

CENTER FOR EUROPEAN ECONOMICS RESEARCH

RESEARCH PROPOSAL SUMMER SCHOOL REVEALED PREFERENCES

REVEALED PREFERENCES UNDER FRAMING: USERS VALUATION OF PRIVACY

Johannes Walter

Abstract

Diese Dokumentation enth"alt eine sortierte Liste der wichtigsten L^AT_EX-Befehle. Die einzelnen Listeneintr"age sind untereinander durch viele Querverweise verkettet, die ein Auffinden inhaltlich zusammengeh"origer Informationen erheblich erleichtern.

Summer School Revealed Preferences
Nicolai Kuminoff

July 11, 2020

1 Introduction

In stated preference studies people tend to express a high valuation for privacy, yet observed behavior is typically at odds with stated preferences. This phenomenon has come to known in the literature as privacy paradox. Some have argued that this observation is merely illusory, since one can state a high valuation of privacy in general, but in a specific situation a cost-benefit analysis might lead people to discount privacy concerns (Acquisti, Brandimarte, and Loewenstein 2015, p. 2). This refutation falls short insofar, as it is a well documented result in behavioral economic research that people’s decision making capabilities are only in part rational. Especially in situations when people are uncertain about the consequences of their actions, when they are unsure about their preferences or when they are under time pressure, people often search for clues in their surrounding to provide guidance (Acquisti, Brandimarte, and Loewenstein 2015, p. 3). One possible source of orientation can stem from the specific way in which a decision is framed. Deriving revealed preferences in the presence of framing can be problematic, if people’s preference is not consistent over the set of framings. Goldin and Reck (2020) propose a new method to recover consistent population preferences even when framing effects are present.

Their methodology rests on the insight that even though framing-consistent preferences cannot be observed on the *individual level* in typical datasets, on the *population level* the fraction of consistent decision-makers can be identified. From the subgroup of consistent decision-makers their approach extrapolates to the entire population. Their method is described in detail in section 2.

This research draft proposes two innovations. First, it is proposed to collect an online dataset recording people’s decisions about browser cookie settings via a type of data collection akin to the one in Levy (2020). Goldin and Reck (2020)’s method can subsequently be applied to the analysis of this dataset. The second innovation is methodological in nature. The dataset collected online will contain repeated observations of the same individual both under an ad-tracking cookie opt-in framing and an ad-tracking cookie opt-out framing. This additional information can be used to overcome the central limitation in most datasets. This means that the necessity for Goldin and Reck (2020)’s model vanishes. Instead, the subset of framing-consistent people can be observed on the individual level. Once these individuals are identified, one can again extrapolate to the entire population, as is done in Goldin and Reck (2020)’s method. Finally, the results from the first approach and the second approach can be compared, yielding an empirical test of Goldin and Reck (2020)’s method.

To understand the data GDPR introduced mandatory choice

What would ideal results of this research look like? If the proposed research succeeds it find whether Goldin and Reck (2020)’s new method generates estimates that are qualitatively similar and quantitatively close to estimates based on repeated observations of the same individual. By doing so, it will produce estimates of people’s privacy preferences which are robust with respect to framing. Here two scenarios are conceivable: Either it will be found that From a policy point of view, the latter result could possibly be interpreted to imply that the rules governing cookie default settings would have to be further tightened, such that websites default setting is to only use technically necessary cookies. Ad-tracking cookies would then mandatorily be opt-in.

Policy implications

This research lies in the intersection of three strands of literature, each of which it will attempt to contribute to. First, it will contribute to the literature on measuring preferences for and attitudes towards privacy. Both in the social and behavioral sciences There is a large body of literature that examines the nature of privacy preferences. This research will contribute by uncovering preferences for online privacy settings in particular. Second, by carrying Goldin and

Reck (2020)'s model into a new setting and confronting it with an empirical test, this research will contribute to the literature on revealed preferences under framing. And finally, there is a growing literature on the effects of the GDPR. The change in laws regarding cookie settings is what enables this research in the first place. The results of this study will have practical implications for the question whether even stricter privacy regulations seem justified.

mention the types of framing effects and that the relevant one here is a default type mention that their model does it with a 401k plan that doesn't allow for repeated observation. Our setting and method allows for repeated observation though

This causes a problem for revealed pref analy if decisions depend on framing paper by dings in context of digital bums what is their overall idea: How I'm gonna use it data I need and how to get contribution to the literature

methodological innovation with my idea

Idea:

- Setting: GDPR introduced choice on cookies. Cookies choice depends on framing. This new model in the JPE can factor in framing. We can then derive people's revealed preference for privacy
- Data: Collect browsing behavior from random sample. Requirement: Browser add-on. I can ref
- Create appendix chapter with a table detailing the different possible browser cookie settings and images of a couple of cookie setting screenshots.

Literatures touched: revealed preferences under framing, privacy evaluation, effects of GDPR. Previous literature suggests that there must be framing effects -j Madrian and Shea Literature also belongs to how does the GDPR affect people and firms? Policy Implications: GDPR wants to privacy protect citizens. If design fails, implication is that all sites should default to only technically necessary What would ideal results look like?

Acquisti, Brandimarte, and Loewenstein 2015 mention online example even though they go through with the 401k plan example

2 Economic Model

The description of this model follows chapter one of Acquisti, Brandimarte, and Loewenstein (2015). A fully detailed description of the model can be found there and is beyond the scope of this research proposal. Here, only the main concepts necessary for understanding this proposal are presented.

A decision maker i chooses from a binary decision set $\mathbf{S} = \{0, 1\}$ and two possible frames $D_i \in \{0, 1\}$. In the context of this research proposal, the decision coded with a 0 could refer to an individual's decision to accept only technically necessary cookies and the decision coded with a 1 could refer to an individual's decision to allow for non technically necessary cookies. Frame 0 could be the situation where the default setting is such that technically necessary cookies are the only pre-chosen ones and the user would have to actively engage in clicking on all non-technically necessary cookies she wishes to allow. Frame 1 could then refer to the situation where in the default setting both types of cookies are pre-chosen and the user accepts both types with one click. We can adapt a notation that is akin to what many researchers like to use in a potential outcome setting. $Y_i(0)$ and $Y_i(1)$ denotes then the decision individual i makes under frame $D_i = 0$ and $D_i = 1$, respectively. Decision makers are assumed to have strict ordinal preferences over the set of available options. $Y_i^* \in \{0, 1\}$ denotes the most preferred option. Each decision maker is characterized by a vector of random variables

$(Y_i(0), Y_i(1), D_i, Y_i^*)$, which are drawn from an underlying population distribution. For each i , the researcher observes the pair (Y_i, D_i) , where $Y_i = Y_i(0)D_i + Y_i(1)(1 - D_i)$. Goldin and Reck (2020) assume for their model that the researcher does not observe Y_i^* and only observes one of $Y_i(0)$ and $Y_i(1)$, depending on the frame D_i .

The data that is suggested to be collected in this proposal deviates in a crucial way from this assumption. In contrast to the datasets Goldin and Reck (2020) have in mind, the dataset in this proposal will contain *repeated* observations of the same individual. The number of observations for each i can be denoted with a subscript $k \in \{1, \dots, K\}$ such that $Y_{i,k}$ is observed. Importantly, here it is assumed that K is sufficiently large to allow for at least one observation of each framing $D_i \in \{0, 1\}$.

Goldin and Reck (2020) continue their description of the model as follows: Each decision maker can choose either consistently, i.e. the same choice under each frame, or choose in a way that is responsive to the frame. Consistency is denoted by $C_i = 1 \text{ iff } Y_i(0) = Y_i(1) \text{ else } C_i = 0$. Again, they assume that each i is observed only under one frame, such that C_i is not observed. In the context of this proposal, as described above, it is assumed that each i is observed under both frames, such that C_i is identified. This difference is crucial for one of the main contributions of this proposal. The additional information in this dataset allows to calculate two sets of results. For the first one, the additional information in this dataset is not used. The analysis will proceed under the assumption of Goldin and Reck (2020). The second set of results will be obtained using the full amount of information in the dataset. As such, the "ground-truth" that remains hidden under Goldin and Reck (2020)'s assumptions, is observed. In other words, with the dataset proposed here, it will be able to identify all individuals who make consistent choices, independent of the frame. Having these two sets of results, one can compare the estimates from both to arrive at an assessment of the quality of Goldin and Reck (2020)'s model accuracy.

Four +1 assumptions real quick

- Assumption A1 (Frame separability) For all i , Y_i^* does not depend on D .

3 Data and Methods

Subsection 1) Data collection describe how he did it and what kind of data you need to collect (i.e. recognize when asked for cookies and what the default cookie setting is)

Also describe a little bit how cookies actually work

What kind of data is required?

- Data needs to have an opt-in and an opt-out possible framing. This should be given in the cookie context:
 - What does the GDPR say exactly?
 - Either it says all cookies always must be disabled by default ...
 - Or it varies only by how the website presents the choice. What are the options?
- I need to observe whether individual i chooses only the technically necessary cookies or all the cookies
- I also need to observe how the choice is presented:

A) Data Komplettes Deaktivieren der Cookies in Firefox oder Cookies deaktivieren für nutzungsbasierte Online-Werbung <https://www.bild.de/wa/ll/bild-de/privater-modus-unangemeldet-54578900.bild.html>

By using our site, you acknowledge that you have read and understand our Cookie Policy, Privacy Policy, and our Terms of Service <https://tex.stackexchange.com/questions/823/remove-ugly-borders-around-clickable-cross-references-and-hyperlinks>

I observe whether individual i accepts all cookies or only the technically necessary ones (Y_i), and whether the default is opt-in ($D_i = 0$) or opt-out ($D_i = 1$) at the date of visiting the website. Through an additional survey, we also observe age, sex and race for each employee.

B) Recovery of Consistent Preferences

- Under assumptions A1 - A4, proposition 1 allows us to identify the preferences of the consistent visitors
- A1) requires that the preferences over the cookie choice do not depend on opt-in or opt-out.
- A2) Frame exogeneity

Under A1 to A4), we identify the preferences of the consistent decision makers. What you get is the preference of the consistent decision-makers.

C) Recovery of the Population Preferences

Subsection 2) Identification and identification of population preferences

Explain how the original model arrives at estimates and how it extrapolates. If this takes too long and is too difficult, just make it top level and refer to the paper for details

Cookies:

Technically necessary cookies	Technically non-necessary cookies
<ul style="list-style-type: none"> • I 	<ul style="list-style-type: none"> • Sample Text
<ul style="list-style-type: none"> • Sample Text 	<ul style="list-style-type: none"> • Sample Text
<ul style="list-style-type: none"> • Sample Text 	<ul style="list-style-type: none"> • Sample Text

References

- Acquisti, Alessandro, Laura Brandimarte, and George Loewenstein (2015). "Privacy and human behavior in the age of information". In: *Science* 347.6221, pp. 509–514.
- Goldin, Jacob and Daniel Reck (2020). "Revealed-Preference Analysis with Framing Effects". In: *Journal of Political Economy* 128.7, pp. 000–000.
- Levy, Ro'ee (2020). *Social Media, News Consumption, and Polarization: Evidence from a Field Experiment*. <https://www.roeelevy.com/research>. [Online; accessed 09-July-2020].