Privacy with Information Externalities and Complexity*

Sunny Qingyi Wang Eytan Adar Yan Chen

October 17, 2025

Abstract

Privacy decisions in social communication are complex in that they often affect not only the decision maker but also those in their social network. This study investigates people's information sharing behavior with the presence of externalities, using an information sharing game in the lab, in real life involving friends, and on a social media platform. We find that participants over-share relative to the Bayesian Nash equilibrium prediction in the game and that externalities increase the likelihood of sharing in both the game and real-life information sharing involving friends. When we exogenously reduce computational complexity in the game by providing an average payoff matrix, participants are 19.4 pp more likely to best respond, leading to a 24.4 pp increase in efficiency. Our results highlight the potential to improve users' welfare by reducing the complexity in their privacy decisions.

Keywords: privacy, information sharing, externalities, complexity, social network

JEL Codes: C91, D83, D91

*We thank Gonzalo Arrieta, Björn Bartling, Roberto Weber, Malte Dold, and audiences at the University of Michigan, University of Zürich, the 2024 North American ESA Meeting (Columbus, OH), the SABE 2025 Annual Conference (Trento Italy), and the 17th Maastricht Behavioral and Experimental Economics Symposium (Maastricht, Netherlands) for helpful discussions and comments. Yuri Chen and Ally Vern provided excellent research assistance. Financial support from the National Science Foundation through grant no. 2209507 to Chen and Adar is gratefully acknowledged. This research was approved by the IRB of a large public university in the US (HUM#00236045) and pre-registered at the AEA RCT Registry (AEARCTR-0012875). Wang, Adar and Chen: School of Information, University of Michigan, 2200 Hayward Street, Ann Arbor, MI 48109. Emails: wqingyi@umich.edu, eadar@umich.edu, yanchen@umich.edu.

1 Introduction

In 2024, an average social media user spends 2.4 hours per day across seven unique social platforms (Aridor et al., 2024; Dixon, 2024). Every click, keystroke, and file upload reveals personal information to the platform or a third party. Control of this information, when available, requires users to navigate a maze of privacy settings, disclosure choices, and trust judgments. Not only are users faced with *more* privacy decisions, but those decisions are also becoming *increasingly complex*. One reason for the complexity is the externalities associated with individual information sharing. As distributed data becomes more central to digital platforms, data collected from a single user can reveal information not only about that user but also about others with similar characteristics or behaviors (Bergemann et al., 2022).

Personal information sharing often imposes negative externalities on others. When an individual discloses personal genetic data, it can inadvertently reveal information about their relatives. For example, if a user submits her DNA to 23andMe and learns that she carries a pathogenic BRCA1/2 mutation associated with hereditary breast cancer, that finding would also have implications for her daughters, siblings, and first cousins, even though they themselves never share any of their genetic information.

Poorly calibrated sharing decisions in environments with negative externalities can have serious consequences. For example, oversharing on social media reveals locations and routines that cybercriminals exploit to track and harass victims, making these platforms a common tool for cyberstalking. From 2010 to 2020, 43% of federal cyberstalking cases involved stalkers using these platforms (Adamson et al., 2023).

Information sharing can also create positive externalities: in product recommendation communities, collaborative filtering uses a user's ratings and those of similar "neighbors" to generate personalized recommendations, improving quality for all (Ricci et al., eds, 2022). Thus, ratings function as an impure public good (Chen et al., 2010).

Personal information sharing might generate both additional risks and benefits for others. However, do people account for this in privacy decisions? Understanding when and why they do (not) is essential for better policy and design.

One challenge users face in accounting for externalities is that it complicates the computation of the benefits and costs of information sharing. Prior research shows that even in environments without externalities, users face cognitive limitations that affect their ability to manage trade-offs associated with sharing personal information (Solove, 2012; Acquisti and Grossklags, 2007). As privacy decisions encompass more information and involve more entities, the cognitive load on users increases further.

While it is widely acknowledged that complexity poses significant challenges to privacy decision-making, there is limited empirical research assessing its impact of complexity on privacy decisions when externalities are present. This is what we set out to do.

Social media is a natural environment to study this question, where personal information disclosure might impact one's friends, and the implications of different privacy settings are complex to navigate. Among our experiment participants, for example, 45.7% use the default setting on Instagram, which exposes their activity information on third-party platforms to Instagram for ads personalization. Why participants pick the default is multifaceted. Among other reasons, they may undervalue privacy or struggle to estimate downstream costs to themselves and friends.

To disentangle the effects of computational complexity and externalities in privacy decision-making, we run a lab experiment with three modules: (i) a survey on Instagram; (ii) a friends decision making module in which each participant shares real-life information about themselves and their friends; and (iii) an abstract information sharing game that exogenously varies computational complexity and the magnitude of negative externalities to assess their impact on sharing decisions. Moving from realistic modules to the abstract game increases control and inferential precision, while cross-module correlations sharpen our ability to identify predictors of social media privacy choices.

The main module of the study is an information sharing game designed based on a simplified version of the theoretical model of Acemoglu et al. (2022). Participants are randomly matched into pairs at the beginning of each round to make information sharing decisions. Their type (private information) is randomly drawn from a joint normal distribution and correlated within each pair. Participants decide simultaneously and independently whether to share their private information with a social networking platform, represented by a robot in the experiment. To study the role of computational complexity, we implement a within-subject treatment variation which simplifies the computation of the benefit and cost of sharing. The controlled setting allows us to define optimal sharing decisions and measure the impact of computational complexity, which is challenging in the real world.

We find that, in the control condition, when participants compute benefits and costs on their own, they over-share relative to the equilibrium prediction. Our comparative statics analysis reveals that, in the control condition, participants react strongly to the benefit of sharing while ignoring the cost coming from type correlations, i.e., externalities. By contrast, the reduction in computational complexity in the matrix treatment increases their likelihood of best response by 19.5 percentage points (pp), leading to a 20.5 pp increase in efficiency. Indeed, with reduced computational complexity in the treatment, the type correlation has a sizable and significant impact on their sharing decision, suggesting that participants are able to take the externalities of their sharing decision into consideration in a less complex environment.

To explore whether privacy preferences measured in the abstract information sharing game predict individual's privacy preferences in more realistic settings, we implement a friends decision making module, with a similar underlying type correlation structure to that of the information sharing game. In this module, participants are asked to share private and sensitive information about themselves and their friend, such as body weight and the number of sexual partners, with other participants in the lab. Similar to the information sharing game in the control condition, the benefits of sharing are easy to grasp (direct monetary compensation for sharing the information) whereas the cost of sharing on a participant and their friend is more complex. In this module, consistent with theoretical predictions and the results from the information sharing game, participants are more willing to share when externalities (type correlation) are present.

Investigating cross-module correlations, we find that sharing decisions in the control condition of the information sharing game strongly predict privacy preferences in the friends decision making module. The former likely reflects an individual's true ability to manage the trade-offs of information sharing, as they have to compute the benefit and cost without any assistance to reduce computational complexity. In other words, the decision environment in the control condition is closer to real-world privacy decision making environments.

Lastly, we investigate the predictive power of information sharing behaviors in the lab on social media privacy choices. We do so by collecting participants' Instagram privacy settings through an incentivized survey. Again, fully understanding the terms and future consequences of various privacy settings takes time and effort, whereas ignoring them or using the default setting has immediate benefits and uncertain costs in the future. We find that those who are more willing to share in the friends decision making module tend to have more lenient privacy settings on Instagram. These cross-module correlations suggest that sharing decisions elicited in neutral contexts can serve as a reliable proxy for understanding users' privacy preferences, especially when direct observation of real-world choices is difficult.

The study contributes to our understanding of the impact of computational complexity on user privacy decisions, particularly in environments where one person's sharing decisions have negative externalities on others. Our results demonstrate that (1) without any assistance to reduce computational complexities, users over-share as they ignore the cost of sharing coming from externalities; (2) reducing computational complexity significantly improves the quality of their information sharing decisions; and (3) strong cross-module correlations suggest that a reduction in computational complexity in real-world privacy decisions is likely to improve the quality of such decisions as well.

These findings suggest that platform designers or regulators need to be explicit about both the benefits and especially the costs of sharing personal information, including pointing out the negative externalities that potential shared information could impose on others. One such example in a different context is the Surgeon General's Warning on cigarette packs.[1] These measures have the potential to improve the quality of privacy decisions and protect users.

2 Literature Review

Our research builds upon and contributes to two streams of literature, the economics of privacy and the role of complexity in decision-making. For a review of the economic analysis on privacy, see Acquisti et al. (2016). See also Goldfarb and Que (2023) and Acquisti et al. (2020) for reviews of privacy research with a focus on the digital economy.

The theoretical foundation of privacy as an economic concept begins with models proposed by scholars from the Chicago School. Early work by Stigler (1961, 1962, 1980) conceptualize privacy as a public good, where information can be simultaneously used and disseminated by owners, buyers, and third parties. Posner (1978, 1981) extend this perspective by emphasizing spillover effects of individual sharing decisions, arguing that individuals' concealment of information may impose harms on others and thus lead to inefficiencies. These early models recognize the externalities of information sharing, but focus primarily on how one party's transmission of information affects other entities in the communication chain.

While the models point out that the positive externalities of information sharing may be underappreciated, later research underscores that bad consequences may also arise due to negative externalities. For example, firms may sell consumer data to third parties, resulting in price discrimination against consumers (Odlyzko, 2003). Yet neither firms nor consumers fully internalize these negative externalities when making sharing decisions (Swire and Litan, 2010). One proposed solution is to create information markets in which such externalities can be internalized through pricing mechanisms (Laudon, 1997). These studies highlight the challenges that consumers and firms face as a result of externalities. However, the literature at this stage continues to concentrate primarily on interactions between individual consumers, firms, and third parties, with relatively little attention to how consumer decisions affect peers.

Several later studies turn their attention to the spillover effects of one consumer's disclosure on others. Goh et al. (2015) provide early empirical evidence using the U.S. National Do Not Call (DNC) Registry, showing that when some consumers opt out of telemarketing, those outside the Registry receive more such calls, which in turn leads to more subsequent opting out. This illustrates how decision makers adjust their behavior in response to reduced information sharing by others. Conversely, when others disclose more information to gain a competitive advantage in a three-player distribution game, decision makers may feel pressured to reveal more of their

[1] One version of the Surgeon General's Warning on cigarette packs points out the negative externalities: Smoking By Pregnant Women May Result in Fetal Injury, Premature Birth, And Low Birth Weight.

own information (Ackfeld and Güth, 2023). Moreover, when the negative externalities of one's information sharing behavior on others are made explicit, individuals become less willing to sell their data (Friehe et al., 2025). While these studies demonstrate important spillovers of individual actions, they primarily emphasize the direct effects of sharing behavior, rather than the correlations across individuals' information.

Building on these insights, more recent contributions develop richer theoretical frameworks for understanding privacy with externalities, placing greater emphasis on the role of correlations across individuals' information. Fairfield and Engel (2015) develop a comprehensive framework for analyzing privacy applying public goods theory. They propose strategies for privacy protection that allow groups to maintain cooperation and preserve collective privacy. Two more recent models characterize the conditions under which excessive information sharing occurs and explore policy interventions to improve market outcomes (Choi et al., 2019; Acemoglu et al., 2022). Complementing these perspectives, Bergemann et al. (2022) incorporate the role of data intermediaries alongside platforms and users, showing how intermediation alters the nature of externalities. Distinct from these studies, Ichihashi (2021) solves an information design problem, in which the firm chooses what information to buy from consumers, balancing the value and price of information.

There is a large empirical literature on privacy, primarily published outside of economics (Smith et al., 2011). Earlier privacy research often relies on self-reported survey measures, such as the Harris-Equifax Consumer Privacy Surveys (for a review, see Kumaraguru and Cranor (2005)). As research on the privacy paradox expanded (see Kokolakis (2017) for a review), scholars increasingly recognized a gap between individuals' stated privacy attitudes and their actual behaviors. In response, more recent studies begin to incentivize real-world information sharing behavior to better capture privacy preferences. One common approach is to compare consumers' willingness to share under different levels of perceived privacy risks. This can be implemented by manipulating either the level of privacy protection provided for shared information (Tsai et al., 2011), the intrusiveness of survey questions (Brandimarte et al., 2013; John et al., 2011; Moon, 2000), or the amount of personal information required to access a product or service (Acquisti et al., 2013; Biener et al., 2020).

Another approach is to directly measure participants' monetary valuation of private information. A key challenge of doing so in controlled settings is to simulate privacy concerns that feel meaningful to participants. The pioneering study by Huberman et al. (2005) addresses this challenge by using a second-price auction to elicit individual valuations for personal data. Participants submit bids indicating the minimum amount they would accept to reveal sensitive personal information (e.g., body weight) to the other participants in the same session. Building on this approach, subsequent studies elicit individuals' willingness to share personal information with strangers (Schudy and Utikal, 2017; Lee and Weber, 2025) or with third-party companies (Benndorf and Normann, 2018; Lin, 2022; Prince and Wallsten, 2022), primarily using the Becker-DeGroot-Marschak mechanism (Becker et al., 1964).

Our paper contributes to a growing literature in economics that investigates how computational complexity shapes human decision-making. Early theoretical work highlights that the difficulty of computing expected utility can prevent individuals from making optimal choices (Gilboa, 1988; Ben-Porath, 1990). Several recent studies demonstrate that complexity may distort the elicitation of preferences (Bernheim and Sprenger, 2020; Enke and Shubatt, 2023; Frydman and Jin, 2022; Oprea, 2024; Puri, 2025). When decisions become cognitively demanding, individuals may rely on heuristics or exhibit inconsistent choices, so complexity becomes the primary driver of observed behavior rather than true preferences. As a result, elicited preferences become noisier and less reliable. Complexity also leads to more suboptimal decisions in strategic contexts. The inability to identify optimal responses under complexity can lead to systematic deviations from equilibrium behavior (Enke and Graeber, 2023; Li, 2017; Shubatt and Yang, 2024).

Our study contributes to the literature in the following ways. First, while there is a growing body of theoretical work modeling privacy decisions under externalities, empirical evidence remains limited. This paper seeks to fill the gap. Second, this study builds on prior work by employing incentivized measures to elicit individuals' privacy preferences. In addition, we collect information sharing decisions across both neutral and realistic contexts. This variation in contextual realism allows us to assess how well behavior observed in lab experiments predicts privacy decisions in real-world environments. Third, we extend the study of complexity in decision making to the domain of privacy by quantifying its impact on strategic information sharing behaviors. This contributes to the complexity literature by providing new evidence from a context that has received relatively little attention to date.

3 A Theoretical Framework

In this section, we present a simplified theoretical model of privacy based on Acemoglu et al. (2022), where information sharing on social media might create negative externalities for others. In the original model, when a user shares their private data with an online platform, they reveal information about others. In such a setting, information externalities depress the price of data, as a user has less incentive to protect their data once their information is leaked by others. The authors characterize the user equilibrium in the data market and characterize conditions for data sharing.

Our simplified model focuses on a scenario with only two users, providing the basis for our experimental design and hypotheses. For a more general framework, including social welfare analysis, and market regulations, we refer the reader to the original theory paper.

The simplified model consists of two users $i \in \mathcal{V} = \{1, 2\}$ and a platform. Each user $i \in \mathcal{V}$ has a type denoted by θ_i. The vector $\mathbf{\Theta} = (\theta_1, \theta_2)$ has a joint normal distribution, which is common knowledge:

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right).$$

The correlation between user types is denoted by $\rho \in [0, 1]$. For example, ρ might represent the extent to which a user's music preferences can be inferred from their friend's preferences, as close friends might have similar tastes in music. When a user shares a song on a social networking platform, it is likely that their friend also likes this song. As their preferences for songs become more aligned, there is more overlap in their playlists.

In this framework, each user has some personal data, X_i, which is informative about their type, θ_i. Assume $X_i = \theta_i + \epsilon_i$, where ϵ_i is an independent random variable with a standard normal distribution $\epsilon_i \sim \mathcal{N}(0, 1)$. Users independently decide whether to share their personal data with the platform. Such data may consist of both information generated through user activity and engagement on the platform, as well as additional details voluntarily provided about preferences, contacts, and connections. Each user has an intrinsic valuation of their personal data, denoted by v_i.

The platform offers users benefits for sharing their personal data, represented by a price vector $\mathbf{p} = (p_1, p_2)$. Upon receiving the price offer, each user can choose between two actions: withholding or sharing information with the platform, denoted by $a_i \in \{0, 1\}$. User action profile is denoted by $\mathbf{a} = (a_1, a_2)$. A signal $\mathbf{X_a}$ of the users' personal data is shared with the platform under action profile $\mathbf{a}$: $\mathbf{X_a} = (X_j : j \in \mathcal{V}$ such that $a_j = 1)$. For example, if both share ($a_1 = a_2 = 1$), the signal becomes $\mathbf{X_a} = (X_1, X_2)$; if only user 1 shares ($a_1 = 1$, $a_2 = 0$), the signal is $\mathbf{X_a} = X_1$; if neither shares ($a_1 = a_2 = 0$), the signal is null, $\mathbf{X_a} = \emptyset$.

The platform infers a user's true type, θ_i, based on the received signal, $\mathbf{X_a}$. Its inference of θ_i, denoted by $\hat{\theta}_i$, is generated based on the rules described below:

$$\hat{\theta}_i = \begin{cases} \frac{2-\rho^2}{4-\rho^2} X_i + \frac{\rho}{4-\rho^2} X_j, & \text{if both } i \text{ and } j \text{ share, or } \mathbf{X_a} = (X_i, X_j); \\ \frac{1}{2} X_i, & \text{if only } i \text{ shares, or } \mathbf{X_a} = X_i; \\ \frac{\rho}{2} X_j, & \text{if only } j \text{ shares, or } \mathbf{X_a} = X_j; \\ 0, & \text{if neither shares, or } \mathbf{X_a} = \emptyset. \end{cases} \tag{1}$$

Intuitively, the platform's inference is the most accurate if both share their data. It is moderate if only one shares. If neither shares, the platform will guess the user's type to be zero, which is the mean of the joint normal distribution.

A user's utility is determined by the benefit they receive from the platform if they share their personal data, and the cost of sharing their data:

$$u_i(\mathbf{a}, \mathbf{p}) = \underbrace{p_i \cdot \mathbb{1}(a_i = 1)}_{\text{benefit of sharing}} + \underbrace{v_i \times Var(\hat{\theta}_i | X_a)}_{\text{preference for privacy}}, \tag{2}$$

where the first term, $p_i \cdot \mathbb{1}(a_i = 1)$, reflects the benefit of data sharing, such as a direct monetary payment or better recommendations from the platform, and the second term, $v_i \times Var(\hat{\theta}_i|X_a)$, reflects the cost of sharing one's personal data. The latter enables the platform to infer a user's type more accurately, leading to a reduced variance and a reduction in privacy. Therefore, a rational user chooses the optimal action to maximize their utility.

The utility for the platform is the opposite of the sum of the utilities for both users: $u_p(\mathbf{a}, \mathbf{p}) = -u_1(\mathbf{a}, \mathbf{p}) - u_2(\mathbf{a}, \mathbf{p})$. In this experiment, we focus on the users' information sharing decisions, and thus treat the platform's pricing decisions as exogenous.

The model yields several predictions on how the benefits offered by the platform, captured by the price p_i, the information sharing externalities, captured by the correlation coefficient ρ, and user beliefs influence their sharing decisions. Proofs are provided in Appendix A. The first prediction characterizes price thresholds at which users change their sharing decisions.

Prediction 1 (Price Effect) *Users choose to share if the benefit offered by the platform is sufficiently high. More specifically:*

(a) Only user i chooses to share their personal data if and only if $p_i \geq \frac{1}{2}v_i$ and $p_j < \frac{1}{2}v_j$.

(b) Both users share their personal data if and only if $p_k \geq \frac{(2-\rho^2)^2}{2(4-\rho^2)}v_k$, for $k = i, j$.

Prediction 1 reflects the assumption that users value their personal data. As a result, they need to be compensated to share their personal data. Users will only share their personal data if the compensation at least offsets the cost of data sharing. Therefore, in Prediction 1(a), given that the other user chooses not to share, the minimum price to induce a user i to share is $p_i = \frac{1}{2}v_i$.

However, since user types are correlated, the platform gets more information about both users' types if either of them shares. That is, if one user chooses to share, the other user's type can also be more accurately inferred even if they choose not to share. The stronger the correlation between the user types, the greater the externality will be. Given that the platform can make more accurate inferences as long as one of the users shares, a user will accept a lower price as compensation. Consequently, in 1(b), the minimum price to induce sharing decreases as the correlation coefficient increases. When the user types are not correlated at all ($\rho = 0$), the cutoff for user i is $p_i = \frac{1}{2}v_i$. As correlation increases, the minimum price to induce sharing decreases. When user types are perfectly correlated ($\rho = 1$), the cutoff becomes $p_i = \frac{1}{6}v_i$. The next prediction discusses how users' sharing decisions vary with the correlation coefficient ρ.

Prediction 2 (Externalities) *When $\rho^2 \geq 0.72$, the platform pays less to induce both users to share than the price to induce one user to share.*

As discussed in Prediction 1, each user demands a lower price to share as the correlation coefficient increases, given that the other user chooses to share. Consequently, the minimum total price required to induce sharing decreases as the correlation coefficient rises. When the correlation coefficient becomes sufficiently large, the minimum total price required to induce both users to share falls below the minimum total price needed to induce only one user to share. The latter is identified in Prediction 1(a). Specifically, when $\rho^2 \geq 0.72$, it is cheaper to induce both users to share than to induce just one.

Prediction 3 (Beliefs) *A user is more likely to share if they believe that the other user decides to share.*

In Prediction 1, we characterize two important cutoffs that affect user sharing decisions. The first cutoff suggests the optimal decision when a user believes the other user decides not to share. The second cutoff applies when they believe the other user decides to share. Given that $\rho^2 \in [0, 1]$, it always holds that $\frac{1}{2}v_i \geq \frac{(2-\rho^2)^2}{2(4-\rho^2)}v_i$. This implies that users always demand a lower or equal price when believing the other user decides to share. In other words, given the price offered by the platform, a user is more likely to share when they believe that the other user decides to share.

4 Experimental Design

We design a laboratory experiment to test the key predictions in our simplified model of Acemoglu et al. (2022), to evaluate how externalities and complexity affect participants' information sharing decisions, and to predict users' real-world privacy decisions using their decisions in economic games. The experiment is approved by the IRB of a large public university (HUM#00236045) and pre-registered at the AEA RCT Registry with the unique identifying number AEARCTR-0012875 (Wang et al., 2024). Appendix B contains a summary of our pre-analysis plan.

Figure 1: Overview of the Experimental Design

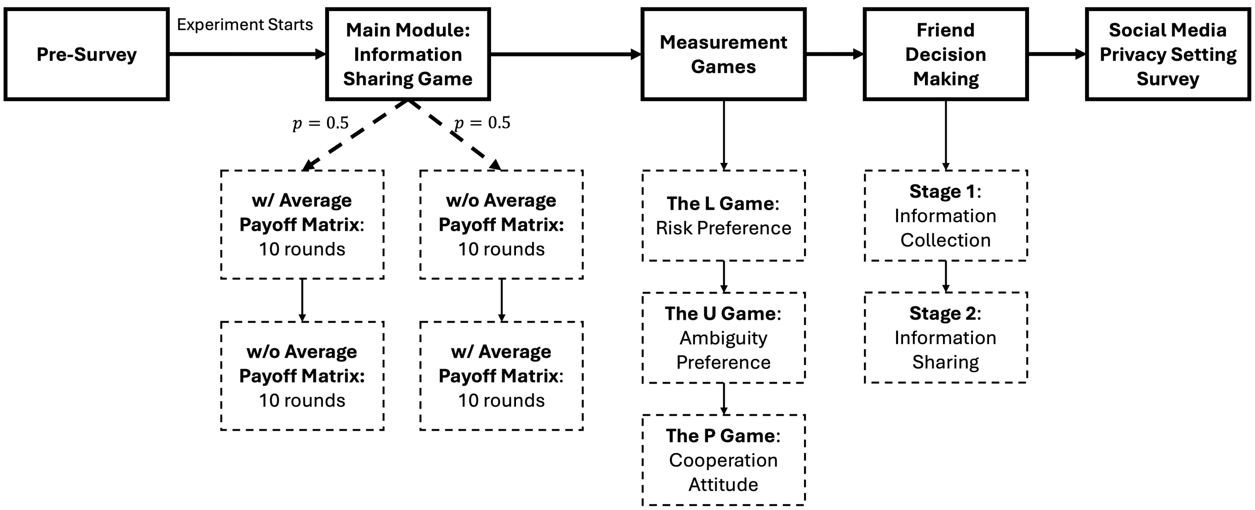


Figure 1 presents an overview of the experimental design. As part of the enrollment process, participants complete a pre-survey online to sign the consent form and nominate a friend within their university. The friend does not need to be physically present in the lab. However, they must be available to fill out an online survey during the two-hour window when the participant is in the lab (specifically, for a module that requires information from both the participant and the friend).

The experiment consists of four modules. The main module is an *information sharing game*, where we test the predictions of the simplified Acemoglu et al. (2022) model. In the *measurement games* module, we measure participants' risk, ambiguity and cooperation preferences through a series of classic economic games. To gauge information sharing decisions in real-world settings, we assign specific real-world scenarios to participants based on the model predictions in the *friends decision making* module. In this module, we collect private information about each participant and the friend they nominate, and elicit their willingness-to-accept to share sensitive information that might impose externalities on their friend. Lastly, we use a survey to collect participants' privacy settings on Instagram.[2] These four modules are designed to measure participants' privacy preferences in the presence of externalities from an abstract neutral setting to progressively more realistic settings. In what follows, we explain each module in detail.

4.1 The Information Sharing Game

We now translate our simplified theoretical model in Section 3 into our experiment setting, and explain our main design choices, including necessary deviations from the model. First, the model assumes continuous distributions for the correlation coefficient ρ and the type parameter θ, whereas

[2]Among US adults under 30, Instagram is the most widely used social media platform (Gottfried 2024).

our experimental implementation relies on discrete parameter spaces. Second, instead of relying on endogenous correlations within pairs, we exogenously assign correlation coefficients. This enables us to identify the causal impact of user similarity. Third, the benefit offered by the platform, p, is randomized at the pair level, which simplifies participants' decision process. Lastly, we choose the price interval around the key cutoffs based on the theoretical predictions, which improves the efficiency of the experiment design.

In the experiment, the information sharing game consists of two participants, $k \in \{1, 2\}$, and a robot representing a social networking platform. The participant pair is assigned a correlation coefficient ρ, which is randomly drawn from the set, $\{0.1, 0.2, \ldots, 0.9\}$. We focus on the interior correlation coefficients, as the decisions under $\rho = 0$ or 1 are straightforward. Furthermore, we use a set of realizations to simulate each participant's private type, $\theta_i \in \mathbb{R}$, which are modeled as random variables following a joint normal distribution. To avoid participants focusing on one specific number, we discretized the types into sets of three real numbers, independently drawn from the joint normal distribution characterized by ρ. The three numbers in each participant's set are noted as θ_{ik}, where $k \in \{1, 2, 3\}$. While $\{\theta_{ik}\}_{k=1}^{3}$ is only observable to participant i, ρ is disclosed to both participants and the platform. The type distribution is common knowledge.

Recall that each participant privately observe their type, θ_{ik}, where $k \in \{1, 2, 3\}$. Participants are informed that the platform offers to pay them a price p for sharing their private information, with p randomly drawn from the interval $[\frac{1}{10}v_i, \frac{3}{5}v_i]$. We exclude prices that always induce sharing or withholding, based on the theoretical predictions. We further set $v_i = v_j = 300$ points, for $i = 1$ and 2. As a result, $p_i = p_j = p \in [30, 180]$.

The participants then simultaneously and independently decide whether to share their personal data, $X_{ik} = \theta_{ik} + \epsilon_{ik}$, which is informative about their type, θ_{ik}. If a participant chooses to share their data, their entire set of three secret numbers will be shared with the platform. Following the specification of the theory model, each number in the set will be perturbed by adding a number independently and randomly drawn from a standard normal distribution, which represents the noise, or $X_{ik} = \theta_{ik} + \epsilon_i$. The platform observes both participants' decisions and makes inferences about their secret numbers based on the available information. The inference rules used by the platform, as described in Equation (1) in Section 3, are common knowledge to the platform and the participants.

Each participant makes sharing decisions to maximize their payoff. A participant's payoff consists of two components: (1) the benefit of sharing, which equals the price paid by the platform if the participant decides to share their personal data, and (2) the cost of sharing which equals the sum of squared differences between the platform's guesses and the corresponding true type, regardless of the participant's decision. Specifically, a participant's payoff u_i is given by the following equation:

$$u_i = \underbrace{p \times \mathbb{1}\{\text{Share}\}_i}_{\text{benefit of sharing}} + \underbrace{v \times \frac{1}{3}\sum_{k=1}^{3}(\hat{\theta}_{ik} - \theta_{ik})^2}_{\text{preference for privacy}}, \tag{3}$$

where θ_{ik} is the kth number in participant i's set, and $v = 300$ is the participant's value of privacy. Note that the term $\frac{1}{3}\sum_{k=1}^{3}(\hat{\theta}_{ik} - \theta_{ik})^2$ represents the accuracy of the platform's inference. Intuitively, a participant's payoff increases with the price offered by the platform if they decide to share, and decreases as the platform's guesses get closer to their true type. While the benefit of sharing is straightforward and easy to understand, the cost requires more effort to compute.

Based on the model predictions, we derive the following comparative statics.

Hypothesis 1 (Price Effects) *In the information sharing game, participants are more likely to share their information with the platform as prices increase.*

Hypothesis 1 is based on Prediction 1 that a participant decides to share when the price sufficiently compensates for the cost of data sharing. Though we are unable to test the exact cutoffs characterized by Prediction 1 at the individual pair level due to the discretization in the experiment design, at the session level we still expect to see more participants share as the prices increase.

Hypothesis 2 (Externalities) *In the information sharing game, given the price offered by the platform, participants are more likely to share their information with the platform as their type correlation, ρ, increases.*

Hypotheses 2 is based on Prediction 2. As the correlation coefficient gets larger, it becomes cheaper for the platform to induce both participants to share than to induce just one. This implies that, given the price offered by the platform, participants are more likely to share as the correlation coefficient increases.

If Hypotheses 1 and 2 are true, we should observe an increase in the likelihood of sharing, measured by the dummy variable indicating whether the participant decides to share, as ρ and p_i increase.

Hypothesis 3 (Beliefs) *In the information sharing game, given the price offered by the platform p and the correlation coefficient ρ, participants are more likely to share their information with the platform if they believe their match decides to share.*

Hypotheses 3 is based on Prediction 3. Given the correlation coefficient (ρ), participants demand a lower price to share if they believe their match shares.

4.1.1 Treatment

Our experiment employs a within-subject intervention to assess the impact of computational complexity on privacy-related decision-making.

In our setting, participants must carefully evaluate the potential benefits and costs of sharing. However, the presence of externalities complicates the computation of expected payoffs related to different outcomes. To reduce this complexity, we introduce a within-subject treatment variation in which participants see their match's expected type and an average payoff matrix before making decisions. This allows participants to focus on comparing the payoffs in the average payoff matrix without having to compute their expected payoffs, while keeping the information condition identical between the treatment and control conditions.

Figure 2: Interface for the Control Condition

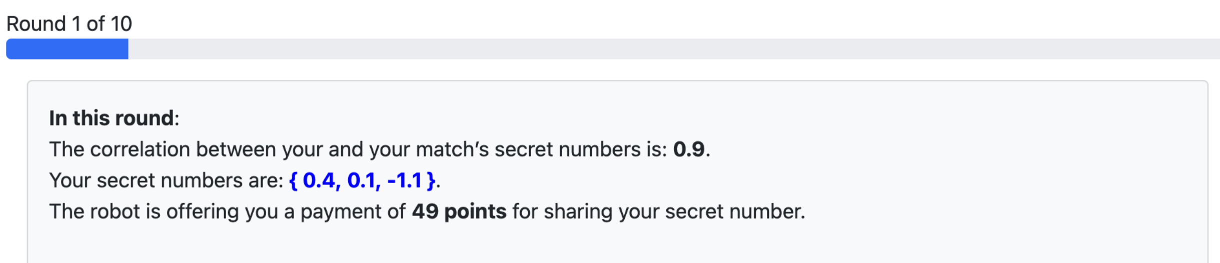


Figure 2 presents the interface for the control condition, where a participant has all the necessary information to compute their expected payoff based on Equations (3) and (1). A physical calculator is provided to each participant throughout the session, enabling them to perform the calculations themselves.

Figure 3: Interface for the Matrix Treatment

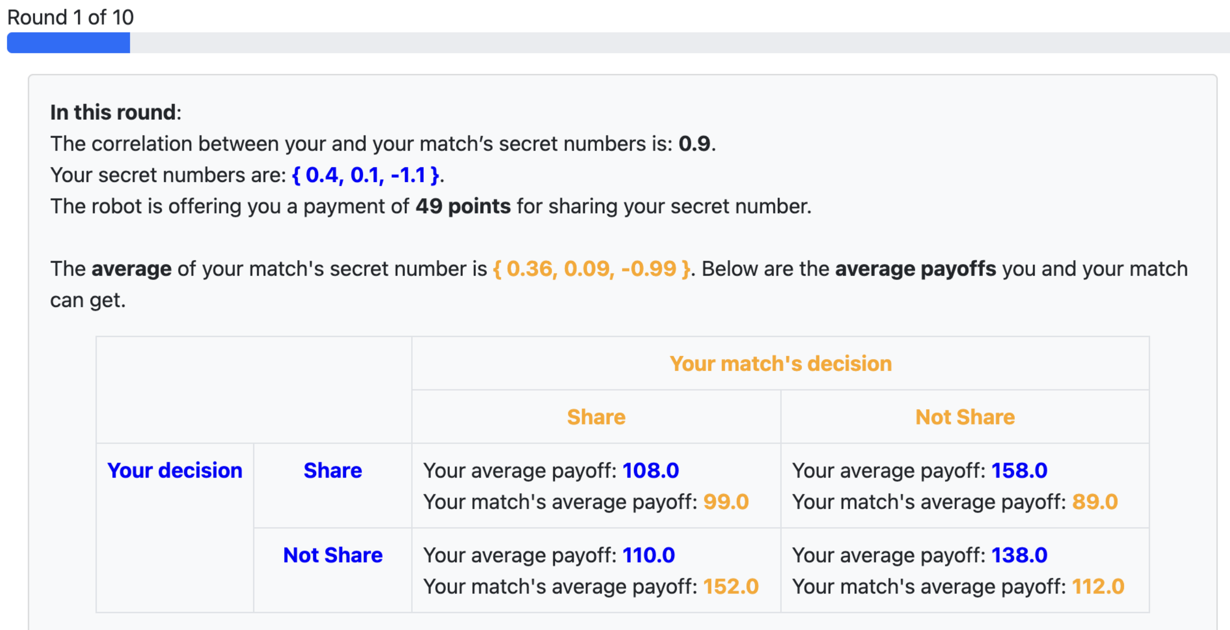


Figure 3 presents the interface for the matrix treatment condition. In addition to the information given in the control condition, an average matrix displays the expected payoffs for each participant and their match. The average payoffs are computed based on the participant's information, without knowing the match's exact secret numbers. Specifically, given a participant's secret number θ_{ik} and the correlation coefficient ρ, the average of the match's corresponding secret number is $\rho \times \theta_{ik}$. Using these values, one can compute each pair's respective average payoffs. As shown in Figure 3, the averages of the match's secret numbers are also provided to help the participants understand how the average payoffs are computed. However, this information is already incorporated in the matrix and should reveal no additional information beyond the matrix itself.

Each participant makes two types of decisions given the same setup, with or without the average payoff matrix. To minimize the likelihood that participants memorize the best responses, the two types of decision are separated into two blocks, each of which consists of 10 rounds. The average payoff matrix is provided in the treatment block, but not in the control block. The order in which the two blocks are played is randomized at the session level.

At the beginning of each round $t \in \{1, 2, \cdots, 10\}$ in the first block, participants are randomly rematched into pairs, to minimize repeated game effects and to enable learning about the game. The computer randomly draws the correlation coefficient (ρ^t) and the price for information sharing (p^t) for each pair, and one set of three secret numbers for each participant ($\{\theta_{ik}^t\}_{k=1}^3$). These parameters are stored and reused for the same round t in the second block. In other words, round t in the treatment and control blocks has identical parameters. Participants remain anonymous throughout the game.

In each round, we elicit a participant's strategies and beliefs by asking the following four questions:

1. If your match decides to share, what would your decision be?
2. If your match decides not to share, what would your decision be?
3. Do you decide to share or not to share your secret numbers?
4. If your match knew your decision, would they decide to share or not to share? Please indicate your prediction.

Participants' beliefs about their match's decisions are incentivized using a simple scheme. If a participant correctly predicts their match's decision, they receive a bonus of 20 points for that round, and zero otherwise. This scheme is inspired by recent research showing that simplified descriptions of the incentive mechanism are more effective in eliciting true belief reports than complex formulations (Danz et al., 2022).

After both participants submit their decisions, they receive feedback on their decisions, their match's decisions, the platform's inference of their type(s) based on Equation (1), and the payoff for each of them. One round is randomly chosen at the end of each module to determine the final payoffs. We anticipate the following treatment effect.

Hypothesis 4 (Complexity) *Participants are more likely to choose best responses in the matrix treatment, where they are provided with the average payoff matrix.*

If Hypothesis 4 holds, we also expect the likelihood of mutual best responses and efficiency to be higher under the matrix treatment.

4.2 Friends Decision Making

This module is designed to measure information sharing decisions in real-life settings. In this module, participants make decisions about what to share with and without externalities. Unlike the information sharing game where pairs and private information are both randomly assigned, participants in this module are paired with a friend and make sharing decisions concerning genuine private and sensitive information about themselves and their friend.

This module comprises three real-world scenarios. Each scenario consists of two questions about the participant or their friend's behavior or opinions on sensitive topics. In each scenario, the first question does not involve externalities ($\rho = 0$). Here, the participant's answer reveals information solely about the participant themselves. In contrast, answering the second question reveals information about both the participant and their friend ($\rho > 0$). The three scenarios are as follows:

Scenario 1: Body weight. Participants are first asked to share their own body weight information ($\rho = 0$). Then they are asked to share the average weight between their friend and themselves ($\rho > 0$).

This scenario is adapted from Huberman et al. (2005) who used the second-price auction to elicit participants' willingness-to-accept to announce their own body weight in front of other participants in the lab. Body weight information is sensitive, as it involves both self-presentation and perception by others. First, individuals want to have a desirable self-presentation (Goffman et al., 2002). For example, online daters often lie about characteristics such as weight, height, and age to appear more desirable (Toma et al., 2008; Hancock et al., 2007). This suggests that less favorable truths are considered private, as people try to hide them. Second, weight-related information can carry a stigma, leading to negative comments or discriminatory behavior (Goffman, 2009). Fearing stigmatization and embarrassment, participants may be unwilling to share such information (Huberman et al., 2005).

Scenario 2: Intimate relationships. Participants are first asked how many sexual partners they themselves have had ($\rho = 0$). Then they are asked the total number of sexual partners they and their friend have had ($\rho > 0$).

These questions are adapted from Brandimarte et al. (2013). The original question asks about participants' marital and relationship status, including whether they are single or married, and whether they have a girlfriend/boyfriend. Their participants rated this question moderately intrusive, indicating the discomfort in sharing such information. Building on the idea that relationship-related questions evoke a sense of privacy, we focus instead on sexual relationships to heighten sensitivity.

This sensitivity is also supported by prior research showing that questions about sexual history are especially prone to social desirability bias. Participants tend to report their sexual behaviors in ways that appear more socially desirable when they are not pressured to be fully truthful (Fisher, 2013, 2009, 2007). Thus, when they are incentivized to answer truthfully but face the risk of negative judgment, they become less willing to disclose their sexual history.

Scenario 3: Opinions on intimate relationships. Participants are first asked to rate the extent to which having more than four sexual partners in one semester is acceptable ($\rho = 0$). Then they are asked the average opinion of their friend and themselves on the same issue ($\rho > 0$).

Although the first two scenarios deal with highly sensitive information, this scenario more closely resembles user content sharing actions on social networking platforms. Users may not frequently share sensitive information on these platforms, but they often feel comfortable commenting on sensitive topics, which might reveal private information about themselves. For example, a user may not be comfortable sharing that they had COVID in a public post. However, when their friend posts about having COVID, the user may self-disclose with the response, "It was terrible when I had it." Differences in self-disclosure among these channels may be due to the system's affordances (e.g., easily allowing for quick 'emoji' reactions) or for social reasons (e.g., perceptions of the size and makeup of the audience (Litt, 2012) or potential social benefits and costs (Vitak, 2012)). Even different forms of response (e.g., *liking* or *sharing/retweeting* or *commenting*) may disclose different information (Kim and Yang, 2017). The questions in this scenario also complement those in the second scenario by eliciting perceived social norms regarding sexual relationships.

Procedures. After participants enter the waiting room, they are asked if they would like to step on a scale and receive their weight information. Participants could choose to skip this part. For those who agree to be weighed, a research assistant records their body weight individually and without identifiable information. Each participant also receives a sticky note with their weight written on it.

When participants enter the lab, they each randomly draw an ID card and sit at the corresponding computer terminal. The experimenter reads an overview of the study aloud. Then everyone is instructed to turn off their phones and place them in an envelope on their desk (visible to the experimenters) to ensure that they do not communicate with their friends during the experiment.

The experimenter then displays on the screen the survey questions that will be shown to each participant's friend, cautions about the sensitive nature of these questions, and clarifies that answering them is optional for their friend. Participants are asked to confirm whether they wish to continue sending the survey. Choosing not to send does not affect their role in the study, though neither the participant nor their friend will receive the payoffs for the friends decision making module.

Participants submit their decisions on a sticky note that they fold and place in a non-transparent box. To opt in, they write down their own university 'unique name' (effectively, their email address). To opt out, they either leave the note blank or write a random string. Across all sessions, 81.25% of participants sent the survey to their friends.

After all opt-in decisions are collected, a research assistant sends the online survey to the nominated friends of those who opt in, who then have two hours to complete it. Friends receive a completion bonus of $2 if they respond to the survey during the two-hour window of the session, paid in virtual gift cards after all survey responses are collected.

Based on the three scenarios, this module is made up of two stages: an information collection stage and a information sharing stage.

1. Information collection stage. In this stage, participants who opt in provide answers to the questions in the three scenarios. Meanwhile, their friend answer the same questions in the online survey sent to them. Participants (and their friends) have the option to skip any of the questions. In the instructions, they are told that their responses are accessible to the experimenters only. Those who opt out are told to skip all questions.

To ensure truthful information revelation, we verify participants' answers to the questions. For the second question ($\rho > 0$) in all scenarios, we compare the participants' responses with those of their friends. An answer is verified if the two answers are within 10% of each other. For each verified answer, the participant, as well as their friend, receive \$1. If no answer is provided, or the answer is unverified, they receives no reward. The amount is added to the virtual gift cards. With the \$2 bonus, a friend can earn up to \$5. In addition, for the body weight scenario, we validate participants' answers to the first question by comparing them with their weights recorded in the waiting room.

2. Information sharing stage. We use the Becker–DeGroot–Marschak mechanism (BDM) to elicit the participant's willingness-to-accept (WTA) to share their answers with other participants in the same experiment session (Becker et al., 1964). For each question, participants indicate the minimum number of points they are willing to accept, ranging from $[0, 100]$, to share their answers. The lower the number they indicate, the more willing they are to share. A value of 0 means they are always willing to share, while 100 means they are not willing to share at all. Adapted from Huberman et al. (2005), sharing involves publicly announcing their answer in front of the other participants in the lab. This procedure may cause some embarrassment for the participants thereby creating a cost for revealing their personal information.

After indicating their WTA for each question, a number, between $[1, 100]$, is drawn randomly for each participant. If the random number is greater than their WTA, the participant needs to announce their answer in front of others in the lab, and their payment equals the random number. If the random number is less than or equal to their WTA, the participant does not need to announce their answer or receive any payment.

If a participant chooses to skip a question in the first stage, their WTA to answer that question is set to the maximum value, 100, in our data analysis.

At the end of the module, participants see a summary page listing all the answers they need to announce in front of others. They are then called sequentially by their ID number and stand up to read the answers aloud.

Based on Prediction 2, we obtain the following hypothesis.

Hypothesis 5 (WTA and Externalities) *Within each scenario of the friends decision making module, compared to the first question ($\rho = 0$), participants are more willing to share in the second question ($\rho = 1$), or $WTA_{\rho=0} > WTA_{\rho=1}$.*

Lastly, Huberman et al. (2005) find that the female participants in their study demand a higher price to share their own body weight information. Thus, we hypothesize that women will demand a higher WTA to reveal sensitive private information.

Hypothesis 6 (Gender) *In the friends decision making module, women are less willing to announce their answers, or, $WTA_w > WTA_m$.*

4.3 Social Media Privacy Settings

To gain further insight into participants' real-world information-sharing decisions, our last module asks participants to disclose their privacy settings on Instagram. Instagram is one of the most popular social media platforms among college students, with 76% of US adults between 18 and 29 indicating they used the platform (Pew Research Center, 2024).

Participants are first asked to check the list of advertisers from which they see ads on Instagram (Figure 4). Instagram utilizes user activity information from these ad partners to customize ads. One common placement for these ads is in Instagram Stories, a temporary collection of photos or videos that disappears after 24 hours from users' profiles and feeds. When users view stories published by their friends, ads are inserted between different friends' stories. Instagram Stories has 500 million daily active users, about one third of Instagram's total audience (Tafradzhiyski, 2025). Therefore, the likelihood that users encounter these ads is high, and users may notice that the content is closely aligned with their recent browsing behaviors. We use the instruments of Lee and Kobsa (2016) to elicit participants' feelings after reviewing this list from three aspects: comfort, perceived risk, and appropriateness of monitoring. Following this, participants are directed to review a specific privacy setting: whether they enable activity information sharing from ads partners (Figure 5). Finally, we ask participants when was the last time they adjusted their privacy settings on Instagram and their perception of the necessity of making changes to these settings.

At the end of the study, one third of the participants are randomly chosen to show their Instagram settings to the experimenters to verify their answers for this part of the survey.

Figure 4: List of Advertisers An User Sees Ads from on Instagram

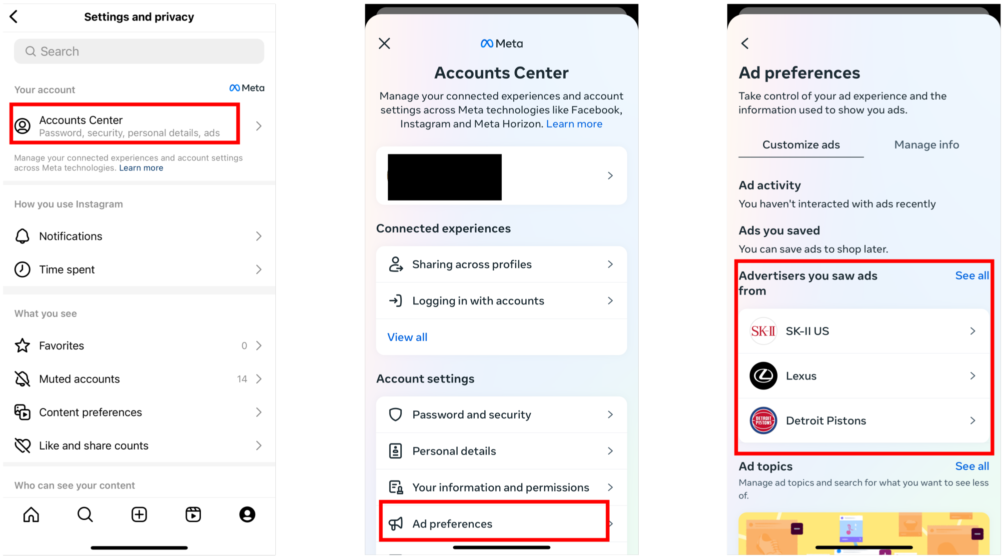
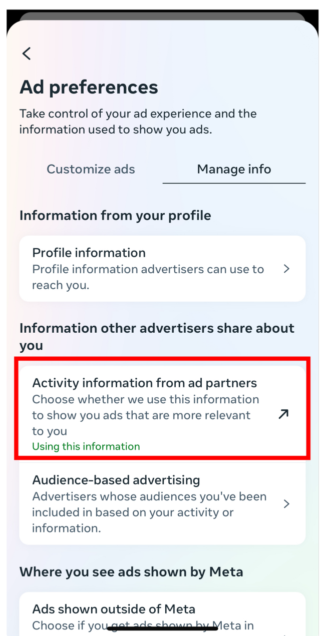


Figure 5: Ad Setting on Instagram



Our settings in the information sharing game and the friends decision making module simulate situations where users actively and voluntarily share information on social networking platforms. In reality, however, information sharing with platforms also occurs passively, such as browsing behaviors. Rather than examining specific content that participants post on social media, we collect their preferences for personalized ads, which provide a more general measure of sharing.

Users face similar cost-benefit analyses on social networking platforms. Shared information is often leveraged by social networking platforms to create better personalized recommendations, such as targeted advertisements. This benefit is represented as monetary payoffs in the information sharing game and the friends decision making module. However, users may still be reluctant to share their private information, perhaps due to intrinsic privacy valuations or social desirability concerns. Given the connections between our experiment setup and the social networking platforms, we expect to see a correlation between participants' information-sharing behaviors in these simulated contexts and their real-world actions.

Hypothesis 7 (Correlation) *Willingness to share in the information sharing game, in the friends decision making module, and on Instagram are correlated with each other.*

If Hypothesis 7 is true, we should observe that participants who are more inclined to share in the information sharing game also demands less monetary compensation to share in the friends decision making module, and are also more inclined to enable sharing on Instagram.

4.4 Measuring Risk, Ambiguity and Cooperation Preferences

Between the information sharing game and the friends decision making module, we measure participants' risk and ambiguity preferences, as well as their attitudes towards cooperation.

Risk Preference. We elicit participants' risk preferences using the standard lottery choice game with the double price list (Holt and Laury, 2002). Given the structure of the double price list, a risk averse expected utility maximizer would start by selecting Option A (the safer lottery with lower variance) in the first lottery and switch to Option B (the riskier lottery with higher variance) at some point. We capture their risk preference as the switching point between the two options. To simplify the measurement, a single switching point is enforced.

While the theoretical model assumes risk neutrality for analytical tractability, risk preferences matter in privacy decisions. Based on our equilibrium analysis comparing a risk-neutral and a risk averse user with a parameterized utility function in Appendix C, the risk averse individual is less likely to share their private information than their risk neutral counterpart. We anticipate that the result is more general, and formulate the following hypothesis.

Hypothesis 8 (Risk) *More risk averse individuals are less likely to share their private information in the information sharing game, less likely to share sensitive personal information in the friends decision making module, and less likely to enable sharing on Instagram.*

Ambiguity Preference. While participants know the distribution of other's types in the information sharing game, in many real-world privacy decision-making scenarios such as on social media, it is unrealistic to assume that they know the distribution of various states of the world when they share sensitive personal information. For example, enabling personalized ads on Instagram could lead to harmful outcomes in the event of data breaches or misuse. However, users typically do not know the probability of data breaches. This uncertainty about the probability distributions of the outcomes is commonly referred to as *ambiguity*.

We use the Ellsberg urn design to measure ambiguity preference (Ellsberg, 1961), adapted from the protocol in Li et al. (2018). We present participants with two bags, one containing a known number of purple and yellow balls (adding up to 100), called the known bag. The other bag contains 100 balls in an unknown combination of purple and yellow and is called the unknown bag. Participants first decide a winning color to bet on, purple or yellow. They then decide from which bag a ball will be randomly drawn. If the drawn ball is in the winning color, they earn 500 points; otherwise, they earn nothing.

Choices are elicited for all 101 possible compositions of the known bag using choice lists. Each row presents a choice between the known and unknown bags. The first choice list contains 11 scenarios with 0, 10, $\cdots$, 100 balls of the winning color in the known bag. Ambiguity-averse participants tend to choose the known bag, switching to the unknown bag only when the known bag's composition becomes sufficiently unbalanced. Again, we enforce a single switching point for each participant.

After completing the first list, participants proceed to a more refined choice list that focuses on the two values from the first list between which the preference change occurred. This second list contains nine additional scenarios to precisely locate the switching point. As before, a single switching point is enforced, and this switching point identified in the refined list serves as our measure of ambiguity preference.

We anticipate that participants' ambiguity preferences measured from this task predict their information sharing decisions in real-world settings. In particular, more ambiguity averse individuals hold more pessimistic beliefs about the state of the world (Gilboa and Schmeidler, 1989). Therefore, we formulate the following hypothesis.

Hypothesis 9 (Ambiguity) *More ambiguity-averse individuals are less likely to share sensitive personal information in the friends decision making module and less likely to enable sharing on Instagram.*

Cooperation Attitude. In our environment, information sharing imposes externalities on others. We use a linear public goods game to capture the extent to which participants consider the externalities their decision might have on others.

We use a classic linear public goods game with voluntary contributions (VCM) to measure behavior in an environment with positive externalities. Our design is adapted from that of Charness et al. (2014). This game consists of 5 rounds. In each round, participants are randomly matched with another participant in the session to form a group of two. Each participant is endowed with a private good, w_i, which is set to 200 points. From this endowment, they decide how many points, g_i, to contribute to a group project. The number of points chosen is multiplied by 1.2 and contributed to the project. The total contributions to the project are then shared with the group as a public good. The payoff function for each participant i in each period is given by: $\pi_i = w_i - g_i + 0.6 \times (g_i + g_j)$. Full free-riding ($g_i = 0$) is a dominant strategy, while the sum of the group payoffs is maximized if each member of the group contributes everything, or $g_i = w_i$, for all i.

Participants play this game for five rounds with random rematching each round. We measure participants' cooperation attitude based on their contribution in the first round, which is not affected by others' behavior.

Hypothesis 10 (Cooperation) *Those who contribute more in the public goods game are less likely to share information that could harm others.*

In sum, we anticipate that participants' risk, ambiguity and cooperation preferences will be correlated with their information sharing decision in various modules, and we will control for them in our subsequent analysis.

4.5 Procedures

As summarized in Figure 1, each experimental session consists of four modules: the information sharing game, the measurement games, the friends decision making module, and the social media privacy setting survey module. Each session consists of 12 participants. We conducted a total of 16 sessions at the Behavioral and Experimental Economics Lab at a large public university, with a total of 192 participants. The sessions were conducted between February and April, 2024.

Upon entering the lab, participants have 10 minutes to read the instructions for the first segment of the information sharing game. They then watch a 5-minute video summarizing the instructions. The video is the same for all sessions and does not mention the average payoff matrix. Participants answer a set of review questions designed to check their comprehension. They receive 20 points for each correct answer. Participants answered an average of 7.18 out of eight questions correctly.

Following this, participants play ten rounds of the information sharing game. Upon completion, the experimenter reads the instructions aloud for the second segment. For sessions in which the segment without the matrix is played first, participants answer the remaining review questions about the average payoff matrix before playing another ten rounds of the information sharing game. For sessions where the segment with the matrix is played first, participants already answer all review questions in the first segment and proceed directly to the decisions after the experimenter reads the instructions.

After the information sharing game, we implement the measurement games. Each participant follows the same sequence: the lottery choice game, the Ellsberg urn experiment, and the public goods game. The experimenter reads aloud the instructions for each game before it starts. Immediate feedback regarding their payoffs is provided to participants after the completion of each game.

Moving forward, all participants advance to the friends decision making module. The experimenter reads aloud the instructions for the information collection stage. The participants then provide their answers to the questions in the three real-world scenarios.

To help participants distinguish between the first stage (sharing with the experimenters) and the second stage (sharing with everyone else in the session), they complete a demographic survey as an interim activity.

Next, the experimenter reads aloud the instructions for the second stage, the information sharing stage. Before making decisions, participants first answer a set of review questions to check their understanding of the BDM mechanism. Each correct answer is again rewarded with 20 points. Participants answered an average of 3.78 out of five questions correctly. After every participant in the session finishes this stage, the experimenter calls the participants in the order of their ID number to check what information they need to announce. Those who need to announce stand up and read aloud their answers. Those who do not need to announce anything remain seated.

After wrapping up all four modules, participants complete the social media privacy setting survey.

All participants are paid in private. The exchange rate between points and cash is 40 points = \$1. Each experimental session lasts approximately 120 minutes. The average payment is \$50.02, including a \$5 show-up fee. The experiment is programmed in oTree (Chen et al., 2016).

Our participants are students from a large public university, recruited using ORSEE (Greiner, 2015). In the first eight sessions of the study, the weighing process is announced and explained after all participants arrive in the waiting room, and then the participants are individually weighed. However, we find that only two of the participants voluntarily stepped on the scale. To increase this ratio, we changed the protocol for the last eight sessions. In the last eight sessions, participants are asked to be weighed privately when they arrive and check in individually. The proportion of participants who were weighed increased to 40.6%. Each of the eight sessions is balanced: four treatment sessions and four control sessions. A coding error occurred in one session, in which all participants played only one round of the public goods game. This does not affect our main analysis, as we use the first-round unconditional contribution amount as a proxy for their cooperation attitude.

Experimental instructions for the treatment group are included in Appendix D. Data are available from the authors upon request.

5 Results

In this section, we report participant information sharing behavior in three contexts: the information sharing game, the friends decision making module as well as on Instagram. The summary statistics of the key variables and robustness checks are in Appendix E.

5.1 Information Sharing Game

Using the induced value method in experimental economics (Smith, 1982), our information sharing game presents the decision-making environment with neutral language in an abstract setting. This setting enables us to investigate the extent to which participants best respond to their beliefs and the underlying economic incentives, the resulting equilibrium outcomes, and efficiency.

5.1.1 Best Response and Equilibrium Outcomes

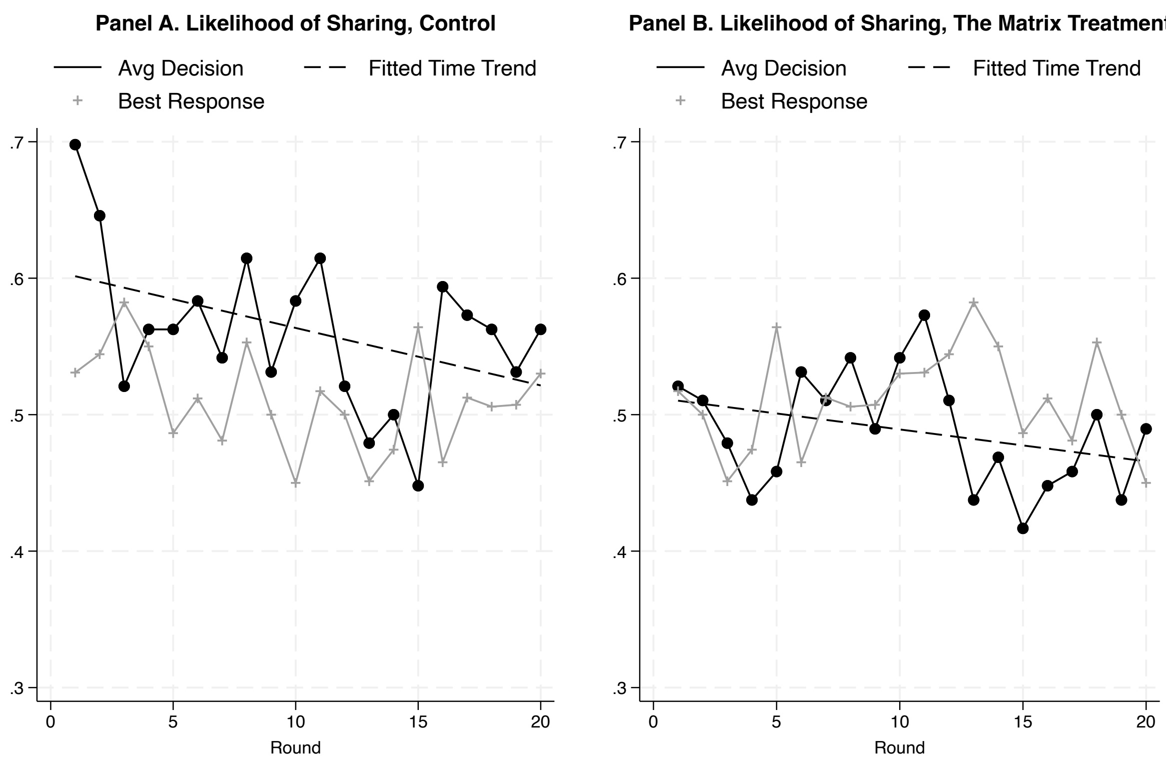
While instructions in the control condition present all the necessary information for participants to make *ex ante* optimal information sharing decisions, our matrix treatment additionally provides an average payoff matrix to reduce computational complexity. We first present the treatment effect on best responses.

Figure 6 presents the proportion of participants who choose to share their private information with the platform each round in the control (Panel A) and treatment (Panel B), respectively. The solid black line represents the actual proportion of sharing each round, whereas the solid gray line represents the proportion of sharing if participants best respond.

From Figure 6, we see that the participants share significantly more in the control condition compared to the treatment (56.1% vs 48.8%, $p < 0.001$,two-sided t-test). Their sharing decisions are further away from the best responses in the control condition. For example, in the first round under the control condition, 70.0% of the participants decide to share, while 52.4% should share if they best respond.

We use the following within-subject panel logit specification to assess the treatment effect on participants' likelihood of choosing best responses. The dependent variable, Y_{it}, is the likelihood that participant i chooses the best response in round t.

Figure 6: Likelihood of sharing private information in the control (panel A) and treatment (panel B) conditions



Note: The solid black line represents the actual proportion of sharing each round, whereas the solid gray line represents the proportion of sharing if participants best respond.

$$P\left(Y_{it}=1\right)=\frac{1}{1+exp(-(\beta_1\cdot\mathbb{1}\{\text{Treatment}\}_{it}+\gamma\cdot t+\mathbf{C}_i'\alpha},$$

where $\mathbf{C}_i$ include demographics (female, ethnicity, age), whether they have participated in economic or psychological studies before, whether they have taken introductory statistics courses before, and whether they have correctly answered both calculation questions in the review questions.[3] In addition, we also control for participants' risk preferences measured by the lottery choice game and their cooperation attitudes measured by the public goods game. Each measure is standardized using the sample mean and standard deviation.

Table 1 presents three panel logit specifications investigating the treatment effects on the likelihood of best responses. We see that the treatment effect is economically sizeable (19.5 percent-

[3]We focus on these two calculation questions (questions 3 and 4) because participants make the most mistakes on them. While over 90% of participants answer all other review questions correctly, only 61.5% answer question 3 correctly and 80.2% answer question 4 correctly (see Appendix Table E.3).

Table 1: Treatment Effects on Likelihood of Choosing the Best Response; Panel Logit Regression

	Choosing Best Responses		
	(1)	(2)	(3)
Matrix Treatment	0.194***	0.194***	0.195***
	(0.012)	(0.012)	(0.012)
Round		0.001	0.001
		(0.001)	(0.001)
Risk Aversion (Standardized)			0.019**
			(0.009)
Cooperation (Standardized)			-0.005
			(0.012)
Controls	No	Yes	Yes
# of Observations	3840	3840	3840
# of Participants	192	192	192

[a] Standard errors in parentheses are clustered at the session level.
[b] Coefficients are marginal effects.
[c] Controls include a participant's gender, age, ethnicity, and indicators of whether they have participated in economics or psychological studies before, whether they have taken introductory statistics courses and whether they have answered both calculation questions correctly.
[d] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

age points, pp hereafter), statistically significant ($p < 0.01$) and stable in the three specifications. Meanwhile, we do not observe any significant learning effect on the likelihood of choosing the best response (0.01, $p > 0.10$). Among the covariates, we find that participants who are one standard deviation more risk averse are 1.9 pp more likely to choose the best responses. In comparison, cooperation attitudes do not correlate with the likelihood of best response. We summarize the results below.

Result 1 (Treatment Effect on Best Response) *Participants in the matrix treatment are 19.5 pp more likely to choose the best responses compared to their control counterparts whose best response rate is 55.9%. This effect represents a 35% increase over the control condition.*

By Result 1, we reject the null in favor of Hypothesis 4, suggesting that simplifying the computation significantly improves participants' likelihood of choosing best responses. This suggests that computational complexity may be a key factor preventing participants from making optimal information sharing decisions.

To further assess to what extent the effect of the treatment is due to the reduction in complexity, we adapt the objective complexity measure from Puri (2025) and Enke and Shubatt (2023), who use the error rate, which is the likelihood that a participant chooses an option with a lower expected payoff. Similarly, we use the likelihood that a participant chooses the strictly dominated strategy given by the realized payoff matrix as a proxy for the complexity of the decision problem. The results are presented in Table 2. We see that participants' likelihood of choosing a strictly dominated strategy decreases from 37.1% to 16.5%, representing a 20.6 pp reduction. The reduction in error rates closely mirrors the increase in the likelihood of choosing best responses, providing further evidence that complexity is a key factor driving participants' failure to respond optimally.

Table 2: Objective Complexity: Likelihood of Choosing the Strictly Dominated Strategy

	Mean	SD	Obs
Control Condition	0.371	0.483	1920
Matrix Treatment	0.165	0.371	1920

Next, we examine the treatment effects on equilibrium outcomes. As participants become better best responding in the treatment condition, we expect a corresponding improvement in equilibrium outcomes.

Table 3: Treatment Effects on Likelihood of Choosing the Mutual Best Response and Prediction Accuracy; Panel Logit Regression

	Choosing Mutual Best Responses			Predicting Correctly		
	(1)	(2)	(3)	(4)	(5)	(6)
Matrix Treatment	0.283***	0.283***	0.283***	0.022*	0.022*	0.022*
	(0.019)	(0.018)	(0.039)	(0.013)	(0.013)	(0.013)
Round		0.002	0.002		0.000	0.000
		(0.002)	(0.003)		(0.001)	(0.001)
Risk Aversion (Standardized)			0.017*			-0.005
			(0.010)			(0.007)
Cooperation (Standardized)			-0.005			-0.009
			(0.007)			(0.006)
Controls	No	Yes	Yes	No	Yes	Yes
# of Observations	3840	3840	3840	3840	3840	3840
# of Participants	192	192	192	192	192	192

[a] Standard errors in parentheses are clustered at the session level.
[b] Coefficients are marginal effects.
[c] Controls included are identical as in Table 1
[d] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3 presents six panel logit specifications. Columns (1) - (3) present the treatment effects on mutual best responses, i.e., equilibrium outcomes. The likelihood of choosing the mutual best response is 28.3 pp higher in the treatment condition, representing a 95% improvement compared to the control mean of 0.299. Part of the improvement might be due to increased prediction accuracy about their match's choices. Specifications (4) - (6) investigate this possibility. The accuracy of participants' prediction of their match's decision is 2.2 pp marginally higher in the treatment condition, representing a 4% improvement over the control mean of 0.523. We summarize the results below.

Result 2 (Treatment Effects on the Likelihood of Mutual Best Responses) *Participants in the matrix treatment are 28.3 pp more likely to choose mutual best responses than their counterparts in the control condition, with a mean likelihood of 29.9%.*

By Result 2 we reject the null in favor of Hypothesis 4 that the average payoff matrix helps participants achieve better equilibrium outcomes. Furthermore, as participants' predictions are only marginally more accurate, the improvement in equilibrium outcomes appears to be primarily driven by participants gaining a better understanding of their own potential payoffs.

5.1.2 Determinants of Sharing Decisions: Comparative Statics

To better understand why complexity matters in this decision scenario, we examine the comparative statics to understand the determinants of participants' sharing decisions. According to our theoretical predictions, participants' willingness to share in a given round should increase with the price offered by the platform p_{it} (Prediction 1) and the correlation coefficient ρ_{it} (Prediction 2). They are also expected to be more likely to share when they believe that their match chooses to share (Prediction 3). We use the following panel logit specification to analyze these relationships.

$$P(\mathbb{1}\{\text{Share}\}_{it} = 1) = \frac{1}{1 + exp(-(\mathbf{X}'_{it}\beta + \gamma \cdot t + \mathbf{C}'_i\alpha))}$$

Table 4 presents the results from five logit specifications investigating factors affecting participants' sharing decisions. Independent variables include the matrix treatment dummy, the price, externality (ρ), the latter's interaction with the treatment dummy, the average absolute value of the secret numbers, whether a participant believes that their match will share, round, and the standardized measures risk aversion and cooperation. Arguably, the benefit of sharing (the price effect) is more intuitive than the cost of sharing (the externality effect). This is indeed what we find. As predicted in Hypothesis 1, the likelihood of sharing increases with the payment offered by the robot (0.096, $p < 0.01$, columns 3-5).

Result 3 (Price Effect) *A participant's likelihood of sharing increases by 9.6 pp as the price offered by the robot increases by \$1.*

Table 4: Determinants of Participant's Binary Sharing Decisions; Panel Logit Regression

	Choosing to Share				
	(1)	(2)	(3)	(4)	(5)
Matrix Treatment	-0.073***	-0.102**	-0.235***	-0.235***	-0.236***
	(0.018)	(0.048)	(0.031)	(0.032)	(0.031)
Price	0.123***	0.123***	0.096***	0.097***	0.096***
	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
Externality (ρ)	-0.041	-0.070	-0.045	-0.045	-0.046
	(0.048)	(0.080)	(0.058)	(0.058)	(0.058)
Externality $\times$ Matrix Treatment		0.057	0.171***	0.172***	0.173***
		(0.088)	(0.060)	(0.060)	(0.060)
Avg Absolute Value of Secret Numbers			-0.397***	-0.396***	-0.396***
			(0.016)	(0.016)	(0.015)
Prediction = Match Shares			0.329***	0.331***	0.332***
			(0.017)	(0.018)	(0.018)
Round			-0.002**	-0.002**	-0.002**
			(0.001)	(0.001)	(0.001)
Risk Averse (Standardized)					-0.018**
					(0.009)
Cooperation (Standardized)					-0.006
					(0.007)
Externality$\times$Matrix Treatment + Externality = 0 (p-values of χ^2 Test)		0.779	0.000	0.000	0.000
Controls	No	No	No	Yes	Yes
# of Observations	3840	3840	3840	3840	3840
# of Participants	192	192	192	192	192

[a] Standard errors in parentheses are clustered at the session level.

[b] Coefficients are marginal effects.

[c] Predict That Match Will Share is a dummy variable indicating a participant's prediction of their match's decision, taking the value 1 when a participant predicts their match will share.

[d] Controls included are identical as in Table 1

[e] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

By Result 3, we reject the null in favor of Hypothesis 1. It shows that participants respond to the changes in price (or benefits more broadly) offered by the platform.

Next we examine how participants respond to externalities. We find that the likelihood of sharing increases with the correlation coefficient, but only in the matrix treatment (0.127, Externality + Externality $\times$ Matrix Treatment, $p < 0.001$, columns 3-5).

Result 4 (Externality Effect) *A participant's likelihood of sharing increases by 1.27 pp as the correlation coefficient increases by 0.1 in the matrix treatment. By contrast, participants do not respond to changes in the correlation coefficient in the control condition.*

By Result 4, we reject the null in favor of Hypothesis 2 only in the matrix treatment. That is, the correlation coefficient influences privacy decisions only with a reduction in computational complexity. In the control condition, participants do not adequately account for the externalities when making cost-benefit evaluations on their own. This finding highlights an important mechanism underlying the impact of complexity: externalities complicate the cost-benefit analysis. Participants often fail to account for these externalities, leading them to underestimate the costs of sharing. Consequently, they tend to over-share compared to the optimal level.

Next, we examine the role of beliefs and find that participants are more likely to share if they believe that their match will share (0.332, $p < 0.01$, columns 3-5). The result is summarized below.

Result 5 (Beliefs) *When a participant predicts that their match will share, they are 33.2 pp more likely to share.*

By Result 5, we reject the null in favor of Hypothesis 3. Intuitively, since one's information can be inferred from their match's information, the platform gains more information as long as one participant in a pair decides to share. Therefore, the benefit of preserving one's own information decreases if the match shares, and vice versa. As a result, when one participant is expected to share, it is likely that the other participant will also decide to share.

Furthermore, we observe that the likelihood of sharing decreases with the average absolute value of the secret numbers (-0.396, $p < 0.01$, columns 3-5). The farther the secret numbers are from zero, the less likely participants are to share. Since the platform's guess is zero when no one shares, if a participant's secret numbers are close to zero, the platform is likely to guess their secret numbers fairly closely even if they choose not to share. This implies a lower cost of sharing, as measured by the sum of squared differences between the platform's guesses and actual secret numbers. Consequently, a lower price is sufficient to compensate for the cost, making the participant more likely to share. However, the cost increases as the secret numbers are farther away from zero. Participants therefore demand a higher price. As a result, they are less likely to share.

In other words, outliers are less likely to share. Our simulation results reported in Appendix C visualize this trend.

Lastly, column (5) of Table 4 presents the results for the correlation between sharing decisions and risk preferences as well as cooperation attitudes. We find that a participant who is one standard deviation more risk averse is 1.8 pp less likely to share ($p < 0.05$). This finding also explains why risk averse participants are more likely to best respond, as shown in Table 1. While participants tend to overshare, risk averse participants share less, bringing them closer to the best response. By this finding, we reject the null in favor of Hypothesis 8 that risk preferences is correlated with sharing decisions.

By contrast, sharing decisions do not correlate with cooperation attitudes (-0.006, $p > 0.10$). This suggests that participants may not perceive the environment as cooperative.

5.1.3 Welfare Analysis: Payoffs and Efficiency

We next examine the treatment effects on welfare, using two measures - participant payoffs and normalized efficiency. The latter is calculated for each pair in each round as follows:

$$\text{Normalized Efficiency} = \frac{\text{Actual Sum of Pair Payoffs - Minimum Sum of Pair Payoffs}}{\text{Maximum Sum of Pair Payoffs - Minimum Sum of Pair Payoffs}},$$

where the *Maximum Sum of Pair Payoffs* is the sum of payoffs for both participants for the Pareto efficient allocation(s) in a round. Likewise, the *Minimum Sum of Pair Payoffs* is the pair's sum of payoffs for the worst allocation in that round. Because of this normalization, this measure always lies between zero and one, inclusive.

Table 5 presents four OLS specifications investigating the treatment effects on participant payoffs and efficiency. Compared with mean payoffs of \$6.77 in the control condition (column 1), the \$1.49 increase in payoffs in the treatment condition represents a 21.9% increase. The 24.4 pp increase in efficiency is a 47.5% increase compared with the mean efficiency of 51.4% in the control condition (column 3). The results are summarized below.

Result 6 (Treatment Effect on Efficiency) *The matrix treatment increases efficiency by 24.4 pp compared with the control mean of 51.4%.*

In sum, our treatment condition reduces the computational complexity of the information sharing problem, leading to a substantial increase in best responses and equilibrium outcomes, leading to a significant increase in efficiency. Notably, while participants are responsive to prices in both conditions, the comparative statics with respect to externalities only work as expected in the matrix treatment.

Table 5: Treatment Effects on Payoff and Efficiency; Panel Linear Regression

	Payoff		Efficiency	
	(1)	(2)	(3)	(4)
Matrix Treatment	1.485***	1.485***	0.244***	0.244***
	(0.121)	(0.117)	(0.023)	(0.023)
Round		0.029**		0.002
		(0.014)		(0.002)
Risk Aversion (Standardized)		0.032		0.011
		(0.074)		(0.010)
Cooperation (Standardized)		0.046		-0.006
		(0.116)		(0.008)
Constants	6.773***	6.887***	0.514***	0.428***
	(0.125)	(0.561)	(0.016)	(0.039)
Controls	No	Yes	No	Yes
# of Observations	3840	3840	3840	3840
# of Participants	192	192	192	192

[a] Standard errors in parentheses are clustered at the session level.
[b] Controls included are identical as in Table 1
[c] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.2 Friends Decision Making

We first examine whether there is a difference in participants' willingness-to-accept (WTA) to announce their answers between the two questions of the same scenario. Theory predicts that participants are more willing to share their answers when those answers imply externalities for their friend ($\rho > 0$) than when the answers pertain only to themselves ($\rho = 0$). As a result, they should indicate a lower WTA to announce their answer.

Table 6 presents the test results of participant WTA in the pair of questions in each scenario. We find that participants demand lower payments to share answers with externalities ($\rho > 0$) than to share answers without externalities ($\rho = 0$) across all scenarios. However, the difference is significant only in the body weight scenario and when aggregating across all scenarios. Participants demand 3.80 points less to share the average body weight between themselves and their friends ($\rho > 0$) compared to their own body weight ($\rho = 0$), reflecting a 7.1% reduction. Across all scenarios, participants demand significantly lower payments to share their answers to questions with externalities ($\rho > 0$) compared to those without ($\rho = 0$). This difference represents 3.8% reduction. Results are summarized below.

Table 6: Difference in WTA in Each Pair of Questions

WTA (in Points)	w/o Externalities	w/ Externalities	Diff	t
Body Weight	53.16	49.36	3.80***	2.60
Number of Sexual Partners	62.82	61.18	1.64	1.16
Opinions on Intimate Relationships	52.69	51.76	0.94	1.11
Total	168.68	162.30	6.38**	2.42

[a] WTA is the minimum number a participant is willing to accept to announce their answer to a question.

[b] All analysis in this section is based on the recoded WTA, where participants who opted out of sending their friend the survey are assigned a WTA of 100.

Result 7 (WTA and Externalities) *A participant demands 7.1% less to share the average body weight of themselves and their friend compared to their own body weight. Considering all three scenarios, a participant demands 3.8% less to share the answers to questions with externalities compared to those without.*

By Result 7, we reject the null in favor of Hypothesis 5 regarding WTA and externalities. This result also align with our finding in the information sharing game, i.e., as the correlation coefficient increases, participants are more willing to share their information in exchange for the benefit. Consequently, they can be induced to share with a lower payment amount.

Next, we investigate determinants of participants' WTA. Since the elicited WTA is censored below at 0 and above at 100, we use the following Tobit model to analyze the determinants of the observed WTA, denoted by $\overline{WTA}_i$:

$$\overline{WTA}_i^* = X_i'\beta + \varepsilon_i, \text{ and } \overline{WTA}_i = \max\{0, \min\{\overline{WTA}_i^*, 100\}\},$$

where ε_i is normally distributed; X_i is a vector of independent variables and controls; and only a censored version of the latent variable $\overline{WTA}_i^*$ is observed, denoted by $\overline{WTA}_i$. In the data, $\overline{WTA}_i$ is participant i's WTA from each scenario. Among the WTA decisions across the three scenarios, between 46.3% and 55.7% are at the boundaries. Appendix Table E.11 shows the proportion of participants that are censored in each scenario.[4] Our independent variables include risk, ambiguity and cooperation preferences, and demographics (female, age, race and ethnicity). Furthermore, we control for whether they have participated in economic or psychology experiments before, whether they have taken introductory statistics courses, their performance in the review questions, and how long they and their friend have known each other.

Table 7 presents the results of the Tobit regressions for all scenarios. We find that participants'

[4]In a robustness check, we find similar results in linear models (see Table E.13 in Appendix E), as well as logit models where decisions are coded as willing to share ($WTA < 100$) or not ($WTA = 100$) (see Table E.14 in Appendix E).

Table 7: WTA to Announce Private Information; Tobit Regression

WTA to Announce:	Body Weight		# of Sexual Partners		Opinion	
	Own (1)	Avg (2)	Own (3)	Total (4)	Own (5)	Avg (6)
Likelihood of Sharing w/o Matrix	-30.670	-56.998**	-54.566*	-63.198**	-33.269	-37.313*
	(18.825)	(24.000)	(32.431)	(26.190)	(22.742)	(19.208)
Risk Aversion (Standardized)	-4.063	-6.929	0.560	-0.395	-4.732	-3.647
	(5.406)	(5.887)	(6.562)	(6.470)	(5.492)	(6.643)
Ambiguity Aversion (Standardized)	6.597	9.022	10.123**	8.355	1.606	2.857
	(7.105)	(6.455)	(4.872)	(5.547)	(5.999)	(5.419)
Cooperation (Standardized)	0.688	-0.770	0.196	1.829	5.351	3.726
	(4.465)	(4.931)	(7.716)	(7.223)	(5.249)	(5.491)
Female	7.105	2.455	5.888	11.387	10.222	11.206
	(11.281)	(12.740)	(12.503)	(12.543)	(13.385)	(13.770)
Proportion of Female	18.061	42.468	-8.010	15.948	19.175	25.371
	(33.621)	(31.694)	(37.083)	(45.200)	(38.862)	(42.170)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
# of Participants	192	192	192	192	192	192

[a] Standard errors in parentheses are clustered at the session level.
[b] The outcome variables are participant's WTA to announce their answers. The first two regressions represent the Body Weight scenario. The second two regressions represent the Intimate Relationships scenario. The last two regressions represent the Opinions on Intimate Relationships scenario.
[c] WTA is measured in points and is chosen on the interval [0, 100]. A higher WTA means being less willing to share.
[d] Controls include a participant's age, ethnicity, and indicators suggesting whether they have participated in economics or psychological studies before, whether they have taken introductory statistics courses, whether they correctly answered two key questions (which over 25% of participants answered incorrectly), and how long a participant and their friend have known each other.
[e] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

decisions in the control condition of the information sharing game (without matrix) predict their decisions in the friends decision making module. Decisions in the control condition are made without any assistance, which reflects their own ability to evaluate the costs and benefits of sharing. Participants who are more likely to share in these decisions have lower WTA, i.e., they are also more likely to share in the friends decision making module. They demand significantly lower amounts to announce their answers in all scenarios. In general, a participant who is 10% more likely to share in the information sharing game demands 3.1–6.3 points ($0.08–$0.16) less to share their private information. The results are summarized below.

Result 8 (Correlations Between Sharing Decisions in Different Contexts) *Participants who are more likely to share in the control condition of the information sharing game demand a lower payment to announce their answers in the friends decision making module.*

By Result 8, we reject the null in favor of Hypotheses 7. This indicates that privacy preferences measured in neutral contexts can predict information-sharing decisions in more concrete settings where real personal information is involved.

In Table 7, we also note that participants who are more ambiguity averse demand higher payments to announce their answers, that is, they are less likely to share. However, the correlations are only significant for the question about one's own number of sexual partners. Specifically, a participant who is one standard deviation more ambiguity averse demands 10.12 more points ($p < 0.05$, column 3), or \$0.25, to announce the number of sexual partners they have ever had. Participants may be uncertain about the potential negative consequences of sharing their personal information. Greater ambiguity aversion is correlated with a lower willingness to share.[5]

We also explore whether there is any gender difference in information sharing decisions. Across all scenarios in the module, female participants demand a higher payment to share their private information publicly compared to their non-female counterparts. While these effects are not significant in the full sample, they become significant in the restricted sample and after controlling for question-specific factors (see Table E.12 in Appendix E). Furthermore, we control for the proportion of female participants in the session, assuming that a higher share of female participants may create a more comfortable environment for revealing information on sensitive topics. However, we find neither gender nor gender composition effect (columns 1-6, Table 7).

5.3 Sharing on Instagram

Towards the end of each session, we collect participants' privacy settings on Instagram. In our sample, 91.1% are Instagram users. The privacy setting we focus on is whether they enable their activity information on ads partner platforms to be used by Instagram. Among Instagram users, 45.7% enable this option, which is the default, while the remaining 54.3% disable this option.

When asked to review this setting as well as the list of advertisers they saw ads from, 82.5% of those who enabled sharing report that it is at least somewhat necessary to adjust their settings. This need for adjustment does not appear to stem from a lack of awareness, since 81.71% of Instagram users in our sample indicate at least some knowledge that Instagram uses private information for

[5] Note we present the full sample analysis in Table 7. Additional analysis restricting the sample to those who answer all questions are presented in Appendix Table E.12.

ad personalization. A potential explanation is that participants struggle to compute the costs of sharing. While the benefits of keeping the default are straightforward, as users avoid the effort of adjusting settings and receive more relevant ads, the costs are more complex to evaluate. Participants may not fully recognize how Instagram uses their data and how this may lead to longer-term risks, such as data leakage to third parties or inadvertent revelation of their friends' preferences. As a result, they are prone to underestimating the true costs of sharing and may end up disclosing more than they would want to.

Table 8: Influencing Factors of Privacy Setting on Instagram; Logit Regression

	Sharing Enabled			
	(1)	(2)	(3)	(4)
Cooperation (Standardized)	0.140***	0.143***	0.143***	0.142***
	(0.038)	(0.039)	(0.039)	(0.039)
Risk Aversion (Standardized)	-0.044	-0.051	-0.049	-0.049
	(0.036)	(0.035)	(0.037)	(0.036)
Ambiguity Aversion (Standardized)	0.012	0.016	0.017	0.017
	(0.040)	(0.039)	(0.039)	(0.039)
Likelihood of Sharing w/o Matrix	-0.006			
	(0.208)			
Total WTA for the Three $\rho = 0$ Questions (in \$)		-0.028**		
		(0.011)		
Total WTA for the Three $\rho > 0$ Questions (in \$)			-0.020*	
			(0.011)	
Total WTA for Six Questions (in \$)				-0.013**
				(0.006)
Controls	Yes	Yes	Yes	Yes
# of Participants	175	175	175	175

[a] Standard errors in parentheses are clustered at the session level.
[b] Marginal effects are reported in the table.
[c] The sample is comprised of the 175 Instagram users.
[d] Controls include a participant's gender, age, ethnicity, and indicators of whether they are aware that their behaviors on the ad partners are shared with Instagram.
[d] $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table 8 presents four logit specifications investigating factors influencing a participant's Instagram privacy setting. Independent variables include standardized cooperation attitude, risk and ambiguity aversion, the likelihood of sharing in the control condition of the information sharing game (column 1), (iteratively adding) WTA for $\rho = 0$ (column 2), WTA for $\rho > 0$ (column 3), and total WTA (column 4) in the friends decision making module. We find that participants who are one standard deviation more more cooperative in the first round of the public goods game are 14 pp more likely to share their activity information with ad partners on Instagram ($p < 0.01$, columns 1-4). By this finding, we reject the null in favor of Hypothesis 10. This result is consistent with a finding in an online field experiment that prosocial users are more likely to contribute rare movie rating data to the recommender algorithm (Chen et al., 2010). By contrast, we do not observe significant correlations between privacy settings and risk or ambiguity preferences.

We further find that the sharing decisions in the two real-world contexts, i.e., the friends decision making module and Instagram, are strongly correlated. Specifically, for every additional $1 (or 40 points) participants demand in the friends decision making module, they are 1.3 pp less likely to enable sharing on Instagram ($p < 0.05$, column 4). In other words, participants who tend to share more in the friends decision making module also have more lenient privacy settings on Instagram. The results are summarized below.

Result 9 (Correlation between Sharing Decisions in Realistic Contexts) *Participants who demand a lower payment to share in the friends decision making module are significantly more likely to enable sharing on Instagram.*

By Result 9, we reject the null in favor of Hypothesis 7. While we are only able to predict participants' sharing decisions on Instagram using their decisions in the friends decision making module and fail to observe a direct connection with decisions in the neutral information sharing game, the friends decision making module itself is grounded in the information sharing game and is strongly predicted by it. This establishes a full chain from abstract, lab-based decisions to real-world behavior, suggesting that privacy preferences measured in the lab can serve as a reliable proxy for real-world privacy preferences with comparable computational complexity.

6 Conclusion

This study investigates individuals' information sharing decisions in environments with externalities, simulating their behaviors on social networking platforms (e.g., Instagram). As accounting for externalities complicates the computation of benefits and costs of sharing, we focus on how this increased complexity affects individuals' information sharing decisions.

Our survey of Instagram privacy settings suggests that the current settings of many participants may not be optimal. Although a large share of participants are aware that Instagram uses their personal information for ad personalization, 45.7% retain the default option that enables activity sharing cross platforms. Moreover, 82.5% of those who keep the default later report that their settings *should* be adjusted, indicating that their privacy settings were not optimal.

To better understand how far users are from making optimal information sharing decisions and factors affecting these decisions, we turn to a controlled laboratory environment that allows us to investigate the comparative statics. As suggested by our simplified framework based on Acemoglu et al. (2022), deriving the equilibrium typically requires knowledge of the direct benefit of sharing, the correlation between a user's information and that of others (i.e., the externality), and the intrinsic value users assign to privacy. However, it is difficult to directly obtain reliable estimates of these parameters in the real world. The laboratory setting enables us to manipulate these factors and isolate their effects on sharing decisions.

Our friends decision making module simplifies the social media context by asking participants to share genuine private information about themselves and their friends, but in a more controlled setting. In this module, we manipulate the benefit of sharing and the externality, but allow participants to make decisions based on their own intrinsic valuation of privacy. The results align with theoretical predictions: participants respond to negative externalities by sharing more when such externalities are present. Since individual valuations of privacy remain unobserved, we cannot directly determine whether these decisions are optimal.

By contrast, the information sharing game enables us to estimate the effects of computational complexity on information sharing decisions under externalities, as we exogenously vary all three key parameters. Unlike Instagram and the friends decision making module, where the cost of sharing is difficult to pin down (Chen et al., 2021; Reidenberg et al., 2015), the game clearly specifies how costs are determined, enabling us to isolate the impact of computational complexity directly.

In the information sharing game without assistance, we find that participants struggle to resolve the trade-offs associated with sharing and tend to over-share. Reducing computational complexity significantly increases the likelihood that participants make optimal sharing decisions, which in turn enhances their payoffs and efficiency. Our findings further suggest that even if costs and benefits are explicitly described, cognitive limitations still hinder participants' ability to compute and compare them.

Participants' difficulty in computation likely stems from insufficient attention to the externalities that their sharing decisions impose on others. In this study, we focus on negative externalities, thereby observing a general pattern of oversharing. In contrast, in contexts where sharing generates positive externalities, benefits are usually more complex than costs. While costs are typically the individual burden of sharing, users may fail to recognize the broader, long-term benefits that spill over to others, which can just as easily result in undersharing. These patterns suggest that helping users recognize the broader impact of their actions could promote decision making. Policymakers might achieve this by incorporating reminder messages that explicitly highlight the potential effects on others when users make privacy decisions. This solution is supported by the findings of Friehe et al. (2025), which show that reminding individuals of the negative externalities their decisions impose on others decreases the willingness to sell their data.

We observe cross-module correlations between participants' sharing decisions, which support the predictive power of laboratory games for related real-world behaviors (Karlan, 2005; Dohmen et al., 2011). Moreover, in practice, defining optimal sharing decisions in real-world settings is challenging, and the best practice to share personal information remains uncertain. Our results suggest one potential strategy: classify users based on preferences elicited in neutral contexts and then recommend strategies adopted by similar users. Such personalized intervention could help promote better privacy decision-making.

Overall, this study highlights the critical role of computational complexity in shaping individuals' privacy decisions in environments with externalities. As information sharing becomes increasingly pervasive across digital platforms, simplifying decision environments can help users make more informed choices and improve user welfare. Future research can build on these insights to design interventions that promote more optimal information sharing on real-world social networking platforms.

References

Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asu Ozdaglar, "Too much data: Prices and inefficiencies in data markets," *American Economic Journal: Microeconomics*, 2022, *14* (4), 218–256.

Ackfeld, Viola and Werner Güth, "Personal information disclosure under competition for benefits: Is sharing caring?," *Games and Economic Behavior*, 2023, *140*, 1–32.

Acquisti, Alessandro and Jens Grossklags, "What can behavioral economics teach us about privacy?," in "Digital privacy," Auerbach Publications, 2007, pp. 363–378.

_ , Curtis Taylor, and Liad Wagman, "The economics of privacy," *Journal of economic Literature*, 2016, *54* (2), 442–492.

_ , Laura Brandimarte, and George Loewenstein, "Secrets and likes: The drive for privacy and the difficulty of achieving it in the digital age," *Journal of Consumer Psychology*, 2020, *30* (4), 736–758.

_ , Leslie K John, and George Loewenstein, "What is privacy worth?," *The Journal of Legal Studies*, 2013, *42* (2), 249–274.

Adamson, David M., Annie Brothers, Sasha Romanosky, Marjory S. Blumenthal, Douglas C. Ligor, Karlyn D. Stanley, Peter Schirmer, and Julia Vidal Verástegui, *Cyberstalking: A Growing Challenge for the U.S. Legal System*, Santa Monica, CA: RAND Corporation, 2023.

Aridor, Guy, Rafael Jiménez-Durán, Ro'ee Levy, and Lena Song, "The Economics of Social Media," *Journal of Economic Literature*, December 2024, *62* (4), 1422–74.

Becker, Gordon M, Morris H DeGroot, and Jacob Marschak, "Measuring utility by a single-response sequential method," *Behavioral science*, 1964, *9* (3), 226–232.

Ben-Porath, Elchanan, "The complexity of computing a best response automaton in repeated games with mixed strategies," *Games and Economic Behavior*, 1990, *2* (1), 1–12.

Benndorf, Volker and Hans-Theo Normann, "The willingness to sell personal data," *The Scandinavian Journal of Economics*, 2018, *120* (4), 1260–1278.

Bergemann, Dirk, Alessandro Bonatti, and Tan Gan, "The economics of social data," *The RAND Journal of Economics*, 2022, *53* (2), 263–296.

Bernheim, B Douglas and Charles Sprenger, "On the empirical validity of cumulative prospect theory: Experimental evidence of rank-independent probability weighting," *Econometrica*, 2020, *88* (4), 1363–1409.

Biener, Christian, Martin Eling, and Martin Lehmann, "Balancing the desire for privacy against the desire to hedge risk," *Journal of Economic Behavior & Organization*, 2020, *180*, 608–620.

Brandimarte, Laura, Alessandro Acquisti, and George Loewenstein, "Misplaced confidences: Privacy and the control paradox," *Social psychological and personality science*, 2013, *4* (3), 340–347.

Charness, Gary, Ramón Cobo-Reyes, and Natalia Jiménez, "Identities, selection, and contributions in a public-goods game," *Games and Economic Behavior*, 2014, *87*, 322–338.

Chen, Daniel L, Martin Schonger, and Chris Wickens, "oTree—An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 2016, *9*, 88–97.

Chen, Rex, Fei Fang, Thomas Norton, Aleecia M McDonald, and Norman Sadeh, "Fighting the fog: Evaluating the clarity of privacy disclosures in the age of CCPA," in "Proceedings of the 20th Workshop on Workshop on Privacy in the Electronic Society" 2021, pp. 73–102.

Chen, Yan, F. Maxwell Harper, Joseph Konstan, and Sherry Xin Li, "Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens," *The American Economic Review*, September 2010, *100* (4), 1358–1398.

Choi, Jay Pil, Doh-Shin Jeon, and Byung-Cheol Kim, "Privacy and personal data collection with information externalities," *Journal of Public Economics*, 2019, *173*, 113–124.

Danz, David, Lise Vesterlund, and Alistair J Wilson, "Belief elicitation and behavioral incentive compatibility," *American Economic Review*, 2022, *112* (9), 2851–2883.

Dixon, Stacy Jo, "Average daily time spent on social media worldwide 2012-2024," Apr 10 2024.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G Wagner, "Individual risk attitudes: Measurement, determinants, and behavioral consequences," *Journal of the european economic association*, 2011, *9* (3), 522–550.

Ellsberg, Daniel, "Risk, ambiguity, and the Savage axioms," *The quarterly journal of economics*, 1961, *75* (4), 643–669.

Enke, Benjamin and Cassidy Shubatt, "Quantifying lottery choice complexity," Technical Report, National Bureau of Economic Research 2023.

_ **and Thomas Graeber**, "Cognitive uncertainty," *The Quarterly Journal of Economics*, 2023, *138* (4), 2021–2067.

Fairfield, Joshua AT and Christoph Engel, "Privacy as a public good," *Duke LJ*, 2015, *65*, 385.

Fisher, Terri D, "Sex of experimenter and social norm effects on reports of sexual behavior in young men and women," *Archives of sexual Behavior*, 2007, *36*, 89–100.

_ , "The impact of socially conveyed norms on the reporting of sexual behavior and attitudes by men and women," *Journal of Experimental Social Psychology*, 2009, *45* (3), 567–572.

_ , "Gender roles and pressure to be truthful: The bogus pipeline modifies gender differences in sexual but not non-sexual behavior," *Sex Roles*, 2013, *68* (7), 401–414.

Friehe, Tim, Leonie Gerhards, and Franziska Weber, "Keep them out of It! How information externalities affect the willingness to sell personal data online," *Journal of Economic Psychology*, 2025, p. 102830.

Frydman, Cary and Lawrence J Jin, "Efficient coding and risky choice," *The Quarterly Journal of Economics*, 2022, *137* (1), 161–213.

Gilboa, Itzhak, "The complexity of computing best-response automata in repeated games," *Journal of economic theory*, 1988, *45* (2), 342–352.

_ **and David Schmeidler**, "Maxmin expected utility with non-unique prior," *Journal of mathematical economics*, 1989, *18* (2), 141–153.

Goffman, Erving, *Stigma: Notes on the management of spoiled identity*, Simon and schuster, 2009.

_ **et al.**, "The presentation of self in everyday life. 1959," *Garden City, NY*, 2002, *259.*

Goh, Khim-Yong, Kai-Lung Hui, and Ivan PL Png, "Privacy and marketing externalities: Evidence from do not call," *Management Science*, 2015, *61* (12), 2982–3000.

Goldfarb, Avi and Verina F Que, "The economics of digital privacy," *Annual Review of Economics*, 2023, *15* (1), 267–286.

Greiner, Ben, "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, 2015, *1* (1), 114–125.

Hancock, Jeffrey T, Catalina Toma, and Nicole Ellison, "The truth about lying in online dating profiles," in "Proceedings of the SIGCHI conference on Human factors in computing systems" 2007, pp. 449–452.

Holt, Charles A and Susan K Laury, "Risk aversion and incentive effects," *American economic*

review, 2002, *92* (5), 1644–1655.
Huberman, Bernardo A, Eytan Adar, and Leslie R Fine, "Valuating privacy," *IEEE security & privacy*, 2005, *3* (5), 22–25.
Ichihashi, Shota, "The economics of data externalities," *Journal of Economic Theory*, 2021, *196*, 105316.
John, Leslie K, Alessandro Acquisti, and George Loewenstein, "Strangers on a plane: Context-dependent willingness to divulge sensitive information," *Journal of consumer research*, 2011, *37* (5), 858–873.
Karlan, Dean S, "Using experimental economics to measure social capital and predict financial decisions," *American Economic Review*, 2005, *95* (5), 1688–1699.
Kim, Cheonsoo and Sung-Un Yang, "Like, comment, and share on Facebook: How each behavior differs from the other," *Public relations review*, 2017, *43* (2), 441–449.
Kokolakis, Spyros, "Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon," *Computers & security*, 2017, *64*, 122–134.
Krupka, Erin L and Roberto A Weber, "Identifying social norms using coordination games: Why does dictator game sharing vary?," *Journal of the European Economic Association*, 2013, *11* (3), 495–524.
Kumaraguru, Ponnurangam and Lorrie Faith Cranor, "Privacy indexes: a survey of Westin's studies," 2005.
Laudon, Kenneth, "Extensions to the theory of markets and privacy: Mechanics of pricing information," 1997.
Lee, Hosub and Alfred Kobsa, "Understanding user privacy in Internet of Things environments," in "2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)" IEEE 2016, pp. 407–412.
Lee, Yi-Shan and Roberto A Weber, "Revealed privacy preferences: Are privacy choices rational?," *Management Science*, 2025, *71* (3), 2657–2677.
Li, Shengwu, "Obviously strategy-proof mechanisms," *American Economic Review*, 2017, *107* (11), 3257–3287.
Li, Zhihua, Julia Müller, Peter P Wakker, and Tong V Wang, "The rich domain of ambiguity explored," *Management Science*, 2018, *64* (7), 3227–3240.
Lin, Tesary, "Valuing intrinsic and instrumental preferences for privacy," *Marketing Science*, 2022, *41* (4), 663–681.
Litt, Eden, "Knock, knock. Who's there? The imagined audience," *Journal of broadcasting & electronic media*, 2012, *56* (3), 330–345.
Moon, Youngme, "Intimate exchanges: Using computers to elicit self-disclosure from consumers," *Journal of consumer research*, 2000, *26* (4), 323–339.
Odlyzko, Andrew, "Privacy, economics, and price discrimination on the internet," in "Proceedings of the 5th international conference on Electronic commerce" 2003, pp. 355–366.
Oprea, Ryan, "Decisions under risk are decisions under complexity," *American Economic Review*, 2024, *114* (12), 3789–3811.
Pew Research Center, "Social Media Fact Sheet," November 2024. Pew Research Center, Washington, D.C. (November 13, 2024). Retrieved from https://www.pewresearch.org/internet/fact-sheet/social-media/.
Posner, Michael I, *Chronometric explorations of mind.*, Lawrence Erlbaum, 1978.
Posner, Richard A, "The economics of privacy," *The American economic review*, 1981, *71* (2), 405–409.

Prince, Jeffrey T and Scott Wallsten, "How much is privacy worth around the world and across platforms?," *Journal of Economics & Management Strategy*, 2022, *31* (4), 841–861.

Puri, Indira, "Simplicity and risk," *The Journal of Finance*, 2025, *80* (2), 1029–1080.

Reidenberg, Joel R, Travis Breaux, Lorrie Faith Cranor, Brian French, Amanda Grannis, James T Graves, Fei Liu, Aleecia McDonald, Thomas B Norton, Rohan Ramanath et al., "Disagreeable privacy policies: Mismatches between meaning and users' understanding," *Berkeley Tech. LJ*, 2015, *30*, 39.

Ricci, Francesco, Lior Rokach, and Bracha Shapira, eds, *Recommender Systems Handbook*, 3 ed., New York, NY: Springer, 2022.

Schudy, Simeon and Verena Utikal, "'You must not know about me'—On the willingness to share personal data," *Journal of Economic Behavior & Organization*, 2017, *141*, 1–13.

Shubatt, Cassidy and Jeffrey Yang, "Tradeoffs and comparison complexity," *arXiv preprint arXiv:2401.17578*, 2024.

Smith, H Jeff, Tamara Dinev, and Heng Xu, "Information privacy research: an interdisciplinary review," *MIS quarterly*, 2011, pp. 989–1015.

Smith, Vernon L, "Microeconomic systems as an experimental science," *The American economic review*, 1982, *72* (5), 923–955.

Solove, Daniel J, "Introduction: Privacy self-management and the consent dilemma," *Harv. L. Rev.*, 2012, *126*, 1880.

Stigler, George J, "The economics of information," *Journal of political economy*, 1961, *69* (3), 213–225.

_ , "Information in the labor market," *Journal of political economy*, 1962, *70* (5, Part 2), 94–105.

_ , "An introduction to privacy in economics and politics," *The Journal of Legal Studies*, 1980, *9* (4), 623–644.

Swire, Peter P and Robert E Litan, *None of your business: world data flows, electronic commerce, and the European privacy directive*, Rowman & Littlefield, 2010.

Tafradzhiyski, Nayden, "Instagram Revenue and Usage Statistics (2025)," https://www.businessofapps.com/data/instagram-statistics 2025. Accessed: 2025-08-25.

Toma, Catalina L, Jeffrey T Hancock, and Nicole B Ellison, "Separating fact from fiction: An examination of deceptive self-presentation in online dating profiles," *Personality and social psychology bulletin*, 2008, *34* (8), 1023–1036.

Tsai, Janice Y, Serge Egelman, Lorrie Cranor, and Alessandro Acquisti, "The effect of online privacy information on purchasing behavior: An experimental study," *Information systems research*, 2011, *22* (2), 254–268.

Vitak, Jessica, "The impact of context collapse and privacy on social network site disclosures," *Journal of broadcasting & electronic media*, 2012, *56* (4), 451–470.

Wang, Qingyi, Yan Chen, and Eytan Adar, "Adapting economic games to personalize privacy nudges," February 2024. Title of the publication associated with this dataset: AEA Randomized Controlled Trials.

Appendices

Table of Contents

A Derivations of Model Predictions

First, we introduce two facts regarding the properties of joint normal distributions, which are useful for later derivations.

Fact 1 *For two variables Z_1, Z_2 which follow the joint normal distribution*

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$$

the conditional distribution of Z_1 given $Z_2 = z_2$ is $\mathcal{N}(\hat{\mu}, \hat{\Sigma})$, where

$$\hat{\mu} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(z_2 - \mu_2)$$
$$\hat{\Sigma} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

Fact 2 *The robot's posterior belief about θ conditional on signal realization $X = x$ is normally distributed with mean*

$$\mathbb{E}(\theta|X = x) = (\frac{\sigma_\epsilon^2}{\sigma_\theta^2 + \sigma_\epsilon^2})\mu + (\frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\epsilon^2})x$$

and variance

$$Var(\theta|X = x) = \frac{\sigma_\theta^2 \sigma_\epsilon^2}{\sigma_\theta^2 + \sigma_\epsilon^2}$$

Based on $\begin{pmatrix} \theta_i \\ \theta_j \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$, we also know that $\mathbb{E}[\theta_1] = \mathbb{E}[\theta_2] = 0$, $Cov(\theta_1, \theta_2) = \rho$, and $Var(\theta_1) = Var(\theta_2) = 1$.

Therefore, we have $\mathbb{E}[\theta_1^2] = \mathbb{E}[\theta_2^2] = 1$, and $\mathbb{E}[\theta_1\theta_2] = \rho$.

We now start to derive the model predictions by solving for player i's best responses given player j's actions. In particular, we aim to prove the following propositions:

Proposition 1 *Player i's best response to $a_j = 1$ is $a_i = 1$ if and only if $p_i \geq \frac{(2-\rho^2)^2}{2(4-\rho^2)}v_i$.*

Proposition 2 *Player i's best response to $a_j = 0$ is $a_i = 1$ if and only if $p_i \geq \frac{1}{2}v_i$.*

Proposition 3 *It is cheaper for the robot to induce two players share than one if and only if $\rho^2 \geq 0.72$.*

PROOF OF PROPOSITION 1

Proof.
Given player j chooses to share, i.e., $a_j = 1$.

We first prove the "only if" part, i.e., player i's best response to $a_j = 1$ is $a_i = 1 \Rightarrow p_i \geq \frac{(2-\rho^2)^2}{2(4-\rho^2)} v_i$.

If player i does not share, the robot's posterior belief follows

$$\begin{pmatrix}\theta_i \\ X_j\end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix}\mathbb{E}[\theta_i] \\ \mathbb{E}[\theta_j + \epsilon_j]\end{pmatrix}, \begin{pmatrix}Var(\theta_i) & Cov(\theta_i, \theta_j + \epsilon_j) \\ Cov(\theta_i, \theta_j + \epsilon_j) & Var(\theta_j + \epsilon_j)\end{pmatrix}\right) = \mathcal{N}\left(\begin{pmatrix}0 \\ 0\end{pmatrix}, \begin{pmatrix}1 & \rho \\ \rho & 2\end{pmatrix}\right)$$

By Fact 1, the conditional distribution of θ_i given $X_j = x_j$ therefore has mean

$$\mathbb{E}(\theta_i | X_j = x_j) = 0 + \rho \times \frac{1}{2} \times (x_j - 0) = \frac{\rho}{2} x_j$$

and variance

$$Var(\theta_i | X_j = x_j) = 1 - \rho \times \frac{1}{2} \times \rho = 1 - \frac{\rho^2}{2}.$$

The robot's best strategy of guessing is to guess the posterior mean. Therefore, when $a_i = 0, a_j = 1$, i.e., $\boldsymbol{X_a} = X_j$, the robot's guess of θ_i is

$$\hat{\theta}_i = \mathbb{E}(\theta_i | X_j = x_j) = \frac{\rho}{2} x_j$$

Player i's payoff is

$$u_i = v_i \times Var(\theta_i | X_j = x_j) = (1 - \frac{\rho^2}{2}) v_i$$

If player i chooses to share, by Fact 2, the robot's posterior mean becomes

$$\mathbb{E}(\hat{\theta}_i | X_i = x_i) = (\frac{1}{1 + 1 - \frac{\rho^2}{2}} \times \frac{\rho}{2} x_j + \frac{1 - \frac{\rho^2}{2}}{1 + 1 - \frac{\rho^2}{2}} x_i) = \frac{2 - \rho^2}{4 - \rho^2} x_i + \frac{\rho}{4 - \rho^2} x_j$$

the variance of $\hat{\theta}_i$ reduces to

$$Var(\hat{\theta}_i | X_i = x_i) = \frac{1 - \frac{\rho^2}{2}}{1 + 1 - \frac{\rho^2}{2}} = \frac{2 - \rho^2}{4 - \rho^2}$$

The robot's guess becomes

$$\hat{\hat{\theta}}_i = \mathbb{E}(\hat{\theta}_i | X_i = x_i) = \frac{2 - \rho^2}{4 - \rho^2} x_i + \frac{\rho}{4 - \rho^2} x_j$$

Player i's payoff is

$$u_i = v_i \times Var(\hat{\theta}_i | X_i = x_i) + p_i = \frac{2 - \rho^2}{4 - \rho^2} v_i + p_i$$

Therefore, player i's best response to $a_j = 1$ is $a_i = 1$ only if $u_i(a_i = 1|a_j = 1) \geq u_i(a_i = 0|a_j = 1)$, which is equivalent to

$$p_i \geq \frac{(2-\rho^2)^2}{2(4-\rho^2)} v_i \tag{4}$$

Now we prove the "if" part, which is: player i's best response to $a_j = 1$ is $a_i = 1 \Leftarrow p_i \geq \frac{(2-\rho^2)^2}{2(4-\rho^2)} v_i$. Given the price offered by the robot, $p_i \geq \frac{(2-\rho^2)^2}{2(4-\rho^2)} v_i$, player i needs to compare between two actions: not to share ($a_i = 0$) or to share ($a_i = 1$).

If player i chooses not to share, as calculated above, the robot's posterior belief has mean

$$\mathbb{E}(\theta_i|X_j = x_j) = 0 + \rho \times \frac{1}{2} \times (x_j - 0) = \frac{\rho}{2} x_j$$

and variance

$$Var(\theta_i|X_j = x_j) = 1 - \rho \times \frac{1}{2} \times \rho = 1 - \frac{\rho^2}{2}.$$

Player i's payoff is

$$u_i = v_i \times Var(\theta_i|X_j = x_j) = (1 - \frac{\rho^2}{2}) v_i$$

If player i chooses to share, as calculated above, the robot's posterior belief has mean

$$\mathbb{E}(\hat{\theta}_i|X_i = x_i) = (\frac{1}{1 + 1 - \frac{\rho^2}{2}} \times \frac{\rho}{2} x_j + \frac{1 - \frac{\rho^2}{2}}{1 + 1 - \frac{\rho^2}{2}} x_i) = \frac{2-\rho^2}{4-\rho^2} x_i + \frac{\rho}{4-\rho^2} x_j$$

and variance

$$Var(\hat{\theta}_i|X_i = x_i) = \frac{1 - \frac{\rho^2}{2}}{1 + 1 - \frac{\rho^2}{2}} = \frac{2-\rho^2}{4-\rho^2}$$

Player i's payoff is

$$u_i = v_i \times Var(\hat{\theta}_i|X_i = x_i) + p_i = \frac{2-\rho^2}{4-\rho^2} v_i + p_i \geq (\frac{2-\rho^2}{4-\rho^2} + \frac{(2-\rho^2)^2}{2(4-\rho^2)}) v_i = (1 - \frac{\rho^2}{2}) v_i$$

Therefore, $u_i(a_i = 1|a_j = 1) \geq u_i(a_i = 0|a_j = 1)$, $a_i = 0$ is weakly dominated by $a_i = 1$. Player i's best response to $a_j = 1$ should be $a_i = 1$. ■

PROOF OF PROPOSITION 2

Proof.
Given player j chooses not to share, i.e., $a_j = 0$.

We first prove the "only if" part, i.e., player i's best response to $a_j = 0$ is $a_i = 1 \Rightarrow p_i \geq \frac{1}{2} v_i$.

If player i does not share, the variance of robot's posterior belief about θ_i is its prior 1. Its optimal guess is also to guess its prior mean value 0. Player i's payoff is $u_i = v_i$.

If player i chooses to share, by Fact 2, the robot's posterior mean becomes

$$\mathbb{E}(\theta_i|X_i = x_i) = (\frac{1}{1+1} \times 0 + \frac{1}{1+1} x_i) = \frac{1}{2} x_i$$

the variance of θ_i becomes

$$Var(\theta_i|X_i = x_i) = \frac{1 \times 1}{1+1} = \frac{1}{2}$$

The robot's guess becomes

$$\hat{\theta}_i = \mathbb{E}(\theta_i|X_i = x_i) = \frac{1}{2} x_i$$

Player i's payoff is

$$u_i = v_i \times Var(\theta_i|X_i = x_i) + p_i = \frac{1}{2} v_i + p_i$$

Therefore, player i's best response to $a_j = 0$ is $a_i = 1$ if and only if $u_i(a_i = 1|a_j = 0) \geq u_i(a_i = 0|a_j = 0)$, which is equivalent to

$$p_i \geq \frac{1}{2} v_i \tag{5}$$

Now we prove the "if" part, which is: player i's best response to $a_j = 0$ is $a_i = 1 \Leftarrow p_i \geq \frac{1}{2}$. Given the price offered by the robot, $p_i \geq \frac{1}{2} v_i$, player i needs to compare between two actions: not to share ($a_i = 0$) or to share ($a_i = 1$).

If player i chooses not to share, as discussed above, the robot's posterior belief has mean 0 and variance 1. Player i's payoff is 1.

If player i chooses to share, as calculated above, the robot's posterior belief has mean

$$\mathbb{E}(\theta_i|X_i = x_i) = (\frac{1}{1+1} \times 0 + \frac{1}{1+1} x_i) = \frac{1}{2} x_i$$

and variance

$$Var(\theta_i|X_i = x_i) = \frac{1 \times 1}{1+1} = \frac{1}{2}$$

Player i's payoff is

$$u_i = v_i \times Var(\theta_i|X_i = x_i) + p_i = \frac{1}{2} v_i + p_i \geq \frac{1}{2} v_i + \frac{1}{2} v_i = v_i$$

Therefore, $u_i(a_i = 1|a_j = 0) \geq u_i(a_i = 0|a_j = 0)$, $a_i = 0$ is weakly dominated by $a_i = 1$. Player i's best response to $a_j = 0$ should be $a_i = 1$. ∎

PROOF OF PROPOSITION 3
Proof.

We first prove the "only if" part, i.e., it is cheaper for the robot to induce two players share than one $\Rightarrow \rho^2 \geq 0.72$.

To prove this, we need to compare the conditions in (1) and (2). To make it cheaper for the robot to induce two players to share than one, we need the sum of prices paid to both players to be smaller than the price paid to only one player, which is equivalent to

$$2 \times \frac{(2-\rho^2)^2}{2(4-\rho^2)} v_i \leq \frac{1}{2} v_i$$

Therefore, we have $\rho^2 \geq 0.72$.

Now we prove the "if" part, which is: it is cheaper for the robot to induce two players share than one $\Leftarrow \rho^2 \geq 0.72$.

To induce one player to share, the minimum price offered by the robot should be $\frac{1}{2}$ according to Proposition 2.

To induce both players to share, the minimum price offered to each player by the robot should be $\frac{(2-\rho^2)^2}{2(4-\rho^2)} v_i$ according to Proposition 1. Therefore, the total minimum price it offers to both players should be

$$\frac{(2-\rho^2)^2}{4-\rho^2} v_i$$

Since this price decreases with ρ^2, the maximum of the minimum price is reached when $\rho = 0.72$. The maximum is $\frac{(2-0.72)^2}{4-0.72} = 0.50 = \frac{1}{2}$.

Therefore, it is always cheaper to induce two players to share than one. ∎

B Pre-Analysis Plan

Trail Link: https://www.socialscienceregistry.org/trials/12875

The overlaps between the pre-analysis plan and the theoretical framework and experimental design sections are omitted here.

B.1 Summary of Outcome Variables

Primary outcomes:

- In the information sharing game, the primary outcome is the likelihood that participants choose the best responses given by the realized payoff matrices. The analysis will be based on the following logit regression function:

$$logit(a_i) = \beta_0 + \beta_1 p_i + \beta_2 \rho_i + \beta_3 \cdot \boldsymbol{C}_i + \epsilon_i$$

  where $logit(a_i) = ln(a_i/(1 - a_i))$, p_i is the price offered by the robot for sharing, ρ_i is the correlation coefficient that specifies the joint normal distribution where the secret numbers are drawn from, and $\boldsymbol{C}_i$ is the set of covariates.

- In the friends decision making module. the primary outcome is the willingness-to-accept to announce answers elicited by the BDM method. The analysis will be based on the following regression function:

$$WTA_i = \beta_0 + \beta_1 \cdot \mathbb{1}(\rho > 0) + \beta_2 \cdot \boldsymbol{C}_i + \epsilon_i$$

Secondary outcomes:

- In the information sharing game, the secondary outcomes are (1) the likelihood that participants choose the mutual best responses given by the realized payoff matrices; (2) payoffs; (3) efficiency

B.2 Pre-Registered Hypotheses

The main variable of interest is ρ, which we vary exogenously in both our original game and modified game. We hypothesize that as predicted in the theoretical framework, when the correlation coefficient (ρ) between users' types is high, then users will have a higher willingness to share and also demand a lower compensation for sharing.

Hypothesis 1 *In the information sharing game, given p_i, participants are more likely to share their information with the robot as ρ increases.*

Hypothesis 2 *Within each scenario in the modified game, compared with the $\rho = 0$ question, participants' are more willing to share in the $\rho > 0$ question.*

If Hypothesis 1 is true, we should observe the likelihood of sharing, measured by the dummy variable which indicates whether the participant choose to share, increases with ρ.

If Hypothesis 2 is true, we should observe that participants' willingness-to-accept measured by the switching amount is lower in the $\rho > 0$ questions compared with the $\rho = 0$ questions. As a result, they are more likely to announce their answers to the $\rho = 0$ questions, compared with the $\rho > 0$ questions.

In our original game, we are also interested in validating the effects of lacking calculating ability.

Hypothesis 3 *Participants are better at choosing the best responses in the original game when they are provided with the average payoff matrices, compared with no matrices.*

If Hypothesis 3 is true, we should observe that participants' likelihood of choosing the best response is higher in the rounds with the average payoff matrices. Correspondingly, the likelihood that they choose the mutual best responses, their payoff, and their efficiency ($= \frac{\text{actual payoff} - \text{min payoff}}{\text{max payoff} - \text{min payoff}}$) should also be higher.

In the other measurement games we use, we hypothesize that participants' privacy preferences are related to their risk preference, ambiguity preference and cooperation preference elicited.

Hypothesis 4 (Risk) *More risk averse individuals are less likely to disclose sensitive personal information.*

Hypothesis 5 (Ambiguity) *More ambiguity averse individuals are less likely to disclose sensitive personal information.*

Hypothesis 6 (Cooperation) *Those who are more altruistic are more likely to be nudged into sharing sensitive personal information to benefit the society.*

If the above hypotheses are true, we should observe that: participants who (1) switch to the option B (which has a higher variance) earlier in the lottery choice game; (2) switch to the unknown bag earlier as the composition in the known bag becomes more unbalanced in the Ellsberg urns experiment; and (3) contribute more to the public account in the public goods game, tend to have a more lenient privacy setting and are more willing to share their information with more people in the real world. In particular, we focus on the three types of sharing decisions elicited in our study: (1) sharing decisions in the original game (The information sharing game); (2) sharing decisions in the modified game (The friends decision making Module); (3) sharing decisions on Instagram.

Hypothesis 7 (Correlation) *Sharing decisions in the original game (The information sharing game), in the modified game (The friends decision making Module), and on Instagram are correlated with each other.*

If Hypothesis 7 is true, we should observe that participants who are more likely to choose to share in the original game are also more likely to have low willingness-to-accept in the modified game, and are also more likely to enable sharing on Instagram.

Hypothesis 8 (Gender Differences) *In the modified game, female participants are on average less willing to announce their answers.*

Hypothesis 8 incorporates insights from Huberman et al. (2005), where the authors find that female participants demand a higher price to share their own body weight information. If Hypothesis 8 is true, we should observe that female participants on average indicate a larger number as their WTA in all rounds in the modified game. As a result, they should on average announce fewer answers compared with the male participants.

C Equilibrium Analyses of the information sharing game with risk averse Participants

The first part of Appendix C presents theoretical derivations using continuous utility functions, while the second part includes simulations based on discrete utility functions.

C.1 Derivations

The utility function in Acemoglu et al. (2022) assumes risk neutrality.

$$u_i(\mathbf{a}, \mathbf{p}) = \underbrace{p_i \cdot \mathbb{1}(a_i = 1)}_{\text{benefit of sharing}} + \underbrace{v_i \times Var(\hat{\theta}_i | X_a)}_{\text{preference for privacy}},$$

To compare the decisions between users with different risk preferences, we assume a risk averse user's payoff is determined by the following utility function:

$$u_i(\mathbf{a}, \mathbf{p}) = \underbrace{\sqrt{p_i} \cdot \mathbb{1}(a_i = 1)}_{\text{benefit of sharing}} + \underbrace{v_i \times Var(\hat{\theta}_i | X_a)}_{\text{preference for privacy}},$$

Given player j chooses to share, i.e., $a_j = 1$.

According to Appendix A:

- If player i does not share, the variance of the robot's guess of θ_i is

$$Var(\theta_i | X_j = x_j) = 1 - \rho \times \frac{1}{2} \times \rho = 1 - \frac{\rho^2}{2}.$$

  Player i's payoff is

$$u_i = v_i \times Var(\theta_i | X_j = x_j) = (1 - \frac{\rho^2}{2}) v_i.$$

- If player i shares, the variance of the robot's guess is

$$Var(\hat{\theta}_i | X_i = x_i) = \frac{1 - \frac{\rho^2}{2}}{1 + 1 - \frac{\rho^2}{2}} = \frac{2 - \rho^2}{4 - \rho^2}$$

  Player i's payoff is

$$u_i = v_i \times Var(\hat{\theta}_i | X_i = x_i) + \sqrt{p_i} = \frac{2 - \rho^2}{4 - \rho^2} v_i + \sqrt{p_i}$$

Therefore, player i's best response to $a_j = 1$ is $a_i = 1$ if and only if $u_i(a_i = 1 | a_j = 1) \geq u_i(a_i = 0 | a_j = 1)$, which is equivalent to

$$p_i \geq (\frac{(2 - \rho^2)^2}{2(4 - \rho^2)})^2 v_i^2$$

Given that the minimum price to induce a risk-neutral user to share is $p_i^{neutral} \geq \frac{(2-\rho^2)^2}{2(4-\rho^2)} v_i$. Let d_i denote the difference between the two price cutoffs:

$$d_i = p_i - p_i^{neutral} = (\frac{(2-\rho^2)^2}{2(4-\rho^2)})^2 v_i^2 - \frac{(2-\rho^2)^2}{2(4-\rho^2)} v_i$$

$d_i > 0$ when $v_i > \frac{2(4-\rho^2)}{(2-\rho^2)^2}$. That is, when $v_i > \frac{2(4-\rho^2)}{(2-\rho^2)^2}$, a risk averse user always demand a higher price to share compared with a risk-neutral user.

Given player j chooses not to share, i.e., $a_j = 0$.

According to Appendix A:

- If player i does not share, the variance of the robot's guess of θ_i is the prior variance, 1

  Player i's payoff is v_i

- If player i shares, the variance of the robot's guess is $\frac{1}{2}$

  Player i's payoff is $\frac{1}{2} + \sqrt{p_i}$

Therefore, player i's best response to $a_j = 0$ is $a_i = 1$ if and only if $u_i(a_i = 1|a_j = 0) \geq u_i(a_i = 0|a_j = 0)$, which is equivalent to

$$p_i \geq \frac{1}{4} v_i^2$$

Given that the minimum price to induce a risk-neutral user to share is $p_i^{neutral} \geq \frac{1}{2} v_i$. Let d_i denote the difference between the two price cutoffs:

$$d_i = p_i - p_i^{neutral} = \frac{1}{4} v_i^2 - \frac{1}{2} v_i$$

$d_i > 0$ when $v_i > 2$. That is, when $v_i > 2$, a risk averse user always demand a higher price to share compared with a risk-neutral user.

C.2 Simulations

To gain a deeper understanding of the role of risk preferences, we run several simulations to compare sharing decisions under the discretized utility functions:

The utility function in our experiment,

$$u_i = \underbrace{p \times \mathbb{1}\{\text{Share}\}_i}_{\text{benefit of sharing}} + \underbrace{v \times \frac{1}{3} \sum_{k=1}^{3} (\hat{\theta}_{ik} - \theta_{ik})^2}_{\text{preference for privacy}}, \tag{6}$$

is discretized from the theoretical framework which assumes risk neutrality.

The corresponding utility function for risk averse users is:

$$u_i = \underbrace{\sqrt{p} \times \mathbb{1}\{\text{Share}\}_i}_{\text{benefit of sharing}} + \underbrace{v \times \frac{1}{3}\sum_{k=1}^{3}(\hat{\theta}_{ik} - \theta_{ik})^2}_{\text{preference for privacy}}, \tag{7}$$

We conduct several simulations to compare the minimum prices required to induce users to share their private information under different correlation coefficients, based on the two payoff functions described above. As in the experiment, we normalize the value of sharing, v, to 300 points (equivalent to $7.50) in all simulations.

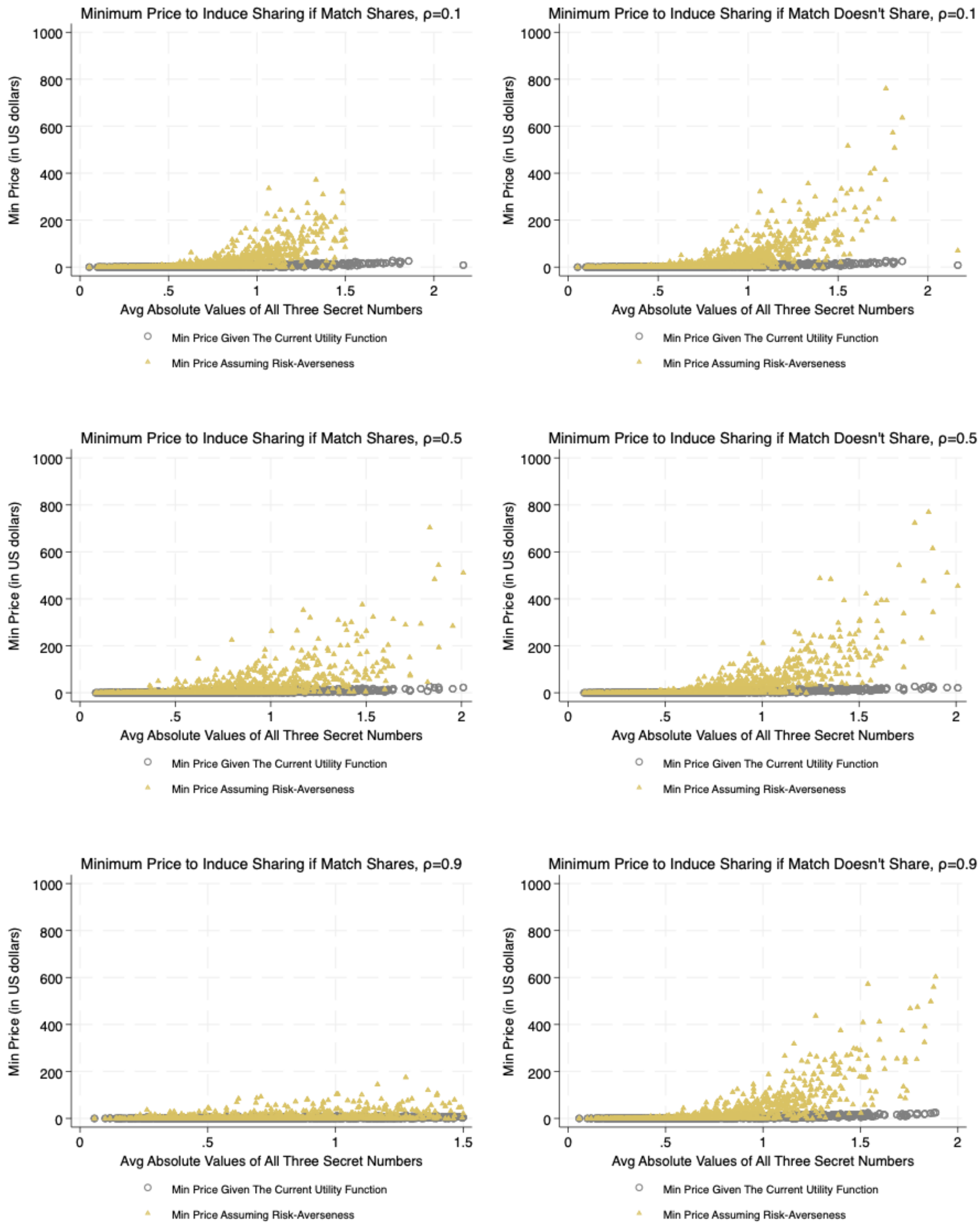
The figures below present the simulation results for correlation coefficients of 0.1, 0.5, and 0.9, respectively. The left panels show cases where users believe their match chooses to share, while the right panels represent cases where users believe their match chooses not to share.

Across both payoff functions, the minimum prices increase with the average absolute value of the secret numbers. That is, whether users are risk-neutral or risk averse, they demand higher compensation when their true types (the secret numbers) are farther from zero. This occurs because it gets harder for the platform to infer a user's private types through random guesses when the absolute value of the secret numbers are larger, as such values are less likely to occur by chance under the joint normal distribution. Therefore, the marginal increase in the platform's inference accuracy is greater, making it more costly to preserve one's privacy.

risk averse users, being more sensitive to potential information leakage, respond more strongly to increases in the absolute value of their secret numbers. As a result, they demand higher prices for sharing compared to risk-neutral users. Furthermore, the price gap widens as the average absolute value of the secret numbers increases.

When a user believes their match chooses not to share, their sharing decisions are independent of the correlation coefficient. Therefore, the right panels across all three figures exhibit similar patterns. In contrast, when users believe their match chooses to share, the price gap between risk averse and risk-neutral users is influenced by the sharing externalities (i.e., correlation coefficients). Comparing the left panels, as the correlation increases, the match's information shared becomes more informative of the user's private information. This amplifies the data leakage due to sharing externalities even if users choose not to share, prompting them to accept lower compensation to avoid missing out on any potential benefits. Consequently, the price gap between risk averse and risk-neutral users narrows as the correlation coefficient increases.

Figure C.1: Comparison of Minimum Prices Under Different ρ



D Experimental Instructions (With Matrix First)

Instructions - Information Sharing Experiment
(Please turn off your cell phone. Thank you.)

This is an experiment in the economics of decision making. In this experiment, we simulate an information sharing scenario on a social network platform. The procedure, interactions and payment rules are described below. The amount of money you earn will depend upon the decisions you make and on the decisions other people make. Do not communicate with each other during the experiment. If you have questions at any point during the experiment, raise your hand and the experimenters will help you.

Overview:

- There are 12 participants in this experiment.
- The experiment consists of four modules:
  1. In the first module which consists of 10 rounds, you will be randomly matched with another participant each round, and play an information sharing game.
  2. In the second module, you will play the same information sharing game as the first module. The only difference is that we vary the amount of information in this module.
  3. In the third module, you will first make decisions independently in two tasks. Then you will be matched with randomly selected participants in this session to play a game.
  4. In the last module, which consists of 3 rounds, you will independently make a series of decisions regarding whether to share information about your daily life with the friend you nominated with other participants in the session.
- **Your total payoff** is equal to the sum of your payoffs for each module. Your earnings are given in points.

  At the end of the experiment you will be paid based on the exchange rate,

  40 points = $1.

  In addition, you will be paid $5 for participation. Everyone will be paid in private, and you are under no obligation to tell others how much you earn.

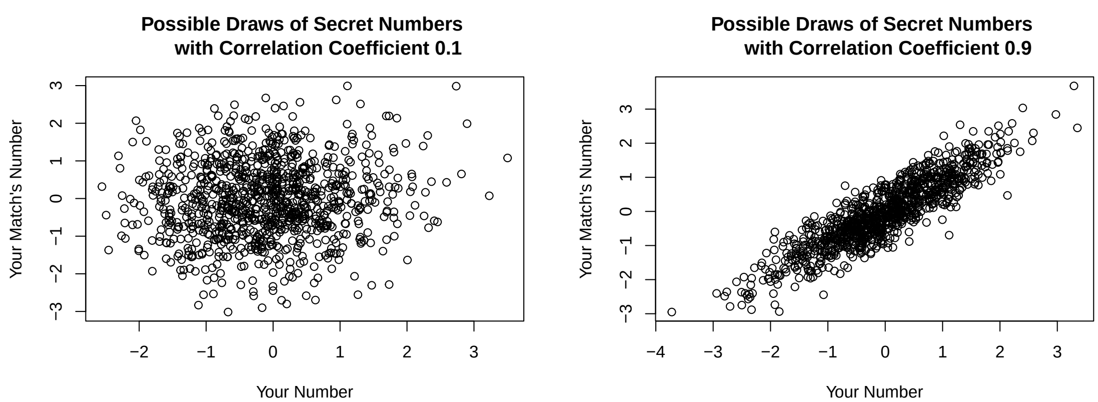
What questions do you have so far?

Module 1. Information Sharing Game - Stage 1

This module consists of 10 rounds of an information sharing game. In each round, you will be randomly matched with another participant in the room. You and your match will interact with a robot who represents a social networking platform. This game simulates a situation where users try to balance the benefits of using a social networking platform with the preservation of their private information.

In this game, you and your match will make decisions independently, but your earnings will be affected both by your and your match's sharing decisions and the robot's guesses of your private information. At the end of this module, one round will be randomly chosen for your final payment.

Your secret numbers. At the beginning of each round, the computer will randomly draw three pairs of real numbers from a joint normal distribution. Each pair will include one number designated for you and another number designated for your match. As a result, both you and your match will have a set of three numbers. Your numbers are private information which are only known to you. However, the distribution where your and your match's secret numbers are drawn from is publicly known. Your and your match's secret numbers are correlated with coefficient ρ. Below are two plots showing random pairs of numbers drawn under different correlation coefficients. The values on the x-axis represent your secret numbers, while the values on the y-axis represent your match's secret numbers. In general, a larger correlation coefficient implies a stronger association between the two sets of numbers. More details about joint normal distributions are explained in the Appendix.



(a) Low correlation: $\rho = 0.1$

(b) High correlation: $\rho = 0.9$

Figure D.1: Examples of random draws under different correlation coefficients: $\rho = 0.1, 0.9$.

At the beginning of each round, the computer randomly draws a correlation coefficient, $\rho \in \{0.1, 0.2, \ldots, 0.9\}$, independently for each pair of participants, and announces this coefficient to each pair of participants. The three pairs of secret numbers assigned to you and your match share the same correlation coefficient. In other words, you and your match will know ρ at the beginning of each round.

What questions do you have so far?

The robot. The social networking platform is represented by a robot. The robot's goal is to guess everyone's secret numbers as accurately as possible, based on the information shared.

- At the beginning of each round, the robot offers you and your match a payment for all three of your secret numbers. The payment is randomly drawn from a uniform distribution on the interval [30, 180]. The payment is the same for you and your match.
- You and your match independently decide whether to share your secret numbers with the robot or not.
  - If you decide to share your secret numbers, each number in your set will be delivered to the robot with some noise, which is a number drawn from a normal distribution with mean 0 and variance 1. Each secret number will have an independent noise added to it when the information is delivered. That is to say, the robot is unlikely to know exactly the secret numbers you have. Instead, it will know a set of numbers that are relatively close to the true numbers. Here is an example which can help you understand the noise: Your secret numbers represent your music preferences. The robot can observe the number of times you play a song. However, that count might not perfectly reflect your real preferences. The noise simulates the difference between the true value (your actual preference) and what is observed (the count of plays).
  - If you decide not to share, your numbers will not be delivered to the robot.
- If you decide to share, you will get the payment from the robot. Otherwise, you will not receive any payment from the robot.
- Depending on your and your match's decisions, the robot will try to guess each number in your set and your match's set as accurately as possible.

What questions do you have so far?

Your payoff. Your payoff is comprised of two parts: (1) the sum of squared differences between each of the robot's guesses and your corresponding secret numbers, regardless of your decision, and (2) the payment you get from the robot if you decide to share your secret numbers. More specifically, the following equation denotes your payoff:

$$\text{Your Payoff} = \begin{cases} 100 \times \text{Sum of (Robot's guess - Your secret number)}^2, & \text{if you don't share;} \\ 100 \times \text{Sum of (Robot's guess - Your secret number)}^2 + \text{Payment}, & \text{if you share.} \end{cases}$$

For example, if

- **Payment**: The robot is offering 50 points to you for sharing your whole set of numbers
- Your set of secret numbers are $\{-0.1, 0, 0.1\}$
- If you decide not to share your numbers:

- The robot's corresponding guesses are $\{0.1, 0.2, -0.2\}$
- Your payoff will be $100 \times [(0.1 - (-0.1))^2 + (0.2 - 0)^2 + (-0.2 - 0.1)^2] = 17$

- If you decide to share your numbers:
  - The robot's corresponding guesses are $\{0, 0.1, -0.1\}$
  - Your payoff will be $100 \times [(0 - (-0.1))^2 + (0.1 - 0)^2 + (-0.1 - 0.1)^2] + 50 = 56$

As you can see from the equation, your payoff decreases as the robot's guesses get closer to your secret numbers. It also increases with the payment offered by the robot if you decide to share your secret numbers.

Your and your match's sharing decisions affect both parts of your payoff. On the one hand, you will only earn the payment from the robot if you decide to share your secret numbers. On the other hand, your sharing decision as well as that of your match's will affect the accuracy of the robot's guesses in the following way:

- If you decide to share your numbers, each number will be delivered to the robot with some noise. Therefore, the robot's guesses will be quite close to your secret numbers.
  - If you decide to share, but your match decides not to share, the robot's guesses will only depend on your numbers (with noise). Specifically, the robot's guess of each number in your set will be
    $$\frac{1}{2} * \text{Your number with noise}$$
  - If both you and your match decide to share your numbers, the robot will make guesses based on both sets of numbers with noises. As a result, the robot's guesses will become more accurate. The robot's guess of each number in your set will be a weighted sum of your number with noise and your match's number with noise:
    $$\frac{2-\rho^2}{4-\rho^2} * \text{Your number with noise} + \frac{\rho}{4-\rho^2} * \text{Your match's number with noise}$$
    When ρ is close to 0, the robot's guess of your secret number weights your number with noise more than your match's in the calculation. When ρ gets close to 1, the weights assigned to your number with noise and your match's number with noise will become more equal.

- If you decide not to share your numbers, but your match decides to share their numbers, the robot can still guess your numbers based on your match's numbers (with noise). In this case, the robot can make some inference about your secret numbers, but not as accurate as when you decide to share yourself. More specifically, the robot's guesses of each number in your set will only depend on your match's numbers with noise:
  $$\frac{\rho}{2} * \text{You match's number with noise}$$
  The larger ρ is, i.e., the more correlated your and your match's numbers are, the closer the robot will be able to guess your number.

- If neither of you decides to share, the robot's guesses for all numbers in both of your sets will be 0, which is the mean value of the distribution from which your and your match's numbers are drawn. It is likely to be far off from your true numbers.

What questions do you have so far?

Your decision. In each round, you will see your secret numbers, the correlation coefficient and the payment offered by the robot. In addition, to simplify the calculations for you, you will see a table of average payoffs that you can potentially get based on the information you receive, without knowing your match's exact secret numbers. More specifically, given each of your secret number X and the correlation coefficient ρ, the average of your match's corresponding secret number will be $\rho * X$. Therefore, taken each of your match's number as $\rho * X$, we can calculate your and your match's respective average payoffs.

		Your match's decision	
		Share	Not Share
Your decision	Share	Your average payoff: The match's average payoff:	Your average payoff: The match's average payoff:
	Not Share	Your average payoff: The match's average payoff:	Your average payoff: The match's average payoff:

Table D.1: Average Payoffs of You and Your Match

The top left grid of Table 1 shows the average payoffs you and your match can get if both of you decide to share. The top right grid shows the payoffs if you decide to share but your match decides not to share. The bottom left grid shows the payoffs if you decide not to share but your match decides to share. The bottom right grid shows the payoffs if neither you nor your match decides to share.

Based on this table, we will ask you four questions:

1. If your match decides to share, what would your decision be?
2. If your match decides not to share, what would your decision be?
3. Do you decide to share or not to share your secret numbers?
4. If your match knew your decision, would they decide to share or not to share? Please indicate your prediction.

Note that the table here only helps you understand what potential payoffs you can get. Your actual payoff will still be determined by your and your match's decisions, your respective secret numbers and the robot's actual guesses.

Then we will calculate your and your match's respective payoffs based on the decisions you and your match make. In addition, if you correctly predict your match's decision in the fourth question, you will receive a bonus of 20 points.

What questions do you have so far?

Feedback. At the end of each round, you will know your and your match's sets of secret numbers, your respective decisions, the robot's guesses, and your respective earnings for that round.

Matching Protocol. The process repeats for 10 rounds, but you will be matched with a different participant each round. That is to say, you will never be matched with the same participant more than once. At the end of ten rounds, one round will be randomly chosen to determine your payoff.

What questions do you have so far?

Review Questions. To make sure that everyone understands the game, please answer the following review questions. Feel free to refer to the experimental instructions before you answer any questions. Each correct answer is worth 20 points, and will be added to your total earnings. When everyone is finished with them, we will go through the answers together.

1. True or false: You will be matched with the same participant each round.

2. True or false: You will earn a payment from the robot if you choose to share your information.

3. Your payoff _____ (increases/decreases) with the accuracy of the robot's guess.

4. True or false: Your and your match's sharing decisions do not affect each other's payoff.

5. As the correlation coefficient becomes closer to 1, it becomes _____ (easier/harder) for the robot to guess your number based on your match's shared information if you choose not to share your number.

6. Please calculate your payoff in the round described in the following example:

   - Your set of secret numbers is: $\{1, 2, 3\}$
   - The correlation coefficient between your number and your match's number is: 0.8
   - The price for sharing offered by the robot is: 50
   - You sharing decision is: share
   - Your match's sharing decision is: not share
   - The robot's guesses of your secret numbers are: $\{3, 2, 4\}$

   Your payoff in this round is: _____ (Please enter your answer).

7. If you know your secret number $\{x_1, x_2, x_3\}$ and the correlation coefficient ρ, on average, your match's corresponding secret numbers are going to be $\rho x_1, \rho x_2, \rho x_3$. Based on this information, please do the following calculations. The bonus is given only when you answer all three questions correctly. Suppose one of your secret numbers is $x_1 = 1$

   - If the correlation coefficient $\rho = 0.1$, your match's corresponding secret number on average is: _____

- If the correlation coefficient $\rho = 0.5$, your match's corresponding secret number on average is: _____

8. Please look at the average payoff matrix below. If you decide to share while your match decides not to share:

    - The average payoff you can get is: _____
    - The average payoff your match can get is: _____

		Your match's decision	
		Share	Not Share
Your decision	Share	Your average payoff: 100 Your match's average payoff: 100	Your average payoff: 150 Your match's average payoff: 50
	Not Share	Your average payoff: 75 Your match's average payoff: 125	Your average payoff: 200 Your match's average payoff: 200

We are now ready to start the 10 rounds of the information sharing game. Please compare carefully between the benefits of keeping your secret number and the payments you can get from sharing. Try to earn as much as you can.

What questions do you have so far?

Answers to Review Questions

1. True or false: You will be matched with the same participant each round.

   Answer: False. You will be randomly re-matched with a different participants each round.

2. True or false: You will earn a payment from the robot if you choose to share your information.

   Answer: True.

3. Your payoff _____ (increases/decreases) with the accuracy of the robot's guess.

   Answer: decreases

4. True or false: Your and your match's sharing decisions do not affect each other's payoff.

   Answer: False. Your earnings will be affected by both of your decisions.

5. As the correlation coefficient becomes closer to 1, it becomes _____ (easier/harder) for the robot to guess your number based on your match's shared information if you choose not to share your number.

   Answer: easier.

6. Payoff Calculation Question.

   Answer: $100 * [(3-1)^2 + (2-2)^2 + (4-3)^2] + 50 = 550$.

7. Match's Number Calculation Question:

   Answer: $0.1 \times 1 = 0.1$; $0.5 \times 1 = 0.5$

8. Please look at the payoff matrix below. If you decide to share while your match decides not to share:

		Your match's decision: Share	Your match's decision: Not Share
Your decision	Share	Your average payoff: 100 Your match's average payoff: 100	Your average payoff: 150 Your match's average payoff: 50
	Not Share	Your average payoff: 75 Your match's average payoff: 125	Your average payoff: 200 Your match's average payoff: 200

   - The average payoff you can get is: _____
   - The average payoff your match can get is: _____

   Answer:

   - The average payoff you can get is: 150
   - The average payoff your match can get is: 50

Module 2. Information Sharing Game - Stage 2

This module consists of 10 rounds of an information sharing game same as the first module. In each round, you will be matched with the same participant as in the first module. You and your match will again interact with the robot who represents a social networking platform. The secret numbers, correlation coefficients, payments, the robot's guessing rules, and the payoff calculation functions will be kept the same as in the first module. If you do not remember the details, please refer to the Module 1 instructions. The only difference is that you will not see the average payoff matrices in this module. Instead, you need to make evaluations yourself. At the end of this module, one round will be randomly chosen for your final payment.

Your decision. In each round, you will see your secret numbers, the correlation coefficient and the payment offered by the robot. You need to evaluate the information carefully, and answer the following four questions:

1. If your match decides to share, what would your decision be?
2. If your match decides not to share, what would your decision be?
3. Do you decide to share or not to share your secret numbers?
4. If your match knew your decision, would they decide to share or not to share? Please indicate your prediction.

Then we will calculate your and your match's respective payoffs based on the decisions you and your match make. In addition, if you correctly predict your match's decision in the fourth question, you will receive a bonus of 20 points.

What questions do you have so far?

Feedback. At the end of each round, you will know your and your match's sets of secret numbers, your respective decisions, the robot's guesses, and your respective earnings for that round. At the end of ten rounds, one round will be randomly chosen to determine your payoff.

We are now ready to start the 10 rounds of the information sharing game. Please compare carefully between the benefits of keeping your secret numbers and the payments you can get from sharing. Try to earn as much as you can.

What questions do you have so far?

Module 3. Measurement Games

This module consists of 3 different tasks. In the first and second tasks, each participant will make decisions individually. In the last task, you will be randomly matched with other participants in this session and play a game. Your payoff in this module equals the sum of your payoffs in each task. We encourage you to earn as much cash as you can in each of the tasks.

The L Game

Making ten decisions. On your screen, you will see a table with 10 decisions in 10 separate rows. For each row, you will be making a selection (option A or option B) by clicking on the buttons in the middle.

The money prizes are determined by the computer equivalent of throwing a ten-sided dice. Each outcome, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, is equally likely. If you choose Option A in the row shown below, you will have a 1 in 10 chance of earning 200 points and a 9 in 10 chance of earning 160 points. Similarly, Option B offers a 1 in 10 chance of earning 385 points and a 9 in 10 chance of earning 10 points.

As you move down the list, Option B gets better because it has a higher expected payoff. If you choose Option B in a row, the computer automatically selects Option B for you in all the rows below as Option B gets better in those rows. Similarly, if you choose Option A in a row, the computer automatically selects Option A for all rows above. You may make these choices in any order and change them as much as you wish until you press the 'Next' button at the bottom.

Decision	Option A	Your Choice	Option B
1	200 points if the dice is 1 160 points if the dice is 2-10	A or B	385 points if the dice is 1 10 points if the dice is 2-10

The relevant decision. After you submit, one of the rows is selected at random. The option (A or B) that you chose in that row will be used to determine your earnings. Note: Please think about each decision carefully, since each row is equally likely to end up being the one that is used to determine payoffs.

For example, suppose that you make all ten decisions and the throw of the dice is 9. In this case, your decision for row 9 would be used.

Decision	Option A	Your Choice	Option B
9	200 points if the dice is 1-9 160 points if the dice is 10	A or B	385 points if the dice is 1-9 10 points if the dice is 10

Determining the payoff. After one of the decisions has been randomly selected, the computer will generate another random number that corresponds to the throw of a ten-sided dice. The number is equally likely to be 1, 2, 3, ... 10. This random number determines your earnings for the Option (A or B) that you previously selected for the decision being used.

For example, in Decision 9 below, a throw of 1, 2, 3, 4, 5, 6, 7, 8, or 9 will result in the higher payoff for the option you chose, and a throw of 10 will result in the lower payoff.

Decision	Option A	Your Choice	Option B
9	200 points if the dice is 1-9 160 points if the dice is 10	A or B	385 points if the dice is 1-9 10 points if the dice is 10
10	200 points if the dice is 1-10	A or B	385 points if the dice is 1-10

For decision 10, the random dice throw will not be needed, since the choice is between amounts of money that are fixed: 200 points for Option A and 385 points for Option B.

What questions do you have so far?

The U Game

Two bags. There are two bags both contained 100 balls with possibly 2 colors: purple and yellow. There is a known bag K, where the composition of balls is known; and an unknown bag U, where the composition of balls is unknown. The unknown bag U is prepared beforehand by a robot. Therefore, the experimenters do not know its composition either during the experiment.

Bet on colors. In this experiment, you first choose a color between the two colors as the winning color. This means that if a ball randomly drawn from the bags is in the color you bet on, then you win the bet.

Bet on bags. After choosing the winning color, you will choose from which bag, U or K, a ball will be randomly drawn. You will indicate your preference in choice lists, with different compositions of balls in each row.

Payoff in the experiment. If the color of the ball drawn from your chosen bag is a winning color, then you will win 500 points; otherwise you win nothing.

Two choice lists. You will see two choice lists in this task. Below is the first choice list. Each row represents a choice scenario with two options Bag K and Bag U. The left two columns describe the compositions of the colored balls in Bag K. The first left column says the number of balls that are in the color you bet on and that is the key information you should pay attention to. The two columns to the right describe the compositions of colored balls in Bag U. The compositions are always unknown for Bag U. The total number of balls in Bag U is however always 100. You need to indicate your preferences between Bag K and Bag U for each row by clicking one of the two options in the middle.

Number of Balls in Bag K		Choice		Number of Balls in Bag U	
500 points	0			500 points	0
Chosen color	The other color			Chosen color	The other color
0	100	Bag K	Bag U	Unknown	100 - Unknown
10	90	Bag K	Bag U		
20	80	Bag K	Bag U		
30	70	Bag K	Bag U		
40	60	Bag K	Bag U		
50	50	Bag K	Bag U		
60	40	Bag K	Bag U		
70	30	Bag K	Bag U		
80	20	Bag K	Bag U		
90	10	Bag K	Bag U		
100	0	Bag K	Bag U		

Table D.2: Step 1 Choice List

As you move down the list, Bag K gets better because it contains more balls of color that you bet on. If you choose Bag K in a row, the computer automatically selects Bag K for you in all the rows below as Bag K gets better in those rows. Similarly, if you choose Bag U in a row, the computer automatically selects Bag U for all rows above.

You can change your choices as many times as you like. Once you are satisfied with your choices, submit them by clicking the "Next" button. Then a similar but more refined choice list will appear which contains all the choice scenarios of this question. Again, you have to indicate your choices between Bag K and Bag U for every row.

After you submit your first choice list, you will see the second choice list to further indicate your preference. Below is an example. If in the first choice list,

$$\text{Your Choice} = \begin{cases} \text{Bag U} & \text{if there are 50 or fewer balls in your chosen color in Bag K;} \\ \text{Bag K} & \text{if there are 60 or more balls in your chosen color in Bag K} \end{cases}$$

you need to further indicate your preferences in the below choice list, between row 51 and 59.

<table>
<tr><th colspan="2">Number of Balls in Bag K</th><th colspan="2" rowspan="3">Choice</th><th colspan="2">Number of Balls in Bag U</th></tr>
<tr><th>500 points</th><th>0</th><th>500 points</th><th>0</th></tr>
<tr><th>Chosen color</th><th>The other color</th><th>Chosen color</th><th>The other color</th></tr>
<tr><td>0</td><td>100</td><td colspan="2">Bag U</td><td rowspan="15">Unknown</td><td rowspan="15">100 - Unknown</td></tr>
<tr><td>...</td><td>...</td><td colspan="2">Bag U</td></tr>
<tr><td>50</td><td>50</td><td colspan="2">Bag U</td></tr>
<tr><td>51</td><td>49</td><td>Bag K</td><td>Bag U</td></tr>
<tr><td>52</td><td>48</td><td>Bag K</td><td>Bag U</td></tr>
<tr><td>53</td><td>47</td><td>Bag K</td><td>Bag U</td></tr>
<tr><td>54</td><td>46</td><td>Bag K</td><td>Bag U</td></tr>
<tr><td>55</td><td>45</td><td>Bag K</td><td>Bag U</td></tr>
<tr><td>56</td><td>44</td><td>Bag K</td><td>Bag U</td></tr>
<tr><td>57</td><td>43</td><td>Bag K</td><td>Bag U</td></tr>
<tr><td>58</td><td>42</td><td>Bag K</td><td>Bag U</td></tr>
<tr><td>59</td><td>41</td><td>Bag K</td><td>Bag U</td></tr>
<tr><td>60</td><td>40</td><td colspan="2">Bag K</td></tr>
<tr><td>...</td><td>...</td><td colspan="2">Bag K</td></tr>
<tr><td>100</td><td>0</td><td colspan="2">Bag K</td></tr>
</table>

Table D.3: Step 2 Choice List

What questions do you have so far?

The P Game

There will be a total of 5 rounds in this game. In each round, you will be randomly matched with another participant in this session.

Allocations. You will be a member of a group of 2 people. Each member has to decide on the division of 200 points. You can put these 200 points into a private account or you can invest them fully or partially into a project. Each point you do not invest into the project will be transferred to your private account.

Your income from the private account. You will earn one point for each point you put into your private account. For example, if you put 200 points into your private account (i.e., do not invest into the project) your income will be exactly 200 points. If you put 6 points into your private account, your income from this account will be 6 points. No one except you earns something from your private account.

Your income from the project. Each group member will profit equally from the amount you invest into the project. Your contribution to this project will be shared with the other group member. This means that you will also get a payoff from the other group member's investments into the project. The income for each group member will be determined as follows:

$$\text{Income from the project} = \text{sum of all contributions} \times 0.6$$

If, for example, the sum of all contributions to the project is 400 points, then you and the other member of your group each earn 400 × 0.6 = 240 points out of it. If two members of the group contribute a total of 10 points to the project, you and the other member of your group each earn 10 × 0.6 = 6 points.

Your total income. Your total income is the sum of your income from your private account and that from the project
Total Income =

$$\underbrace{200 - \text{Contribution to the Project}}_{\text{Income from Private Account}} + \underbrace{0.6 \times \text{Sum of Contributions to the Project}}_{\text{Income from the Project}}$$

At the end of the study, one round will be randomly chosen for your payment in this game.

Feedback. At the end of each round, you will know the total group contribution, the per person return from the project, and your payoff for that round.

What questions do you have so far?

Module 4. Friends Decision Making

The module consists of 3 rounds of decision making, each comprising two stages. You will make decisions individually throughout this module. The first stage is the information collection stage. You will answer a series of questions asking for information about the daily life and opinions on some topics of you and the friend you nominated in the pre-survey. The second stage is the information sharing stage. You need to decide for how many points you are willing to share each answer you provided in the first stage with other participants in this room.

Stage 1

Decision Scenarios. In each round, you will face two questions. The first question asks for information about your own daily life or your own opinions. The second question asks about your experiences with your friend or the average opinion between you two. Therefore, in the second question, your responses may reveal information about both your and your friend's daily lives and opinions.

The information asked in each of the 3 decision scenarios is listed below. You will see the detailed questions later.

1. Body weight of you and your friend;
2. Intimate relationships of you and your friend;
3. Your and your friend's opinions on intimate relationships.

In each round, both questions are entirely optional and you have the choice of not answering if you feel uncomfortable. Your responses are accessible only to the experimenters and will not be disclosed to anyone else.

What questions do you have so far?

Verification. To gauge the authenticity of the information you share, we will verify the answers you choose to provide. We have sent an online survey to the friend you nominated as the session begins. In the survey, your friend faces the same three decision scenarios as you. That is, they are presented with two questions in each scenario, one asking about their own daily life or opinions, the other one asking about their experiences with you or the average opinion between you two. All questions are optional to them. They can decide whether to answer each question or not. Since your answers to the second questions reveal information about both your and your friend's daily lives and opinions, to verify your answers, we will compare your answers about your friend with your friend's if both of you provide your answers to the second questions. Note that your answers only help us understand to what extent you are telling the truth. Your and your friend's answers will remain confidential and will not be disclosed to each other.

What questions do you have so far?

Your Payoff in Stage 1. Both you and your friend will receive \$2 in the form of a virtual gift card if your friend completes the survey within the two hours of your enrolled session. In addition, for the second question in each round, if your answer about your friend is no more than 10% different from your friend's answer, you will receive a verification bonus of \$1. Therefore, the maximum total verification bonus you can earn is \$3. The amount will be added to the gift card. As a result, the maximum amount of bonus you can earn is \$5. Your friend faces similiar incentives.

If your friend does not respond to the survey within the two-hour time slot, it will not affect your cash payment in the study. The only difference is that neither of you will receive the \$2 bonus or the additional verification bonus. The gift card will be sent to you online after we collect all survey responses.

What questions do you have so far?

After you finish answering questions in all three rounds, you will be asked to complete a survey that collects demographic information (i.e., age, gender, major, etc.) about you. We will proceed to Stage 2 after everyone completes the survey.

We are now ready to start the 3 rounds of the first stage. What questions do you have so far?

Below are the questions for the 3 rounds:

1. Body weight
   - **Q1**: What is your weight in pounds?
   - **Q2**: What is the average weight in pounds for you and your friend? If your answer is no more than 10% different from your friend's answer, you will get a verification bonus of $1.
2. Intimate relationships
   - **Q1**: How many sexual partners have you ever had?
   - **Q2**: How many sexual partners have you and your friend ever had in total? If your answer is no more than 10% different from your friend's answer, you will get a verification bonus of $1.
3. Opinions on Intimate Relationships
   - **Q1**: To what extent do you think it is acceptable to have more than 4 sexual partners within one semester?
     - 1 Very unacceptable
     - 2 Somewhat unacceptable
     - 3 Somewhat acceptable
     - 4 Very acceptable
   - **Q2**: To what extent do you and your friend on average think it is acceptable to have more than 4 sexual partners within one semester? If your answer is no more than 10% different from your friend's answer, you will get a verification bonus of $1.

     - 1 Very unacceptable
     - 2 Somewhat unacceptable
     - 3 Somewhat acceptable
     - 4 Very acceptable

Survey Questions

Before going to the next stage, please answer the following survey questions. Your answers will be used for this study only. Individual data will not be exposed.

1. What is your age?
2. What gender do you most closely identify with?
    - Female
    - Male
    - Other
    - Prefer not to say
3. What is your major at the University of Michigan?
4. Are you an undergraduate or graduate student?
    - Undergraduate Student
    - Graduate Student
5. Which year are you in your program?
6. What is the primary ethnicity or race you identify with?
    - White/European American
    - Black/African American
    - American Indian or Alaska Native
    - Asian/Asian American
    - Hispanic/Latino
    - Other
7. Have you ever participated in any economics or psychology experimental studies before?
    - Yes
    - No
8. How long have you and your friend known each other?
    - Less than 1 month
    - 1 - 2 months
    - 2 - 3 months
    - 3 - 4 months
    - 4 - 12 months
    - More than 1 year

9. How did you and your friend meet?
    - We used to take the same class
    - We used to live in the same building
    - We attended the same social event
    - We were introduced by mutual friends
    - Other (Please specify): ______
10. How strong is your relationship with this friend (0–10 slider)?
    - 0 barely know them
    - 10 we are very close
11. How would you feel asking this friend to loan you $50 or more (0–10 slider)?
    - 0 would never ask
    - 10 very comfortable

Stage 2

Willingness to accept to announce your answer. For each question we ask, you need to indicate the minimum number of points within $[0, 100]$ you are willing to accept in order to share the information with other participants in this room. More specifically, "willing to share" means you are willing to announce your answer in front of other participants in this room. For example, if you are willing to receive any number of points between $[50, 100]$ to announce your answer to a question in front of other participants in this room, you should put 50 as your minimum number. If no number of points below 100 is acceptable for you to announce your answer, or if you chose not to answer this question, you should put 100 as your minimum number. This applies to both questions in each scenario.

After you tell us how many points you are willing to accept to share the information, the computer will randomly draw a number between 1 and 100. This number is drawn independently for each participant. It represents the payment that we offer you for sharing your information. Whether you need to share or not depends on the relationship between the number you choose and the random number the computer draws:

- If the randomly drawn number$>$ the minimum number you choose: You need to announce your answer in front of other participants and the payment is whatever number we draw.
- If the randomly drawn number$\leq$ the minimum number you choose: You will not receive any payment and do not need to announce your answer.

The intuition behind the above payment rules is as follows: If the payment we offer you for sharing (the randomly drawn number) is larger than the lowest payment you are willing to accept (the minimum number you choose), the payment should be satisfying enough so that you are willing to exchange your information for the payment. However, if the payment we offer you for sharing (the randomly drawn number) is lower than or equal to the lowest payment you are willing to accept (the minimum number you choose), it means that we are offering too low a payment to you. Therefore, you are not willing to exchange your information for our payment. You do not need to share your answer, and we do not need to pay you anything.

To make sure that the benefit of sharing and preserving private information is optimally balanced, your best strategy is to truthfully report the minimum number of points you are willing to accept in order to share your information with other participants.

Answer announcement. At the end of this module, we will call each participant to check what information they need to announce. For example, if you are willing to accept a minimum of 23 to share your own body weight information and the random number is 35, which is greater than 23, you need to share your own body weight with other participants in this room. We will then call you to stand up and announce your own body weight. You will receive a payment of 35. If the random number is 6, which is less than 23, you do not need to share your body weight information or receive any payment. In contrast, if you are only willing to accept a minimum number of 100 in order to share your own body weight information, then the randomly drawn number will certainly be smaller than or equal to the minimum number you choose. Therefore, you never need to share your own body weight with other participants in this room regardless of which random number is drawn. Consequently, you will not earn any payment for the announcement. We will calculate and share the results for this module with you.

What questions do you have so far?

Your Payoff in Stage 2. For each of your answers to the questions, there is a payment for announcement which depends on the minimum number you choose for sharing your answer and the random number we draw. Your payoff in this stage will be the sum of points you earn from announcing your answers.

What questions do you have so far?

Review Questions. To make sure that everyone understands the task, answer the following review questions. Feel free to refer to the experimental instructions before answering any questions. Each correct answer is worth 20 points, and will be added to your total earnings. When everyone is done, we will go through the answers together.

1. True or false: You do not need to announce anything if you chose not to answer the questions in Stage 1.
2. True or false: Your decision may reveal some information about your friend.
3. True or false: If the drawn number is less than your chosen number, you do not need to share the information and will receive the drawn number of points.
4. If for one question, you choose 70 as the minimum number of points you are willing to accept in order to announce your answer, and the random number drawn by the computer is 80:
   - The payment we offer you for sharing in this round is: _____ points
   - Will you need to announce your answer in front of other participants at the end of the study? Yes or no

We are now ready to start the 3 rounds of the second stage. What questions do you have so far?

Answers to Review Questions

1. True or false: You do not need to announce anything if you chose not to answer the question in Stage 1.

   Answer: True

2. True or false: Your decision may reveal some information about your friend.

   Answer: True

3. True or false: If the drawn number is smaller than your chosen number, you do not need to share the information and will receive the drawn number of points.

   Answer: False. You will not receive any payment in this case.

4. The payment we offer for sharing is 80 points. It is larger than the minimum number you choose. Therefore, you will need to announce your answer in front of other participants at the end of the study.

Last, we would like to know your real-world privacy settings on social media platforms. Answer the following questions. Feel free to use your cell phone to check your privacy settings.

1. Do you use Instagram?

   - Yes
   - No

2. (Conditional on the answers to the first question) How often do you use Instagram?

   - Several times a year
   - Several times a week
   - About once a day
   - Several times a day
   - Almost constantly

3. (Conditional on the answers to the first question) We would like to know your ad preferences on Instagram. You can find your settings by clicking: **Settings and privacy - Accounts Center - Ad preferences - Customize ads**.

   Please check this tab: **advertisers you saw ads from**.

   To what extent are you aware that your behaviors on these ad partners are shared with Instagram?

   - I have no idea at all
   - I have heard about this but don't know much about it
   - I have some awareness, but I'm not entirely sure how much of my behavior is shared
   - I am fully aware

4. How did you feel when you saw the list? Answer the following questions.
   How comfortable are you with your behavior on these ad partners' platforms shared with Instagram?

   - Very uncomfortable
   - Somewhat uncomfortable
   - Somewhat comfortable
   - Very comfortable

   How risky do you think it is to share your behavior on these ad partners' platforms with Instagram?

   - Very risky
   - Moderately risky
   - Slightly risky

- Not risky at all

How appropriate do you think it is to share your behavior on these ad partners' platforms with Instagram?

- Very inappropriate
- Somewhat inappropriate
- Somewhat appropriate
- Very appropriate

5. (Conditional on the answers to the first question) Now, please go to the **Manage info** tab to check your ad preferences settings. Let us know your setting by answering the following questions. After all responses are collected, there is a 1/3 probability that you will be selected to show your setting page to the experimenters.

   Check this tab: **Activity information from ad partners**. Do you allow Instagram to use this information to show you ads that are more relevant to you? Your choice is:

   - Using this information
   - Not using this information

6. (Conditional on the answers to the first question) When was the last time you adjusted your ad preferences on Instagram?

   - Within one week
   - Within one month
   - Within one year
   - I have never adjusted my ad preferences settings

7. (Conditional on the answers to the first question) To what extent do you believe it is necessary to modify your ad preferences on Instagram?

   - Not at all necessary
   - Slightly necessary
   - Moderately necessary
   - Very necessary

8. (Optional) Do you have any comments or suggestions you would like to share with the researchers who designed this study? Is there anything you found unclear or confusing? Are there questions you had wished we asked?

E Additional Results

E.1 Summary Statistics

Table E.1: Summary Statistics: Information Sharing Game (Main Module)

	Full Sample	Control	Treatment
	(1)	(2)	(3)
Likelihood of Sharing	0.525	0.561	0.488
	(0.499)	(0.496)	(0.500)
Choosing Best Response	0.659	0.559	0.758
	(0.474)	(0.497)	(0.428)
Choosing Mutual Best Response	0.449	0.299	0.600
	(0.498)	(0.458)	(0.490)
Payoff (in US Dollars)	7.516	6.773	8.258
	(4.902)	(4.349)	(5.297)
Efficiency	0.636	0.514	0.758
	(0.394)	(0.403)	(0.344)
Prediction of Match's Decision (0 = Not Share, 1 = Share)	0.630	0.537	0.723
	(0.483)	(0.499)	(0.447)
Predicting Match's Decision Correctly	0.534	0.523	0.545
	(0.499)	(0.500)	(0.498)
# of Observations	3840	1920	1920
# of Participants	192	192	192

[a] This table reports means and standard deviations (in parentheses) for the variables in the analysis of the Information Sharing Game.

[b] Column (1) presents the data of the full sample. Column (2) and (3) present data from rounds without and with the average payoff matrix, respectively.

Table E.2: Summary Statistics: Other Modules

	Full Sample (1)	On Instagram (2)
WTA for Announcing Own Body Weight	53.16	53.38
	(41.23)	(41.30)
WTA for Announcing Avg Body Weight	49.36	49.28
	(41.62)	(41.60)
WTA: Own Number of Sexual Partners	62.82	62.97
	(41.48)	(41.29)
WTA: Total Number of Sexual Partners	61.18	60.86
	(41.07)	(41.11)
WTA: Own Opinion on Sexual Relationships	52.69	52.99
	(41.83)	(41.41)
WTA: Avg Opinion on Sexual Relationships	51.76	51.51
	(41.53)	(41.22)
Lottery Choice Switching Point (Risk Preference) (0 = Very Risk Loving, 11 = Very Risk Averse)	6.615	6.594
	(1.875)	(1.863)
Unknown Urns Switching Point (Ambiguity Preference) (0 = Very Ambiguity Averse, 100 = Very Ambiguity Loving)	47.77	48.01
	(12.84)	(12.67)
Contribution Amount in Public Goods Game (Cooperation) (0 = Not Cooperative, 200 = Very Cooperative)	89.09	88.52
	(69.67)	(69.46)
On Instagram	0.911	1
	(0.285)	(0)
Enable Sharing on Instagram (Default)		0.457
		(0.500)
Age	20.72	20.52
	(3.475)	(2.854)
Female	0.714	0.726
	(0.453)	(0.447)
Undergraduate Student	0.844	0.863
	(0.364)	(0.345)
Year in Program (0-7)	2.240	2.200
	(1.336)	(1.282)
Asian/Asian American	0.406	0.394
	(0.492)	(0.490)
Black/African American	0.0781	0.0686
	(0.269)	(0.253)
Hispanic/Latino	0.0625	0.0686
	(0.243)	(0.253)
White/European American	0.417	0.429
	(0.494)	(0.496)
Participated in Experiments Before	0.490	0.469
	(0.501)	(0.500)
# of Participants	192	175

[a] This table reports means and standard deviations (in parentheses) for the variables in the Measurement Games Module, the Friend Decision Making Module, and the Social Media Privacy Setting Survey.

[b] Column (1) presents the data of the full sample. Column (2) presents the data of participants who use Instagram.

E.2 Information Sharing Game

Review Questions

Among all review questions for the information sharing game, more than 90% of the participants answer Questions 1-5 and 8 correctly. Participants make more mistakes in Questions 6 and 7, which are the calculation questions. Therefore, whether they answer the calculation questions can work as a proxy for how good they are at comparing the benefits and costs of sharing in the information sharing game.

Table E.3: Summary Statistics of Information Sharing Game Review Questions

	Mean	SD	Min	Max	N
Review Question 1	0.984	0.124	0	1	192
Review Question 2	0.995	0.0722	0	1	192
Review Question 3	0.927	0.261	0	1	192
Review Question 4	0.984	0.124	0	1	192
Review Question 5	0.958	0.200	0	1	192
Review Question 6	0.615	0.488	0	1	192
Review Question 7	0.802	0.399	0	1	192
Review Question 8	0.917	0.277	0	1	192
Total Number of Correct Answers	7.182	1.035	3	8	192

BR Given by Average vs Realized Payoff Matrices

We analyze the extent to which participants follow the best responses indicated by the average payoff matrix and how this influences their likelihood of choosing the best responses in the realized payoff matrix. Most participants adhere to the best and mutual best responses suggested by the average payoff matrix, with only a small proportion choosing to deviate. The average payoff matrix does not always guide them to choose the best responses given by the realized payoff matrix, but it is still better than when participants make decisions by themselves.

Table E.4: Summary Statistics of Outcomes Given the Realized vs Average Payoff Matrix

	Mean	SD	Min	Max	N
Choosing Best Response	0.758	0.428	0	1	1920
Choosing Best Response in the Average Payoff Matrix	0.888	0.315	0	1	1920
Choosing Mutual Best Response	0.600	0.490	0	1	1920
Choosing Mutual Best Response in the Average Payoff Matrix	0.508	0.500	0	1	1920

Participants' Decision Types

We compare our participants with the following types of decision-makers: (1) decision-makers who always share (extremely open); (2) decision-makers who never share (extremely conservative); (3) decision-makers who randomly decide whether to share or not (naive) (4) decision-makers who always choose the best responses given by the theoretical predictions (derived from continuous cases); and (5) decision-makers who always choose the best responses given by the realized payoff matrix (decision-makers with perfect information). Table E.5 presents the rankings of different types of decision-makers based on the simulation results of their likelihood of choosing the best response.

Decision-makers with perfect information perform the best, followed by participants using the average payoff matrix. Decision-makers who follow the theoretical cutoffs rank third. Participants making their own decisions come next. Naive, extremely open, and extremely conservative decision-makers have similar performances, all worse than our actual participants.

Table E.5: Simulated Likelihood of Choosing Best Response by Different Types of Subjects

	Mean	SD	Min	Max	N
Perfect Information	1	0	1	1	1920
With Average Payoff Matrix	0.758	0.428	0	1	1920
Theoretical Cutoff	0.589	0.492	0	1	1920
Without Average Payoff Matrix	0.559	0.497	0	1	1920
Naive	0.506	0.500	0	1	1920
Extremely Open	0.501	0.500	0	1	1920
Extremely Conservative	0.499	0.500	0	1	1920

E.3 Friends Decision Making

Review Questions

Among all review questions for the friends decision making module, most participants make mistakes in the second question, which asks them to judge the statement ”Your decision may reveal some information about your friend.”. This was also the question that many participants ask about during the sessions. Questions 3-5 are about the BDM method. It seems that most participants have a good understanding of it.

Table E.6: Summary Statistics of Friend Decision Making Module Review Questions

	Mean	SD	Min	Max	N
Review Question 1	0.906	0.292	0	1	192
Review Question 2	0.380	0.487	0	1	192
Review Question 3	0.714	0.453	0	1	192
Review Question 4	0.906	0.292	0	1	192
Review Question 5	0.875	0.332	0	1	192
Total Number of Correct Answers	3.781	0.918	1	5	192

Summary Statistics

Though participants who opted out of sending the survey to their friends were told to skip questions in this module, some of them still answered the questions and announced some of their answers. In the main analysis, taking opting out as a signal of privacy protection, we recode the data by assigning 100 to participants who opted out of the survey as the lowest payment they are willing to accept in all questions. Table E.7 presents the summary statistics of participants' original willingness-to-accept reported in the study.

Table E.7: Summary Statistics of Willingness to Accept

	Mean	SD	Min	Max	N
Male & Other					
Own Body Weight	37.07	38.08	0	100	55
Avg Body Weight	35.58	38.41	0	100	55
Own Number of Sexual Partners	55.75	43.19	0	100	55
Total Number of Sexual Partners	53.38	42.05	0	100	55
Own Opinion on Sexual Relationships	39.80	39.46	0	100	55
Avg Opinion on Sexual Relationships	39.44	38.79	0	100	55
Total WTA	261.0	208.8	0	600	55
WTA for Q1 (Self)	132.6	106.6	0	300	55
WTA for Q2 (Both)	128.4	105.5	0	300	55
Female					
Own Body Weight	49.26	38.81	0	100	137
Avg Body Weight	45.81	39.48	0	100	137
Own Number of Sexual Partners	60.66	40.37	0	100	137
Total Number of Sexual Partners	59.68	40.74	0	100	137
Own Opinion on Sexual Relationships	47.96	39.95	0	100	137
Avg Opinion on Sexual Relationships	47.91	39.90	0	100	137
Total WTA	311.3	205.9	0	600	137
WTA for Q1 (Self)	157.9	103.8	0	300	137
WTA for Q2 (Both)	153.4	107.1	0	300	137
Total					
Own Body Weight	45.77	38.90	0	100	192
Avg Body Weight	42.88	39.35	0	100	192
Own Number of Sexual Partners	59.26	41.14	0	100	192
Total Number of Sexual Partners	57.88	41.11	0	100	192
Own Opinion on Sexual Relationships	45.62	39.88	0	100	192
Avg Opinion on Sexual Relationships	45.48	39.67	0	100	192
Total WTA	296.9	207.4	0	600	192
WTA for Q1 (Self)	150.6	105.0	0	300	192
WTA for Q2 (Both)	146.2	107.0	0	300	192

Table E.8 summarizes the sharing outcomes in the friends decision making module. Participants are very responsive in the first stage of this module. They on average answered more than 5 out of 6 questions. The number of answers they announce on average, however, is only around 3 out of 6. As female participants demand a higher payment for announcing, they also announce fewer questions than the other participants on average.

Table E.8: Summary Statistics of Answers and Announcements

	Mean	SD	Min	Max	N
Male & Other					
Total Number of Questions Answered	5.236	1.710	0	6	55
Number of Q1(Self) Answered	2.709	0.809	0	3	55
Number of Q2(Both) Answered	2.527	0.979	0	3	55
Total Number of Answers Announced	3.327	2.253	0	6	55
Number of Q1(Self) Answers Announced	1.636	1.253	0	3	55
Number of Q2(Both) Answers Announced	1.691	1.184	0	3	55
Female					
Total Number of Questions Answered	5.591	1.380	0	6	137
Number of Q1(Self) Answered	2.825	0.674	0	3	137
Number of Q2(Both) Answered	2.766	0.740	0	3	137
Total Number of Answers Announced	2.861	2.223	0	6	137
Number of Q1(Self) Answers Announced	1.416	1.180	0	3	137
Number of Q2(Both) Answers Announced	1.445	1.194	0	3	137
Total					
Total Number of Questions Answered	5.490	1.486	0	6	192
Number of Q1(Self) Answered	2.792	0.715	0	3	192
Number of Q2(Both) Answered	2.698	0.820	0	3	192
Total Number of Answers Announced	2.995	2.236	0	6	192
Number of Q1(Self) Answers Announced	1.479	1.202	0	3	192
Number of Q2(Both) Answers Announced	1.516	1.193	0	3	192

Correlations Between WTA Across Scenarios

We observe significant correlations between participants' WTA reported across the three scenarios. Table E.9 and E.10 shows the pairwise correlations with p-values.

Table E.9: Correlations Between WTA for the First Question Across Scenarios

WTA	Own Body Weight	Own # of Sexual Partners	Own Opinion
Own Body Weight	-	-	-
Own # of Sexual Partners	0.698***	-	-
Own Opinion	0.713***	0.755***	-

[a] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[b] p-value for (the correlation between Total # of Sexual Partners and Avg Opinion) = (the correlation between Avg Body Weight and Avg Opinion): 0.375

Table E.10: Correlations Between WTA for the Second Question Across Scenarios

WTA	Avg Body Weight	Total # of Sexual Partners	Avg Opinion
Avg Body Weight	-	-	-
Total # of Sexual Partners	0.745***	-	-
Avg Opinion	0.770***	0.834***	-

[a] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[b] p-value for (the correlation between Total # of Sexual Partners and Avg Opinion) = (the correlation between Avg Body Weight and Avg Opinion): 0.079

Proportion of Participants Censored

Table E.11: Proportion of Participants Whose WTA is Censored at Upper/Lower Bounds

Proportion Censored (in %)	Question	Left-Censored (WTA = 0)	Right-Censored (WTA = 100)	Total
Body Weight	Own	16.1	30.2	46.3
	Joint	19.3	30.2	49.5
Number of Sexual Partners	Own	13.0	42.7	55.7
	Joint	12.5	41.1	53.6
Opinion on Intimate Relationships	Own	17.2	30.2	47.4
	Joint	15.6	30.7	46.3

Robustness Checks

Table E.12 presents the results after controlling for question-specific characteristics. Following the method of Krupka and Weber (2013), we use the reported body weight, number of sexual partners, and opinions of all other participants as the social norms. We standardize participants' answers against these social norms to measure how much they deviate from the average. The further their answers differ from the norm, the more stigmatization they may perceive. Furthermore, since a participant's friend's information can also be inferred if both questions in a scenario are answered, we also standardize the inferred friend's information for the regressions on the $\rho > 0$ questions to measure how much a participant thinks their friend deviates from the social norms.

In the body weight scenario, we find that a participant demands a higher payment to share their own body weight as their weight increases. This is likely because being overweight often incurs negative judgments, making participants less willing to share this information.

In contrast, in the intimate relationship scenario, participants demand a lower payment to share their own number of sexual partners as the number increases. Participants with more sexual partners may be perceived as more attractive and are thus more confident to share their experiences. They may also be more open to discussing such topics in front of others.

We observe no significant correlations between WTA and the responses provided for the $\rho > 0$ questions in these two scenarios, nor for the scenario asking about participants' opinions on intimate relationships.

We use whether a participant's answer is consistent with their friend's as a proxy for their confidence in the answer and the extent to which they are telling the truth. Generally, we find that participants demand a lower payment if their answers are consistent with their friend's answers. This may be because when participants are more certain about their answers, they anticipate others to have less misunderstanding of themselves and their friends.

We use Tobit regressions in the main analysis to analyze the influencing factors of participants' WTA. Here, we present the results from linear regressions in Table E.13. In addition, since participants' WTA are concentrated at the two extremes across all scenarios, we also recode participants' WTA into binary terms and run logit regressions. The results are presented in Table E.14.

Table E.12: WTA in All Scenarios; Tobit Regression

	Own Weight (1)	Avg Weight (2)	Own # of Partners (3)	Total # of Partners (4)	Own Opinion (5)	Avg Opinion (6)
Likelihood of Sharing w/o Matrix	-75.500***	-90.262***	-83.310**	-100.561***	-63.223**	-62.184***
	(22.632)	(28.618)	(32.242)	(35.099)	(30.877)	(22.753)
Risk Aversion (Standardized)	4.147	8.956	17.866***	9.794	6.879	8.144**
	(5.359)	(5.862)	(4.037)	(6.627)	(4.285)	(3.400)
Ambiguity Aversion (Standardized)	0.542	3.788	9.584	4.159	-1.786	1.487
	(7.055)	(5.171)	(6.341)	(6.543)	(6.447)	(5.267)
Cooperation (Standardized)	-5.474	-5.759	0.662	2.788	1.895	0.383
	(3.844)	(3.871)	(9.227)	(8.826)	(5.614)	(5.320)
Female	40.427***	27.539*	24.008*	30.070**	25.132*	18.758
	(14.623)	(16.247)	(13.519)	(14.645)	(13.964)	(14.230)
Own Information (Standardized)	16.151***	-2.365	-16.549**	-18.489**	-8.286	-5.931
	(5.469)	(8.210)	(6.377)	(7.174)	(6.421)	(5.842)
Own Information Is Verified	-23.154**		-20.834		-1.320	
	(9.747)		(13.926)		(15.813)	
Inferred Friend's Info (Standardized)		4.139		-2.898		-6.730
		(3.735)		(11.216)		(6.011)
Aggregate Information Is Verified		-18.057		-22.577***		-12.306
		(13.119)		(7.000)		(12.293)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
# of Participants	115	112	110	105	116	116

[a] Standard errors in parentheses are clustered at the session level.

[b] The outcome variables are participant's WTA to announce their answers. The first two regressions represent the Body Weight scenario. The second two regressions represent the Intimate Relationships scenario. The last two regressions represent the Opinions on Intimate Relationships scenario.

[c] WTA is measured in points and is chosen on the interval [0, 100]. A higher WTA means being less willing to share.

[d] Own and inferred friend's information are standardized based on the mean and standard deviation of all other participants' reported own information.

[e] If a participant's answer is no more than 10% different from their friend's answer, this answer is verified.

[f] Controls included are identical as in Table 7

[g] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table E.13: WTA in All Scenarios; Linear Regression

	Own Weight (1)	Avg Weight (2)	Own # of Partners (3)	Total # of Partners (4)	Own Opinion (5)	Avg Opinion (6)
Likelihood of Sharing w/o Matrix	-18.026	-26.654*	-27.378	-30.052**	-18.609	-21.789*
	(11.815)	(14.786)	(17.152)	(13.625)	(13.365)	(11.641)
Risk Aversion (Standardized)	-2.236	-3.222	-0.873	-0.692	-1.855	-1.272
	(3.088)	(2.923)	(3.085)	(3.356)	(3.294)	(3.648)
Ambiguity Aversion (Standardized)	2.509	3.524	1.835	1.082	-0.143	0.863
	(3.925)	(3.603)	(2.618)	(3.054)	(3.731)	(3.562)
Cooperation (Standardized)	0.318	0.397	0.883	1.887	2.644	2.047
	(2.779)	(2.843)	(3.844)	(3.504)	(3.250)	(3.331)
Female	6.687	4.516	8.131	10.349	8.868	8.825
	(5.969)	(6.622)	(6.326)	(6.786)	(7.160)	(7.359)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
# of Participants	192	192	192	192	192	192

[a] Standard errors in parentheses are clustered at the session level.

[b] The outcome variables are participant's WTA to announce their answers. The first two regressions represent the Body Weight scenario. The second two regressions represent the Intimate Relationships scenario. The last two regressions represent the Opinions on Intimate Relationships scenario.

[c] WTA is measured in points and is chosen on the interval [0, 100]. A higher WTA means being less willing to share.

[d] Controls included are identical as in Table 7.

[e] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table E.14: Willing to Share in All Scenarios; Logit Regression

	Own Weight	Avg Weight	Own # of Partners	Total # of Partners	Own Opinion	Avg Opinion
	(1)	(2)	(3)	(4)	(5)	(6)
Likelihood of Sharing w/o Matrix	0.059	0.345**	0.161	0.311*	0.146	0.137
	(0.129)	(0.147)	(0.209)	(0.176)	(0.151)	(0.123)
Risk Aversion (Standardized)	0.016	0.055	-0.013	0.005	0.060*	0.051
	(0.031)	(0.035)	(0.038)	(0.035)	(0.032)	(0.043)
Ambiguity Aversion (Standardized)	-0.048	-0.055	-0.091***	-0.090***	-0.022	-0.029
	(0.039)	(0.035)	(0.034)	(0.035)	(0.036)	(0.030)
Cooperation (Standardized)	-0.034	-0.011	0.008	0.001	-0.056**	-0.042
	(0.032)	(0.031)	(0.043)	(0.040)	(0.026)	(0.030)
Female	0.055	0.057	0.089	0.026	0.023	-0.007
	(0.053)	(0.061)	(0.055)	(0.060)	(0.060)	(0.075)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
# of Participants	192	192	192	192	192	192

[a] Standard errors in parentheses are clustered at the session level.

[b] The outcome variables are indicators of whether participant's WTA to announce their answers is smaller than 100 points. The first two regressions represent the Body Weight scenario. The second two regressions represent the Intimate Relationships scenario. The last two regressions represent the Opinions on Intimate Relationships scenario.

[d] Controls included are identical as in Table 7.

[e] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We collect participants' height to calculate their BMI, which is a measure of whether a participant is overweight. We include this as a control variable to examine if we can observe the same effect as when we include the social norm controls. Table E.15 presents the Tobit regression results. We do observe participants with higher BMI demand a higher payment to share their own body weight.

Table E.15: WTA in the Body Weight Scenario; Tobit Regression

	WTA for Own Body Weight	
	(1)	(2)
Likelihood of Sharing w/o Matrix	-30.670	-38.769
	(18.825)	(30.226)
Risk Aversion (Standardized)	-4.063	-9.194
	(5.406)	(5.799)
Ambiguity Aversion (Standardized)	6.597	5.585
	(7.105)	(7.902)
Cooperation (Standardized)	0.688	-3.704
	(4.465)	(5.052)
Female	7.105	46.328***
	(11.281)	(15.654)
BMI		4.053***
		(1.052)
Controls	Yes	Yes
# of Participants	192	122

[a] Standard errors in parentheses are clustered at the session level.

[b] The outcome variable is a participant's WTA to announce their own body weight.

[c] WTA is measured in points and is chosen on the interval [0, 100]. A higher WTA means being less willing to share.

[d] Controls include a participant's age, ethnicity, and indicators suggesting whether they have participated in economics or psychological studies before, whether they have taken introductory statistics courses, whether they correctly answered two key questions (which over 25% of participants answered incorrectly), and how long a participant and their friend have known each other.

[e] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table E.16 presents regression results investigating influencing factors of total WTA, measured by the sum of the minimum amounts participants are willing to accept across all scenarios.

Table E.16: Total WTA Across All Scenarios; Tobit Regression

	(1) Total WTA	(2) WTA for Q1 ($\rho = 0$)	(3) WTA for Q2 ($\rho > 0$)
Likelihood of Sharing w/o Matrix	-221.738***	-93.562*	-127.751***
	(84.813)	(47.715)	(46.190)
Risk Aversion (Standardized)	-0.053	-7.738	-6.714
	(20.103)	(12.061)	(13.738)
Ambiguity Aversion (Standardized)	14.570	8.341	11.763
	(20.987)	(11.930)	(13.364)
Cooperation (Standardized)	5.651	6.806	5.930
	(19.982)	(11.869)	(13.248)
Female	51.946	27.545	26.022
	(44.801)	(25.840)	(28.496)
Controls	Yes	Yes	Yes
# of Participants	192	192	192

[a] Standard errors in parentheses are clustered at the session level.

[b] The outcome variable in the first regression is the sum of a participant's WTA in all 6 questions. The outcome variable in the second regression is the sum of a participant's WTA in all $\rho = 0$ questions. The outcome variable in the last regression is the sum of a participant's WTA in all $\rho > 0$ questions.

[c] WTA is measured in points. Total WTA is chosen on the interval [0, 600]. WTA for Q1 and Q2 are chosen on the interval [0, 600]. A higher WTA means being less willing to share.

[d] Controls included are identical as in Table 7.

[e] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table E.17 presents regression results on announcement outcomes.

Table E.17: Announcement Outcomes Across All Scenarios; Logit Regression

	Announced	For Q1 ($\rho = 0$)	For Q2 ($\rho > 0$)
	(1)	(2)	(3)
Likelihood of Sharing w/o Matrix	0.335**	0.474***	0.402**
	(0.155)	(0.156)	(0.157)
Risk Aversion (Standardized)	-0.031	-0.027	0.001
	(0.031)	(0.030)	(0.031)
Ambiguity Aversion (Standardized)	-0.057*	-0.065	-0.071**
	(0.032)	(0.042)	(0.034)
Cooperation (Standardized)	0.020	0.039	0.014
	(0.028)	(0.025)	(0.032)
Female	0.027	-0.039	0.022
	(0.067)	(0.075)	(0.080)
Controls	Yes	Yes	Yes
# of Participants	192	192	192

[a] Standard errors in parentheses are clustered at the session level.

[b] The outcome variable in the first regression is whether a participant announces at least an answer. The outcome variable in the second regression is whether a participant announces at least an answer to any of the $\rho = 0$ questions. The outcome variable in the last regression is whether a participant announces at least an answer to any of the $\rho > 0$ questions.

[c] Controls included are identical as in Table 7.

[d] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

E.4 Social Networking Platform Privacy Settings

Over 90% of our participants use Instagram, and among those, the majority are frequent users, with 88% checking Instagram at least once a day.

Table E.18 presents summary statistics for some of the survey questions. The status quo appears unsatisfactory for many participants. When instructed to check the list of ad partners sharing their activity information with Instagram, most participants perceive this as risky and inappropriate, and they do not feel comfortable with it. This finding is consistent with our results from the information sharing game, where participants tend to over-share.

Table E.18: Summary Statistics of the Social Network Platform Privacy Settings Survey

	Mean	SD	Min	Max	N
Comfortable with Sharing	1.343	0.763	0	3	175
Think Sharing is Risky	1.657	0.763	0	3	175
Think Sharing is Appropriate	1.269	0.796	0	3	175
Think Adjusting Settings Is Necessary	1.371	0.880	0	3	175

Note: The comfortableness, risk, appropriateness and necessity questions are reported on a scale of 0-3.

68.6% of participants have never adjusted their privacy settings on Instagram before. There is no significant difference in adjustment behaviors between those who enable and those who disable sharing. Additionally, 84% of participants believe it is somewhat necessary to adjust their privacy settings on Instagram, with no significant difference in attitudes between those who enable and those who disable sharing.

Table E.19: Adjusting Behaviors And Attitudes on Instagram; Linear Regression

	Last Time Adjusted	Think Necessary to Adjust
Enable Sharing	-0.264	-0.063
	(0.165)	(0.134)
Constants	0.789***	1.400***
	(0.112)	(0.090)
# of Participants	175	175

[a] Standard errors in parentheses are clustered at the session level.
[b] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

In the follow-up survey, we also asked participants how they adjusted their privacy settings in the two months following the lab study. Among the 125 participants who completed the follow-up survey, approximately 75% did not make any adjustments to their privacy settings on any social networking platforms. Meanwhile, 23.2% of participants adjusted their settings on Instagram to be more conservative, and only one participant adjusted their settings on Instagram to be more open.

Table E.20 presents the regression results. Those who believe it is necessary to adjust their privacy settings are more likely to make their settings more conservative.

Table E.20: Adjusting Behaviors After the Study; Linear Regression

	Adjusted on Instagram	Adjusted on Other Platforms
Enable Sharing	0.088	0.097
	(0.084)	(0.088)
Think Necessary to Adjust	0.114^{**}	0.125^{***}
	(0.046)	(0.043)
Constants	0.047	0.004
	(0.075)	(0.068)
# of Participants	115	112

[a] Standard errors in parentheses are clustered at the session level.

[b] The outcome in Column (1) is how participants adjust their privacy settings on Instagram. The outcome in Column (2) regression is how participant adjust their privacy settings on Other Platforms.

[c] The outcome variables take value 0 if no adjustment incurred, 1 if adjusted to be more conservative, -1 if adjusted to be more open.

[d] * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

E.5 Payoffs

Table E.21: Summary Statistics of Payoffs by Module

	Mean	SD	Min	Max	N
Total Cash Payment (in US dollars)	50.02	10.62	30	85	192
Total Payoffs	1779.8	424.1	987	3169	192
Information Sharing Game (Module 1)	308.8	176.5	54	981	192
Information Sharing Game (Module 2)	330.9	206.6	5	1189	192
The L Game	223.8	123.0	10	385	192
The U Game	309.9	243.4	0	500	192
The P Game	214.4	48.34	120	320	192
Friend Decision Making	172.7	126.4	0	471	192