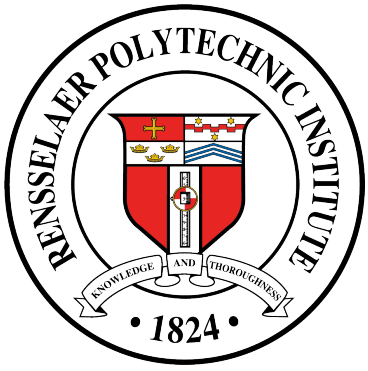
Pricing Privacy

An Experimental Design

Max Troeger

ECON 6340 - Behavioral Financial Economics
Dr. Billur Aksoy



Department of Economics
Rensselaer Polytechnic Institute
Troy, NY, USA
28 Apr. 2025

Contents

1 Introduction & Motivation

Breaches of privacy are costly. In the 2010s, Rao and Reiley (2012) estimate that unrestricted access to email addresses led to an annual productivity loss of approximately $20 billion from spam alone. Pricing privacy *itself*, however, is not trivial (Regner and Riener, 2017). The valuation of privacy has important legal and policy implications as, for example, whether or not the costs of enforcing HIPAA outweigh the benefits depends on individual valuations of privacy (Alessandro Acquisti and Loewenstein, 2013).

Nevertheless, Alessandro Acquisti and Loewenstein (2013) demonstrate that privacy valuations, far from being impossible to determine, "are affected not only by endowment [...] but also by the order in which different privacy options are described," and that subjects in their experimental design were five times more likely to reject offers for their data *if they believed their data would otherwise be protected* than if they believed otherwise.

According to Malgieri and Custers (2018), "personal data in the modern digital economy can be used, instead of money, to pay for digital content." In the era of Big Data, Zuiderveen Borgesius and Poort (2017) write that privacy is no longer about qualitative assessments of the value of privacy (i.e., how having your privacy taken away feels) but rather the quantitative assessment of what access to your private data brings in as revenue to data aggregating firms. From a policy perspective, informing people of the value of their data therefore gives them a greater understanding of the power they hold in digital markets as the potential for online firms to price discriminate based on private spending data is now possible. A monetary assessment of the value of privacy is critical now more than ever.

In the following experimental design we contribute a method of eliciting the value of privacy and provide a framework for piecing out the heterogeneous demographic effects of social preferences on this evaluation. The data collected in our experiment would also indicate preferences regarding digital privacy in general.

2 Research Questions

Our literature review motivates the following research questions:

1. How much do people value their privacy monetarily?

2. Are social preferences present in people's "undervaluing" or "overvaluing" of their privacy?

3 Initial Hypotheses

Moreover, we test the following hypotheses:

H_1 People have a significant, positive valuation of privacy.

H_2 There is a strong endowment effect in privacy valuations.

H_3 Social preferences have no impact on privacy cost evaluations.

4 Research Methodology

We generally follow the social norms elicitation experimental design outlined by Krupka and Weber (2013) and the peer informational nudge performed by Bursztyn et al. (2020). We illustrate our experimental design in Figure 1.

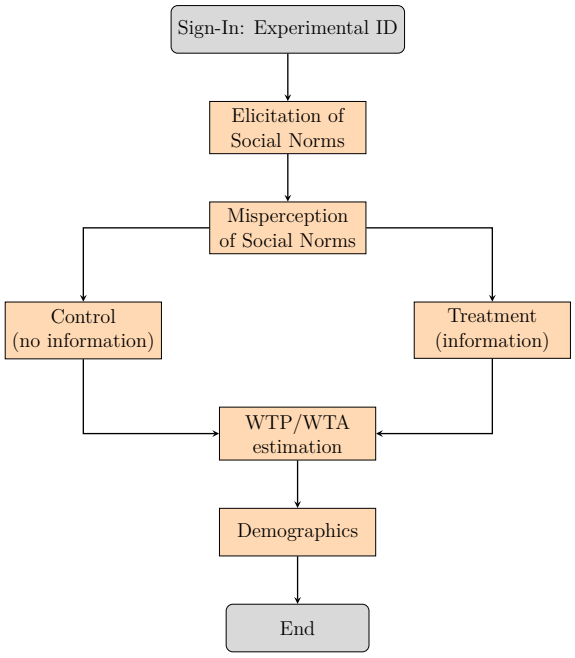


Figure 1: Experimental design.

4.1 Participants

We will draw our sample from college campuses by sending out digital and mail communications to students. Although college students generally have higher tech literacy than older generations—particularly their parents—the choice of college students in randomized experiments is common despite external validity concerns (Dincer, 2012; Peterson, 2001; Svahnberg et al., 2008). Randomization of students into our control and treatment groups will assuage most of these.

4.2 Procedure

As subjects enter the testing center, we randomly assign them to a *control* or *treatment* group. We do this by handing study participants a unique, 5 character experimental ID made up of the last 4 digits of their phone number and a letter A if the subject is in the control group or a B if in the treatment group.

For example, an arbitrary participant in the treatment group with a randomly generated phone number of 716-877-2383 would have an experimental ID of "2383B". Assigning experimental IDs in this way preserves subject anonymity (Bursztyn et al., 2020).

Following the procedure of Krupka and Weber (2013), we now pose a series of survey questions where we ask subjects to rank the "social appropriateness" of a series of statements. In particular, participants may choose "very socially inappropriate", "somewhat socially inappropriate", "somewhat socially appropriate", to "very socially appropriate". We present the following instructions:

> On the following screens, you will be presented with a series of situations A–D in which an "Individual A" must make a decision. Each situation will have a brief description and a list of the possible decisions available to "Individual A".
>
> After you read the description of the situation, you will be asked to evaluate the different possible choices available to "Individual A" and to decide, for each possible action, whether taking that action would either be, "socially appropriate" and "consistent with moral or proper social behavior" or "socially inappropriate" and "inconsistent with moral or proper social behavior."
>
> **By socially appropriate, we mean behavior that most people agree is the "correct" or "ethical" thing to do**; moreover, if "Individual A" selected a socially inappropriate choice, then someone else might be angry at "Individual A" for doing so.
>
> Please answer as completely as possible in each of your responses, based on your opinions of what constitutes socially appropriate or socially inappropriate behavior.

A situation as described above would, for example, look like the following:

Situation A

Individual A accesses the internet with a web browser and has the option of using an ad-blocker, a piece of software which removes advertisements (ads) from websites.

The websites Individual A visits will not be able to know if any ad-blocker was used.

Please evaluate the social acceptability of the following ad-blocker use decisions Individual A can make:

1. To block targeted marketing ads on shopping websites
2. To block ads on small-scale, advertising-funded video content
3. To block ads on major news sites
4. To remove ads that disrupt website accessibility

This task is made incentive compatible by selecting one of the situations at random. In the selected situation, one of the presented decisions will also be chosen at random. From this selected situation-decision pair, we will determine which evaluation was most commonly chosen by the other experimental subjects. If subjects give the same response as that most frequently given by other people, they will receive \$10. This social appropriateness elicitation task is a *coordination game* aimed at identifying the social norms of the subject pool (Krupka and Weber, 2013). Presenting subjects with a series of these social norm coordination games also serves to "jog" their memory about social norms, but it is imperative that the statements we present subjects in these games not be too specific to privacy concerns so that we may limit priming effects (Bursztyn et al., 2020).

After all subjects have completed the social norm elicitation task, we move on to a group estimation task. Subjects are presented with the following instructions:

On each of the following pages, you will be presented with a statement.

First, you will be asked whether you personally agree with the statement.

You will then be asked to guess how many of the other n participants in the room agree with the statement.

Payment: Participants whose estimate of the number of other participants that agree with the statement within a ± 1 range of the true number will receive \$10. *For example, if your guess is 10 and the true number is in the range 9–11, you will receive \$10.*

Following Bursztyn et al. (2020) we recommend presenting subjects with two such statements. As an illustration, for one statement subjects would be presented with the following prompts (where n is the number of participants):

Do you agree with the following statement?
In my opinion, digital privacy is important and people's data should not be freely accessible to companies.

☐ Yes

☐ No

If you had to guess, how many people among the other n study participants in the room do you think agree with the following statement?
In my opinion, digital privacy is important and people's data should not be freely accessible to companies.

<Enter a positive whole number>

Because of the strong incentive structure, we be-

lieve that subjects would give sufficiently complete answers.

We now split the subjects into the control and treatment groups. Participants in the control group would be moved, in silence, to a separate room or testing facility. Subjects in the treatment group would remain in the original setting of the experiment. To emphasize, it is imperative that—for the coming informational nudge—that the treatment and control groups remain separated for the remainder of the experiment. Any intermixing of the two groups introduces the possibility of spillover effects that would diminish the treatment effect.

Once the treatment and control groups are separated we implement our experimental treatment: we present subjects in the treatment group to an information nudge. Namely, we show the number and percentage of experimental participants that answered "Yes" and "No" to the group estimation task questions. We anticipate that, if subject estimates of group opinion deviate systematically from the sample proportion, the direction of under or over-estimation of the number of experiment participants who agree will materially affect answers to the coming willingness-to-pay (WTP) and willingness-to-accept (WTA) questions.

Keeping the groups segregated, we now ask a series of WTP and WTA questions regarding privacy. Subjects receive the following instructions:

> In this section you will be instructed to assess a series of statements.
>
> You will be asked to estimate how much your peers' would be willing to *spend or receive* on a scale of \$0–\$100 *per month* for the following data privacy privileges.
>
> **Payment**: *for each response that is within 10% of the average group response, you will receive \$2.*

For example, a pair of WTP and WTA statements would look like the following:

> 1. How much money per month would a company *have to pay* your peers for them to be willing *to allow* a company access to data on their shopping preferences?
>
>   <Enter a number between \$0 and \$100>
>
> 2. How much money per month would your peers be *be willing to spend* in order *to prevent* a company from accessing data on their shopping preferences?
>
>   <Enter a number between \$0 and \$100>

We conclude the experiment by collecting demographic information. We do this at the end of the experiment to reduce the effect of "stereotype threat" in responses because, for example, we anticipate that computer science majors care considerably more about privacy as a group than economics majors (Fernandez et al., 2016; Ziegenfuss et al., 2021):

> **Exit Survey**
>
> What is your gender?
>
> ☐ Female
>
> ☐ Male
>
> ☐ Other <Write in>
>
> What is your age?
> <Enter a positive whole number>
>
> What is your college major?
> <Enter a string>
>
> What is your college graduation year?
> <Enter a positive whole number>

5 Discussion & Concluding Remarks

We believe that the direction of the information nudge in the treatment will be consistent with difference in the mean WTP and WTA responses between the treatment and control groups. As indicated by the literature in privacy pricing, we expect to see an endowment effect (i.e., that people are willing to spend less to preserve privacy than they would to receive to break it) in the form of consistently higher WTA responses than WTP responses. If social preferences do affect WTA and WTP evaluations, we anticipate that these responses will deviate systematically between the treatment and control group. We expect that, for an information nudge which indicates that the participant sample cares more about privacy than expected, we will observe substantially lower WTP numbers and substantially higher WTA numbers compared to the control. Conversely, if the sample cares less about privacy than expected, we expect to see lower WTP *and* WTA estimations compared to the control. The collected demographic information would allow further study into the heterogeneous effect of the treatment.

For further research, we recommend conducting this experiment on older generations. In particular, we suggest using a pool of employed baby boomers because it is not immediately clear if the observed lower tech literacy of baby boomers (which negatively affects privacy cost evaluations) overpowers their generally higher precaution (which would positively affect privacy) on the internet (Obal and Kunz, 2013).

References

Alessandro Acquisti, L. K. J., & Loewenstein, G. (2013). What is privacy worth? *The Journal of Legal Studies*, (249). https://advance.lexis.com/api/document?collection=analytical-materials%5C&id=urn%5C%3acontentItem%5C%3a598N-VXD0-00KD-K0FJ-00000-00%5C&context=1519360%5C&identityprofileid=R6FPSM51533.

Bursztyn, L., González, A. L., & Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in saudi arabia. *American Economic Review*, *110*(10), 2997–3029. https://doi.org/10.1257/aer.20180975

Dincer, S. (2012). A study of the relationship between pupils and parents' computer literacy level and use. *Procedia - Social and Behavioral Sciences*, *46*, 484–489. https://doi.org/10.1016/j.sbspro.2012.05.146

Fernandez, T., Godwin, A., Doyle, J., Verdin, D., Boone, H., Kirn, A., Benson, L., & Potvin, G. (2016). More comprehensive and inclusive approaches to demographic data collection. *2016 ASEE Annual Conference & Exposition Proceedings*. https://doi.org/10.18260/p.25751

Krupka, E. L., & Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, *11*(3), 495–524.

Malgieri, G., & Custers, B. (2018). Pricing privacy – the right to know the value of your personal data. *Computer Law & Security Review*, *34*(2), 289–303. https://doi.org/10.1016/j.clsr.2017.08.006

Obal, M., & Kunz, W. (2013). Trust development in e-services: A cohort analysis of millennials and baby boomers. *Journal of service management*, *24*(1), 45–63.

Peterson, R. A. (2001). On the use of college students in social science research: Insights from a second-order meta-analysis. *Journal of Consumer Research*, *28*(3), 450–461. https://doi.org/10.1086/323732

Rao, J. M., & Reiley, D. H. (2012). The economics of spam. *Journal of Economic Perspectives*, *26*(3), 87–110. https://doi.org/10.1257/jep.26.3.87

Regner, T., & Riener, G. (2017). Privacy is precious: On the attempt to lift anonymity on the internet to increase revenue. *Journal of Eco-*

nomics & Management Strategy, *26*(2), 318–336. https://doi.org/10.1111/jems.12192

Svahnberg, M., Aurum, A., & Wohlin, C. (2008). Using students as subjects-an empirical evaluation. *Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement*, 288–290.

Ziegenfuss, J. Y., Easterday, C. A., Dinh, J. M., JaKa, M. M., Kottke, T. E., & Canterbury, M. (2021). Impact of demographic survey questions on response rate and measurement: A randomized experiment. *Survey Practice*, *14*(1), 1–11. https://doi.org/10.29115/sp-2021-0010

Zuiderveen Borgesius, F., & Poort, J. (2017). Online price discrimination and eu data privacy law. *Journal of Consumer Policy*, *40*(3), 347–366. https://doi.org/10.1007/s10603-017-9354-z