Measuring the Value of Genetic Data Privacy

Zeeshan Samad[a,2], **Myrna Wooders**[a], **Bradley Malin**[b], **and Eugene Vorobeychik**[c]

[a]Department of Economics, Vanderbilt University; [b]Departments of Biomedical Informatics, Biostatistics, & Computer Science, Vanderbilt University; [c]School of Engineering, Washington University in St. Louis

This manuscript was compiled on January 28, 2020

How can people's value of genetic data privacy be estimated? We address this question using methods of behavioral/experimental economics. Through decision-making tasks, in two different frames involving risk of loss of privacy of data, in one case genetic data and in the other case financial data, we elicit people's risk attitudes towards their genetic data relative to their risk attitudes towards their financial data. We find that people are more willing to risk a loss of genetic data privacy for health benefits than they are willing to risk a loss of financial data privacy for financial benefits. We conduct related experiments to better understand people's motivations for willingness to share their genetic data and undertake the risk of doing so. The results of these experiments suggest that individuals are primarily motivated by private benefits, whether they are health or monetary and little to none by altruism and trust.

Genetic | Data | Privacy | Experiment | Laboratory | Risk

Genetic data is used in a variety of settings: To name a few, biomedical researchers use it to look for ways to prevent and treat diseases, clinical physicians use it to provide better care to patients, and consumers use it to learn about risk of genetic diseases and to get insights into their geographical ancestry. In most cases, genetic data is not discarded after serving its initial purpose and is instead stored in databases. These databases are often sold or shared with other organizations who use them for different purposes, creating a cause for concern for genetic data privacy.

The primary goal of this research is to evaluate the risk attitudes of individuals toward loss of privacy of genetic data relative to their attitudes toward loss of financial data. We use financial data as a basis for comparison as we live in a society where risk, costs, and benefits are commonly measured in monetary terms (e.g. insurance). Another part of our goal is to better understand the conditions under which individuals are willing to share their genetic data in a healthcare setting. To answer these questions, we use a behavioral economics approach and conduct a laboratory experiment to observe how people make decisions in a two-player game. This approach allows us to elicit privacy concerns through actions, as opposed to opinions, of individuals.

Although several previous investigations have studied the extent to which people are concerned about privacy and security of their genetic data, they elicit privacy concerns through survey responses.[*] One important limitation of this methodology is a phenomenon, which has come to be known as the *privacy paradox*, that reported attitudes about privacy are not consistent with people's true concerns about privacy, as revealed by their actions and measures taken to protect their data.[†] In an attempt to minimize the bias caused by the privacy paradox, we elicit privacy concerns through actions as opposed to stated attitudes.

In a base experiment, we divide subjects into pairs and ask each pair to participate in a two-player game that involves risks and benefits. We use framing effects to separately elicit privacy concerns about genetic data and financial data. That is, we present the same risks and benefits in two different ways, or *frames*. About half of the subjects are told that the game represents an interaction between a patient and a physician and the other half are told that the game represents an interaction between an investor and a money manager. In the patient-physician frame, which we call the "genetic frame", the subject who is assigned the role of a patient needs to decide whether to get a genetic test which has potential health benefits but puts her genetic data at risk. In the investor-money manager frame, the subject who is assigned the role of an investor needs to decide whether to make a monetary investment which has potential monetary benefits but puts her financial data at risk. In both frames, all payoffs and risks are in terms of money and in identical amounts; the only difference is what the amounts are meant to represent. If subjects who receive the genetic frame make systematically different choices than subjects who receive the investment frame, then it would mean that people assess risks to their genetic data differently from risks to their financial data.[‡]

[*] Some examples of such studies are Sanderson et al. (1), Kaufman et al. (2), Condit et al. (3), Duquette et al. (4), Rauscher et al. (5), Botkin et al. (6), Edwards et al. (7), Lemke et al. (8). See Clayton, Halverson, Sathe, and Malin (9) for a comprehensive literature review on individuals' perspectives on genetic data privacy.

[†] See Barth and De Jong (10) for a literature review of the privacy paradox

[‡] If patients (i.e. subjects who receive the genetic frame) make different choices than investors (i.e. subjects who receive the investment frame), we observe framing effects. The only plausible explanation for framing effects is that subjects' attitudes/concerns about genetic data privacy are different than those about financial data privacy. However, if we do not observe a framing effect, we would not be able to conclude anything about privacy concerns for genetic data relative to privacy concerns for financial data. This is because, although there is only one reason that can explain framing effects, there are several possible reasons that can explain the absence of framing effects. For example, the frames may not be effective if the context/story in each frame is very short and/or if there are only extremely minor differences across the frames.

Significance Statement

When asked "Are you worried about genetic privacy," the general public tends to say "yes." It appears to also be the case that when asked whether they are concerned about privacy of financial data, the public tends to answer "yes" or even "YES!". Indeed, people have many concerns. We develop a novel experiment to assess the depth of concern of individuals about privacy of genetic data relative to their concern about privacy of financial data. We find that individuals are more willing to take risks for possible health benefits from genetic testing than they are to take risks for possible financial benefits from investment of money and quantify that difference, shedding new light on the depth of concern about privacy of genetic data.

[2]To whom correspondence should be addressed. E-mail: zeeshan.samad@vanderbilt.edu

Our secondary question is about the conditions under which individuals are willing to share their genetic data. To disentangle the different motivations behind subjects' willingness to risk privacy of their data, we ask subjects to participate in three additional versions of the two-player game, where each version removes or adds a single feature of the game. This allows us to isolate and control competing explanations and incentives. Thus, we ask the subjects to play a total of four treatments of the same two-player game, where each treatment is a slightly modified version of another treatment. These small modifications allow us to isolate and control competing explanations and incentive.

The first treatment is designed so that subjects are motivated only by personal benefit. The second treatment allows subjects to be motivated by personal benefit as well as altruism. The third treatment allows subjects to be motivated by personal benefit, altruism, and trust. Observing the differences across treatments allows us to decompose the factors that motivate subjects to share their genetic data. For example, in the third treatment, patients may trust their physician to take measures to protect their genetic data but such an option does not exist in the second treatment. So, if we find that patients are more likely to share their genetic data in the third treatment than in the second treatment, it suggests that trust in one's physician increases the likelihood of sharing one's genetic data.

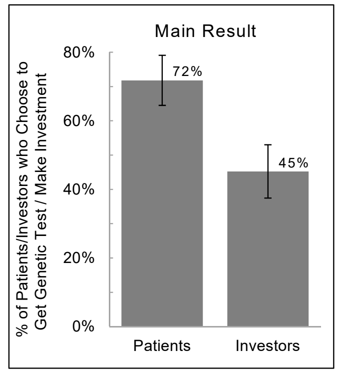
Even though we are primarily interested in the choices made by the party that shares the data (e.g. patient of investor), we nonetheless design a two-player game – as opposed to designing an individual decision problem) for two reasons. First, an interaction with another party makes the scenario a bit more similar to the "real world" where any data-sharing choice involves an interaction with clinicians, hospitals, government agencies, direct-to-consumer companies, etc. Such interactions make trust an important factor for voluntarily sharing data in the real world and thus a crucial issue that needs to be studied. Second, a two-player game allows us to study the incentives of second party. For example, do physicians care about protecting their patients' data because they are altruistic or because they feel beholden to do so?

Results

Subjects who are assigned the genetic frame are more likely to take a risk for a potential benefit than subjects who are assigned the investment frame. As Figure 1 shows, about 72% of the subjects in the genetic frame choose the risky option compared to 45% of subjects in the investment frame. In both frames, subjects choose between a safe option, which is to receive \$6 with certainty, and a risky option, which is to receive \$60 with a probability of 5%, \$4 with a probability of 71%, or \$0 with the remaining probability of 24%. The only difference between the two frames is what the numbers are meant to represent. In the genetic frame, subjects choose whether or not to undergo genetic testing which might result in a health benefit but would put their genetic data at risk. In the investment frame, subjects choose whether or not to make an investment that might result in a monetary benefit but would put their financial data at risk.

Figure 1 shows our main finding: people are more willing to risk their genetic data for the sake of health benefits than they are willing to risk their financial data for the sake of monetary benefits. We also collect some demographic information of subjects, such as age, race, gender, religiousness, and whether subjects have ever been a victim of fraud or identity theft. Table 1 shows results of regressions that include these control variables. The OLS column shows results for a linear specification and the Probit column shows the results for a probit model. In both specifications, the dependent variable is the % of subjects who choose to get a genetic test / make an investment, which is the same variable on the vertical axis in Figure 1. Our key independent variable is Genetic Framing, which is a binary variable indicating whether a subject was given the genetic frame or investment frame. This variable has a coefficient of 0.254 in the OLS regression, implying that patients are about 25.4% on average more likely to get a genetic test than investors are about making an investment. For the probit results, we do a post-estimation analysis (output not shown) to find that the probability that a patient chooses to get a genetic test is 68.6%, whereas the probability that an investor chooses to make an investment is 42.5%. This means that patients are about 26.1% more likely to get a genetic test than investors are to make an investment. We do not find any correlation between willingness to take a risk and any demographic variable.

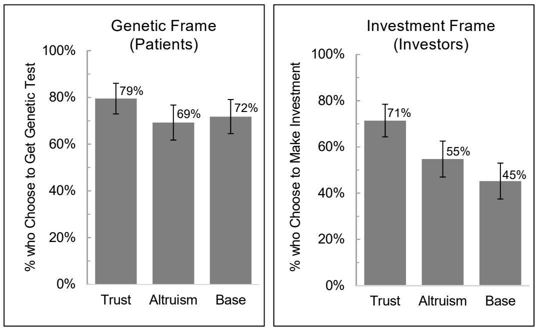
Fig. 1. Framing Effects in Base Treatment



A secondary goal of this study is to understand the factors that affect subjects' willingness to put their genetic data at risk. In particular, we are interested in the effect of other-regarding preferences on the willingness to put one's genetic data at risk. To do this, we conduct four treatments of the experiment, where each treatment potentially contains additional motivations depending on subjects' other regarding preferences. Our main finding, discussed above, results from the first treatment of the experiment which we call "base treatment". In this treatment, subjects are assigned the role of either a patient or an investor and are not paired with any other subject. Thus, they make choices without any regard for other subjects.

In a second treatment, which we call "altruism treatment", subjects assigned the role of a patient (investor) are randomly and anonymously paired with another subject who is assigned the role of a physician (money manager). The only difference between the altruism and base treatments is that in the altruism treatment, the risky option also benefits the physician, allowing subjects to be potentially influenced by their altruism. In this treatment, patients are told that a genetic test would additionally help their physician with his medical research. So, if patients are altruistic towards their physician or medical

Fig. 2. Treatment Effects for Patients and Investors



research in general, they should be more willing to undergo genetic testing than they are in the base treatment.

A third treatment, which we call "trust treatment", adds one more feature to the altruism treatment by allowing the physician (or, in the case of subjects who are presented the investment framing, money manager) to reduce the risk to the patient's genetic data by strengthening the security of his database. This allows patients (investors) to be influenced by how much they trust their physician (money manager). If they believe that their physician (money manager) is considerate enough of their well-being, they have more reason to choose the risky option.

Figure 2 compares the results of these treatments. We do not find a significant difference across treatments in the genetic framing, but we find differences across treatments in the investment framing. In the investment framing, a significantly greater number of investors are willing to put their financial data at risk when their money manager is able to reduce their risk, indicating that investors trust their money manager to have their best interest at heart. This difference also suggests that the treatment design works as it was intended to, because having trust increases one's willingness to take a risk. The fact that there is no statistically significant difference in the genetic framing suggests that patients do not trust their physicians to have their best interest at heart. In fact, we observe that fewer – even though by a very small margin – subjects put their data at risk in the altruism treatment relative to the base treatment. Although the difference is very small, it is interesting because it is in the opposite direction from what we expected, indicating that patients are averse, if anything, towards benefiting their physicians.

One of the unique features about our study design is the interaction between a patient/investor and physician/money manager (in the genetic frame patients interact with physicians whereas in the investment frame investors interact with money managers). The fact that physicians/money managers also have decisions to make allows us to gain insights into their preferences, particularly what motivates them to protect their patient's/client's data. Physician/money manager subjects are given some money and told that their database is currently at risk of a data breach and they can reduce this risk by spending some of their money. Although a data breach would not directly affect their own payoff, it would significantly reduce the payoff of their patient/client. If we observe a framing effect, i.e. if physicians systematically spend more or less than money managers, it would imply that people asses risk to genetic data privacy differently that risk to financial data privacy, even when the risk is to *other* people's data.

All physician/money manager subjects participate in two treatments of the experiment, an "altruism treatment" and a "reciprocity treatment". In both treatments, subjects are given \$4.00 and faced with the same decision problem; the only difference between the two treatments is *where* their money comes from. In the reciprocity treatment, physicians/money managers receive money from their patients/investors, whereas in the altruism treatment they get it from the experimenter.[§] This allows us to decompose the motivations of physicians/money managers. In particular, if we find that subjects spend more money to improve their database security in the reciprocity treatment compared to the altruism treatment, it must be because they feel beholden towards their patient/investor for giving them their \$4.00.

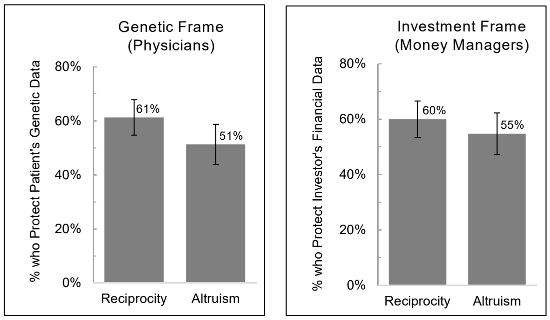
Figure 3 presents the responses of physicians and money managers. In the reciprocity treatment, 61% of physicians and 60% of money managers spend a positive amount to reduce the risk to their patient's/client's data. In the altruism treatment, 51% of physicians and 55% of money managers spend a positive amount to reduce the risk to their patient's/client's data. Since neither of these differences are significant at the 5% level, we conclude that there are no framing effects for physicians and money managers. That is, when it comes to *other* people's data, subjects do not exhibit different levels of concern for different types of data. In particular, subjects are no more concerned about other people's genetic data privacy any more (or less) than they are about their financial data privacy. Figure 3 also shows treatment effects within each frame. Within the genetic frame, 51% of physicians in the altruism treatment and 61% of physicians in the reciprocity treatment spend some positive amount of money to protect their patient's genetic data. This difference is marginally significant ($p = 0.1033$ in a two-tailed test and $p = 0.0516$ in a one-tailed test) at the 10% level. In the investment frame, 55% of money managers in the altruism treatment and 60% of money managers in the reciprocity treatment spend some positive amount of money to protect their client's financial data; and these two are not statistically different (two-tailed test $p = 0.3256$). This suggests that part of the reason physicians spend money to protect their patient's genetic data is because they feel beholden to the patient for choosing to get a genetic test. By contrast, money managers do not feel beholden to their clients for choosing to make an investment.

Discussion

Our main finding is that individuals are more willing to put their genetic data at risk for a health benefit than they are willing to put their financial data at risk for the sake of a monetary benefit. This would happen if people have greater privacy concerns about their financial data than about their genetic data. Another possible explanation may be that subjects place a greater value on health benefits than on monetary

§In the reciprocity treatment, if a patient/investor chooses the risky option (i.e. gets a genetic test/makes an investment), their physician/money manager receives \$4.00. Since about 79% of patients choose to get a genetic test and 71% of investors choose to make an investment, that is how many physicians and money managers receive \$4.00 in the reciprocity treatment. By contrast, in the altruism treatment all physicians/money managers are endowed with \$4.00 and receive nothing if their patient/investor chooses the risky option.

Fig. 3. Framing and Treatment Effects for Physicians and Money Managers



benefits. However, this is rather unlikely because both health and monetary benefits are presented by the same amount of \$6.00, and there is no reason for any subject to value this benefit at either greater than or less than \$6.00. Privacy concerns, on the other hand, are presented in terms of probability of a data breach. Unlike a constant dollar amount, probability is subject to over- or under- weighting, depending on the environment in which individuals make decisions.[¶] That is, individuals tend to incorrectly perceive probabilities due to a wide range of factors, particularly their current emotional state or decision-making environment. So, individuals tend to exhibit greater willingness to take risks in environments where they feel fearless (or even adventurous, comfortable, etc.) by under-weighting the probability of losses. That is, they respond to any probability (e.g. 25%) in that environment the way they would respond to some lower probability (say 20%) in another environment.

Instead of using framing effects, we could have approached our research question by presenting the problem only as the decision to get a genetic test (i.e. removing the investment frame). One could argue that this approach would give us twice the amount of observations at the same cost, and we could compare subjects' willingness to put their genetic data at risk with the theoretical willingness of a risk-neutral individual. For example, given that either choice in the base treatment results in an expected payoff of \$6.00, a risk-neutral individual would be indifferent between getting a genetic test and not getting a genetic test. Based on this theoretical expectation, we could then set our null hypothesis that about 50% of subjects will choose to get a genetic test. And if we ended up observing that significantly more (fewer) subjects choose to get a genetic test, we could interpret it to mean that people are generally more (less) willing to put their genetic data at risk, relative to other types of data. The biggest problem with such an approach is the assumption that people have risk-neutral preferences when it comes to putting their data at risk.

Our current experiment design involving framing effects allows us to address the research question without making any assumptions about subjects' general risk preferences. By comparing results from two groups, we allow subjects to maintain heterogeneous levels of concerns about data privacy in general and measure their attitudes towards genetic data privacy relative to privacy of other types of data.

¶ There is an extensive amount of literature demonstrating a cognitive bias with regards to perception of probabilities. Some seminal papers in this literature include Wu and Gonzalez (11), Loewenstein, Weber, Hsee, and Welch (12), Kahneman and Tversky (13); for more comprehensive surveys of this literature, see Olson (14), Hertwig and Erev (15), Prietzel (16).

Materials and Methods

We design a game in which two participants interact with one another and need to make risky decisions. Our specific goal is to determine if people perceive risks to genetic data privacy differently from risks to financial data privacy. We achieve this goal by conducting a laboratory experiment where subjects play this game with one another.[‖] In each experiment session, we first divide subjects into pairs randomly and anonymously – that is, subjects do not get to choose their partner nor ever find out their partner's identity.[**] Next, each subject pair is randomly assigned to one of two frames, the "genetic frame" or the "investment frame". Each subject in the genetic frame is asked to assume the role of a patient or a physician, and each subject in the investment frame is asked to assume the role of an investor or a money manager.

The underlying decision problem – in terms of monetary payoffs, risks, and choices – is the same for all subjects; the only difference is what the numbers are meant to represent. In the genetic frame, the decision problem is said to represent a situation where a patient needs to decide whether or not to get a genetic test. In the investment frame, the decision problem is said to represent a situation where an investor needs to decide whether or not to make a risky investment. For further clarity in this discussion, subjects who are assigned the role of a patient are the ones who need to decide whether to undergo genetic testing, and subjects who are assigned the role of an investor need to decide whether to invest money in a risky asset. We pair patients with "physicians" and investors with "money managers". Thus, in the genetic frame we have pairs of patients and physicians and in the investment frame, we have pairs of investors and money managers. Experiment subjects are randomly assigned one of these four roles, which they keep throughout the experiment.

For illustration of the underlying game, it will be useful to speak of Player 1, who could be either a patient or an investor, and Player 2, who could be either a physician or a money manager. Our experiment design consists of four treatments that allow us to decompose the various motivations that individuals may have to get a genetic test. We hypothesize that individuals choose to get a genetic test (which is represented by spending money in the experiment) for three reasons: (i) They get utility from any potential health benefits that may result from the genetic test; (ii) they want to benefit their providers/institutions by contributing to research; and (iii) they trust their provider (the physician or institute where they take the genetic test) to protect their genetic data from getting stolen or misused. All subjects participate in four versions, or treatments, of the game. We call these "base treatment", "altruism treatment", "trust and reciprocity treatment", and "Player 2's altruism treatment". These treatments are explained below.

Base Treatment. In a base treatment, Player 1 is endowed with \$6.00 and has to choose whether to "invest" or "not invest" \$2 out of this \$6 endowment. If Player 1 invests, she can win \$60 with a 5% chance but also lose her remaining \$4 with a 25% chance.[††]

‖ We avoid the use of the words like "game", "lottery", or "win" with experiment participants. The actual experiment instructions are provided in Appendix 1. Moreover, we eschew the use of any game theoretic equilibrium concepts such as Nash equilibrium because, as is now well understood, the motivations of individuals can be much more complicated than those reflected in monetary payoffs and thus such equilibrium concepts do not well reflect behavior except, perhaps, in the most stylized of situations.

** Subjects interact with their partners through computer terminals and while they can see the choices made by the other person, they do not know the identity of the person. However, they know that it is another participant, just like themselves, and is currently sitting in the same room as them and using one of the 30 available computers in the room.

†† Patients are told that the \$60 represents potential health benefits and the loss of \$4 represents the loss of privacy of their genetic data. Investors are told that \$60 represents a monetary gain (i.e. a dollar earned merely represents a dollar earned) while \$4 represents loss of privacy of their financial data. If a person's genetic data gets hacked and misused, she could potentially lose her health insurance, and in the worst case, even her job and family ties, depending on the information revealed through the genetic test. To capture this extreme case, we represent the cost of hacking as the loss of all wealth. In order to get meaningful results, we deliberately give participants a higher risk of their data getting hacked than what it is in the real world. The fact that we do not use real-world numbers does not create a problem because our goal is not to capture the proportion of people who are willing to get a genetic test given a certain amount of risk. Instead, our objective is to capture the proportion of people who treat genetic data differently from financial data. Although we ask subjects to think about their financial data, we expect this to represent any "other" data. This is because the actual experiment payoffs are in monetary terms and using financial benefits and financial risks is a natural representation of risks and benefits in that setting.

There is no Player 2 (i.e. physician/money manager) in the base treatment, which means that Player 1 considers only her own benefit in making this decision. Thus, even if she does have other-regarding preferences – such as altruism or trust – they will not influence her decision.

Altruism Treatment. In any (real world) investment or genetic testing situation, the involvement of money managers and physicians makes the decision more complex.[‡‡] Moreover, a number of factors may influence the players, including altruism, trust, and reciprocity. In order to decompose these factors, we introduce another treatment called the Altruism treatment. In this treatment, Player 1's decision problem is the same as in the base treatment except for one difference: if Player 1 chooses to invest her $2, Player 2 (who is either a physician or a money manager) receives $4. This $4 represents either the physician's benefit from advancing his research or the money manager's benefit. Note that while Player 2 gets money, he does not have a decision to make. This treatment includes two motivations for Player 1 to choose to invest: (i) her potential benefit (which is also present in the base treatment) and (ii) her desire to benefit Player 2 or altruism. We hypothesize that if altruism is a motivation for Player 1, she should be more likely to invest in the altruism treatment than she is in the base treatment.

Trust and Reciprocity Treatment. The trust and reciprocity treatment experiment resembles a trust game (17). In such a game, Player 1 may send some amount of her money to Player 2. The experimenter then gives Player 2 some multiple of the money sent by Player 1, typically three times as much, so if Player 1 sends $5 to Player 2, for example, Player 2 receives $15. Player 2 can then send back some money to Player 1. Both players know the entire game. The interpretation given is that if Player 1 trusts Player 2 to send back at least as much money as Player 1 sent, then Player 1 will send some positive amount to Player 2. For the example, Player 1 may think that Player 2 will "share the wealth" so that both players end up with $10. In the trust and reciprocity treatment, Player 1 trusts Player 2 to invest in the security of Player 1's data (either their financial data or their genetic data, depending on the frame) and thus reduce the probability of a bad outcome for Player 1. Player 1 has funds that can be invested and also some wealth, which cannot be invested but may be lost. Our game involves treachery in that Player 2 may not invest any of the money invested by Player 1 in data security and Player 1 will not only loose the money invested but also their wealth.

The trust and reciprocity treatment is different from the altruism treatment in only one way: in this treatment, Player 2 (who is either a physician or money manager) may spend some of his $4 to reduce the risk of loss to Player 1 (who is a patient or investor). We give Player 2 three options: spend nothing ($0), spend half of his earnings ($2), or spend all of his earnings ($4). If Player 2 spends all of his $4, Player 1 does not experience any risk; if Player 2 spends $2 then Player 1 loses her wealth with a 15% chance; finally, if Player 2 does not spend any money, Player 1 loses her wealth with a 25% chance, which is the same chance as in the other treatments.[§§] Therefore, Player 1 now has three motivations to invest: (i) her own benefit (which is also present in the previous two treatments), (ii) her desire to benefit Player 2, aka her altruism (which is also present in the altruism treatment), and (iii) whether she trusts Player 2 to protect her wealth (which is not present in the the previous treatments). If trust is a motivation for Player 1, then Player 1 should be more likely to invest in the trust treatment than the altruism treatment. Player 2 has two motivations to spend money: (i) altruism: he wants to protect Player 1 from a loss; and (ii) reciprocity: he understands that his income is a result of Player 1's decision and wants to reciprocate by spending some money to help Player 1.

[‡‡] In actuality, Player 2 may represent medical centers, hospitals, brokerage firms, financial advisers, etc.

[§§] In the previous two treatments, Player 1's expected payoff was $6 from investing and $6 from not investing. Now Player 1's expected payoff depends on Player 2's decision. If Player 2 spends all of his $4, Player 1's expected payoff is $7 (her $4 plus 0.05 times $60). If Player 2 spends $2, Player 1's expected payoff is $6. If Player 2 does not spend any money, Player 1's expected payoff is $6, which equals Player 1's payoff if she did not invest. Therefore, if Player 1 is risk-neutral and looking only to maximize her expected payoff, she should invest. If, instead, Player 1 is risk-averse, she should invest only if she believes that Player 2 will spend a positive amount to protect her wealth. Her decision depends on her belief about Player 2 as well as on her degree of risk aversion. For example, if she is very risk averse, she might find not investing to be utility-maximizing even if she believes that Player 2 will spend all of his $4. By contrast, if she is only slightly risk averse, she might find investing to be utility-maximizing even if she believes that Player 2 will spend $2.

Player 2's Altruism Treatment. The purpose of this treatment is to decompose the different motivations for Player 2 (i.e. physicians and money managers) to spend money for the security of Player 1's wealth. This treatment is the same as the trust and reciprocity treatment, except for one change: Player 2's earning of $4 is not a consequence of Player 1's decision to invest. Instead, we endow Player 2 with $4 regardless of Player 1's choice. If Player 1 makes the investment, Player 2 is told that Player 1's personal data, and hence wealth, is at risk and needs to decide whether to spend some money to protect it. Thus, Player 2 only has one motivation to spend money to protect Player 1's wealth: altruism. We hypothesize that if reciprocity is a motivation for Player 2, he should be more likely to invest in the trust and reciprocity treatment than Player 2's altruism treatment.

Implementation. We conducted a laboratory experiment at Vanderbilt University with a total of 162 subjects who were all undergraduate students. After random assignment of frames, we ended up with 78 subjects (39 pairs) in the genetic frame and 84 subjects (42 pairs) in the investment frame. All subjects were paid a participation fee of $5, and any earnings from the experiment were in addition to this. The average payment made to subjects was about $20. Each subject spent about 40 minutes on average (including all administrative tasks such as introduction, explanation of tasks, and payments) to participate in an experiment session.

Table 2 provides the summary statistics of the subject population and a test for whether assignment to the genetic and investment groups is random. The results support perfect randomization as there is no statistically significant difference across the two groups in terms of the following: average age, proportion of individuals who have been a victim of identity theft, proportion of females, proportion who are white, and how religious the subjects are on average. The variable *real-life investment* is statistically different across the two groups but that is because this question is different for each group. For the genetic group, the question asks whether the respondent has got a genetic test in real life; for the investment group, the question asks if respondent has made any investments in real life.

Each subject participated in multiple versions, or treatments, of the experiment. In order to avoid any effect of priming or preferences for consistency, these treatments were presented to them in different orders. We do not find any evidence of order effects (see Appendix B). Results from OLS regression and probit regression were similar. In both regressions, only the treatment effect variable, namely genetic, was significant. We also conducted Lasso and step-wise regressions to identify any significant variables with a 20% significance cutoff, both of which confirmed that none demographic control variables are significant.

Lastly, we replicated this experiment using the online platform Amazon Mechanical Turk. The results of the replication, which are very similar to the results of the lab experiment discussed above, are shown in Appendix B. We do not pool the experiment results because the subject populations and experiment environments are obviously different. Instead, we treat this replication as a gateway to consider online surveys as substitutes for laboratory experiments in universities.

ACKNOWLEDGMENTS. We thank the directors, faculty, researchers, and external advisory board of the Center for Genetic Privacy and Identity in Community Settings (GetPreCiSe) for their helpful feedback and discussion during the development of this manuscript. We also thank the NIH for its generous support of this project through grant # 4-50-713-7813 and the DASPY Grant 4-22-420-0073 FF_280102, Task 30, and also the Douglas Grey Fund for Research in Economics.

Table 1. Regression Results

Variables	OLS	Probit
Genetic Framing	0.254**	0.675**
	(0.125)	(0.325)
Failed control questions	0.0243	0.0768
	(0.124)	(0.327)
Age	-0.0389	-0.111
	(0.0403)	(0.108)
Victim of ID theft	0.174	0.493
	(0.298)	(0.828)
% Female	0.0522	0.123
	(0.123)	(0.321)
% White	0.169	0.466
	(0.127)	(0.335)
% Hispanic	-0.216	-0.584
	(0.280)	(0.749)
Religiousness (0-3 scale)	-0.0101	-0.0302
	(0.0546)	(0.144)
Constant	1.089	1.716
	(0.801)	(2.138)
Observations	73	73

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2. Summary Statistics

Variable	Investment Frame	Genetic Frame	Difference
Age	19.71***	20.03***	-0.326
	(0.157)	(0.301)	(0.314)
Victim of ID theft	0.0759**	0.0290	0.0330
	(0.0300)	(0.0203)	(0.0361)
Real-life Investment/Genetic Test	0.519***	0.116***	0.409***
	(0.0566)	(0.0388)	(0.0660)
% Female	0.506***	0.406***	0.133*
	(0.0566)	(0.0595)	(0.0783)
% White	0.544***	0.580***	-0.00366
	(0.0564)	(0.0599)	(0.0787)
% Hispanic	0.101***	0.130***	-0.0304
	(0.0342)	(0.0408)	(0.0525)
Religiousness (0-3 scale)	1.329***	1.304***	-0.0604
	(0.115)	(0.126)	(0.163)

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes*:* This table reports means and standard deviations (in parenthesis) in the total sample and in frames Genetic and Investment. The last column displays p-values for the null hypothesis of perfect randomization. These are p-values from two-sample t-tests (we get the same conclusions with χ^2 test and Mann-Whitney test). *Age* is the individual's age in years. The variables *Victim*, *Female*, and *White* are dummy variables indicating female subjects, victim of identity theft or fraud, and subjects who classify their race as white. The variable *Real-life Investment* is also a dummy variable, but represents different things for the two frames. For the investment frame, it represents whether the subject has made personal investments in real life; for the genetic framing, it represents whether the subject has got a genetic test in real life. The variable *Religious* measures how religious the individual is and ranges from 0 to 3, with 0 indicating not religious at all and 3 indicating very religious.

1. Saskia C Sanderson, Kyle B Brothers, Nathaniel D Mercaldo, Ellen Wright Clayton, Armand H Matheny Antommaria, Sharon A Aufox, Murray H Brilliant, Diego Campos, David S Carrell, John Connolly, et al. Public attitudes toward consent and data sharing in biobank research: a large multi-site experimental survey in the us. *The American Journal of Human Genetics*, 100(3):414–427, 2017.
2. David Kaufman, Juli Murphy, Lori Erby, Kathy Hudson, and Joan Scott. Veterans' attitudes regarding a database for genomic research. *Genetics in Medicine*, 11(5):329, 2009.
3. Celeste M Condit, Lijiang Shen, Karen L Edwards, Deborah J Bowen, Diane M Korngiebel, and Catherine O Johnson. Participants' role expectations in genetics research and re-consent: Revising the theory and methods of mental models research relating to roles. *Journal of health communication*, 21(sup2):16–24, 2016.
4. D Duquette, C Langbo, J Bach, and M Kleyn. Michigan biotrust for health: public support for using residual dried blood spot samples for health research. *Public Health Genomics*, 15 (3-4):146–155, 2012.
5. Emily A Rauscher, Colin Hesse, Stacey Miller, Whitney Ford, and Erin L Youngs. Privacy and family communication about genetic cancer risk: Investigating factors promoting women's disclosure decisions. *Journal of Family Communication*, 15(4):368–386, 2015.
6. Jeffrey R Botkin, Erin Rothwell, Rebecca Anderson, Louisa Stark, Aaron Goldenberg, Michelle Lewis, Matthew Burbank, and Bob Wong. Public attitudes regarding the use of residual newborn screening specimens for research. *Pediatrics*, 129(2):231, 2012.
7. Karen L Edwards, Amy A Lemke, Susan B Trinidad, Susan M Lewis, Helene Starks, Katherine W Snapinn, Mary Quinn Griffin, Georgia L Wiesner, and Wylie Burke. Genetics researchers' and irb professionals' attitudes toward genetic research review: a comparative analysis. *Genetics in medicine*, 14(2):236, 2012.
8. Amy A Lemke, Wendy A Wolf, Jennifer Hebert-Beirne, and Maureen E Smith. Public and biobank participant attitudes toward genetic research participation and data sharing. *Public Health Genomics*, 13(6):368–377, 2010.
9. Ellen W Clayton, Colin M Halverson, Nila A Sathe, and Bradley A Malin. A systematic literature review of individuals' perspectives on privacy and genetic information in the united states. *PloS one*, 13(10):e0204417, 2018.
10. Susanne Barth and Menno DT De Jong. The privacy paradox - investigating discrepancies between expressed privacy concerns and actual online behavior - a systematic literature review. *Telematics and Informatics*, 34(7):1038–1058, 2017.
11. George Wu and Richard Gonzalez. Curvature of the probability weighting function. *Management science*, 42(12):1676–1690, 1996.
12. George F Loewenstein, Elke U Weber, Christopher K Hsee, and Ned Welch. Risk as feelings. *Psychological bulletin*, 127(2):267, 2001.
13. Daniel Kahneman and Amos Tversky. Prospect theory: An analysis of decision under risk. In *Handbook of the fundamentals of financial decision making: Part I*, pages 99–127. World

Scientific, 2013.
14. Kenneth R Olson. A literature review of social mood. *The Journal of Behavioral Finance*, 7 (4):193–203, 2006.
15. Ralph Hertwig and Ido Erev. The description–experience gap in risky choice. *Trends in cognitive sciences*, 13(12):517–523, 2009.
16. Tobias Thomas Prietzel. The effect of emotion on risky decision making in the context of prospect theory: a comprehensive literature review. *Management Review Quarterly*, pages 1–41, 2019.
17. Joyce Berg, John Dickhaut, and Kevin McCabe. Trust, reciprocity, and social history. *Games and economic behavior*, 10(1):122–142, 1995.