unobserved minimum acceptable price. Second, for those who selected into participating, it estimates the average changes to the observed, reservation price they were willing to accept. The observed minimum acceptable price when  $Participation_{it} = 1$  is



$$Price_{it} = \begin{cases} 0.01, & y_{it}^* \leq 0.01, \\ 0.49, & 0.01 < y_{it}^* \leq 0.49, \\ 0.99, & 0.49 < y_{it}^* \leq 0.99, \\ 1.99, & 0.99 < y_{it}^* \leq 1.99, \text{ and} \\ 2.99, & 1.99 < y_{it}^* \leq 2.99. \end{cases} \quad (2)$$

Since each individual  $i$  made four data-sharing decisions across  $t$  decision-periods, I use a panel random effects model that corrects for the non-independence of multiple responses from a single individual (Liang and Zeger 1986):

$$[Participation_{it}, Price_{it}] = \alpha + \beta \cdot \mathbf{Condition}_{it} + \delta \cdot \mathbf{T}_t + \gamma \cdot \mathbf{Y}_i + \theta_i + u_{it} \quad (3)$$

where  $\mathbf{Condition}_{it}$  is a set of categorical variables indicating the conditions of interest to be compared with a “leave-out” condition. A set of decision-period controls are denoted as  $\mathbf{T}_t$ . These are included to capture effects specific to each decision period that can presumably affect all individuals uniformly. The variable  $\mathbf{Y}_i$  is a set of participant-specific characteristics. The first characteristic includes the sample number of the individual, which is used when pooling results from all three samples. The second characteristic controls for whether the individual was randomly assigned to no exploitation information ( $N$ ) in the first two decision periods and any exploitation information ( $P$ ,  $F$ , or  $F'$ ) in the last two decision periods. Finally, as set of optional characteristics are included as an opportunity to analyze variation in privacy behavior across different types of participants. The participant-specific random effect is denoted by  $\theta_i$  and  $u_{it}$  is the error term. Since all individuals experienced conditions randomly assigned to a decision period, the estimates on the data-sharing conditions are uncorrelated with the observed ( $\mathbf{Y}_i$ ) and unobserved individual differences and error term ( $\theta_i$  and  $u_{it}$ ).

## 6. Data & Results

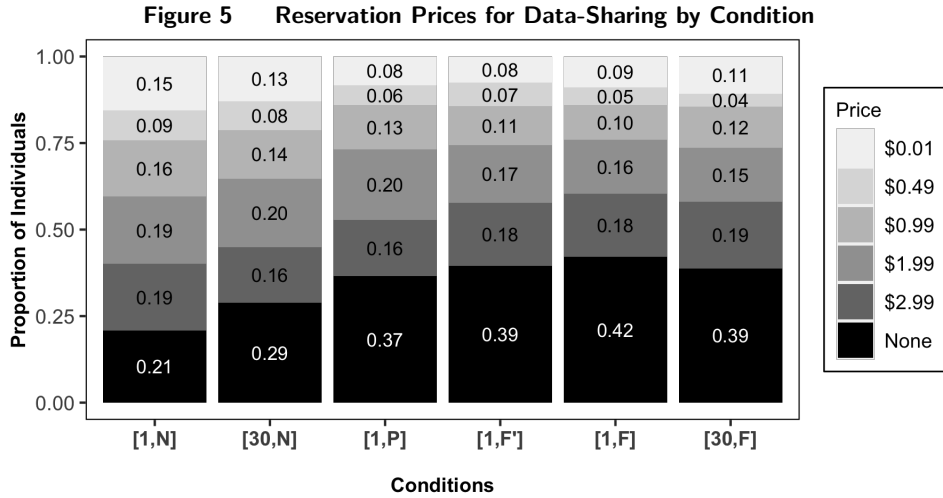
The data was collected from 1,188 participants in three samples spanning 16 months; each sample was conducted in four consecutive, weekly sessions. The first sample in March and April 2019 included 413 subjects; the second sample in October and November 2019 included 420 subjects; and the third sample in May and June 2020 included 355 subjects. The experiment collected self-reported activities with three major online platforms (Facebook, Twitter, and LinkedIn), and approximately 79% of individuals self-reported using Facebook with a non-anonymous account and accessed the platform at least once per week.<sup>16</sup> Data collected from these questions provide an

<sup>16</sup> Following this same definition of platform usage, 50.8% and 24.3% of participants self-reported as users of LinkedIn and Twitter, respectively. A significant minority of those surveyed reported active engagement with these platforms. For example, 19% were Facebook users that posted or shared content on the platform at least once a week; 39.1% were LinkedIn users who responded to requests for connections within a week; and 12.3% were Twitter users who posted or shared content at least weekly.

approximate understanding of the survey respondents' out-of-experiment engagement with online platforms that collect user-generated data for internal or external business uses. The experimental subjects were predominantly female, in their early-20s, and students.<sup>17</sup>

### 6.1. Descriptive Summary of Privacy Choices

After individuals disclosed personal information in the self-assessment, they displayed privacy-seeking behavior over the data created from those assessments in downstream choices, descriptively supporting Hypothesis 1 (e.g., see the proportion of individuals' minimum price they were willing-to-accept by experimental condition in Figure 5). For example, approximately 60 percent of individuals' minimum willingness-to-accept was greater than \$0.99 for sharing data with *one* other study participant and in the *absence* of any information about secondary market data exploitation (i.e., condition  $[1, N]$ ), and 21 percent rejected any price in the list. In Figures 5 and 6, descriptive evidence also suggests that—as predicted by Hypothesis 2—privacy-seeking behavior was more prevalent when exposed to more data recipients. Finally, support for Hypothesis 3 is most visible at this descriptive stage: Privacy-seeking behaviors were most prevalent when there were information provisions about a data recipient's secondary market exploitation abilities.



As shown in Figure 6, 23 percent of subjects were willing to share personal data with *one* recipient (in return for prices less than or equal to \$2.99) in the  $[1, N]$  condition, while unwilling to share when treated with information about their data recipient's secondary market exploitation abilities in the  $[1, F]$  condition. To compare with the effect that greater exposure has on privacy behavior, 10 percent of individuals were willing to participate in the  $[1, N]$  market but not willing

<sup>17</sup> This study's demographic make-up is not unusual compared to other studies conducted at the University maintained subject pools. Over 72% of participants self-identified as female. Their mean reported age was 23.8 years ( $SD = 6.95$ ). Over 70% reported to be students as opposed to employed (full or part-time).

**Figure 6 Data Market Participation and Non-Participation Across Conditions**

Non-Participants	All Samples							Sample 1						
	[30,F]													
	[1,F]	0.42 (501)	0.23 (270)	0.18 (212)			0.00 (0)	0.45 (187)	0.24 (100)	0.19 (80)	0.12 (48)		0.00 (0)	
	[1,F']													
	[1,P]							0.37 (151)	0.16 (65)	0.12 (50)	0.00 (0)		0.03 (12)	
	[30,N]	0.29 (343)	0.10 (121)	0.00 (0)			0.05 (54)	0.31 (127)	0.10 (40)	0.00 (0)	0.06 (26)		0.05 (20)	
	[1,N]	0.21 (248)	0.00 (0)	0.02 (26)			0.01 (17)	0.23 (93)	0.00 (0)	0.01 (6)	0.02 (7)		0.01 (6)	
	Total	1.00 (1188)	0.79 (940)	0.71 (845)			0.58 (687)	1.00 (413)	0.77 (320)	0.69 (286)	0.63 (262)		0.55 (226)	
	Sample 2							Sample 3						
	[30,F]	0.39 (163)	0.23 (96)	0.17 (72)			0.05 (20)							
Participants	[1,F]	0.37 (156)	0.21 (88)	0.17 (70)			0.00 (0)	0.45 (158)	0.23 (82)	0.17 (62)		0.08 (29)	0.00 (0)	
	[1,F']							0.39 (140)	0.18 (63)	0.13 (47)		0.00 (0)	0.03 (11)	
	[1,P]													
	[30,N]	0.24 (101)	0.09 (37)	0.00 (0)			0.04 (15)	0.32 (115)	0.12 (44)	0.00 (0)		0.06 (22)	0.05 (19)	
	[1,N]	0.18 (74)	0.00 (0)	0.02 (10)			0.01 (6)	0.23 (81)	0.00 (0)	0.03 (10)		0.01 (4)	0.01 (5)	
	Total	1.00 (420)	0.82 (346)	0.76 (319)			0.63 (264)	1.00 (355)	0.77 (274)	0.68 (240)		0.61 (215)	0.55 (197)	
	Total	[1,N]	[30,N]	[1,P]	[1,F']	[1,F]	[30,F]	Total	[1,N]	[30,N]	[1,P]	[1,F']	[1,F]	[30,F]

Note: Proportions out of sample and number of individuals in parentheses.

to participate in the  $[30, N]$  market to share data with *thirty* recipients. In fact, 18 percent of subjects were willing to enter the data market under  $[30, N]$  but not under  $[1, F']$ —in contrast, only 5 percent of subjects exhibited the reverse behavior.

When observing each sample separately and leveraging the data-sharing conditions unique to each sample, I find further evidence that privacy choice changes were responsive to exploitation provisions and more prevalent than privacy responses to increased exposure. Providing partial information about a recipient's "ability to make money" (i.e., excluding the secondary market information on "selling to a third party") changed data-market participation for 16 percent of subjects, who entered the data market under  $[1, N]$  but did not enter under  $[1, P]$ . In addition, 12 percent of participants entered under  $[30, N]$  but not  $[1, P]$ —contrasted with 6 percent of subjects who exhibited the reverse behavior. Provisioning secondary market activities of data recipients still increased the privacy-seeking behavior of those willing to expose their data to many recipients—e.g., 17 percent of subjects participated under  $[30, N]$  but not under  $[30, F]$ . Finally, privacy-seeking

behavior in response to provisions about a data recipient’s ability to sell data was also descriptively robust to changing “third party” to another person in the study, where 18 percent participated in  $[1, N]$  but not  $[1, F']$ .

## 6.2. Regression on Participation and Prices

As previously mentioned, inferences are made in a two-part fashion—as opposed to assuming a natural censoring on the dependent variable—to appropriately acknowledge that the decision to select any price may follow a different decision process than the price chosen. Estimates of individuals’ participation and conditional prices include three specifications of random effects panel regressions. The results tables present the likelihood of data market participation and price changes (conditional on participation), each containing three specifications.

The first, baseline model, includes variables indicating the condition a privacy choice was made under. The second specification extends the first by including a control for  $N$  in  $t \in \{1, 2\}$  to indicate whether the individual experienced no exploitation information conditions in the first two of four decisions and information treatments  $F$ ,  $P$ , or  $F'$  in the latter two decisions. The third specification includes optional controls for individual-specific characteristics.<sup>18</sup> All specifications include controls for decision-period, and regressions with pooled data control for sample. All regression specifications are compared using a Wald Test, to determine whether the inclusion of regressors have meaningful explanatory power.

Table 1 summarizes the estimated impact on data market participation and minimum acceptable prices of  $[1, F]$  and  $[30, N]$  conditions relative to the baseline condition  $[1, N]$ .<sup>19</sup> Both  $[30, N]$  and  $[1, F]$  conditions demonstrate lower willingness-to-share data, confirming Hypotheses 2, Hypothesis 3, and—by construction—Hypothesis 1. Relative to condition  $[1, N]$ , the odds of not entering the data market under  $[30, N]$  were 1.56 times greater<sup>20</sup> and under condition  $[1, F]$  were 2.83 times greater.<sup>21</sup> A similar privacy-seeking pattern follows for those who participated in the data market, where prices demanded under  $[30, N]$  were approximately \$0.10 higher ( $p < 0.0001$ ) and  $[1, F]$  were approximately \$0.39 higher ( $p < 0.0001$ ) relative to  $[1, N]$ . The results suggest that exposure to more data recipients and salient exploitation signals motivated stronger privacy-seeking behavior, both on the extensive and intensive margins for data-sharing.

<sup>18</sup> Demographic controls include female, marital status, Facebook usage, and employment status. Psychometric controls include whether the individual scored above 30 (on a scale of 10 to 50) in each of the Five Factor traits: extraversion, agreeableness, conscientiousness, emotional stability, and intellectual.

<sup>19</sup> If  $C$  is the focal condition and  $\bar{C}$  is the comparison condition, the odds ratio of  $C$  and  $\bar{C}$  is  $\exp[\hat{\beta}] = \exp[\log(\text{odds}C/\text{odds}\bar{C})] = \text{odds}C/\text{odds}\bar{C}$ , where  $\hat{\beta}$  is the coefficient estimate of the condition of interest.

<sup>20</sup> Odds ratio 1.56 =  $\exp(-0.442)^{-1}$ , where  $\hat{\beta}_{[1, N]}^{[30, N]} = -0.442$  ( $p < 0.0001$ )

<sup>21</sup> Odds ratio 2.83 =  $\exp(-1.04)^{-1}$ , where  $\hat{\beta}_{[1, N]}^{[1, F]} = -1.04$  ( $p < 0.0001$ ).

**Table 1** Data Market Participation and Price Results (Pooled)

Variables Dependent:	Logit Participation			OLS Price (\$)   Participation=1		
	(1a)	(2a)	(3a)	(1b)	(2b)	(3b)
Comparison: $[1, N]$						
(Intercept)	1.289*** (0.113)	1.474*** (0.124)	2.107*** (0.378)	1.499*** (0.061)	1.399*** (0.071)	1.154*** (0.208)
$[30, N]$	-0.433*** (0.054)	-0.434*** (0.054)	-0.442*** (0.055)	0.102*** (0.024)	0.102*** (0.024)	0.102*** (0.024)
$[1, F]$	-1.018*** (0.064)	-1.016*** (0.064)	-1.039*** (0.065)	0.390*** (0.033)	0.393*** (0.033)	0.393*** (0.033)
$N$ in $t \in \{1, 2\}$		-0.369*** (0.108)	-0.387*** (0.110)		0.183** (0.066)	0.196** (0.066)
Sample controls	Yes	Yes	Yes	Yes	Yes	Yes
Decision-period controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographic controls	No	No	Yes	No	No	Yes
Psychometric controls	No	No	Yes	No	No	Yes
Individual clusters	1188	1188	1188	975	975	975
Observations	3564	3564	3564	2472	2472	2472
<b>ANOVA: Wald Test</b>	(1a),(2a)	(2a),(3a)	(1a),(3a)	(1b),(2b)	(2b),(3b)	(1b),(3b)
$Pr(> \chi^2)$	0.001	0.000	0.000	0.006	0.192	0.030

Note: Clustered robust standard errors in parentheses.

<sup>†</sup> $p < 0.1$ ; \* $p < 0.05$ ; \*\* $p < 0.01$ ; and \*\*\* $p < 0.001$

Notably, the magnitude of subjects' aversion to data exploitation—and, simultaneously, their lack of awareness to a recipient's data exploitation abilities—is relatively large when compared to their responses to greater data exposure. The results demonstrate that exploitation signals for just *one* data recipient resulted in stronger privacy-seeking behavior than decisions to expose data to 29 more data recipients—e.g., the odds of not participating in the data market under  $[1, F]$  were 1.82 times greater than  $[30, N]$ .<sup>22</sup>

Evidence of strong privacy responses to exploitation signals, even when compared to high exposure can also be found in different variations of condition  $[1, F]$ . The results shown in Table 2 display the conditions common to all samples ( $[1, N]$ ,  $[30, N]$ , and  $[1, F]$ ) plus a variant condition with exploitation information provisions ( $[1, P]$ ,  $[30, F]$ , or  $[1, F']$ ) that are unique to each sample. All other model specifications are the same as Table 1, with the exclusion of sample controls in these sample-specific results.

First, in Sample 1, I find that salient signals about data monetization alone can motivate privacy-seeking behavior. The odds of not participating in the data market under partial information condition  $[1, P]$  were 2.01 times greater relative to  $[1, N]$ <sup>23</sup> and 1.3 times greater relative to  $[30, N]$ .<sup>24</sup> The results of data market participants' reservation prices follow the same pattern. As shown in

<sup>22</sup> Odds ratio  $1.82 = \exp(-0.597)^{-1}$ , where  $\hat{\beta}_{[1, N]}^{[1, F]} - \hat{\beta}_{[1, N]}^{[30, N]} = -0.597$  ( $p < 0.0001$ ).

<sup>23</sup> Odds ratio  $2.01 = \exp(-0.70)^{-1}$ , where  $\hat{\beta}_{[1, N]}^{[1, P]} = -0.70$  ( $p < 0.0001$ ).

<sup>24</sup> Odds ratio  $1.3 = \exp(-0.26)^{-1}$ , where  $\hat{\beta}_{[1, N]}^{[1, P]} - \hat{\beta}_{[1, N]}^{[30, N]} = -0.26$  ( $p = 0.0007$ ).

**Table 2 Data Market Participation and Price Results (By Sample)**

Variables Dependent:	Logit <i>Participation</i>			OLS <i>Price (\$)   Participation=1</i>		
	(1a)	(2a)	(3a)	(1b)	(2b)	(3b)
<b>Sample 1</b>						
(Intercept)	1.317*** (0.138)	1.431*** (0.160)	1.031 <sup>†</sup> (0.571)	1.472*** (0.070)	1.373*** (0.091)	1.504*** (0.326)
[30, $N$ ]	-0.425*** (0.083)	-0.425*** (0.083)	-0.439*** (0.085)	0.092* (0.039)	0.091* (0.039)	0.090* (0.039)
[1, $P$ ]	-0.682*** (0.097)	-0.678*** (0.097)	-0.699*** (0.100)	0.252*** (0.050)	0.253*** (0.050)	0.253*** (0.050)
[1, $F$ ]	-1.046*** (0.108)	-1.045*** (0.107)	-1.078*** (0.110)	0.428*** (0.058)	0.429*** (0.058)	0.427*** (0.058)
$N$ in $t \in \{1, 2\}$		-0.243 (0.179)	-0.276 (0.182)		0.195 <sup>†</sup> (0.107)	0.190 <sup>†</sup> (0.106)
<b>Sample 2</b>						
(Intercept)	1.554*** (0.149)	1.882*** (0.175)	3.819*** (0.634)	1.412*** (0.064)	1.330*** (0.087)	0.632* (0.294)
[30, $N$ ]	-0.395*** (0.098)	-0.401*** (0.100)	-0.418*** (0.103)	0.069 <sup>†</sup> (0.036)	0.069 <sup>†</sup> (0.036)	0.070 <sup>†</sup> (0.036)
[1, $F$ ]	-1.007*** (0.115)	-1.016*** (0.114)	-1.073*** (0.118)	0.430*** (0.052)	0.431*** (0.052)	0.432*** (0.052)
[30, $F$ ]	-1.079*** (0.118)	-1.090*** (0.119)	-1.153*** (0.123)	0.395*** (0.053)	0.397*** (0.053)	0.397*** (0.052)
$N$ in $t \in \{1, 2\}$		-0.668*** (0.190)	-0.674*** (0.195)		0.166 (0.110)	0.174 (0.109)
<b>Sample 3</b>						
(Intercept)	1.284*** (0.144)	1.364*** (0.172)	1.978** (0.761)	1.551*** (0.071)	1.425*** (0.095)	1.365** (0.415)
[30, $N$ ]	-0.484*** (0.102)	-0.483*** (0.102)	-0.492*** (0.103)	0.168*** (0.048)	0.168*** (0.048)	0.168*** (0.048)
[1, $F'$ ]	-0.789*** (0.103)	-0.785*** (0.102)	-0.800*** (0.104)	0.390*** (0.057)	0.392*** (0.057)	0.391*** (0.057)
[1, $F$ ]	-0.998*** (0.113)	-0.996*** (0.113)	-1.015*** (0.114)	0.362*** (0.062)	0.364*** (0.062)	0.362*** (0.062)
$N$ in $t \in \{1, 2\}$		-0.172 (0.191)	-0.176 (0.194)		0.243* (0.118)	0.243* (0.118)
Decision-period controls	Yes	Yes	Yes	Yes	Yes	Yes
Demographic controls	No	No	Yes	No	No	Yes
Psychometric controls	No	No	Yes	No	No	Yes

Note: Clustered robust standard errors in parentheses.

<sup>†</sup> $p < 0.1$ ; \* $p < 0.05$ ; \*\* $p < 0.01$ ; and \*\*\* $p < 0.001$

Model (3b), condition [1,  $P$ ] had approximately \$0.25 higher prices than [1,  $N$ ] ( $p < 0.0001$ ) and \$0.16 higher prices than [30,  $N$ ] ( $p = 0.001$ ). Notably, there was also a significant increase in the exit behavior and prices under [1,  $F$ ] relative to [1,  $P$ ], suggesting a privacy response to third party sales.

Second, in Sample 3, replacing “third party” with another person in the study was robust to the finding that individuals were more motivated by aversion to data exploitation than data

exposure—e.g., the odds of exiting the data market were 1.36 times greater<sup>25</sup> and prices \$0.22 higher ( $p = 0.0003$ ) compared to  $[30, N]$ . These results suggest that even when the third party is contained within the experiment—and, thus, effectively sharing data with *two* recipients—exploitation-aversion behavior was stronger than aversion to sharing data with *thirty* recipients without salient exploitation abilities. Notably, there was also a small increase in data market exit under  $[1, F]$  relative to  $[1, F']$  but no evidence in price changes.

Finally, privacy responses to data recipients’ exploitation abilities also replicated under a 30 recipient scenario (as opposed to comparing responses with and without signals for one recipient). In Sample 2, the the effect of exploitation signals persisted when individuals considered thirty data recipients—the odds of not participating were 2.09 times greater<sup>26</sup> and prices \$0.33 higher ( $p < 0.0001$ ) under  $[30, F]$  than under  $[30, N]$ . However, in Sample 2, individuals exhibited little change in privacy behavior under  $[1, F]$  versus  $[30, F]$ , where increasing exposure to more data recipients who *each* had exploitation abilities yield similar responses to *just one* data recipient having these abilities. This result suggests that privacy-seeking behavior is diminishing in the number of data recipients who can exploit one’s data for profit.

One interesting effect was found in from the  $N$  in  $t \in \{1, 2\}$  randomization, where the evidence suggests a relationship between the sequence of information provision and privacy decision-making. Subjects who were treated with exploitation signals in the latter half of their decisions exhibited a stronger privacy response than their counterparts. However, as shown in Table 2, this information-sequence effect is primarily attributed to its magnitude and statistical precision in Sample 2, whereas the other two samples show weaker and imprecise estimates of this effect.

### 6.3. Discussion

Throughout three replications and across various versions of the main treatment effect, results show that individuals consistently respond to information signals about a data recipient’s data exploitation abilities. This result begs the question: what is the behavioral mechanism that drives these privacy responses? While further research is needed to disentangle this question, results suggest a behavioral motivation for privacy-seeking behavior beyond extrinsic factors associated with data exposure (e.g., data breaches, surveillance). One obvious one is fairness-related—that is, individuals believe they should split the surplus gained by a recipient from the exchange. Other mechanisms can be specific to personal data characteristics, such as data inalienability and psychological ownership of personal data—despite a transfer of information, individuals may still feel ownership over what is “their” data and demand data royalties in exchange for exploitation.

<sup>25</sup> Odds ratio  $1.36 = \exp(-0.308)^{-1}$ , where  $\hat{\beta}_{[1, N]}^{[1, F']} - \hat{\beta}_{[1, N]}^{[30, N]} = -0.308$  ( $p = 0.002$ ).

<sup>26</sup> Odds ratio  $2.09 = \exp(-0.74)^{-1}$ , where  $\hat{\beta}_{[1, N]}^{[30, F]} - \hat{\beta}_{[1, N]}^{[30, N]} = -0.74$  ( $p < 0.0001$ ).

Furthermore, why have we not seen this evidence in prior experimental works on secondary use such as [Buckman et al. \(2019\)](#)? One possibility is that the consequences of secondary use have previously been focused on risks related to data exposure, such as unwanted surveillance and data breaches. I do not believe my results contradict prior studies; rather, my findings extend our understanding of privacy preferences related to secondary use—specifically, on the ability of others to privately benefit (i.e., monetize in secondary markets) one’s data. As the landscape of privacy research has suggested: privacy-related concerns are indeed wide-ranging and can include non-normative factors. Moreover, my study accommodates for and confirms that “all-or-nothing” privacy responses are prevalent in data-sharing decisions. Valuation elicitation methods that do not include an opt-out choice can miss this critical distinction in privacy responses to notice-and-choice regimes.

There is an open question about why salient exploitation signals lack perfect spillover to later data-sharing choices (e.g., receiving a signal in  $t = 2$  but having no signal in  $t = 3$ ), especially within such a short period of sequential decisions where limited memory capacity is unlikely the explanation. Perfect information spillover is of course the expected behavior for rational, privacy-decision makers, where individuals should remember information provisions from past decisions. However, when the exploitation signals “leave”, on average subjects in my study do not reveal the same privacy-seeking behavior—despite some evidence from the  $N$  in  $t \in \{1, 2\}$  randomization that receiving these signals early can encourage more private behavior later. One possibility is that there is a behavioral, privacy-seeking response to being surprised by the news of a data recipient’s exploitation abilities due to regret over one’s past data-sharing choices. However, the more obvious takeaway is that unless “terms and conditions” are clearly present—and, importantly, salient—at the time of a disclosure choice, individuals incorrectly assume there are regimes to prevent the exploitation of their data in secondary markets. This possibility strengthens the use of for privacy policies that support users’ informed decisions over their personal data, while also supporting the value of governance structures that allow users to have greater control over their personal data in secondary markets.

There are also some additional inferences to be made from the set of controls in the analysis. First, privacy-seeking behavior co-moves with the progression of decisions. This implies a learning and experience effect from data-sharing decision-making on stronger, revealed privacy preferences. Second, Sample 2 (collected in Fall/Winter) has an unexplained, weaker set of privacy-seeking behaviors in comparison to the other two samples collected in Spring and Summer. Finally, there are some patterns in the pooled data that are primarily attributed to effects from Sample 2—i.e., they are not replicated in magnitude or precision in Samples 1 and 3. This includes the effects of  $N$  in  $t \in \{1, 2\}$ , previously discussed in the results section of this paper. In addition, there is a significant effect for those with higher psychometric scores on their Conscientiousness trait in



Sample 2. Thus, it is possible that this personality type may be more vulnerable to a lack of salience about the monetization abilities of recipients. However, rigorous conclusions should not be made from this data given the possibility of unique traits in Sample 2 that do not generalize.

Finally, there are several limitations to the interpretation of my results. First, there is a natural difference between for profit or non-profit studies (i.e, research studies). The expectations of many subjects coming into a research study are disassociated with for-profit data markets. This study cannot conclude whether individuals are aware of data exploitation in these secondary markets in real-world data exchanges the individuals are participating in. However, replicating this study in the field would help understand how these results generalize outside of a controlled environment.

Second, there are known to be strong anchoring effects of price elicitation survey questions. While the relative valuations in this study provides meaningful inferences, the study's average prices should not be generalized to real market prices for persons' psychometric data. Although the multiple price list elicitation method in this study attempts to minimize this issue, the point estimates of reservation prices can vary depending on the lower and upper bounds. One alternative is to directly ask subjects to declare a minimum, acceptable price. However, as mentioned in previous sections of this paper regarding the challenge of measuring privacy preferences, point estimates of reservation prices are unnatural, unstable, and difficult to compute for a decision-maker with no prior experience with explicit prices for their personal data.

Third, I cannot correct for selection bias on the relative valuation of personal data by those who opt-in to the data market in this study. Therefore, I cannot discern between those who are naturally censored by the price range versus those who would exist on an unobserved distribution of prices. For example, it could be the case that \$6.99 is enough to capture all or none of (or somewhere in between) those who rejected \$2.99. Wider ranges of price lists cannot easily correct for this due to the aforementioned anchoring effects. Furthermore, extremely high upper-bound prices can be less believable to decision-makers, which effectively render them opt-out choices and difficult to interpret for researchers.

Finally, I cannot rule out the potential influence of experimenter demand effects, despite employing the most important design choice: the inclusion of real outcomes and incentive-compatibility. Even though an analysis of between-subjects behavior in only the first decision period confirms the main findings, this strategy is not ideal as it rules out the importance of capturing more experienced decisions, which are less prone to errors (i.e., miscalculation of privacy trade-offs).

## 7. Future Research & Implications

This work demonstrates the importance of studying privacy decision-making as a set of choices that depend on secondary data market activities. Privacy behavior responds to salient information

about how parties—who acquire personal data assets—can benefit from trading and exploiting those assets in secondary markets. Even after individuals disclose information to a first party, they have privacy preferences over personal data that remain unrevealed for later, downstream markets. This work has a number of important implications for research into firms’ privacy strategies, individuals’ privacy preferences, and regulators’ privacy policies. A number of extensions of this work should be done to continue the discourse about privacy issues in digital markets.

First, firms in the U.S. and Europe are beginning to implement new consumer privacy policies that adhere to strict guidelines for obtaining opt-in consent for data usage by third parties.<sup>27</sup> Research has already demonstrated GDPR policy impacts on firm strategy, effectively reducing the number of third party web connections, and market power consequences that favor larger firms (e.g., Google) (Batikas et al. 2020). Research on privacy preferences should continue to be done to understand the balance of costs to competition in digital markets and benefits to consumer privacy welfare.

Second, research on how secondary markets influence privacy preferences should be extended beyond the lab. This study was conducted in a research university setting. Individuals’ propensity to share data likely depends on whether the collecting agency is non-profit (e.g., a research university) or for-profit (e.g., a marketing agency). The fact that self-assessment responses were so readily disclosed by those in my study likely relates to the study being marketed as a research project. Future research should examine the impact of this study’s information provisions in a for-profit setting.

Third, this work is only one perspective on how individuals’ privacy preferences depend on secondary market activities of data recipients. This study shows how the exploitation ability of a recipient can matter more than exposure to more recipients. Future research can study other data types—beyond the psychometric data in this study—to see whether aversion to others exploiting one’s data dominates aversion to others experiencing one’s data.

Fourth, there are various mechanisms that potentially explain the finding that data exploitation matters for privacy behavior, which I encourage future researchers to examine both theoretically and empirically. The most compelling and privacy-specific mechanism is the theory that individuals may never fully accept a loss of ownership over their personal data—despite any economic transaction that transfers that ownership, which stems from the *inalienability* feature of personal data (Koutroumpis et al. 2019). In addition, fairness concerns (i.e., “Why should they make so

<sup>27</sup> Many of these new firm strategies are responding to the European GDPR and U.S. CCPA. Moreover, firms are also investing in privacy protecting strategies in general. For example, Apple’s “User Privacy and Data Use” is requiring that users of iOS 14.5 give apps consent for their data to be tracked across apps and websites owned by other companies by the end of 2021. See <https://developer.apple.com/app-store/user-privacy-and-data-use/>.

much money off of my data?”) and hold-out behavior (i.e., “I want the best deal they can offer for my data.”) can also explain the aversion to personal data exploitation.

Fifth, beyond the environmental treatments that are exogenously assigned and examined in this study, there remains many unobserved factors that influence privacy preferences that are unaccounted for in empirical research. Future work in behavioral and psychology research can examine more challenging empirical questions about endogenous determinants to privacy choices. This study provides (non-randomly assigned) data on personality traits and privacy attitudes and, thus, some suggestive evidence that certain characteristics predict more or less privacy behavior.

Sixth, existing data property rights are inadequate and inflexible towards the usage-dependent privacy preferences of individuals. In particular, more flexible exclusionary abilities on behalf of individuals can account for differentiating disclosure and access rights between different kinds of data recipients and for different usages. Regulators should consider policies that adjust for changing downstream privacy choices. The recent data exploiting firms *Cambridge Analytical* (in the case of psychometric data) or *Clearview AI* (in the case of facial imaging data) demonstrate there are strong incentives for firms to create products and services commercialized with personal data in the absence of data licensing agreements with the individuals who supply personal. Recent theoretical work by [Jones and Tonetti \(2020\)](#) found that giving data property rights to firms can result in data hoarding behavior and socially inefficient outcomes, whereas consumer data property rights are more optimal and can balance the economic gains of selling data with privacy concerns. However, the long term consequences to consumer welfare in the absence of sophisticated data licensing structures has not yet been measured.

## 8. Conclusion

Understanding how individuals make privacy trade-offs is critical for firm strategies involving external exploitation of data assets. This project evolves our understanding of the motivations for individual privacy-seeking behavior in the personal data market. As consumers become better-informed agents in digital economies, the way that firms collect, curate, manage, and share data will be priced-into each person’s privacy behavior—not only in their propensity to disclose, but also in their demand for better goods and services. If individuals suffer dis-utility from being exploited in the personal data market, then this motivates privacy regulation and protections for individuals to control their personal data in digital markets. If users are inattentive to how their data are used in secondary markets, information provisions about how firms use and benefit from exploiting data can influence individuals’ (1) decisions related to exiting the data market (e.g., data deletion requests or not consenting to third party data usage) or (2) demand greater benefits in return for data-sharing.

## References


- Acquisti A, Brandimarte L, Loewenstein G (2015) Privacy and human behavior in the age of information. *Science* 347(6221):509–514.
- Acquisti A, Gross R (2009) Predicting social security numbers from public data. *Proceedings of the National Academy of Sciences* 106(27):10975–10980.
- Acquisti A, Grossklags J (2005a) Privacy and rationality in individual decision making. *IEEE Security and Privacy* 3(1):24–33.
- Acquisti A, Grossklags J (2005b) Uncertainty, ambiguity and privacy. *Fourth Workshop on the Economics of Information Security (WEIS05)*, 2–3.
- Acquisti A, John L, Loewenstein G (2013) What is privacy worth? *The Journal of Legal Studies* 42(2):249–274.
- Acquisti A, Taylor C, Wagman L (2016) The economics of privacy. *Journal of Economic Literature* 54(2):442–492.
- Adjerid I, Acquisti A, Brandimarte L, Loewenstein G (2013) Sleights of privacy: Framing, disclosures, and the limits of transparency. *Proceedings of the Ninth Symposium on Usable Privacy and Security, SOUPS '13* (New York, NY, USA: Association for Computing Machinery).
- Adjerid I, Acquisti A, Loewenstein G (2019) Choice architecture, framing, and cascaded privacy choices. *Management Science* 65(5):2267–2290.
- Angst CM, Agarwal R (2009) The Adoption of electronic health records in the presence of privacy concerns: The elaboration likelihood model and individual persuasion. *Management Information Systems Quarterly* 33(2):339–370.
- Athey S, Catalini C, Tucker C (2017) The digital privacy paradox: Small money, small costs, small talk. NBER Working Paper No. 23488, National Bureau of Economic Research.
- Batikas M, Bechtold S, Kretschmer T, Peukert C (2020) European privacy law and global markets for data. CEPR Discussion Papers 14475.
- Becker G, Degroot M, Marschak J (1964) Measuring utility by a single-response sequential method. *Behavioral Science* 9(3):226–232.
- Brandimarte L, Acquisti A, Loewenstein G (2012) Misplaced confidences: Privacy and the control paradox. *Social Psychological and Personality Science* 4(3):340–347.
- Buckman JR, Bockstedt JC, Hashim MJ (2019) Relative privacy valuations under varying disclosure characteristics. *Information Systems Research* 30(2):375–388.
- Culnan MJ (1993) “How did they get my name?”: An exploratory investigation of consumer attitudes toward secondary information use. *MIS Quarterly: Management Information Systems* 17(3):341–361.
- Culnan MJ, Armstrong PK (1999) Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization Science* 10(1):104–115.

- DellaVigna S (2009) Psychology and economics: Evidence from the field. *Journal of Economic Literature* 47(2):315–72.
- Dinev T, Hart P (2006) An extended privacy calculus model for e-commerce transactions. *Information Systems Research* 17(1):61–80.
- Glasgow G, Butler S (2017) The value of non-personally identifiable information to consumers of online services: evidence from a discrete choice experiment. *Applied Economics Letters* 24(6):392–395.
- Goldberg LR (1992) The development of markers for the big-five factor structure. *Psychological Assessment* 4(1):26–42.
- Goldberg LR, Johnson JA, Eber HW, Hogan R, Ashton MC, Cloninger CR, Gough HG (2006) The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality* 40(1):84–96.
- Goldfarb A, Tucker C (2019) Digital economics. *Journal of Economic Literature* 57:3–43.
- Graham-Harrison E, Cadwalladr C (2018) Revealed: 50 million facebook profiles harvested for cambridge analytica in major data breach. *The Guardian* URL <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>.
- Hill K (2020) The secretive company that might end privacy as we know it. *The New York Times* URL <https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html>.
- Hui KL, Teo HH, Lee SYT (2007) The value of privacy assurance: An exploratory field experiment. *Management Information Systems Quarterly* 31(1):19–33.
- John LK, Acquisti A, Loewenstein G (2010) Strangers on a plane: Context-dependent willingness to divulge sensitive information. *Journal of Consumer Research* 37(5):858–873.
- Jones CI, Tonetti C (2020) Nonrivalry and the economics of data. *American Economic Review* 110(9):2819–58.
- Koutroumpis P, Leiponen A, Thomas LDW (2019) The nature of data. *Innovation and Entrepreneurship Working Papers* (Imperial College Business School: London, UK).
- Koutroumpis P, Leiponen A, Thomas LDW (2020) Markets for data. *Industrial and Corporate Change* 29(3):645–660.
- Laufer R, Wolfe M (1977) Privacy as a concept and a social issue: A multidimensional developmental theory. *Journal of Social Issues* 33(3):22–42.
- Liang KY, Zeger SL (1986) Longitudinal data analysis using generalized linear models. *Biometrika* 73(1):13–22.
- List JA, Shogren JF (2002) Calibration of willingness-to-accept. *Journal of Environmental Economics and Management* 43(2):219–233.

- Matz SC, Kosinski M, Nave G, Stillwell DJ (2017) Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences* 114(48):12714–12719.
- McCrae RR, John OP (1992) The five-factor model: issues and applications. *Journal of Personality* 60(2):175–532.
- Miller AR, Tucker C (2018) Privacy protection, personalized medicine, and genetic testing. *Management Science* 64(10):4648–4668.
- O'Donoghue T, Rabin M (1999) Doing it now or later. *American Economic Review* 89(1):103–124.
- Posner RA (1981) The economics of privacy. *American Economic Review* 71(2):405–409.
- Preibusch S (2015) How to explore consumers' privacy choices with behavioral economics. Zeadally S, Badra M, eds., *Privacy in a Digital, Networked World*, chapter 14, 313–341 (Springer International Publishing).
- Schudy S, Utikal V (2017) You must not know about me—on the willingness to share personal data. *Journal of Economic Behavior & Organization* 141:1–13.
- Stigler GJ (1980) An introduction to privacy in economics and politics. *The Journal of Legal Studies* 9(4):623–644.
- Sutanto J, Palme E, Tan CH (2013) Addressing the personalization-privacy paradox. *MIS Quarterly* 37(4):1141–1164.
- Thomas LDW, Leiponen A (2016) Big data commercialization. *IEEE Engineering Management Review* 44(2):74–90.
- Tsai JY, Egelman S, Cranor L, Acquisti A (2011) The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research* 22(2):254–268.
- Varian H (1996) Economic aspects of personal privacy .
- Westin AF (1967) *Privacy and Freedom* (Atheneum), 1st edition.
- Xu H, Teo HH, Tan BC, Agarwal R (2010) The role of push-pull technology in privacy calculus: The case of location-based services. *Journal of Management Information Systems* 26(3):135–174.

## Appendix

**Figure 7     Advertised Study for Individuals Registered with Lab**

Study Information	Approved?	View
<p><b>How well do you know yourself? An economic decision study</b></p> <p>(Amazon gift cards) (Online Study) Seeking participants for an online study on decision making.</p>	<p> Approved</p>	<ul style="list-style-type: none"> <li>• Study Info</li> <li>• Timeslots</li> </ul>

**Figure 8     Information Page for Registered Study Participants.**

### How well do you know yourself? An economic decision study

#### Information Page

Thank you for your interest in participating in an economic decision study. You are now registered to participate in this study as part of the Debra Paget and Jeffrey Berg Business Simulation Laboratory (“The Lab”), which is a Cornell University Institutional Review Board for Human Participants (IRB) approved recruitment system.

This webpage is NOT the study’s survey. **Approximately 3 days before the participation deadline, you will be sent a personal email to participate in the study.** It will be an online survey, and it will take approximately 15 minutes to complete (but plan for a 20 minute period to review the consent form and/or provide [optional] feedback to the researcher).

If at any point you no longer wish to participate in this study, we encourage you to cancel your registration (through the Sona System) at least 4 days before the end of the participation deadline, so that others might be able to sign-up for the study.

#### [OPTIONAL] Questions?

If you have any questions or concerns prior to receiving the link to the study (which will be sent out **3 days** before the participation deadline), please feel free to send a note in the form below. The researcher will respond as soon as possible.

[OPTIONAL] Reply back email (required for a response back to any questions you submit above):

**Figure 9 Informed Consent Form.****Online Consent Form for the Experimental Study “How well do you know yourself?”**

**Background Information:** You are invited to participate in a research study about how individuals and groups of individuals make decisions in a variety of economic contexts. We request that you read the information on this page carefully before agreeing to be in the study. You were selected as a possible participant because you are a part of the Debra Paget and Jeffrey Berg Business Simulation Laboratory (“The Lab”), which is a Cornell University Institutional Review Board for Human Participants (IRB) approved recruitment system.

**Procedures and Compensation:** This survey will last approximately 15-20 minutes, and you will answer a variety of economic questions and take a self-assessment on your personality. After the data collection period for this study, you will receive any earnings you’ve gained in the form of an electronic Amazon gift card, delivered to the email address you used to register for this study. If the survey states you have earned a certain amount of money, then this is real money that will be added towards your gift card. You will be paid \$2 for just participating in this study (once you acknowledge this information). Then, depending on your decisions in the economic questions only, you can earn additional money. Your earnings will not depend on your self-assessment results or any personal information.

**Voluntary Nature of Participation:** Your participation in this study is strictly voluntary. At any point in this survey, you may leave and discontinue participation.

**Risks and Benefits of Participating in the Study:** The risk for participating in this experiment is minimal. You have no greater physical, financial, or psychological risk from the experiment than you would from doing a similar amount of routine paperwork or computer-based activity in any Cornell University classroom. There are no substantial benefits to you from the research, other than the money you will earn by participating. If you are a registered student at Cornell University, there will be no extra-credit or class credit given for participating in this experiment, and your results from the experiment will not impact your performance in any class. By learning more about people’s decision-making, we hope that the research will benefit society by helping economic institutions understand people’s behavior.

**Confidentiality:** All decisions during the experiment can be kept confidential should you choose to remain anonymous. Should you choose to participate in the study, you will have the voluntary choice to reveal some of your decisions to other participants in this study, including your first and last name. However, your email, education, age, sex, and any other personal identifying information will be kept strictly confidential. Please note that if you were recruited for this experiment via e-mail there is a chance that the information you communicated could be read by a third party. De-identified data from this study may be shared with the research community at large to advance science and health. We will remove or code any personal information that could identify you before files are shared with other researchers to ensure that, by current scientific standards and known methods, no one will be able to identify you from the information we share. Despite these measures, we cannot guarantee anonymity of your personal data.

**Contacts and Questions:** After the experiment, Zhouyu Wu (zw369@cornell.edu; 224-715-3153) will be glad to answer any questions that you may have. You may contact the Cornell University Institutional Review Board for Human Participants (IRB) at 607-255-6182. The Cornell University IRB website is <http://www.irb.cornell.edu>. You may also report your concerns or complaints anonymously through Ethicspoint online at [www.hotline.cornell.edu](http://www.hotline.cornell.edu) or by calling toll free at 1-866-293-3077. Ethicspoint is an independent organization that serves as a liaison between the University and the person bringing the complaint so that anonymity can be ensured.

- ☐ I understand the information above and agree to participate in this study.
- ☐ I am 18 years old or older.

*This consent form was approved by IRB on August 27, 2018.*



**Figure 10 Survey Self-Assessment Questions**

On the next page you will take a self-assessment, based on 50 statements. These statements will be spread out over 5 pages. Read all statements with careful consideration and determine how accurately these statements describe yourself.

Your responses for each statement will NOT determine your earnings in this study.

NEXT →

----- Page Break -----

**Part 1 of 5**

	Very Inaccurate	Moderately Inaccurate	Neither Accurate Nor Inaccurate	Moderately Accurate	Very Accurate
I am the life of the party.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel little concern for others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am always prepared.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I get stressed out easily.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have a rich vocabulary.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I don't talk a lot.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am interested in people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I leave my belongings around.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am relaxed most of the time.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have difficulty understanding abstract ideas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

NEXT →

----- Page Break -----

**Part 2 of 5**

	Very Inaccurate	Moderately Inaccurate	Neither Accurate Nor Inaccurate	Moderately Accurate	Very Accurate
I feel comfortable around people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I insult people.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I pay attention to details.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I worry about things.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have a vivid imagination.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I keep in the background.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I sympathize with others' feelings.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I make a mess of things.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I seldom feel blue.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am not interested in abstract ideas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

NEXT →

----- Page Break -----

Figure 11 Survey Self-Assessment Questions (Cont'd.)

	Very Inaccurate	Moderately Inaccurate	Neither Accurate Nor Inaccurate	Moderately Accurate	Very Accurate
I start conversations.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am not interested in other people's problems	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I get chores done right away.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am easily disturbed.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have excellent ideas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have little to say.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have a soft heart.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I often forget to put things back in their proper place.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I get upset easily.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I do not have a good imagination.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

NEXT →

----- Page Break -----

	Very Inaccurate	Moderately Inaccurate	Neither Accurate Nor Inaccurate	Moderately Accurate	Very Accurate
I talk to a lot of different people at parties.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am not really interested in others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I like order.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I change my mood a lot.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am quick to understand things.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I don't like to draw attention to myself.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I take time out for others.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I shirk my duties.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have frequent mood swings.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I use difficult words.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

NEXT →

----- Page Break -----

	Very Inaccurate	Moderately Inaccurate	Neither Accurate Nor Inaccurate	Moderately Accurate	Very Accurate
I don't mind being the center of attention.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I feel others' emotions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I follow a schedule.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I get irritated easily.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I spend time reflecting on things.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am quiet around strangers.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I make people feel at ease.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am exacting in my work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I often feel blue.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I am full of ideas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

NEXT →

**Figure 12 Survey Information on User-Generated Self-Assessment Scores**

**Thank you for completing your self-assessment!**  
Based on your responses, your core personality has been measured across five factors: Extraversion, Agreeableness, Conscientiousness, Emotional Stability, and Intellect. Each factor score is measured on a scale from 10 to 50.

First Name	Last Name	Extraversion	Agreeableness	Conscientiousness	Emotional Stability	Intellect
Jane	Doe	19	30	31	42	48

Note: Each factor is measured on a scale from 10 to 50.

The 50-item questionnaire you completed is a widely used personality measure based on the Five Factor Model. Extensive research has been done to relate these questions to behavioral and psychological phenomena. The Five-Factors (or “Big 5”) are set of essential traits fundamental to your core personality. Each trait is measured across a spectrum of extremes. For example, a low score on extraversion would mean high introversion.

- **Extraversion (or surgency):** Measures assertive, energetic, or outgoing behaviors. A high score indicates high extraversion, and a low score indicates low extraversion.
- **Agreeableness:** Measures empathy, sympathy, and kindness. A low score indicates low agreeableness, and a high score indicates high agreeableness.
- **Conscientiousness:** Measures your sense of responsibility, duty, and foresight. A low score indicates low conscientiousness, and a high score indicates high conscientiousness.
- **Emotional stability (or neuroticism):** Measures irritability and moodiness. High scores indicate high emotional stability (low neuroticism), low scores indicate low emotional stability (high neuroticism).
- **Intellect (or imagination):** Measures inquisitiveness, openness to new experience, thoughtfulness, and propensity for intellectually challenging tasks. High scores indicate high intellect.

**Figure 13 Instruction page for data-sharing choices.**

**You and all other participants in this study will be choosing whether to share personal data from this survey.** This data includes first name, last name, and self-assessment scores generated from the same quiz you just took (displayed in the table below).

First Name	Last Name	Extraversion	Agreeableness	Conscientiousness	Emotional Stability	Intellect
Jane	Doe	19	31	21	42	48

Over the next few pages, **you will be shown different possible scenarios for sharing this data.** Each scenario asks whether or not you would accept certain amounts of money for releasing your data. It is in your best interest to answer honestly, as **one of these scenarios will be randomly selected and made real.**

Any data you and other participants release will be sent together in an email to the recipient(s) approximately 1 week after the survey’s participation deadline.

When you are ready, please continue.

NEXT →

*Note.* Respondent’s name and score from self-assessment populated in data table.

**Table 3** Survey Text Used for Each Data Sharing Condition

Condition	Survey Text
$[1, N]$	“One participant is randomly selected to receive personal data released from you and other participants.”
$[30, N]$	“Thirty participants are randomly selected to receive personal data released from you and other participants.”
$[1, P]$	“One participant is randomly selected to receive personal data released from you and other participants. If this participant has your data, they can use your data to make money. The more data they have from participants in this study, the more money they can make.”
$[1, F]$	“One participant is randomly selected to receive personal data released from you and other participants. If this participant has your data, they can use your data to make money by selling it to a third party. The more data they have from participants in this study, the more money they can make.”
$[1, F']$	“One participant is randomly selected to receive personal data released from you and other participants. If this participant has your data, they can use your data to make money by selling it to another participant. The more data they have from participants in this study, the more money they can make.”
$[30, F]$	“Thirty participants are randomly selected to receive personal data released from you and other participants. If these participants have your data, they can each use your data to make money by selling it to a third party. The more data they have from participants in this study, the more money they can make.”

**Table 4** Randomization Group and Condition Orders

Sample	Group	Conditions (in order)			
		$t = 1$	$t = 2$	$t = 3$	$t = 4$
1	1	$[30, N]$	$[1, N]$	$[1, P]$	$[1, F]$
1	2	$[1, N]$	$[30, N]$	$[1, F]$	$[1, P]$
1	3	$[1, P]$	$[1, F]$	$[30, N]$	$[1, N]$
1	4	$[1, F]$	$[1, P]$	$[1, N]$	$[30, N]$
2	5	$[30, N]$	$[1, N]$	$[30, F]$	$[1, F]$
2	6	$[1, N]$	$[30, N]$	$[1, F]$	$[30, F]$
2	7	$[30, F]$	$[1, F]$	$[30, N]$	$[1, N]$
2	8	$[1, F]$	$[30, F]$	$[1, N]$	$[30, N]$
3	9	$[30, N]$	$[1, N]$	$[1, F']$	$[1, F]$
3	10	$[1, N]$	$[30, N]$	$[1, F]$	$[1, F']$
3	11	$[1, F']$	$[1, F]$	$[30, N]$	$[1, N]$
3	12	$[1, F]$	$[1, F']$	$[1, N]$	$[30, N]$

**Figure 14 Survey Question for Eliciting Reservation Prices for Data-Sharing**

For each of the possible prices below, please indicate whether you would ‘accept’ and release your data, or ‘not accept’ and not release your data.

	I would accept	I would not accept
\$0.01	<input type="radio"/>	<input type="radio"/>
\$0.49	<input type="radio"/>	<input type="radio"/>
\$0.99	<input type="radio"/>	<input type="radio"/>
\$1.99	<input type="radio"/>	<input type="radio"/>
\$2.99	<input type="radio"/>	<input type="radio"/>

If this scenario is made real, the computer will choose one of the prices. If you selected ‘I would accept’ at that price, then we will release your data (under this scenario’s conditions), and you will earn the price. If you selected ‘I would not accept’ at that price, then we will not release your data, and you will not earn the price.

**Table 5 Survey collection details.**

Sample 1	Wave 1	Wave 2	Wave 3	Wave 4	Total
Dates (in 2019)	3/21-3/24	3/28-3/30	4/11-4/14	4/18-4/20	
Start day (time)	Thu (8:30 AM)	Thu (10:09 AM)	Thu (11:31 AM)	Thu (8:55 AM)	
End day (time)	Sat (11:59 PM)	Sat (11:59 PM)	Sat (11:59 PM)	Sat (11:59 PM)	
# Available	200	100	200	100	600
# Registered	200	26	156	91	473
# Incompleted	0	1	0	2	3
# Completed	187	18	128	80	413
Sample 2	Wave 5	Wave 6	Wave 7	Wave 8	Total
Dates (in 2019)	10/31-11/02	11/07-09	11/14-16	11/21-23	
Start day (time)	Thu (8:00 AM)	Thu (8:00 AM)	Thu (8:00 AM)	Thu (8:00 AM)	
End day (time)	Sat (11:59 PM)	Sat (11:59 PM)	Sat (11:59 PM)	Sat (11:59 PM)	
# Available	200	100	200	100	600
# Registered	200	100	147	69	516
# Incompleted	0	0	3	3	6
# Completed	170	87	113	50	420
Sample 3	Wave 9	Wave 10	Wave 11	Wave 12	Total
Dates (in 2020)	5/21-23	5/28-30	6/4-6	6/11-13	
Start day (time)	Thu (8:00 AM)	Thu (8:00 AM)	Thu (8:00 AM)	Thu (8:00 AM)	
End day (time)	Sat (11:59 PM)	Sat (11:59 PM)	Sat (11:59 PM)	Sat (11:59 PM)	
# Available	200	100	200	100	600
# Registered	173	100	70	90	433
# Incompleted	0	1	1	0	2
# Completed	151	90	55	58	355

*Note.* Registration opened between Sundays and Wednesdays prior to survey start time. Email advertisements were sent out by the Lab on Mondays. Incomplete (i.e., responses started but not completed) surveys did *not* receive payments and are removed from the data.

**Table 6** Data Market Participation and Price Results (Pooled, Full Table)

Variables Dependent:	Logit			OLS		
	Participation			Price (\$)	Participation=1	
Comparison: [1, N]	(1a)	(2a)	(3a)	(1b)	(2b)	(3b)
(Intercept)	1.289*** (0.113)	1.474*** (0.124)	2.107*** (0.378)	1.499*** (0.061)	1.399*** (0.071)	1.154*** (0.208)
[30, N]	-0.433*** (0.054)	-0.434*** (0.054)	-0.442*** (0.055)	0.102*** (0.024)	0.102*** (0.024)	0.102*** (0.024)
[1, F]	-1.018*** (0.064)	-1.016*** (0.064)	-1.039*** (0.065)	0.390*** (0.033)	0.393*** (0.033)	0.393*** (0.033)
N in $t \in \{1, 2\}$		-0.369*** (0.108)	-0.387*** (0.110)		0.183** (0.066)	0.196** (0.066)
<i>Extravert</i>			0.098 (0.110)			0.078 (0.066)
<i>Agreeable</i>			-0.399 <sup>†</sup> (0.235)			0.226 <sup>†</sup> (0.127)
<i>Conscientious</i>			-0.571*** (0.158)			0.134 <sup>†</sup> (0.082)
<i>Emostable</i>			0.138 (0.113)			-0.103 (0.067)
<i>Intellectual</i>			-0.156 (0.161)			0.034 (0.093)
<i>Female</i>			-0.126 (0.128)			-0.076 (0.074)
<i>Single</i>			0.449** (0.169)			-0.032 (0.110)
<i>Employed</i>			-0.215 (0.132)			-0.091 (0.081)
<i>Facebooker</i>			0.152 (0.142)			0.012 (0.083)
<i>Sample 2</i>	0.332** (0.129)	0.336** (0.129)	0.314* (0.131)	-0.051 (0.077)	-0.048 (0.077)	-0.036 (0.077)
<i>Sample 3</i>	-0.015 (0.131)	-0.014 (0.132)	0.001 (0.135)	0.000 (0.081)	-0.002 (0.080)	-0.010 (0.080)
<i>Decision t = 2</i>	0.110 (0.068)	0.107 <sup>†</sup> (0.065)	0.111 <sup>†</sup> (0.066)	0.008 (0.031)	0.007 (0.031)	0.007 (0.031)
<i>Decision t = 3</i>	-0.161* (0.074)	-0.160* (0.074)	-0.170* (0.075)	0.006 (0.038)	0.022 (0.038)	0.023 (0.038)
<i>Decision t = 4</i>	-0.197** (0.074)	-0.196** (0.074)	-0.205** (0.076)	-0.003 (0.039)	0.013 (0.040)	0.013 (0.040)
Individual clusters	1188	1188	1188	975	975	975
Observations	3564	3564	3564	2472	2472	2472
<b>ANOVA: Wald Test</b>	(1a),(2a)	(2a),(3a)	(1a),(3a)	(1b),(2b)	(2b),(3b)	(1b),(3b)
$Pr(> \chi^2)$	0.001	0.000	0.000	0.006	0.192	0.030

Note: Clustered robust standard errors in parentheses.

<sup>†</sup> $p < 0.1$ ; \* $p < 0.05$ ; \*\* $p < 0.01$ ; and \*\*\* $p < 0.001$

**Table 7** Participation and Price Results (Sample 1, Full Table)

Variables Dependent:	Logit <i>Participation</i>			OLS <i>Price (\$)</i>   <i>Participation = 1</i>		
	(1a)	(2a)	(3a)	(1b)	(2b)	(3b)
Comparison: $[1, N]$						
(Intercept)	1.317*** (0.138)	1.431*** (0.160)	1.031 <sup>†</sup> (0.571)	1.472*** (0.070)	1.373*** (0.091)	1.504*** (0.326)
$[30, N]$	-0.425*** (0.083)	-0.425*** (0.083)	-0.439*** (0.085)	0.092* (0.039)	0.091* (0.039)	0.090* (0.039)
$[1, P]$	-0.682*** (0.097)	-0.678*** (0.097)	-0.699*** (0.100)	0.252*** (0.050)	0.253*** (0.050)	0.253*** (0.050)
$[1, F]$	-1.046*** (0.108)	-1.045*** (0.107)	-1.078*** (0.110)	0.428*** (0.058)	0.429*** (0.058)	0.427*** (0.058)
$N$ in $t \in \{1, 2\}$		-0.243 (0.179)	-0.276 (0.182)		0.195 <sup>†</sup> (0.107)	0.190 <sup>†</sup> (0.106)
<i>Extravert</i>			-0.200 (0.184)			0.217* (0.107)
<i>Agreeable</i>			-0.213 (0.354)			-0.041 (0.191)
<i>Conscientious</i>			-0.410 (0.265)			0.139 (0.145)
<i>EmoStable</i>			0.326 <sup>†</sup> (0.189)			-0.066 (0.111)
<i>Intellectual</i>			0.402 (0.267)			-0.065 (0.178)
<i>Female</i>			-0.092 (0.218)			-0.078 (0.120)
<i>Single</i>			0.658* (0.276)			-0.076 (0.173)
<i>Employed</i>			-0.018 (0.225)			-0.229 <sup>†</sup> (0.135)
<i>Facebooker</i>			0.428 <sup>†</sup> (0.227)			-0.211 (0.132)
<i>Decision</i> $t = 2$	0.080 (0.075)	0.080 (0.075)	0.085 (0.077)	0.032 (0.042)	0.031 (0.042)	0.031 (0.042)
<i>Decision</i> $t = 3$	-0.142 (0.095)	-0.127 (0.095)	-0.129 (0.098)	0.071 (0.054)	0.074 (0.054)	0.074 (0.054)
<i>Decision</i> $t = 4$	-0.243* (0.100)	-0.231* (0.100)	-0.238* (0.103)	0.039 (0.054)	0.044 (0.055)	0.043 (0.055)
Individual clusters	413	413	413	334	334	334
Observations	1652	1652	1652	1094	1094	1094
<b>ANOVA: Wald Test</b>	(1a),(2a)	(2a),(3a)	(1a),(3a)	(1b),(2b)	(2b),(3b)	(1b),(3b)
$Pr(> \chi^2)$	0.174	0.055	0.054	0.068	0.109	0.057

Note: Clustered robust standard errors in parentheses.

<sup>†</sup> $p < 0.1$ ; \* $p < 0.05$ ; \*\* $p < 0.01$ ; and \*\*\* $p < 0.001$

**Table 8** Participation and Price Results (Sample 2, Full Table)

Variables	Logit			OLS		
	<i>Participation</i>			<i>Price (\$)</i>   <i>Participation=1</i>		
Dependent:						
Comparison: $[1, N]$	(1a)	(2a)	(3a)	(1b)	(2b)	(3b)
(Intercept)	1.554*** (0.149)	1.882*** (0.175)	3.819*** (0.634)	1.412*** (0.064)	1.330*** (0.087)	0.632* (0.294)
$[30, N]$	-0.395*** (0.098)	-0.401*** (0.100)	-0.418*** (0.103)	0.069† (0.036)	0.069† (0.036)	0.070† (0.036)
$[1, F]$	-1.007*** (0.115)	-1.016*** (0.114)	-1.073*** (0.118)	0.430*** (0.052)	0.431*** (0.052)	0.432*** (0.052)
$[30, F]$	-1.079*** (0.118)	-1.090*** (0.119)	-1.153*** (0.123)	0.395*** (0.053)	0.397*** (0.053)	0.397*** (0.052)
$N$ in $t \in \{1, 2\}$		-0.668*** (0.190)	-0.674*** (0.195)		0.166 (0.110)	0.174 (0.109)
<i>Extravert</i>			0.319 (0.195)			-0.085 (0.110)
<i>Agreeable</i>			-1.061* (0.463)			0.555** (0.188)
<i>Conscientious</i>			-1.011*** (0.277)			0.157 (0.123)
<i>EmoStable</i>			-0.040 (0.197)			-0.043 (0.109)
<i>Intellectual</i>			-0.493† (0.279)			0.103 (0.139)
<i>Female</i>			-0.009 (0.227)			-0.114 (0.126)
<i>Single</i>			0.345 (0.290)			0.107 (0.184)
<i>Employed</i>			-0.288 (0.239)			0.027 (0.138)
<i>Facebooker</i>			0.019 (0.256)			0.151 (0.139)
<i>Decision</i> $t = 2$	0.218** (0.083)	0.216** (0.082)	0.226** (0.086)	0.029 (0.034)	0.029 (0.034)	0.029 (0.034)
<i>Decision</i> $t = 3$	-0.115 (0.106)	-0.052 (0.107)	-0.072 (0.111)	0.091† (0.051)	0.095† (0.052)	0.096† (0.052)
<i>Decision</i> $t = 4$	-0.147 (0.106)	-0.085 (0.107)	-0.109 (0.112)	0.053 (0.055)	0.057 (0.056)	0.058 (0.056)
Individual clusters	420	420	420	359	359	359
Observations	1680	1680	1680	1186	1186	1186
<b>ANOVA: Wald Test</b>	(1a),(2a)	(2a),(3a)	(1a),(3a)	(1b),(2b)	(2b),(3b)	(1b),(3b)
$Pr(> \chi^2)$	0.000	0.000	0.000	0.131	0.150	0.100

Note: Clustered robust standard errors in parentheses.

†  $p < 0.1$ ; \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; and \*\*\*  $p < 0.001$



**Table 9** Participation and Price Results (Sample 3, Full Table)

Variables Dependent: Comparison: $[1, N]$	Logit <i>Participation</i>			OLS <i>Price(\$)</i>   <i>Participation=1</i>		
	(1a)	(2a)	(3a)	(1b)	(2b)	(3b)
(Intercept)	1.284*** (0.144)	1.364*** (0.172)	1.978** (0.761)	1.551*** (0.071)	1.425*** (0.095)	1.365** (0.415)
$[30, N]$	-0.484*** (0.102)	-0.483*** (0.102)	-0.492*** (0.103)	0.168*** (0.048)	0.168*** (0.048)	0.168*** (0.048)
$[1, F']$	-0.789*** (0.103)	-0.785*** (0.102)	-0.800*** (0.104)	0.390*** (0.057)	0.392*** (0.057)	0.391*** (0.057)
$[1, F]$	-0.998*** (0.113)	-0.996*** (0.113)	-1.015*** (0.114)	0.362*** (0.062)	0.364*** (0.062)	0.362*** (0.062)
$N$ in $t \in \{1, 2\}$		-0.172 (0.191)	-0.176 (0.194)		0.243* (0.118)	0.243* (0.118)
<i>Extravert</i>			0.115 (0.200)			0.138 (0.116)
<i>Agreeable</i>			-0.022 (0.502)			0.080 (0.289)
<i>Conscientious</i>			-0.282 (0.287)			0.031 (0.156)
<i>EmoStable</i>			0.123 (0.207)			-0.247* (0.122)
<i>Intellectual</i>			-0.234 (0.276)			0.040 (0.167)
<i>Female</i>			-0.412 <sup>†</sup> (0.231)			-0.056 (0.128)
<i>Single</i>			0.175 (0.326)			-0.003 (0.198)
<i>Employed</i>			-0.363 (0.227)			-0.104 (0.138)
<i>Facebooker</i>			-0.065 (0.269)			0.229 (0.161)
<i>Decision</i> $t = 2$	0.133 (0.095)	0.132 (0.094)	0.138 (0.096)	-0.086* (0.043)	-0.086* (0.043)	-0.087* (0.043)
<i>Decision</i> $t = 3$	-0.157 (0.105)	-0.147 (0.104)	-0.150 (0.106)	-0.043 (0.059)	-0.037 (0.059)	-0.037 (0.059)
<i>Decision</i> $t = 4$	-0.232* (0.107)	-0.223* (0.106)	-0.228* (0.108)	-0.120* (0.061)	-0.114 <sup>†</sup> (0.061)	-0.115 <sup>†</sup> (0.061)
Individual clusters	355	355	355	286	286	286
Observations	1420	1420	1420	926	926	926
<b>ANOVA: Wald Test</b>	(1a),(2a)	(2a),(3a)	(1a),(3a)	(1b),(2b)	(2b),(3b)	(1b),(3b)
$Pr(> \chi^2)$	0.368	0.495	0.523	0.040	0.538	0.188

Note: Clustered robust standard errors in parentheses.

<sup>†</sup> $p < 0.1$ ; \* $p < 0.05$ ; \*\* $p < 0.01$ ; and \*\*\* $p < 0.001$