Economic essays on privacy, big data, and climate change

Dengler, Sebastian

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Economic Essays on Privacy, Big Data, and Climate Change

Sebastian Dengler

December 1, 2017

Economic Essays on Privacy, Big Data, and Climate Change

Proefschrift

ter verkrijging van de graad van doctor aan Tilburg University op gezag van de rector magnificus, prof. dr. E.H.L. Aarts, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op vrijdag 1 december 2017 om 14.00 uur door

Sebastian Dengler

geboren op 1 mei 1988 te Worms, Duitsland.

Promotiecommissie:

Promotor:	prof. dr. J.J.M. Potters
Copromotor:	dr. J.O. Prüfer
Overige Leden:	prof. dr. A. Acquisti
	prof. dr. E.E.C. van Damme
	dr. C. Schottmüller
	dr. F. Schütt

To Matthias, my brother, whom I admire more than he thinks.

Acknowledgements

I believe it to be just to first and foremost express my gratitude to my co-promotor Jens Prüfer, under whose supervision this dissertation took shape. Right from the start (literally the first day) of the Research Master program, then in your role as Education Coordinator for the Microeconomics track I had chosen, you emphasized a couple of things: the times ahead might not always be easy, but that the required hard work would eventually pay off and, last but definitely not least, that your door would always be open for advice and support. In each aspect the years that followed and ultimately lead to this dissertation have proven you right. It was not easy, especially in the first year, where you not only pointed out that the usual struggles that students experience were likely amplified by my interdisciplinary background. You also were reassuring that exactly this background may serve me well in the future. Putting in the work resulted not only in a finished Research Master thesis and this dissertation under your supervision, but also many experiences along the way and there is little doubt that my job placement provides supporting evidence for your early prediction about my interdisciplinarity. Most importantly, though, your guidance and support were truly extraordinary, often enough going above and beyond anything I could justifiably expect. Not only was your door always open, you also opened up others by suggesting conferences and workshops to attend, research groups or societies to join, coordinating letters of reference and the dissertation committee. You provided feedback even when I asked you (once again) to be part of a just in time production process and never tired to re-ignite my academic spirit by providing yet another spark when necessary. For all of the big and many more small positive effects you created in my development – not only professionally, but also as a person: Thank you for guiding the way and accompanying me on it!

Likewise, I would like to express my gratitude to my promotor and co-supervisor Jan Potters. Starting off your field course in Behavioral Economics with an in-class experiment of the beauty contest game was a lasting impression. I still recall that it raised the question in my mind how the presence of a single visiting student may have changed my classmates' and my choice in light of the fact that otherwise we could have all been sure that everyone in the room knew this game's theoretical Nash equilibrium solution. While it is probably an overstatement to trace back my use of the same game as part of this dissertation, your enthusiasm for and dedication to research in Behavioral and Experimental Economics profoundly reaffirmed my interest in putting such models to use. So I took a first shot at it in your course trying to combine altruistic and reference-dependent preferences “towards a model of promises among friends”. In the following time you not only encouraged and supported my application to the Behavioral Spring School in San Diego, but also agreed to join Jens as my co-supervisor for both my Research Master thesis and my dissertation. More than once you undertook the effort to view things from a more critically distant view and constructively advised me to perhaps approach things with a (slightly) different perspective than I originally envisioned; and more than once it took me quite some time to see the forest for the trees and realize the value of it. But I could not only rely on your experience and judgment to evaluate plans. When time was severely working to my disadvantage, I could likewise rely on you providing last minute support (and fixes). Thank you for all of it!

Further, I am grateful to Reyer Gerlagh, Gijs van de Kuilen, and Stefan Trautmann, who invited me to collaborate on a project, giving me a chance to broaden my scope both thematically and professionally. Being involved in a project with co-authors distinct from my supervisors provided me with a different perspective on multi-author collaboration and the benefits but also challenges that come along with it, especially when different academic schedules and geographical distance start to factor in. You did not only ask for my input, but also actively for my feedback on your inputs and I am thankful for the appreciation you had for concerns I raised despite me being the most junior member of the team.

Also, I would like to explicitly acknowledge all members of my dissertation committee: Alessandro Acquisti, Eric van Damme, Florian Schütt, and Christoph Schottmüller. I am deeply grateful for your service and feel honored that you invested such a considerable amount of time and effort into providing me with feedback and advice for improvements. Notably, you did so not only as part of the pre-defense and final approval procedure of the actual thesis, but also in the preceding years. The hours long in-depth discussion during the pre-defense brought about some substantial changes to my work and I hope to carry many of your advisory remarks forward to other future work.

I would like to thank all senior faculty who served as discussants or provided written feedback to earlier versions of chapters in this dissertation: Dirk Engelmann, Ricard Gil, Heiko Karle, Marian Moszoro, Valerio Poti, Katja Seim, Giancarlo Spagnolo, and Birger Wernerfelt.

I am thankful to the CentER for Economic Research at Tilburg University for the financial support that allowed me to conduct experimental sessions. Additionally, I would like to thank the George Mason University Law and Economics Center (LEC) for the financial support I received through their Privacy Scholars Fellowship.

Further, I am thankful to several not yet mentioned (former) faculty members of the Department of Economics and the Department of Econometrics and Operations Research, of which several are simultaneously also members of the Tilburg Law and Economics Center (TILEC) or the Experimental Economists group, for their advice and comments on my plans, proposals and papers. I especially benefited from conversations and discussions with Cédric Argenton, Johannes Binswanger, Otilia Boldea, Jan Boone, Elena Cettolin, Patricio Dalton, Sebastian Ebert, Lapo Filistrucchi, Eline van der Heijden, Tobias Klein, Boris van Leeuwen, Wieland Müller, Charles Noussair, Patricia Prüfer, Louis Raes, Martin Salm, Eli Spiegelman, Sigrid Suetens, Moritz Suppliet, Burak Uras, Gonzague Vannoorenberghe, and Bert Willems.

Having been a junior member of the Tilburg Law and Economics Center (TILEC), I have benefited greatly from the many work-in-progress meetings and seminars. Further, I am thankful to have had the chance to get a better understanding of the point of view of legal scholars and would like to especially thank Jan Broulik, Anna Butenko, Victoria Daskalova, Panagiotis Delimatsis, Zlatina Georgieva, Leigh Hancher, Martin Husovec, Olia Kanevskaia, Pierre Larouche, Saskia Lavrijssen, Anna Marhold, Wolf Sauter, Nicolo Zingales, and Gijsbert Zwart.

I also want to decidedly mention that I am grateful to many people in administrative rolese, including officers of the Economics Department, TILEC, and the Graduate School. I would especially like to thank Ank Habraken, Korine Bor, Cecile de Bruijn, and Corine Struis, who were welcoming whenever I came along with a question or request and could always be relied upon to find a quick and satisfactory solution to all parties involved.

Other than those endorsed by name, I would like to thank the organizers, scientific committees and participants of the following conferences and workshops in which I have presented individual chapters of this dissertation: the 13th Session of the European School on New Institutional Economics (ESNIE 2014), the 9th Competition Law and Economics European Network (CLEEN) Workshop, the 19th Annual Conference of the International Society for New Institutional Economics (ISNIE), the Research Roundtable for Economists on Law & Economics of Privacy and Data Security, the 2nd Workshop on Industrial Organization in the Digital Economy, the ENTER Jamboree 2016, the Program on Economics & Privacy: Digital Information Policy Scholars Conference 2016, the 14th ZEW Conference on The Economics of Information and Communication Technologies, the 20th Annual Conference of the Society for Institutional & Organizational Economics (SIOE), the Choice Lab PhD Workshop at Norwegian School of Economics (NHH), the Economic Science Association (ESA) European Meeting, and the 44th Conference of the European Assocation for Research in Industrial Economics (EARIE). I also thank Toulouse School of Economics and Télécom ParisTech for inviting me for a seminar presentation at their faculties.

Throughout the years, many encounters enriched my life in Tilburg and at Tilburg University, but Clemens Fiedler and Víctor González-Jiménez, my office mates each for two years, definitely stand out. I would like to thank you both for your collegiality, your tolerance to my often erratic working hours in our shared office, for your support in spirit as well as substance, for your friendship and countless hours we spent discussing topics ranging from trivia to fundamentally philosophical questions, inside and outside the office, during and after work hours. Among the many others (that I cannot possibly list all) I would like to especially thank Alaa Abi Morshed, Cansu Aslan, Ricardo Barahona, Elisabeth Beusch, Peter Brok, Roxana Fernández Machado, Lenka Fiala, Tomas Fiala, Bas van Heiningen, Dorothee Hillrichs, Mikael Homanen, Nick Huberts, Jan Kabátek, Michał Kobielarz, Maria Lavrutich, Mánuel László Mágó, Abhilash Maji, Emanuel Marcu, Gábor Neszveda, Mario Rothfelder, Gyula Seres, Vatsalya Srivastava, Chen Sun, Mina Vlachaki, Loes Verstegen, YiLong Xu, and Yuxin Yao for being my friends, classmates, intellectual sparring partners or even teachers, but also board game opponents, fellow cyclists of the Rolling Grad School, GSS board members, proofreaders, first-responders, drivers, movers (and movees), or collaborators in solving quests and beating monsters, closing gates, "lynching" werewolves or mafiosi, or advancing the demise of humanity by simple paper cards (or several of the above).

Also, I would like to thank my longstanding friends from places other than Tilburg. Thank you for staying in touch; for all the emails, messages and calls, the postcards, the Lebkuchen and other "survival packages" you sent, your visits here, my visits to you, and our joint travels to other places. It would have been a whole lot less fun and maybe even barely possible to arrive where I am today, if it were not for your continued friendship (although I occassionaly entertained the thought that I might have arrived at particular milestones slightly earlier otherwise). I am deeply grateful to know you and glad to call you my friends: Pierre Frotscher and Tzina Kokkala, Thomas Warbeck and Nancy Heidl, Michael Kirmes, and Christian König and Carina Schneider.

I would like to express my deep gratitude to Renata Rabovič, for so much more than "that conversation that one evening in the library" that turned out to change everything: for every "dirbk" and "mokykis", for showing me what some more discipline could help achieve (but also what too much of it looks like), for your ability to imprint macro models and metrics proofs onto my often reluctant mind by means of constant repetition, for visualizations of OLS and your Lagrangian alphas, somehow also for endless discussions about near-epoch dependence, for converting your studio into a healing station when I hit the ground (figuratively and literally), for every "-ukas" and "-agi" and boats called "Spain", for Vilnius, Caslano, Neos Marmaras, and Sankt Goar, or to put it simpler: for more than five years of your encouragement and support (and sometimes also supervision). Ačiū!

Finally, I would like to use this opportunity to thank my family for all these years and the many ways in which you have helped me to become the person I am today: for fostering my curiosity early on, for shaping my character such that I could dare to think for myself, for not holding on too tight when I started to go my own ways, for letting me know that I could nonetheless rely on you being there, for encouraging me to take the next step, for believing and trusting in me and for your loving support. Danke für alles!

Sebastian Dengler
Tilburg, October 2017

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This doctoral thesis, as apparent from its title *"Economic Essays on Privacy, Big Data, and Climate Change"*, aims to advance our understanding of major topics of concern in the 21st century using theoretical as well as empirical economic methodologies. All three topics do and will continue to affect people's lifes as they can substantially shape the functioning of our societies. The first and second essay are linked thematically and both focus on privacy choices and their consequences in the context of big data algorithms targeting individual consumers. The second and third essay, in contrast, are linked methodologically as both present results from economic laboratory experiments, where the former focuses on cognitive challenges faced by individual decision-makers and the latter on challenges to coordination and cooperation between decision-makers.

Chapter 2, *"Consumers Privacy Choices in the Era of Big Data"* (co-authored with Jens Prüfer), presents results from a theoretical model where consumers can purchase a single non-durable good from a monopolistic seller. Reflecting recent progress in information technologies, the model assumes that the seller can engage in first-degree (or perfect) price discrimination. As first-degree price discrimination can deprive consumers of all surplus from the transaction, they may want to protect their privacy and hide their willingness-to-pay from the seller. Therefore, consumers in our model face a trade-off between a direct sales channel without transaction costs and a privacy-protecting, but costly, anonymous channel. Because consumers have been shown to face cognitive limitations in the complex area of privacy choices, they are at a second disadvantage as compared to the seller. In our model this is reflected by the assumption that the seller, in addition to his ability to price discriminate, also outperforms consumers in terms of strategic sophistication. The main contribution of this chapter is to show under which conditions a costly privacy-protective sales channel is used even if consumers do not have an explicit taste for privacy. We demonstrate that this is the case if consumers are not too strategically sophisticated. Among those who use the anonymous sales channel, some suffer from net losses because prices turn out to be higher than expected, but others receive a net surplus. Thereby, we provide a micro-foundation for consumers' privacy choices as it is rational for some consumers to use the costly anonymization technique even without an exogenous taste for privacy.

Chapter 3, *"Predictive Algorithms and Consumer Behavior"* (single-authored), presents results from an economic laboratory experiment where subjects face a computerized seller individually. This experiment is an implementation of the model developed in Chapter 2. Hence, the computerized seller's algorithm implements perfect price discrimination unless subjects in the role of consumers incur a cost to hide (some) of their valuations. Despite some differences between the experimental implementation and the model which inspired it, the model's central proposition that a costly privacy-protective option is used, even when this implies an expected loss for consumers, receives empirical backing. Because cognitive constraints might play a role in the decision-making process of subjects, I do not restrict the analysis to Nash equilibrium predictions alone and investigate two possible alternative explanations. The choice of the first concept, *level-k thinking*, is motivated by the model in Chapter 2, whereas the second concept, *reinforcement learning*, stems from earlier experimental literature. While understanding how cognitive limitations affect behavior may help to improve economic models, it may be difficult to change. The cost of privacy protection, in turn, seems more actionable from a policy perspective. Therefore, the experiment includes a policy treatment where privacy protection is available for free. In addition to substantial deviations from Nash equilbrium predictions, I find some evidence that level-k thinking measures can explain subjects' hiding choices in the first period. Further, I find evidence for the occurrence of reinforcement learning suggesting that subjects react significantly to having realized losses in the previous period, but not to gains. Finally, the policy treatment leads to a strong increase in hiding behavior: among various Nash equilibria, the majority of subjects chooses the one with maximal hiding behavior. Thus, if consumers understand that information they share can be used to price discriminate against them (despite not exactly understanding how or why), they might simply decide to hide everything as soon as doing so comes for free. This provides an important caveat to policy recommendations. While reducing hiding cost may increase consumer surplus, it might also lead to too little information being disclosed such that efficiency gains made possible by predictive algorithms might get lost in the process.

Chapter 4, *"Climate Policy Commitment Devices"* (co-authored with Reyer Gerlagh, Gijs van de Kuilen, and Stefan Trautmann), presents results from an economic laboratory experiment where subjects are facing a game which mimics the global multi-generation planning problem for climate change and fossil fuel extraction. The contribution of this chapter is threefold. First, we develop a dynamic threshold public good game where players choose their actions sequentially, focusing on intergenerational trade-offs rather than international negotiations. Second, in the context of the intertemporal resource extraction dilemma, we show that a commitment device that reduces future resource demand can help to implement resource conservation. This holds even though the commitment device is costly, meaning that its use is inefficient – a waste of welfare – from a first-best planner's perspective. Yet, in the context of a strategic interaction between generations, it helps to improve the outcome compared to a context where this commitment device is not available. The third major contribution is that we connect subjects' behavior *across* conditions (within-rule choice) with their votes *for* a game condition (rule choice). Our findings suggest that successful cooperation not only needs to overcome a gap between individual incentives and public interests, but also a fundamental heterogeneity between subjects with respect to beliefs and preferences about the way in which this should be achieved.

Chapter 2

Consumers' Privacy Choices in the Era of Big Data

This chapter is based on the identically entitled working paper which is co-authored with Jens Prüfer

Abstract

Recent progress in information technologies provides sellers with detailed knowledge about consumers' preferences, approaching perfect price discrimination in the limit. We construct a model where consumers with less strategic sophistication than the seller's pricing algorithm face a trade-off when buying. They choose between a direct, transaction cost-free sales channel and a privacy-protecting, but costly, anonymous channel. We show that the anonymous channel is used even in the absence of an explicit taste for privacy if consumers are not *too* strategically sophisticated. This provides a micro-foundation for consumers' privacy choices. Some consumers benefit but others suffer from their anonymization.

2.1. Introduction

Two recent technological developments are revolutionizing seller-buyer transactions. First, aided by information and communication technologies (ICTs), sellers have the capability to analyze huge datasets with very detailed information about individual consumers' characteristics and preferences. Second, such data sets are increasingly available, owing to the fact that more economic and social transactions take place supported by ICTs, which easily and inexpensively store the information they produce or transmit.[1] These concurrent developments constitute the rise of big data (Mayer-Schönberger and Cukier 2013). They imply that sellers can make consumers ever more tailored contract offers, which fit their individual preferences or consumption patterns, approaching first-degree (or perfect) price discrimination, as the limit case.[2]

Because first-degree price discrimination can deprive consumers of all surplus from the transaction, they may want to protect their privacy and hide their willingness-to-pay from sellers with market power by employing anonymization techniques. But anonymization is costly: it can come at an explicit cost or at an opportunity cost.[3] Consumers are at a second disadvantage, compared to sellers, because they "will often be overwhelmed with the task of identifying possible outcomes related to privacy threats and means of protection. [...] Especially in the presence of complex, ramified consequences associated with the protection or release of personal information, our innate bounded rationality limits our ability to acquire, memorize and process all relevant information, and it makes

[1]Data analytics firms collect and analyze huge commercial databases on consumers, offering help to marketers. For instance, *Acxiom's* "database contains information about 500 million active consumers worldwide, with about 1,500 data points per person. That includes a majority of adults in the United States" (The New York Times 2012). Smartphone apps with millions of users, such as *Shopkick*, reward users for checking into stores, scanning products, visiting the dressing rooms, and so forth. *Amazon* recently was issued a patent on a novel *Method and System for Anticipatory Package Shipping* (Patent number US008615473 (December 24, 2013), http://pdfpiw.uspto.gov/.piw?docid=08615473). "So Amazon says it may box and ship products it expects customers in a specific area will want – based on previous orders and other factors – but haven't yet ordered" (Wall Street Journal Blog 2014).

[2]Such offers can be made directly, for instance, in online retailing, or indirectly, via selling advertisers access to highly preselected consumer groups. Einav and Levin (2013) provide a list of examples how firms, public administration, and researchers can exploit such novel technological opportunities.

[3]Consumers may need to pay for or install privacy-protective software, experience lower connection speed due to encrypted transmissions, or otherwise increase transaction cost (e.g. by shopping offline with additional costs if done by cash payments). For instance, using a non-tracking search engine decreases the precision of search and hence results in extra time or effort that a consumer must spend to find her preferred product (cf. the literature cited in Argenton and Prüfer (2012) documenting the effect from access to more search log data on the quality of search engines as perceived by users). Some (or a share) of these costs could be considered one-time investments (e.g. installing particular software), while others would persist even in a repeated interaction, (e.g. lower connection speed, even if only used for the final transaction and not for search). However, even the one-time investments could be necessary to be incurred repeatedly if both sides of the market were to engage in a "hide and seek" arms race of tracking and avoiding technologies (as discussed among others by Acquisti and Varian (2005)).

us rely on simplified mental models, approximate strategies, and heuristics" (Acquisti and Grossklags 2007, p.369).

In our model, we study the effects of perfect price discrimination on equilibrium choices and welfare when consumers' anonymization is possible but costly. We explicitly account for the discrepancy between cognitively challenged consumers and a seller whose strategic capabilities outperforms them and investigate how limited strategic sophistication affects equilibrium outcomes. Our main contribution is to show under which conditions a costly privacy-protective sales channel is used even if consumers do not have an explicit taste for privacy and how this equilibrium depends on consumers' sophistication. We thereby provide a micro-foundation for consumers' privacy choices when facing a seller with access to big data.

We construct a model where a mass of consumers with heterogeneous willingness-to-pay for a product is facing a monopolistic seller. Consumers can decide between two channels to buy the product from the seller. The direct channel (D) makes use of all personal information that the seller has about every single consumer. We assume that perfect price discrimination is feasible for the seller in channel D and that this channel economizes on transaction costs, which we normalize to zero. The anonymous channel (A) protects consumers' privacy by hiding individual identities, but comes at a cost, which we denote by s. As a consequence, perfect price discrimination is infeasible for the seller, who responds best by setting a uniform price for this channel.

Our model therefore describes a situation *after* a long period of consumers *not* using anonymization techniques (due to neglect or lack of suitable technologies). Throughout this time, the seller has acquired data shedding light on individual consumers' preferences be it via collecting such information in the past (e.g. Amazon) or via buying such information from an intermediary (e.g. Google, Acxiom). However, the seller can neither directly influence consumers' channel choice nor close down the anonymous channel as the anonymization technique is at the disposal of consumers.

In the three stages of our model, consumers first choose between channel D and channel A. Second, the seller sets prices in both channels. Third, every consumer decides whether to buy for the price offered to her, or not. Our analysis is based on a model of limited strategic sophistication, called *level-k thinking*, which was introduced by Stahl and Wilson (1994; 1995) and Nagel (1995). Models with level-k thinking are defined recursively, starting with, so-called "naïve", level-0 players which employ a "naïve" (often random) strategy. Level-1 players then best respond to the level-0 strategy, level-2 players to the level-1 strategy, and so forth.[4] A sizeable literature has developed that explores

[4]While most of this literature analyzes games with symmetric decisions between (e.g. the beauty contest game (Nagel 1995)), we will adapt the concept slightly to the asymmetric situation of our model where

level-k thinking theoretically and empirically.[5] The literature has found strong experimental support for level-k thinking and suggests values for k of one or two (Camerer, et al. 2004; Crawford and Iriberri 2007b).

In comparison to the behavior-based price discrimination literature where typically either unlimited strategic sophistication or complete naïveté of consumers is assumed, we zoom in and provide an analysis of behavior when players have *some* strategic sophistication.[6] We model consumers' cognitive constraints by their ability to anticipate k strategic iterations and that the seller is able to outperform them in strategic thinking (i.e. has a level of $k+1$) due to superior access to data and computing power. Whether k is relatively low, as suggested by the empirical behavioral literature, or rather high turns out to crucially matter for our results.

We show that the higher consumers' level of sophistication, the *higher* the equilibrium price will be on the anonymized market of channel A. Consumers anonymize if their valuation of the product exceeds the expected price plus the anonymization cost. But when consumers decide about buying, at Stage 3, those anonymization costs are sunk. Hence, the best response of the seller is to increase the price above the one consumers expected. If the level of sophistication rises in the population, consumers will expect to be offered the product for a higher price in the anonymized market. Hence, consumers with medium but not high willingness-to-pay do not choose channel A at Stage 1 anymore, preempting net losses. Consequently, the seller has an incentive to increase the price in the anonymized market even more because he infers that only consumers with high willingness-to-pay have chosen channel A at Stage 1.

We further show that, with any positive cost of anonymization, the anonymized market completely unravels for all sophistication levels $k \geq \bar{k}$, where $\bar{k}$ is a finite number. Hence, unlimited strategic sophistication is not a necessary condition for market unravelling. However, if consumers' k is sufficiently low, only a part of the market unravels and the anonymized sales channel can persist, serving consumers with high willingness-to-pay. Among those who use the anonymous sales channel, some consumers suffer from net losses because prices turn out to be higher than expected, but consumers with a very high willingness-to-pay get some surplus. Thereby, this model offers a micro-foundation for consumers' privacy choices: for some consumers, it is rational to use costly anonymization techniques even without an exogenous taste for privacy. Because a share of the

the seller has a different set of actions than the consumers.

[5]See Ho et al. (1998), Costa-Gomes et al. (2001), Crawford (2003), Camerer et al. (2004), Costa-Gomes and Crawford (2006), Crawford and Iriberri (2007a), and Goldfarb and Yang (2009), among others.

[6]What we call "unlimited strategic sophistication", is often referred to as "perfect rationality". However, players with limited strategic sophistication still act rationally given their (potentially wrong) beliefs, which is why we avoid the terms of "perfect" and "imperfect" rationality.

anonymization cost could be interpreted as a fee that an intermediary can appropriate, this model also suggests that running an anonymous sales channel competing with a channel that tracks individuals and uses all personal data can be a profitable business model when consumers have limited strategic sophistication.

Related Literature: First-degree (or perfect) price discrimination is characterized by complete information of a seller about a specific consumer's willingness-to-pay for a certain product and was introduced into the economics literature by Pigou (1920). However, due to the very high information demand of the seller about consumers' preferences and the rather straightforward allocative and distributional implications, perfect price discrimination has not received a lot of scholarly attention and has mostly been dismissed as a mere theoretical construct.[7]

More prominent are models of so-called "behavior-based price discrimination." Most of this literature focuses on third-degree price discrimination by assuming that a seller learns about the willingness-to-pay of an identifiable (or recognizable) consumer after the first purchase of a good. The idea is that, if a consumer previously bought a product at a certain price, the seller would learn that this particular consumer's willingness-to-pay must have exceeded the price for which she bought the product and consequently raise the price for her. If consumers anticipate behavior-based price discrimination, they will often adjust their behavior in early periods and potentially postpone purchases to avoid future price increases or wait for future price cuts (Villas-Boas 2004). In such cases, firms may find it optimal to have stronger privacy regulations if they lack commitment power to bind themselves to not increase prices after initial purchases (Taylor 2004).[8]

However, lending support to the early conclusion of Odlyzko, "that in the Internet environment, the incentives towards price discrimination and the ability to price discriminate will be growing" (Odlyzko 2003, p.365), online vendors and other retailers have already gone much further and can approximate fully personalized prices more than ever (see Footnote 1). It has been shown empirically that "targeted advertising" techniques increase purchases (Luo et al. 2014), prices (Mikians et al. 2012), and sellers' profits (Shiller 2013). Some consumers, however, feel repelled by this development and want to have control over their personal data back.[9] Many place a value on their privacy (Tsai et al. 2011).

[7]For instance, the standard industrial organization textbook, Tirole (1988), spends three of its more than 1100 pages on perfect price discrimination.

[8]For an overview of this strand of literature, see Fudenberg et al. (2006).

[9]Goldfarb and Tucker (2012) study three million observations between 2001 and 2008 and find that refusals to reveal their income in an online survey have risen over time. Tucker (2014) finds in a field experiment that, when Facebook gave users more control over their personally identifiable information, users were twice as likely to click on personalized ads.

The early theoretical literature about the economics of privacy, being based on the Chicago school argument that more information available to market participants increases the efficiency of markets, has underlined the negative welfare effects of hiding information from sellers (Posner 1978; Stigler 1980; Posner 1981).[10] A lot of progress in our understanding has been made since then. Already Hermalin and Katz (2006, p.229) made clear: "With so many people making extreme claims in discussions of privacy and related public policy, and with so little understanding of the underlying economics, it is important to identify the fundamental forces clearly. A central fact is that, contrary to the Chicago School argument, the flow of information from one trading partner to the other can reduce ex post trade efficiency when the increase in information does not lead to symmetrically or fully informed parties."

Another central theme in the literature are the choices of firms that own some type of personal information about consumers and can decide to disclose it to another firm (Taylor 2004; Acquisti and Varian 2005; Calzolari and Pavan 2006; Casadesus-Masanell and Hervas-Drane 2015). In interactions between an upstream and a downstream firm for whose products consumers' willingness-to-pay is positively correlated, the upstream firm will maintain full privacy of its customers if conditions on the upstream firms preferences about the downstream firm as well as on the downstream relationship itself are met (Calzolari and Pavan 2006). However, if any of the conditions is not met, the upstream firm can find it optimal to disclose the list of its customers to the downstream firm (sometimes even for free), which need not be to the detrimant of consumers but could still yield a Pareto improvement (Calzolari and Pavan 2006).

A core question studied in these papers is, what the welfare consequences of privacy or disclosure are, and who should own the property rights of consumers' personal data (Hermalin and Katz 2006).[11] The answers given have been ambiguous and depend on the specific application of the papers. Recently, the focus has shifted more towards privacy choices of consumers (Conitzer et al. 2012) and the role of platform intermediaries (de Corniere and De Nijs 2014).[12]

With few exceptions, however, cognitive constraints of consumers have not been in-

[10]Even earlier, Warren and Brandeis (1890) study privacy as "right to be let alone", a point later discussed by Varian (1997) in the context of annoyance from telemarketing. This complements our approach of understanding privacy as the absence of a seller's detailed knowledge about a consumer's preferences and characteristics.

[11]On the Internet, for instance, the customer databases of sellers or intermediaries, such as search engines, tracing back the physical address of users on the basis of their IP address (or to clearly identify them as persons on the basis of their registration data or a unique identifier derived from a permanent cookie) was recently qualified as "personal data" (Opinion 1/2008 on Data Protection Issues Related to Search Engines, Advisory Working Party (adopted Apr. 4, 2008) (EC), Data Protection available at `http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2008/wp148_en.pdf`).

[12]For a recent overview of the growing literature on the economics of privacy, see Acquisti et al. (2016).

corporated by theoretical studies of markets driven by big data. Taylor (2004), Acquisti and Varian (2005), and Armstrong (2006) assume the existence of a group of unlimitedly sophisticated consumers and a group of naïve consumers. The latter do not foresee that they may want to trade in the future again and, because of this negligence, ignore the negative effects of disclosing personal data. Hence, if consumers are naïve, a seller may oppose stricter regulations as no commitment device is needed (Taylor 2004). In our model, we allow for a more nuanced, marginal analysis of consumers' sophistication. [13]

The remainder of the paper is organized as follows. In Section 2.2, we construct a model, which is analyzed in Section 2.3. Section 2.4 studies welfare and the payoff consequences of changing the level of sophistication k and the anonymization cost s. Section 2.5 is dedicated to alternative model specifications, covering the beliefs of naïve consumers, heterogeneous costs of anonymization, and the effects of increasing competition. Section 2.6 concludes.

2.2. Model

We consider an economy where a monopolistic seller of a single consumption good faces a unit mass of atomistic consumers who can buy at most one unit of the good and cannot resell it to each other.[14] Abstracting from potential fixed costs, we assume that the monopolist can produce the good at constant marginal cost $c \geq 0$. Consumers have a heterogeneous valuation v for the good, where $v \sim \mathcal{U}[0, 1]$ and can approach the seller in two different ways: directly (referred to as channel D) or after making use of an anonymization technique (referred to as channel A).

Consumers choosing direct channel D incur no cost and the seller perfectly knows their individual valuation. Consumers choosing channel A, on the other hand, incur cost $s > 0$ and their individual valuation is hidden from the seller. We assume that consumers do not have any exogenous taste for privacy and that they choose direct channel D in case of indifference between both channels.

After consumers have made their choice between the channels, the seller sets prices based on the information available. In channel D the seller can set personalized prices $p_i(v)$ conditional on each consumer's valuation, which is known to the seller. However, such personalized pricing is impossible in channel A due to the anonymization technique consumers used. As a consequence, the seller has to set a uniform price p_A for all consumers in channel A.

[13] The need to include cognitive constraints into economic models of privacy is spurred by empirical findings about the so-called privacy paradox: A series of experimental research has shown that consumers' stated and revealed valuations of their own personal data differ highly and depend on the framing of the survey questions. (Acquisti et al. 2009; John, et al. 2009; Jentzsch et al. 2012).

[14] We discuss the case of monopolistic competition in Section 2.5.3.

Finally, consumers decide whether they want to buy the good at the price the seller posted for them and we assume that they choose to buy the product in case of indifference. We assume that outside options yield zero payoff (except for costs incurred within the game before opting out). The timing of the model is summarized as follows:

- Stage 1 (*Anonymizing*): Consumers choose channel D or channel A and incur costs of 0 or s, respectively. Indifferent consumers are assumed to choose channel D.

- Stage 2 (*Pricing*): The seller sets prices $p = \{p_i(v), p_A\}$, where $p_i(v)$ are personalized prices in channel D, and p_A is the uniform price in channel A.

- Stage 3 (*Buying*): Consumers decide whether to buy the good for the offered price. Indifferent consumers are assumed to choose buying the good.

The distribution of v (and hence the demand function), the monopolist's cost structure (and hence the supply function), the cost for anonymization s as well as the timing of the game are common knowledge among all players.

Explicitly modeling consumers' cognitive constraints, we assume that all consumers have the same limited level of strategic sophistication, denoted by $k \in \mathbb{Z}_0^+$. The seller, however, outperforms consumers in terms of sophistication, i.e. has a level of $k + 1$. Due to the limited strategic sophistication of consumers, the game cannot be solved for a Perfect Bayesian Equilibrium. Instead, we solve the game for a subgame perfect Nash equilibrium where players' beliefs about others are modeled with level-k thinking. Following Nagel (1995), players with a level of sophistication $k > 0$ will generally act as if they believe that all other players had a level of strategic sophistication exactly one level below their own level. However, Nagel (1995) considers a setting where all players are symmetric, i.e. have the same set of possible actions. As our model has one player (the seller) whose action set differs from everyone else's, and whose best response is therefore different, we find it useful to adapt the concept slightly.

While we maintain that consumers believe that all other consumers are one level less sophisticated, we deviate in assuming that consumers expect the seller to share their level of sophistication. More formally, consumers form the beliefs $\mathbb{E}_i(k_{j \neq i}) = k_i - 1 = k - 1$ for j being a consumer and $\mathbb{E}_i(k_{j \neq i}) = k_i = k$ for j being the seller. Thus, consumers implicitly think of the seller as simply responding optimally to their believe about the level of sophistication of all other consumers. This assumption is in turn based on the atomistic nature and the resulting insignificance of any individual consumer for the seller's choice. The seller's beliefs, on the other hand, are in line with Nagel (1995), i.e. the seller forms the belief $\mathbb{E}_i(k_{j \neq i}) = k_i - 1 = k$ for i being the seller and j being any consumer (recall that for the seller $k_i = k + 1$).

The difference between our solution concept and a Perfect Bayesian Equilibrium which would result from an interaction of only players with unlimited strategic sophistication is that it does not need to be the case that all expectations about the equilibrium path are eventually confirmed by the resulting equilibrium play. In our setting, this is most notable by the fact that we allow that consumers' beliefs about the prices in both channels do not need to coincide with the prices the seller eventually sets. This means, that we do *not* impose that $\mathbb{E}(p|D) = (p|D)$ and $\mathbb{E}(p|A) = (p|A)$, where $(p|D)$ and $(p|A)$ denote the price after having chosen channel D or channel A, respectively. However, we do assume that players restrict their beliefs about possible prices to the support of the distribution of valuations, i.e. $\mathbb{E}(p) \in [0, 1]$. Their cognitive limitations in belief formation notwithstanding, players still act rationally in the sense that they pursue strategies which maximize their utility *given* their beliefs.

A consumer's strategy in this game is therefore a mapping from her valuation for the good v and her level of strategic sophistication k as well as the game's exogenous parameters s and c to her action space $C \times B$, where $C = \{Channel\ D, Channel\ A\}$ denotes her set of choices in the anonymizing stage (Stage 1) and $B = \{(Buying|p), (NotBuying|p)\}$ denotes her set of choices in the buying stage (Stage 3), where $p = p_i(v)$ after having chosen channel D and $p = p_A$ after having chosen channel A.

The seller's strategy in this game, on the other hand, is a mapping from his level of strategic sophistication $k + 1$ as well as the game's exogenous parameters s and c to a set of prices $p = \{p_i(v), p_A\}$, where $p_i(v)$ are personalized prices he can condition on his knowledge about individual consumers approaching him via channel D, and p_A is a uniform price he has to set for all consumers approaching him via channel A.

In the following section, the game is solved by backward induction. As outlined earlier, models with level-k thinking are best solved recursively. Thus, the analysis starts with the case where consumers have a strategic level of sophistication of $k = 0$ and form a so-called naïve belief. The seller, having a level of sophistication of $k = 1$, believes (correctly) that all consumers have a level of $k = 0$ and therefore his $k = 1$-level best response is also objectively optimal. In later parts of the analysis, consumers have a positive level of strategic sophistication $k > 0$. As they assume that all other consumers have a level of $k - 1$, they effectively expect the seller to employ the k-level best response. This means, that their belief results in a particular price expectation. The seller, then having a level of sophistication of $k + 1$, will again be the only player to have an objectively correct belief and his $k + 1$-level response is again objectively optimal.

2.3. Analysis

Stage 3 – Buying: A utility-maximizing consumer decides to buy the product if the price she has to pay does not exceed her valuation of the good, i.e. if, and only if,

$$v \geq p \in \{p_i(v), p_A\}. \tag{2.1}$$

If she has chosen for channel D, the price will be an individualized price $p_i(v)$, and if she has chosen for channel A, she will receive the same uniform price p_A as all other consumers who have chosen channel D.

Stage 2 – Pricing: A profit-maximizing seller sets individual prices p_i for all consumers approaching him via channel D (denoted by set $\mathcal{C}_D$) and one optimal uniform price p_A for all anonymized consumers in channel A (denoted by set $\mathcal{C}_A$). Knowing v precisely for all consumers in $\mathcal{C}_D$, the seller trivially sets

$$p_i^*(v) = \max\{v, c\} \text{ for all } i \in \mathcal{C}_D, \tag{2.2}$$

where the lower bound c takes into consideration that it is not optimal to sell below marginal cost. Being uninformed about the individual valuations v of all consumers in $\mathcal{C}_A$, the seller can nevertheless infer which consumers are in $\mathcal{C}_A$ due to his higher level of strategic sophistication and set p_A accordingly. We will therefore analyze consumers' general Stage 1 behavior first in order to inform the seller's pricing decision in channel A.

Stage 1 – Anonymizing: Consumers use the anonymization technique of channel A if the expected utility of doing so exceeds the expected utility of the direct channel D, i.e. if, and only if, $\mathbb{E}(u_i(A)) > \mathbb{E}(u_i(D))$, where

$$\mathbb{E}(u_i(D)) = \max\{v - \mathbb{E}(p|D), 0\}, \tag{2.3}$$

$$\mathbb{E}(u_i(A)) = \max\{v - \mathbb{E}(p|A) - s, -s\}. \tag{2.4}$$

The first value in each set in Equations (2.3) and (2.4) reflects the expected payoff the consumer receives if she buys the product at Stage 3. The second value reflects the payoff of subsequently choosing not to buy the product. Although consumers might be limited in their strategic sophistication, we will nonetheless assume that they understand the nature of the two channels, i.e. they understand that the seller has no incentive to decrease the price below their valuation in channel D and that the seller can only ask for a uniform price in channel A. Hence, consumers form the price expectation for channel D

$$\mathbb{E}(p|D) = p_i^*(v) = \max\{v, c\}, \tag{2.5}$$

irrespective of their level of strategic sophistication k. They correctly expect to be left with no surplus when choosing channel D.

With respect to channel A, however, consumers only know that the seller sets a uniform price. Which price exactly they expect depends on their level of strategic sophistication. For now, it is sufficient to replace the expectation $\mathbb{E}(p|A)$ by the expectation of a single uniform price $\mathbb{E}(p_A)$:

$$\mathbb{E}(u_i(D)) = \max\{v - \max\{v, c\}, 0\} = 0, \tag{2.6}$$

$$\mathbb{E}(u_i(A)) = \max\{v - \mathbb{E}(p_A) - s, -s\}, \tag{2.7}$$

This shows that consumers choose channel A if, and only if,

$$\max\{v - \mathbb{E}(p_A) - s, -s\} > 0. \tag{2.8}$$

Since $s > 0$, this can only hold if $v > \mathbb{E}(p_A) + s \equiv \hat{v}$, where $\hat{v}$ denotes the endogenous threshold dividing the population of consumers into $\mathcal{C}_D$ and $\mathcal{C}_A$.

Lemma 2.1 (Anonymization Threshold)**.** *There exists a threshold $\hat{v} = \mathbb{E}(p_A) + s$ that denotes the valuation of a consumer who is indifferent between both channels at Stage 1. Consumers with $v > \hat{v}$ will prefer channel A to channel D; consumers with $v \leq \hat{v}$ prefer channel D to channel A, i.e. $\mathcal{C}_D = [0, \hat{v}]$ and $\mathcal{C}_A = (\hat{v}, 1]$.*

Stage 2 – Pricing (revisited): Having a higher level of strategic sophistication than the consumers, the seller correctly infers $\hat{v}$ and hence knows that $\mathcal{C}_A = (\hat{v}, 1]$. As he further anticipates that consumers will buy the product at Stage 3, if, and only if, $v \geq p_A$, he can easily infer demand $q_A(p_A)$ in channel A:

$$q_A(p_A) = \begin{cases} 0 & \text{if } p_A > 1, \\ 1 - p_A & \text{if } 1 \geq p_A > \hat{v}, \\ 1 - \hat{v} & \text{if } \hat{v} \geq p_A. \end{cases} \tag{2.9}$$

Charging $p_A = \hat{v}$ dominates all prices $p'_A < \hat{v}$ because any price below $\hat{v}$ decreases profits per unit sold without an increase in quantity to counter the loss. Thus by setting $p_A = \hat{v}$, the seller can guarantee himself profits from channel A of:

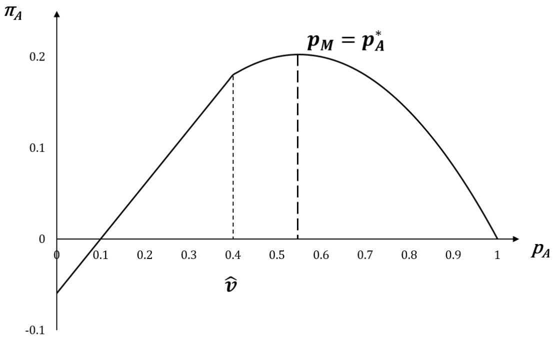
$$\pi_A(\hat{v}) = q_A(\hat{v})(\hat{v} - c) = (1 - \hat{v})(\hat{v} - c). \tag{2.10}$$

However, the seller could also charge a price $p_A > \hat{v}$, depending on where $\hat{v}$ lies exactly. Suppose for the moment that the entire consumer population uses channel A (i.e. $\hat{v} = 0$), which is identical to the case of a monopolist unable to engage in price discrimination. Let us denote the globally profit-maximizing price in this case by $p_M = \frac{1+c}{2}$. Then, there are three different cases for the location of the anonymization threshold $\hat{v}$ (shown in Figure 2.1) compared to p_M:

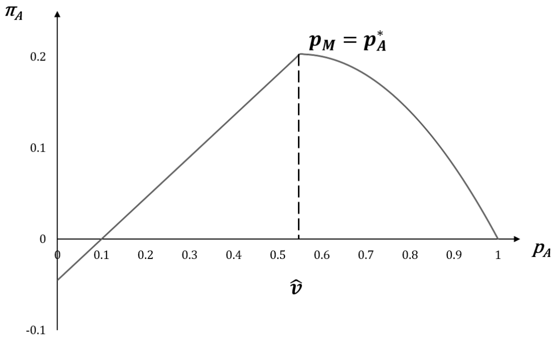
(a) The anonymization threshold is below the standard monopoly price ($\hat{v} < p_M$).

(b) The anonymization threshold is equal to the standard monopoly price ($\hat{v} = p_M$).

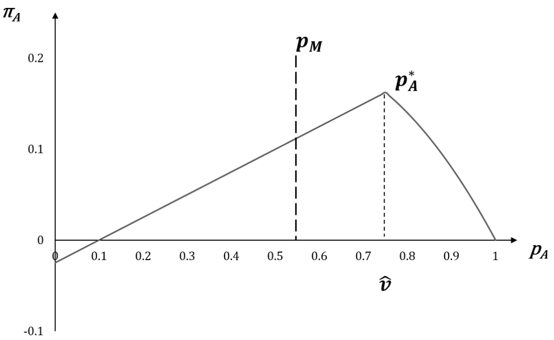
(c) The anonymization threshold is above the standard monopoly price ($\hat{v} > p_M$).



(a) $\hat{v} < p_M$



(b) $\hat{v} = p_M$



(c) $\hat{v} > p_M$

Figure 2.1: Profits in Channel A for Different Locations of $\hat{v}$
with Parameters $v \sim \mathcal{U}[0, 1], c = 0.1$

In cases (a) and (b), the globally profit-maximizing price p_M is in the support of the demand function and hence remains the optimal price to set. The only consumers that are not in $\mathcal{C}_A$ are those that the seller would not have served even if they had anonymized themselves. Only in case (c), where the globally profit-maximizing price p_M is not in the support anymore, a price below the anonymization threshold $\hat{v}$ is at least dominated by setting the price equal to $\hat{v}$. The seller also has no incentive to raise the price above $\hat{v}$

as profits are strictly decreasing to either side of the global maximum at p_M due to the strict concavity of the profit function. Hence, in this case the optimal price p_A^* is equal to $\hat{v}$. The seller's optimal pricing strategy for both channels is summarized in Lemma 2.2.

Lemma 2.2 (Optimal Pricing Strategy)**.** *The optimal strategy of the seller consists of a set of prices* $\{p_i^*(v), p_A^*\}$ *in channel D and channel A, respectively, where* $p_i^*(v) = \max\{v, c\}$ *and* $p_A^* = \max\{\hat{v}, p_M = \frac{1+c}{2}\}$*.*

Note, that this implies that the seller sets a higher price than consumers had expected:

$$p_A^* \geq \hat{v} = \mathbb{E}(p_A) + s > \mathbb{E}(p_A). \tag{2.11}$$

Note additionally that, with unlimited strategic sophistication, it would be required that $\mathbb{E}(p_A) = p_A^*$ in equilibrium, leading to a contradiction. Because only beliefs about off-equilibrium paths can be wrong in a Perfect Bayesian Equilibrium, we can conclude that if all players had unlimited strategic sophistication, channel A would remain unused.

However, with limited strategic sophistication such a discrepancy is possible. This is due to the fact that s will be a sunk cost for consumers at Stage 3, which the seller can exploit via increasing the price by exactly s, compared to their expectations. Consumers, due to their limited strategic sophistication, do not anticipate the seller's strategic response which influences their expectation formation in Stage 1.

Stage 1 – Anonymizing (revisited): The last missing piece to fully characterize equilibrium behavior is the formation of consumers' expectations of the price in channel A, $\mathbb{E}(p_A)$, in Stage 1. As outlined earlier, we capture this by level-k thinking, which is best determined recursively. Thus, we will start with the case of consumers with a strategic level of sophistication of $k = 0$, which are referred to as "naïve" consumers: they naïvely expect the monopolist to engage in regular monopoly pricing[15] in channel A, i.e. $\mathbb{E}_0(p_A) = p_M$, ignoring the fact that the very choice of channel A might be signaling a high willingness to pay to the seller. For channel D, we have already assumed that even the most naïve (but still rational) consumer foresees perfect price discrimination in channel D as it does not require iterative thinking. Lemma 2.3 summarizes equilibrium behavior if consumers are strategically "naïve".

Lemma 2.3 (Equilibrium with Level-0)**.** *For any non-prohibitively high cost of anonymization* $s > 0$ *and cost of production* $c \geq 0$*, and with strategically "naïve" consumers* ($k = 0$)*, there is a unique equilibrium with the following characteristics:*

[15] Alternative assumptions about starting points for naïve consumers are discussed in Section 2.5.1. We chose in favor of expositional simplicity.

- *Consumers form the 0-beliefs* $\mathbb{E}_0(p_D) = p_i^*(v)$ *and* $\mathbb{E}_0(p_A) = p_M = \frac{1+c}{2}$.
- *Consumers anonymize if, and only if,* $v > \hat{v}_0 = \mathbb{E}_0(p_A) + s = p_M + s$, *separating into the sets* $\mathcal{C}_D = [0, \hat{v}_0]$ *and* $\mathcal{C}_A = (\hat{v}_0, 1]$.
- *The seller forms the 1-beliefs* $\mathbb{E}_1(\mathcal{C}_D) = [0, \hat{v}_0]$ *and* $\mathbb{E}_1(\mathcal{C}_A) = (\hat{v}_0, 1]$.
- *The seller sets prices* $p_i^*(v) = \max\{v, c\}$ *and* $p_{A_0}^* = \hat{v}_0 = p_M + s$.
- *All consumers in* $\mathcal{C}_D$ *with* $v \geq c$ *buy the product at the price offered to them.*
- *All consumers in* $\mathcal{C}_A$ *buy the product at the price offered to them.*

Lemma 2.3 shows that naïve consumers in channel A pay a surcharge of s as compared to their expectations ($p_{A_0}^* - \mathbb{E}_0(p_A) = s$). Due to their cognitive constraints, consumers do not anticipate that the seller can infer that only consumers with a valuation of at least $p_M + s$ choose the anonymous channel. Given this lower bound on the valuations in $\mathcal{C}_A$, the seller can ignore that anonymized consumers spent s on top, and extract the lower bound's full willingness-to-pay. This divergence between expected price and realized price, in turn, informs us about the way in which consumers form their price expectation for higher levels of strategic sophistication, $k > 0$.

If instead of being naïve ($k = 0$), all consumers are capable of one iteration of strategic reasoning, they anticipate that the seller's best response to the 0-belief of naïve consumer population is to set $p_{A_0}^* = p_M + s$. Recall, that consumers with a positive level of strategic sophistication believe that all other consumers have a level of strategic sophistication exactly one level below their own level, while assuming that the seller has the same level sophistication as themselves. Therefore, they assume that the seller responds optimally to a population of consumers with $k = 0$ and adjust their expectation. As consumers are atomistic, their own anonymization decision is inconsequential for the seller's best response. Accordingly, they form the 1-belief $\mathbb{E}_1(p_A) = p_{A_0}^* = p_M + s$ leading to $\hat{v}_1 = p_M + 2s$, to which the seller's actual best response is $p_{A_1}^* = p_M + 2s$ (analogue to the reasoning above). This, in turn, would be the expected price in the anonymous channel by consumers with a strategic sophistication level of $k = 2$, thus forming the 2-belief $\mathbb{E}_2(p_A) = p_{A_1}^* = p_M + 2s$, and so forth. More generally we can write $\mathbb{E}_k(p_A) = p_{A_{k-1}}^*$ for all $k > 0$, which in combination with $\mathbb{E}_0(p_A) = p_M$ leads to:

$$\mathbb{E}_k(p_A) = p_M + ks, \tag{2.12}$$

$$p_{A_k}^* = p_M + (k+1)s = \hat{v}_k. \tag{2.13}$$

Thus, at every additional level of strategic sophistication, consumers will incorporate the sunk cost one more time than at the previous level, which induces the seller to raise the price once more. Consequently, $\hat{v}_k$ is increasing in k. $\mathcal{C}_A$ shrinks in size as k increases: The more strategically sophisticated the population of consumers is, the fewer consumers will choose to anonymize, until a point is reached where no consumer does so anymore. Then, channel A remains unused and the anonymous market breaks down completely. This point is reached when the anonymization threshold matches or exceeds even the highest willingness-to-pay of any consumer. We denote the threshold level of strategic sophistication from which onwards this is the case by $\bar{k}$ and define:

$$\bar{k} \equiv \min\{k \in \mathbb{Z}_0^+ | \hat{v}_k \geq 1\}. \tag{2.14}$$

The inequality condition of Equation (2.14) can hold with equality as any consumer indifferent between the two channels opts for channel D, including the one with the maximum valuation for the good $v = 1$. Using Equation (2.13) in (2.14) and solving for $\bar{k}$ yields:

$$\bar{k} \geq \frac{1-c}{2s} - 1 \Rightarrow \bar{k} = \left\lceil \frac{1-c}{2s} - 1 \right\rceil. \tag{2.15}$$

This shows that channel A breaks down at a *finite* level of strategic sophistication, in turn implying that unlimited strategic sophistication, while *sufficient*, is *not necessary* for a breakdown of channel A. Lemma 2.4 summarizes the existence conditions for channel A.

Lemma 2.4 (Usage of Channel A). *For any non-prohibitively high cost of anonymization $s > 0$ and cost of production $c \geq 0$, the anonymous channel is used if, and only if, consumers are not* too *strategically sophisticated, i.e. if $k < \bar{k} = \left\lceil \frac{1-c}{2s} - 1 \right\rceil$.*

That channel A breaks down at a finite level of sophistication $\bar{k}$ has consequences for the belief formation of consumers when $k > \bar{k}$. While belief formation according to Equation (2.12) does not violate that all players restrict price expectations to $p \in [0, 1]$ for $k \leq \bar{k}$, this is not the case for $k > \bar{k}$. Denoting any level of consumer sophistication $k > \bar{k}$ by $\bar{k}^+$, we specify beliefs $\mathbb{E}_{\bar{k}^+}(p_A)$ that meet this condition (Equation 2.16). Additionally, in line with Lemma 2.4, any belief $\mathbb{E}_{\bar{k}^+}(p_A)$ has to render the choice of channel D a Nash strategy for consumers regardless of their valuation (Equation 2.17). This yields:

$$\mathbb{E}_{\bar{k}^+}(p_A) \in [0, 1] \qquad \Rightarrow \mathbb{E}_{\bar{k}^+}(p_A) \leq 1, \tag{2.16}$$

$$\hat{v}_{\bar{k}^+} = \mathbb{E}_{\bar{k}^+}(p_A) + s \geq 1 \qquad \Rightarrow \mathbb{E}_{\bar{k}^+}(p_A) \geq 1 - s. \tag{2.17}$$

Both conditions are satisfied for any belief $\mathbb{E}_{\bar{k}^+}(p_A) \in [1-s, 1]$. Hence, multiple beliefs are possible when $k > \bar{k}$, but all imply that channel A remains unused. For any level of consumer sophistication where channel A remains unused (including $k = \bar{k}$), the seller forms the k+1-beliefs $\mathbb{E}_{k+1}(\mathcal{C}_D) = [0,1]$ and $\mathbb{E}_{k+1}(\mathcal{C}_A) = \emptyset$. Therefore, setting p_A is an off-equilibrium action and the seller can set any price $p^*_{A_{\bar{k}+}} \in [0;1]$ (restricted only by the support of the demand function). Combining the insights of the previous Lemmas, we summarize the analysis with the formulation of the general equilibrium with level-k thinking in Proposition 2.1.

Proposition 2.1 (Equilibrium with Level-k)**.** *For any non-prohibitively high cost of anonymization $s > 0$ and cost of production $c \geq 0$ it holds that:*

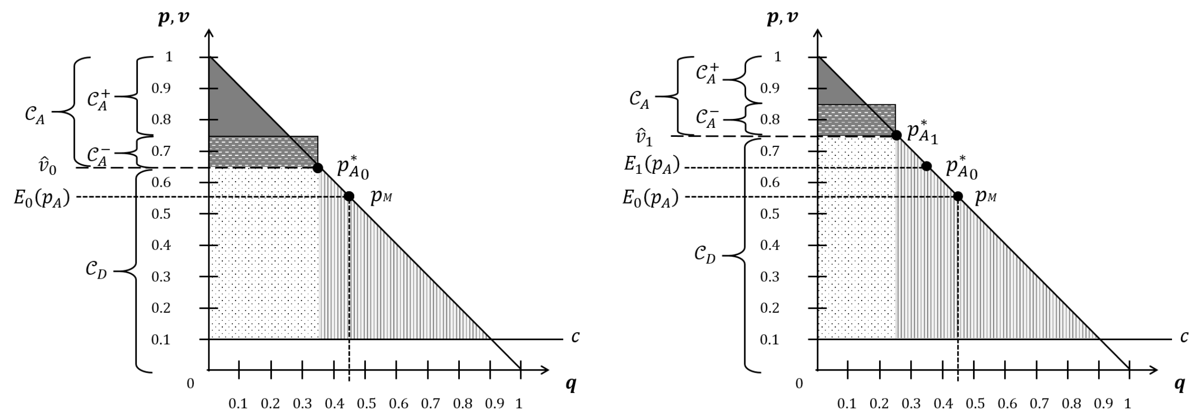
1. *If consumers have a level of strategic sophistication of $k \leq \bar{k} = \left\lceil \frac{1-c}{2s} - 1 \right\rceil$, there is a unique equilibrium with the following characteristics:*
   - *Consumers form the k-beliefs $\mathbb{E}_k(p_D) = p^*_i(v)$ and $\mathbb{E}_k(p_A) = p_M + ks = \frac{1+c}{2} + ks$.*
   - *Consumers anonymize if, and only if, $v > \hat{v}_k = p_M + (k+1)s$, separating into the sets $\mathcal{C}_D = [0, \hat{v}_k]$ and $\mathcal{C}_A = (\hat{v}_k, 1]$ (where $\mathcal{C}_A = \emptyset$ if $k = \bar{k}$).*
   - *The seller forms the k+1-beliefs $\mathbb{E}_{k+1}(\mathcal{C}_D) = [0, \hat{v}_k]$ and $\mathbb{E}_{k+1}(\mathcal{C}_A) = (\hat{v}_k, 1]$.*
   - *If $k < \bar{k}$, the seller sets $p^*_i(v) = \max\{v, c\}$ and $p^*_{A_k} = \hat{v}_k = p_M + (k+1)s$.*
   - *If $k = \bar{k}$, the seller sets $p^*_i(v) = \max\{v, c\}$ and $p^*_{A_{\bar{k}}} \in [0,1]$.*
   - *All consumers in $\mathcal{C}_D$ with $v \geq c$ buy the product at the price offered to them.*
   - *All consumers in $\mathcal{C}_A$ buy the product at the price offered to them.*
2. *If consumers have a level of strategic sophistication of $k > \bar{k} = \left\lceil \frac{1-c}{2s} - 1 \right\rceil$, there are multiple equilibria with the following characteristics:*
   - *Consumers form the k-beliefs $\mathbb{E}_{\bar{k}+}(p_D) = p^*_i(v)$ and $\mathbb{E}_{\bar{k}+}(p_A) \in [1-s, 1]$.*
   - *No consumer anonymizes as $\hat{v}_{\bar{k}+} \in [1, 1+s]$ and hence $v \leq \hat{v}_{\bar{k}+}$ for all v, leading to the sets $\mathcal{C}_D = [0,1]$ and $\mathcal{C}_A = \emptyset$.*
   - *The seller forms the k+1-beliefs $\mathbb{E}_{k+1}(\mathcal{C}_D) = [0,1]$ and $\mathbb{E}_{k+1}(\mathcal{C}_A) = \emptyset$.*
   - *The seller sets $p^*_i(v) = \max\{v, c\}$ and any $p^*_{A_k} \in [0,1]$.*
   - *All consumers in $\mathcal{C}_D$ with $v \geq c$ buy the product at the price offered to them.*
   - *No consumer buys the product via channel A.*

In the equilibrium captured by the first case, consumers with high valuations ($v > \hat{v}_k$) choose the anonymous channel A, consumers with low valuations ($v \leq \hat{v}_k$) choose the direct channel D and are perfectly price discriminated against. Notably, consumers with very low valuations ($v < p_M$) choose the direct channel D irrespectively of k and s as they cannot possibly hope to get a uniform price that is affordable to them via channel A.

Those consumers are the ones that are not served in monopolistic markets, where there is no possibility for perfect price discrimination. The multiplicity of equilibria in the second case of Proposition 2.1 depends only on the multiplicity of possible beliefs about the off-equilibrium path. But all equilibria lead to the same equilibrium behavior, where no consumer anonymizes and the seller charges individualized prices $p_i^*(v)$ to everyone.

2.4. Welfare

As we have shown, different levels of consumer sophistication k lead to different anonymization behavior, which has consequences for consumer surplus (CS), profits (π), and total welfare (W). We will first take a look at consumer surplus and profits for both channels separately. Total welfare, for which we employ the customary definition, $W = CS + \pi$, and hence abstract from preferences by a social planner (or policy-maker) for either side of the market, will only be included in our final aggregate analysis. Throughout the entire section, though, Figure 2.2 might serve as a visualization of the different sets and quantities and illustrates the effects of an increase in k when comparing Figure 2.2a and Figure 2.2b. In the comparative statics analysis of changes in consumer sophistication the discreteness of k is taken into account by calculating changes as differences rather than derivatives. Additionally, due to the potential non-linearity when increasing k from $\bar{k} - 1$ to $\bar{k}$, these differences only hold for $k + 1 < \bar{k}$.[16]



(a) $k = 0$ (b) $k = 1$

Figure 2.2: Welfare Analysis with Parameters $v \sim \mathcal{U}[0, 1]$, $c = 0.1$, $s = 0.1$

[16] Recall that $\bar{k}$ is usually the result of rounding (unless $\frac{1-c}{2} - 1 \in \mathbb{Z}_0^+$) and hence the last change in the composition of $\mathcal{C}_A$ and $\mathcal{C}_D$ is usually of different size than s. When increasing consumer sophistication from $\bar{k} - 1$ to $\bar{k}$, the increase of $\mathcal{C}_D$ is bounded from above by s as *all remaining* consumers switch to channel D. Introducing separate cases in all difference equations is avoided for legibility, but addressed in the text where necessary.

2.4.1. Channel D

Consumer Surplus and Profits in Channel D

As the seller engages in perfect price discrimination for consumers in $\mathcal{C}_D$, it is clear that

$$CS_{D_k} = 0, \tag{2.18}$$

whereas the seller appropriates the entire surplus in channel D:

$$\pi_{D_k} = \frac{(\hat{v}_k - c)^2}{2} = \frac{1}{8}(1-c)^2 + \frac{1-c}{2}(k+1)s + \frac{(k+1)^2}{2}s^2. \tag{2.19}$$

π_{D_k} corresponds to the vertically striped (lower right) triangle in Figure 2.2.

Comparative Statics for k in Channel D

Recalling that $\mathcal{C}_D = [0, \hat{v}_k]$ and $\hat{v} = p_M + (k+1)s$, we note first that increasing k to $k+1$ raises $\hat{v}$ and hence increases the size of $\mathcal{C}_D = [0, \hat{v}_k]$. Let $\Delta CS_{D_k} \equiv CS_{D_{k+1}} - CS_{D_k}$ and $\Delta\pi_{D_k} \equiv \pi_{D_{k+1}} - \pi_{D_k}$ denote the effects of increasing consumer sophistication on consumer surplus and profits in channel D. It can be shown that:

$$\text{For } k < \bar{k} - 1:$$

$$\Delta CS_{D_k} = 0, \tag{2.20}$$

$$\Delta\pi_{D_k} = (\hat{v}_{k+1} - c)s - \frac{s^2}{2} = \frac{1-c}{2}s + \frac{2k+3}{2}s^2. \tag{2.21}$$

Due to perfect price discrimination, consumer surplus in channel D, unsurprisingly, does not change when consumer sophistication increases. Profits in channel D, though, increase because the group of consumers which the seller can perfectly discriminate, $\mathcal{C}_D$, grows. This can also be seen by comparing Figure 2.2a and Figure 2.2b where the larger bracket along the vertical axis shows the increasing size of channel D and the larger striped triangle the increase in profits. Growth of π_D continues when increasing consumer sophistication from $\bar{k}-1$ to $\bar{k}$ (bounded from above by the expression in Equation (2.21)) and comes to a halt from there onwards as all consumers are in $\mathcal{C}_D$.

Lemma 2.5 (Effects of Changing Consumer Sophistication (Channel D)). *Raising the level of strategic sophistication of consumers from k to $k+1$ increases the usage of channel D for all $k < \bar{k}$ (and is maximal for $k \geq \bar{k}$). Consumer surplus in channel D is zero (CS_{D_k}=0) and independent of k (ΔCS_{D_k}=0). The seller's profits from channel D are positive ($\pi_D > 0$) and increasing in k for all $k < \bar{k}$ (and maximal for $k \geq \bar{k}$).*

2.4.2. Channel A

Consumer Surplus and Profits in Channel A

In channel A, consumer surplus consists of two parts: the benefit from consumption of the good after the transaction at Stage 3 (denoted by $CS^+_{A_k}$) and the cost of anonymization incurred at Stage 1 (denoted by $CS^-_{A_k}$):

$$CS^+_{A_k} = \frac{(1-\hat{v}_k)^2}{2} = \frac{(1-c)^2}{8} - \frac{1-c}{2}(k+1)s + \frac{1}{2}(k+1)^2 s^2, \tag{2.22}$$

$$CS^-_{A_k} = (1-\hat{v}_k)s = \frac{1-c}{2}s - (k+1)s^2. \tag{2.23}$$

In Figure 2.2, $CS^+_{A_k}$ corresponds to the solid grey (upper) triangle, whereas the dashed rectangle that partially overlaps this triangle represents the term $CS^-_{A_k}$. Net consumer surplus ($CS_{A_k} \equiv CS^+_{A_k} - CS^-_{A_k}$) in channel A then amounts to:

$$CS_{A_k} = \frac{(1-\hat{v}_k)^2}{2} - (1-\hat{v}_k)s = \frac{1}{8}(1-c)^2 - \frac{1-c}{2}(k+2)s + \frac{(k+1)(k+3)}{2}s^2. \tag{2.24}$$

Additionally, note that only some consumers in channel A end up with positive net surplus (denoted by $\mathcal{C}^+_A = [\hat{v}_k + s, 1]$), whereas others end up with negative net surplus (denoted by $\mathcal{C}^-_A = (\hat{v}_k, \hat{v}_k + s)$).[17] Both sets are shown along the vertical axis of Figure 2.2. The seller's profits in channel A correspond to the dotted white rectangle in Figure 2.2 and are given by

$$\pi_{A_k} = (1-\hat{v}_k)(\hat{v}_k - c) = \frac{1}{4}(1-c)^2 - (k+1)^2 s^2. \tag{2.25}$$

Comparative Statics for k in Channel A

Recalling that $\mathcal{C}_A = (\hat{v}_k, 1]$ and $\hat{v}_k = p_M + (k+1)s$, we note first that increasing k to $k+1$ raises $\hat{v}_k$ and hence decreases the size of $\mathcal{C}_A = (\hat{v}_k, 1]$. Letting $\Delta CS_{A_k} \equiv CS_{A_{k+1}} - CS_{A_k}$ and $\Delta\pi_{A_k} \equiv \pi_{A_{k+1}} - \pi_{A_k}$ denote the effects of increasing consumer sophistication on consumer surplus and profits in channel A, it can be shown that:

For $k < \bar{k} - 1$:

$$\Delta CS_{A_k} = -\left((1-\hat{v}_{k+1})\,s + \frac{s^2}{2}\right) + s^2 = -\frac{1-c}{2}s + \frac{2k+5}{2}s^2, \tag{2.26}$$

$$\Delta\pi_{A_k} = (1-\hat{v}_{k+1})s - (\hat{v}_k - c)s = -(2k+3)s^2. \tag{2.27}$$

While the first term in Equation (2.26) stems from the reduction of consumer surplus from the transaction of the good at Stage 3, the second term stems from the gain from

[17] We include the consumer with a 0 net surplus in the set $\mathcal{C}^+_A$.

fewer consumers incurring the up-front anonymization cost. In Figure 2.2, the first effect is represented by the shrinking area of the dark grey triangle, and the second effect by the shrinking dashed rectangle.[18] Which of these effects dominates determines whether consumer surplus in channel A increases or decreases with increasing k. Denoting the threshold level of consumer sophistication where consumer surplus stops decreasing by $\bar{k}_{\Delta CS}$, we define:

$$\bar{k}_{\Delta CS} \equiv min\{k \in \mathbb{Z}_0^+ | \Delta CS_{A_k} \geq 0\}. \tag{2.28}$$

Using Equation (2.26) in (2.28), solving for $\bar{k}_{\Delta CS}$, and following the same line of reasoning to deal with the discreteness of k as before yields:

$$\bar{k}_{\Delta CS} \geq \frac{1-c}{2s} - \frac{5}{2} \Rightarrow \bar{k}_{\Delta CS} = \left\lceil \frac{1-c}{2s} - \frac{5}{2} \right\rceil. \tag{2.29}$$

To get a better impression of the location of this threshold, recall that $\bar{k} = \lceil \frac{1-c}{2s} - 1 \rceil$ and therefore

$$\bar{k} - \bar{k}_{\Delta CS} = \left\lceil \frac{1-c}{2s} - 1 \right\rceil - \left\lceil \frac{1-c}{2s} - \frac{5}{2} \right\rceil = \left\lceil \frac{1-c}{2s} \right\rceil - \left\lceil \frac{1-c}{2s} - \frac{1}{2} \right\rceil + 1 \in \{1, 2\}, \tag{2.30}$$

which reveals that consumer surplus stops decreasing already one or two levels of sophistication before channel A breaks down. While this seems counterintuitive at first, it is helpful to recall that $\mathcal{C}_A = \mathcal{C}_A^- \cup \mathcal{C}_A^+$ and that $\mathcal{C}_A^-$ is situated below $\mathcal{C}_A^+$. Hence, as k increases, $\mathcal{C}_A^+$ seizes to contain consumers before $\mathcal{C}_A^-$ does, which in turn means that consumer surplus eventually turns negative. Denoting the additional thresholds $\bar{k}_{CS}$, where consumer surplus turns negative, and $\bar{k}_{\mathcal{C}_A^+}$, where no consumer in channel A makes a net surplus from the transaction anymore, we define:

$$\bar{k}_{CS} \equiv min\{k \in \mathbb{Z}_0^+ | CS_{A_k} = 0\}, \tag{2.31}$$

$$\bar{k}_{\mathcal{C}_A^+} \equiv min\{k \in \mathbb{Z}_0^+ | \mathcal{C}_A^+ = \emptyset\}. \tag{2.32}$$

Using Equation (2.24) in (2.31) and the definition of $\mathcal{C}_A^+ = (\hat{v}_k + s, 1]$ in Equation (2.32), solving for the respective thresholds and following the same line of reasoning to deal with the discreteness of k as before yields:

$$\bar{k}_{CS} \geq \frac{1-c}{2s} - 3 \Rightarrow \bar{k}_{CS} = \left\lceil \frac{1-c}{2s} - 3 \right\rceil, \tag{2.33}$$

$$\bar{k}_{\mathcal{C}_A^+} \geq \frac{1-c}{2s} - 2 \Rightarrow \bar{k}_{\mathcal{C}_A^+} = \left\lceil \frac{1-c}{2s} - 2 \right\rceil. \tag{2.34}$$

[18] The dark grey triangle shrinks by a trapezoid composed of the rectangle of area $(1 - \hat{v}_{k+1})s$ and the triangle of area $\frac{s^2}{2}$, whereas the dashed rectangle has height s and shrinks in width by s, making for a decrease in area of s^2.

Similarly, these thresholds can be put in relation to the level of sophistication at which the market for anonymization breaks down:

$$\bar{k} - \bar{k}_{CS} = \left\lceil \frac{1-c}{2s} - 1 \right\rceil - \left\lceil \frac{1-c}{2s} - 3 \right\rceil = \left\lceil \frac{1-c}{2s} \right\rceil - \left\lceil \frac{1-c}{2s} \right\rceil + 2 = 2, \tag{2.35}$$

$$\bar{k} - \bar{k}_{\mathcal{C}_A^+} = \left\lceil \frac{1-c}{2s} - 1 \right\rceil - \left\lceil \frac{1-c}{2s} - 2 \right\rceil = \left\lceil \frac{1-c}{2s} \right\rceil - \left\lceil \frac{1-c}{2s} \right\rceil + 1 = 1. \tag{2.36}$$

Equation (2.35) shows that the combined cost of anonymization incurred by all consumers in $\mathcal{C}_A$ outweighs the combined surplus from the transaction of the good at the penultimate level before the breakdown of channel A, while at the last level before the breakdown of channel A there are no consumers in channel A anymore that make a net surplus, as Equation (2.36) shows. Taken together, they provide the two options derived in Equation (2.30) for the level of sophistication at which consumer surplus stops decreasing. Hence, we can resolve the counterintuitive result that consumer surplus can stop decreasing already at $\bar{k} - 2$ by having shown that this is only possible because consumer surplus is 0, at best, at this point and will be negative at $\bar{k} - 1$ the latest. Due to the discreteness of k, the minimum might be attained at either level (indicated by the result of Equation (2.30)). In any case, raising the level of strategic sophistication from $\bar{k} - 1$ to $\bar{k}$ leads to an increase in consumer surplus as channel A remains unused and consumer surplus jumps to 0 as all consumers are being perfectly price discriminated in channel D. To summarize our discussion of consumer surplus in more plain terms: consumers lose surplus the more strategically sophisticated they become until everyone "gives in" to the seller's price discrimination practices in the direct channel D.

Profits in channel A, however, are generally decreasing in consumer sophistication, as Equation (2.27) shows. Contrary to consumer surplus, there are no thresholds determining a change in this process for profits in channel A as they continue decreasing until channel A is not used by any consumer.

Lemma 2.6 (Effects of Changing Consumer Sophistication (Channel A))**.** *Raising the level of strategic sophistication of consumers from k to $k+1$ decreases the usage of channel A for all $k < \bar{k}$ (and is zero for $k \geq \bar{k}$). Consumer surplus (CS_A) decreases for all $k < \bar{k}_{\Delta CS} = \lceil \frac{1-c}{2s} - \frac{5}{2} \rceil \in \{\bar{k} - 2; \bar{k} - 1\}$ and becomes non-positive at $\bar{k}_{CS} = \lceil \frac{1-c}{2s} - 3 \rceil = \bar{k} - 2$. Additionally, at $\bar{k}_{\mathcal{C}_A^+} = \lceil \frac{1-c}{2s} - 2 \rceil = \bar{k} - 1$ all consumers in channel A incur a net loss. The seller's profits from channel A (π_A) are positive but decreasing in k for all $k < \bar{k}$ (and zero for all $k \geq \bar{k}$).*

2.4.3. Aggregate Market (Channel D & Channel A)

Consumer Surplus, Profits, and Welfare

After the separate analysis of both channels we now return to the bigger picture that consolidates the different effects and allows for an overall welfare analysis. Defining $CS_k \equiv CS_{D_k} + CS_{A_k}$, $\pi_k \equiv \pi_{D_k} + \pi_{A_k}$, and $W_k \equiv CS_k + \pi_k$ leads to the following results (combining Equations (2.18) and (2.24) in (2.37), Equations (2.19) and (2.25) in (2.38), and, ultimately, Equations (2.37) and (2.38) in (2.39):

$$CS_k = \frac{(1-\hat{v}_k)^2}{2} - (1-\hat{v}_k)s = \frac{1}{8}(1-c)^2 - \frac{1-c}{2}(k+2)s + \frac{k^2+4k+3}{2}s^2, \tag{2.37}$$

$$\pi_k = \frac{(\hat{v}_k-c)^2}{2} + (1-\hat{v}_k)(\hat{v}_k-c) = \frac{3}{8}(1-c)^2 + \frac{1-c}{2}(k+1)s - \frac{k^2+2k+1}{2}s^2, \tag{2.38}$$

$$W_k = \frac{(1-c)^2}{2} - (1-\hat{v}_k)s = \frac{1}{2}(1-c)^2 - \frac{1-c}{2}s + (k+1)s^2. \tag{2.39}$$

Like total consumer surplus and total profits, total welfare depends on the strategic level of sophistication of consumers and can be identified graphically in Figure 2.2.[19] The first term in Equation (2.39), $\frac{(1-c)^2}{2}$, corresponds to the whole area between the demand curve and the marginal cost curve in Figure 2.1, while the second term, $(1-\hat{v}_k)s$, corresponds to the dashed rectangle. Although the market outcome of Stage 3 is efficient, because every consumer with a valuation $v \geq c$ buys the product, this shows that total welfare is reduced by the losses stemming from consumers' anonymization behavior as long as $\hat{v}_k < 1$ or, equivalently, $k < \bar{k}$. For any $k \geq \bar{k}$, a fully efficient outcome ensues.

Comparative Statics for k for the Aggregate Market

Similarly as for the two channels before, we derive the effects on the aggregated quantities as differences due to the discrete nature of changes in consumer sophistication:

For $k < \bar{k} - 1$:

$$\Delta CS_k \equiv CS_{k+1} - CS_k = -(1-\hat{v}_{k+1})\,s + \frac{s^2}{2} = -\frac{1-c}{2}s + \frac{2k+5}{2}s^2, \tag{2.40}$$

$$\Delta \pi_k \equiv \pi_{k+1} - \pi_k = (1-\hat{v}_k)\,s + \frac{s^2}{2} = \frac{1-c}{2}s + \frac{2k+3}{2}s^2, \tag{2.41}$$

$$\Delta W_k \equiv W_{k+1} - W_k = s^2 = s^2. \tag{2.42}$$

Since consumer surplus from channel D was equal to zero independent of k, the effect of changing k on aggregate consumer welfare is identical to the already identified effect in channel A, i.e. decreasing as k increases until a certain threshold, $\bar{k}_{\Delta CS}$ is reached. Recognizing the similarity in Equation (2.41), we define an additional threshold level of

[19] An in-depth discussion of the terms of Equations (2.37) and (2.38) can be found in the respective channel's discussions.

consumer sophistication where profits stop increasing $\bar{k}_\pi$:

$$\bar{k}_\pi \equiv \min\{k \in \mathbb{Z}_0^+ | \Delta\pi_k \leq 0\}. \tag{2.43}$$

Substituting Equation (2.41) in (2.43), solving for the threshold level, and again following the same line of reasoning to deal with the discreteness of k as before yields:

$$\bar{k}_\pi \leq \frac{1-c}{2s} - \frac{3}{2} \Rightarrow \bar{k}_\pi = \left\lceil \frac{1-c}{2s} - \frac{3}{2} \right\rceil. \tag{2.44}$$

This threshold is compared to the level of sophistication at which the market for anonymization breaks down:

$$\bar{k} - \bar{k}_\pi = \left\lceil \frac{1-c}{2s} - 1 \right\rceil - \left\lceil \frac{1-c}{2s} - \frac{3}{2} \right\rceil = \left\lceil \frac{1-c}{2s} \right\rceil - \left\lceil \frac{1-c}{2s} - \frac{1}{2} \right\rceil \in \{0, 1\}. \tag{2.45}$$

Equation (2.45) indicates that profits stop increasing either at the last level before the breakdown of channel A or when this happens. Recalling, however, that all comparative statics difference equations (and hence also Equation (2.41) which we used in deriving $\bar{k}_\pi$) are only applicable to $k < \bar{k} - 1$, we have to examine this case closer since $\bar{k}_\pi \in \{\bar{k}-1, \bar{k}\}$. Recall further that $\mathcal{C}_D$ increases until $k = \bar{k}$ and that the seller appropriates all surplus from any consumer in channel D, whereas he only receives a share of the surplus generated from the transaction when selling to consumers in channel . It is straightforward to conclude that profits are still increasing when consumers' sophistication changes from $\bar{k} - 1$ to $\bar{k}$. Hence, we have to adjust Equations (2.44) and (2.45) to (2.46) and (2.47), respectively:

$$\bar{k}_\pi = \left\lceil \frac{1-c}{2s} - 1 \right\rceil, \tag{2.46}$$

$$\bar{k} - \bar{k}_\pi = 0. \tag{2.47}$$

While increasing k has negative effects on consumer surplus and positive effects on profits, welfare is generally increasing in k as Equation (2.42) shows (and it, too, does so including the last change from $\bar{k} - 1$ to $\bar{k}$). A threshold cannot even be determined as the change is independent of k. This result is, of course, driven by the fact that increasing the level of sophistication leads to fewer anonymized consumers, corresponding to smaller cost of anonymization, all the while the surplus from the transaction of the good stays constant at the maximum due to perfect price discrimination in channel D (raising k simply shifts the surplus from consumers to the seller). Combining the insights of the previous Lemmas, we summarize the above analysis in the following propositions.

Proposition 2.2 (Consumer Sophistication and Welfare)**.** *For any non-prohibitively high cost of anonymization $s > 0$ and cost of production $c \geq 0$ and any finite level of* *consumer sophistication k, aggregated consumer surplus (CS_k), profits (π_k), and welfare (W_k) exhibit the following characteristics:*

- $CS_k > 0$ *for* $k < \bar{k}_{CS}$, $CS_k \leq 0$ *for* $\bar{k}_{CS} \leq k < \bar{k}$, *and* $CS_k = 0$ *for* $k \geq \bar{k}$, *where* $\bar{k}_{CS} = \lceil \frac{1-c}{2s} - 3 \rceil$ *and* $\bar{k} - \bar{k}_{CS} = 2$.
- $\pi_k > 0$ *for* $k < \bar{k}$, *and* $\pi_k = W_k$ *for* $k \geq \bar{k}$.
- $W_k > 0$ *for* $k < \bar{k}$, *and* $W_k = W^*$ *for* $k \geq \bar{k}$, *where* $W^* = \frac{(1-c)^2}{2}$ *is the first-best outcome.*

Proposition 2.3 (Effects of Changing Consumer Sophistication)**.** *Raising the level of strategic sophistication of consumers from k to $k+1$ has the following effects on consumer surplus, profits, and welfare (ceteris paribus):*

- $\Delta CS_k < 0$ *for* $k < \bar{k}_{\Delta CS}$, $\Delta CS_k \geq 0$ *for* $\bar{k}_{\Delta CS} \leq k < \bar{k}$, *and* $\Delta CS_k = 0$ *for* $k \geq \bar{k}$, *where* $\bar{k}_{\Delta CS} = \lceil \frac{1-c}{2s} - \frac{5}{2} \rceil$ *and* $\bar{k} - \bar{k}_{\Delta CS} \in \{1, 2\}$.
- $\Delta\pi_k > 0$ *for* $k < \bar{k}$, *and* $\Delta\pi_k = 0$ *for* $k \geq \bar{k}$.
- $\Delta W_k > 0$ *for* $k < \bar{k}$, *and* $\Delta W_k = 0$ *for* $k \geq \bar{k}$.

Corollary 2.1 (Positive and Negative Individual Surplus)**.** *As long as consumers are not* too *sophisticated ($k < \bar{k}_{\mathcal{C}_A^+} = \lceil \frac{1-c}{2s} - 2 \rceil = \bar{k} - 1$), some consumers who anonymize themselves (those in $\mathcal{C}_A^+$) end up with positive net surplus, whereas others (those in $\mathcal{C}_A^-$) end up with negative net surplus.*

Proposition 2.3 predicts that (except for boundary cases) the strategic sophistication of consumers will work to their disadvantage at an aggregated level and can break down the market for anonymous shopping. By contrast, the seller benefits from this stepwise breakdown, a development that would also be appreciated by a total welfare maximizer. The reason for this preference is, interestingly, not based on allocation: Due to perfect price discrimination in the direct channel, all consumers with a valuation for the product above its marginal cost of production get the product independent of the existence of the anonymous channel. Instead, given that big data technologies driving channel D are already in place, it is inefficient to incur the additional cost of anonymization. The smaller channel A, the smaller this inefficiency.

Corollary 2.1 zooms into the set of consumers who choose to invest in their anonymization. See Figure 2.2, where three groups of consumers have been distinguished: $\mathcal{C}_D$, $\mathcal{C}_A^-$, and $\mathcal{C}_A^+$. While the first denotes those consumers who chose channel D, the superscript at the two remaining groups distinguishes those consumers in channel A who make a net loss (because they do not anticipate the seller's incentive to increase the price by s fully) from those ending up with a net benefit of the whole transaction.

Comparative Statics for s

Apart from the effects attributable to changes in strategic sophistication of consumers, we also analyze the effects of changes in the cost of anonymization. This analysis may inform whether policy efforts to make anonymizing techniques available at lower cost are desirable. Before delving into the effects on consumer surplus, profits and welfare, it is useful to first identify the threshold when anonymization becomes prohibitively costly for channel A to be used by any consumer (the equivalent of $\bar{k}$) as our analysis will be limited by this upper bound. Denoting this threshold by $\bar{s}$, we define:

$$\bar{s} \equiv \min\{s \in \mathbb{R}_0^+ | \mathcal{C}_A = \emptyset\}. \tag{2.48}$$

Recalling that $\mathcal{C}_A = (\hat{v}_k; 1]$ and $\hat{v}_k = \frac{1+c}{2} + (k+1)s$, yields:

$$\mathcal{C}_A = \emptyset \Leftrightarrow \hat{v}_k \geq 1 \Leftrightarrow s \geq \frac{1-c}{2}\frac{1}{(k+1)} \Rightarrow \bar{s} = \frac{1-c}{2}\frac{1}{(k+1)}. \tag{2.49}$$

Since s is a continuous variable, we do not need to take further special cases into account and the effects on consumer surplus, profits, and welfare are therefore found by derivatives (instead of differences) using Equations (2.37), (2.38), and (2.39):

For $s \leq \bar{s}$:

$$\frac{\partial CS_k}{\partial s} = -\frac{1-c}{2}(k+2) + (k+1)(k+3)s = (k+1)s - (k+2)(1-\hat{v}_k), \tag{2.50}$$

$$\frac{\partial \pi_k}{\partial s} = \frac{1-c}{2}(k+1) - (k+1)^2 s = (k+1)(1-\hat{v}_k), \tag{2.51}$$

$$\frac{\partial W}{\partial s} = -\frac{1-c}{2} + 2(k+1)s = (k+1)s - (1-\hat{v}_k). \tag{2.52}$$

Defining thresholds in a similar way as in our analysis of the effects of raising consumer sophistication and limiting the analysis to non-prohibitive cost of anonymization, yields:

For $s \leq \bar{s}$:

$$\frac{\partial CS_k}{\partial s} \begin{cases} < 0 & \text{if } s < \frac{1-c}{2}\frac{k+2}{(k+1)(k+3)} \equiv \bar{s}_{CS}, \\ \geq 0 & \text{if } s \geq \frac{1-c}{2}\frac{k+2}{(k+1)(k+3)} \equiv \bar{s}_{CS}, \end{cases} \tag{2.53}$$

$$\frac{\partial \pi_k}{\partial s} \begin{cases} > 0 & \text{if } s < \frac{1-c}{2}\frac{1}{k+1} \equiv \bar{s}_{\pi}, \\ \leq 0 & \text{if } s \geq \frac{1-c}{2}\frac{1}{k+1} \equiv \bar{s}_{\pi}, \end{cases} \tag{2.54}$$

$$\frac{\partial W_k}{\partial s} \begin{cases} < 0 & \text{if } s < \frac{1-c}{2}\frac{1}{2k+2} \equiv \bar{s}_{W}, \\ \geq 0 & \text{if } s \geq \frac{1-c}{2}\frac{1}{2k+2} \equiv \bar{s}_{W}. \end{cases} \tag{2.55}$$

It can further be shown that

$$\bar{s} = \bar{s}_{\pi} > \bar{s}_{CS} > \bar{s}_{W}, \tag{2.56}$$

which reveals that the seller's profits increase in s until the prohibitive level $\bar{s}$ is reached. As it becomes more costly to anonymize, more consumers will choose channel D instead of channel A, where the seller appropriates their entire valuation as profit. The effects on consumer surplus and total welfare, on the other hand, are ambiguous depending on the initial level of s. This is due to the changing effects of raising s on the composition of $\mathcal{C}_A$: At first, $\mathcal{C}_A^-$ increases in size as s increases, but when there are no consumers in $\mathcal{C}_A^+$ anymore, a further increase will reduce the size of $\mathcal{C}_A^-$ again. Taking the respective second derivatives provides further insights and yields:

$$\frac{\partial^2 CS_k}{\partial s^2} = (k+1)(k+3) > 0, \tag{2.57}$$

$$\frac{\partial^2 W_k}{\partial s^2} = 2(k+1) \qquad > 0. \tag{2.58}$$

Thus, consumer surplus as well as total welfare are convex in s, implying that both reach a local maximum at $s = \bar{s}$. Moreover, on the lower end of the distribution, i.e. for $s \to 0$, both consumer surplus and total welfare have a supremum, not a maximum, owing to the fact that we defined $s > 0$ above. Relaxing this constraint for the remainder of this section allows us to study the case of costless anonymization where $s = 0$.[20]

Substituting $s = 0$ and $s = \bar{s}$ in Equations (2.37), (2.38), and (2.39) yields consumer surplus, profits, and welfare at either extreme case:

$$CS_k(s=0) = \frac{1}{8}(1-c)^2, \qquad CS_k(s=\bar{s}) = 0, \tag{2.59}$$

$$\pi_k(s=0) = \frac{3}{8}(1-c)^2, \qquad \pi_k(s=\bar{s}) = \frac{1}{2}(1-c)^2, \tag{2.60}$$

$$W_k(s=0) = \frac{1}{2}(1-c)^2, \qquad W_k(s=\bar{s}) = \frac{1}{2}(1-c)^2. \tag{2.61}$$

For $s = 0$, the difference between consumers' expectations and seller's optimal price disappears. Hence, failing to correctly anticipate the seller's reaction to their anonymization decision does not matter anymore. No consumer in channel A incurs a net loss. As the seller optimally sets the price that consumers expect, there is no change in expectations with increasing consumer sophistication. Hence, if anonymization is costless, the seller's advantage from big data technologies is irrelevant for consumers with high valuation. Those with low valuation get the product but all surplus is extracted by the seller.

[20] If we had assumed $s \geq 0$ already in Section 2.2, we would have had to distinguish among cases for $s > 0$ and $s = 0$ throughout the analysis, sacrificing clarity.

Lemma 2.7 (Equilibrium with Level-k and Costless Anonymization). *With costless anonymization, $s = 0$, and for any non-prohibitive cost of production $c \geq 0$, there is a unique equilibrium with the following characteristics:*

- *Consumers form the k-beliefs $\mathbb{E}_k(p_D) = p_i^*(v)$ and $\mathbb{E}_k(p_A) = p_M = \frac{1+c}{2}$.*
- *Consumers anonymize if, and only if, $v > \hat{v}_k = p_M$, separating into the sets $\mathcal{C}_D = [0, p_M]$ and $\mathcal{C}_A = (p_M, 1]$.*
- *The seller forms the k+1-beliefs $\mathbb{E}_{k+1}(\mathcal{C}_D) = [0, p_M]$ and $\mathbb{E}_{k+1}(\mathcal{C}_A) = (p_M, 1]$.*
- *The seller sets $p_i^*(v) = \max\{v, c\}$ and $p_{A_0}^* = p_M$.*
- *All consumers in $\mathcal{C}_D$ with $v \geq c$ buy the product at the price offered to them.*
- *All consumers in $\mathcal{C}_A$ buy the product at the price offered to them.*

Additionally, channel A does not break down for any level of consumer sophistication and an efficient outcome ensues irrespective of k. Including this boundary case $s = 0$ in our analysis of the effects of anonymization cost, the comparative statics analysis for s is summarized in Proposition 2.4 and visualized in Figure 2.3.

Proposition 2.4 (Anonymization Cost and Welfare). *For any non-prohibitively high level of consumer sophistication $k < \bar{k}$ and cost of production $c \geq 0$, anonymization is prohibitively costly for $s \geq \bar{s} = \frac{1-c}{2}\frac{1}{k+1}$. Then, channel A remains unused. As long as channel A is used, aggregated consumer surplus (CS_k), profits (π_k), and welfare (W_k) exhibit the following characteristics:*

- *CS_k is maximal at $s = 0$, decreases in s to its minimum (which is negative) at $s = \bar{s}_{CS}$, then increases in s back to 0 at $s = \bar{s}$.*
- *π_k is minimal at $s = 0$ and increases in s to its maximum at $s = \bar{s}_{CS} = \bar{s}$.*
- *W_k is maximal at $s = 0$, decreases in s to its minimum at $s = \bar{s}_W$, then increases in s to another maximum at $s = \bar{s}$. Both maxima lead to the first-best outcome W_k^*.*

Proposition 2.4 shows that higher cost of anonymization is negative for consumers despite the fact that consumer surplus increases in s for relatively high values, which becomes apparent from the fact that consumer surplus is maximal when anonymization is costless. The seller, on the other hand, unambiguously benefits from higher cost of anonymization and would prefer a prohibitively high cost of $s = \bar{s}$ as he maximizes his profits when he can perfectly discriminate and extract the entire surplus from all consumers. A total welfare maximizer, focusing on the welfare-deteriorating role of s,

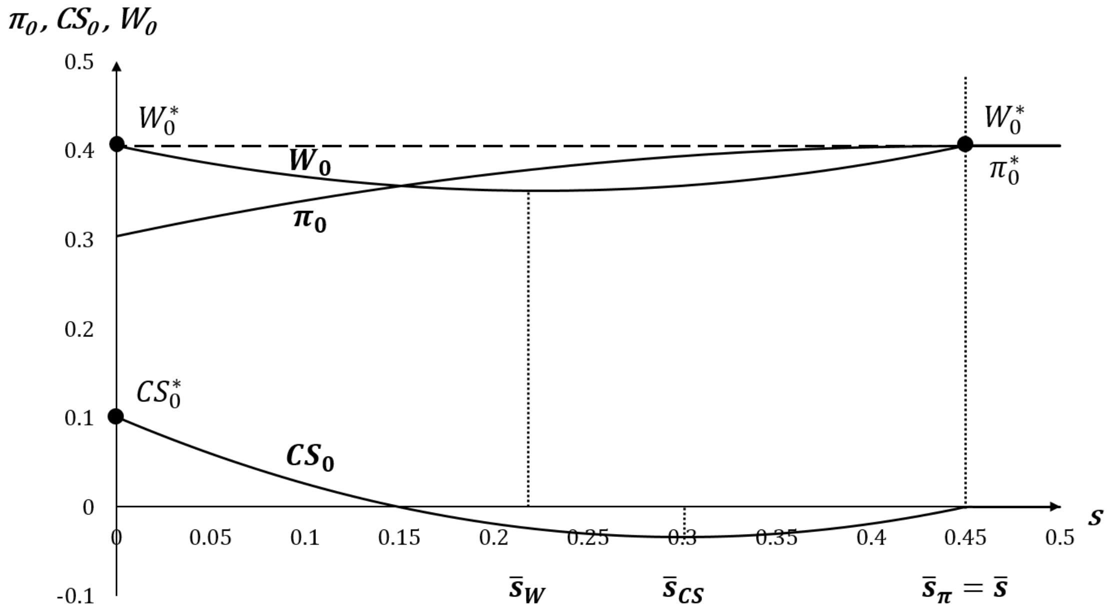


Figure 2.3: Consumer Surplus, Profits and Welfare as Functions of s with Parameters $v \sim \mathcal{U}[0, 1]$, $c = 0.1$, $k = 0$

can choose either extreme to prevent consumers from incurring the cost: To achieve an efficient outcome, anonymization should be costless ($s = 0$) or prohibitively costly ($s = \bar{s}$). Both options are welfare-maximizing, but lead to different allocations of the surplus generated by the market. Note that, while the seller makes positive profits in either welfare-maximizing scenario, consumers only receive positive surplus if $s = 0$.

2.5. Alternative Model Specifications

2.5.1. Beliefs of "Naïve" Consumers

In our model we have assumed that "naïve" consumers will expect the price in channel A to be equal to the unconditional monopoly price p_M. Many other applications of level-k thinking employ a random distribution as a starting point for players with $k = 0$. If the "naïve" consumers in our model were to make their anonymization decision randomly rather than to react to a belief of facing the unconditional monopoly price, the seller would correctly infer this and set the price accordingly. Depending on the location of the valuation of anonymized consumers expected by the seller, three cases can be distinguished:

(a) The expected valuation of anonymized consumers is *below* the unconditional monopoly price, i.e. $\mathbb{E}(v|i \in \mathcal{C}_A) < p_M$.

(b) The expected valuation of anonymized consumers is *equal to* the unconditional monopoly price, i.e. $\mathbb{E}(v|i \in \mathcal{C}_A) = p_M$.

(c) The expected valuation of anonymized consumers is *above* the unconditional monopoly price, i.e. $\mathbb{E}(v|i \in \mathcal{C}_A) > p_M$.

These cases are equivalent to the cases for the anonymization threshold $\hat{v}$ in Section 2.3. As discussed there, the seller's best response in cases (a) and (b) is to charge the unconditional monopoly price p_M. This would require our analysis to include one additional first step of strategic iteration, such that consumers would expect p_M for channel A if they had a level of strategic sophistication of $k = 1$. In case (c), however, the seller's best response is to charge a price equal to $\mathbb{E}(v|i \in \mathcal{C}_A) + s$, essentially responding in the same way as before by increasing the price by s above the cutoff valuation. Depending on the exact distance from p_M, this would reduce the number of steps until the complete breakdown of channel A, but not change the underlying mechanism of iterations from there onwards. Hence, while the choice of p_M as a starting point for our analysis pins the analysis to a particular point, it does not crucially affect the model's analysis.

If anonymization is costless, however, changing the belief of naïve consumers has a larger influence. As the iterative process is suspended, expectations do not change after the initial change from $k = 0$ to $k = 1$, which only changes the seller's best response in case (a), but does so in the same fashion as discussed above for $s > 0$. This allows for any price $p_A \in [p_M, 1]$ to be expected by consumers and to be set by the seller—and hence to become the threshold valuation $\hat{v}$ that influences the surplus distribution. Thus, while the resulting equilibrium is not necessarily efficient for $k = 0$ anymore, it is for any $k \geq 1$ and hence does not constitute a crucial departure from our model either.

2.5.2. Heterogeneous Cost of Anonymization

Our model further assumes that all consumers find it equally costly to use the anonymous channel A. However, it is easy to imagine that some people might find it less cumbersome to discover and make use of privacy-protecting technologies such as deleting cookies, activating "do not track" browser options or installing various plugins.[21] Additionally, heterogeneity in s can stem from differing exogenous tastes for privacy in

[21] For instance, TOR is a "free software and an open network that helps [users] defend against traffic analysis, a form of network surveillance that threatens personal freedom and privacy, confidential business activities and relationships, and state security" (https://www.torproject.org/). A more detailed list can be found in Sellenart's "A paranoid's toolbox for browsing the web": http://pierre.senellart.com/talks/cerre-20160915.pdf (Sellenart 2016, 20).

the consumer population which would then reduce the experienced disutility of a possibly still fixed cost of using channel A. A heterogeneous distribution of s could be seen as the result of aggregating both effects. Assuming that consumers have a heterogeneous cost of anonymization s_i, where $s_i \sim \mathcal{U}[\underline{s}, \bar{s}]$ (we redefine $\bar{s}$ for this section), changes consumers' anonymization stage decision to $v > \mathbb{E}_k(p_A) + s_i$. As this expression now depends on two individually heterogeneous variables, there is no uniform cutoff valuation $\hat{v}$ separating the sets $\mathcal{C}_D$ and $\mathcal{C}_A$. Rather three segments of consumers' valuations v need to be distinguished (cf. Figure 2.4) to get to the composition of $\mathcal{C}_D$ and $\mathcal{C}_A$:

(a) Consumers with $v > \mathbb{E}_k(p_A) + \bar{s}$,

(b) Consumers with $v \in (\mathbb{E}_k(p_A) + \underline{s}, \mathbb{E}_k(p_A) + \bar{s}]$,

(c) Consumers with $v \leq \mathbb{E}_k(p_A) + \underline{s}$.

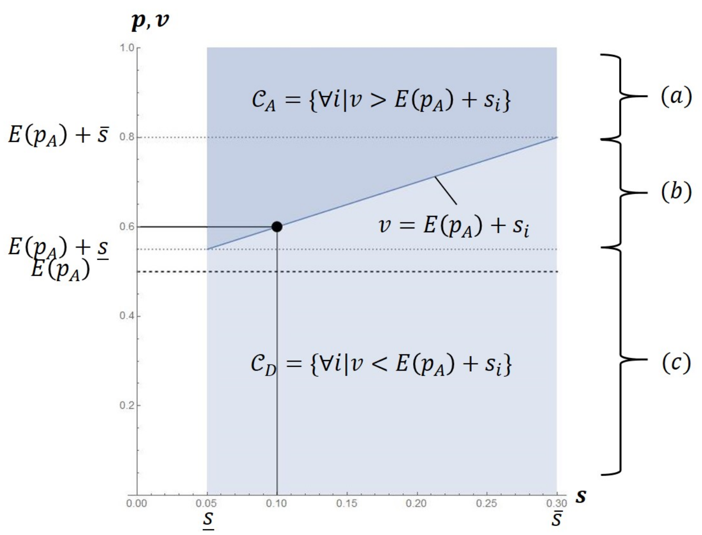


Figure 2.4: Composition of Sets $\mathcal{C}_D$ and $\mathcal{C}_A$ Depending on v and s_i with Parameters $v \sim \mathcal{U}[0, 1]$, $s_i \sim \mathcal{U}[0.05, 0.3]$, $c = 0$, $k = 0$

Given their price expectation, consumers in segment (c) have a valuation v so low that they choose channel D even for the lowest possible cost of anonymization $\underline{s}$. Vice versa, consumers in segment (a) have a sufficiently high valuation v such that they choose channel A even if they face the highest possible cost of anonymization $\bar{s}$. For consumers in segment (b), however, the precise level of their anonymization cost s_i matters for their anonymization choice. Given any valuation $v \in (\mathbb{E}_k(p_A) + \underline{s}, \mathbb{E}_k(p_A) + \bar{s})$ only those whose cost of anonymization s_i is sufficiently low will choose channel A, while others with the same valuation v but a higher cost s_i will choose channel D. Figure 2.4 exemplifies this for consumers with a valuation of $v = 0.6$. At $v = 0.6$ only consumers with anonymization cost $s_i < 0.1$ choose channel A, whereas consumers with $v = 0.6$ but $s_i \geq 0.1$ choose channel D.

The new composition of the sets $\mathcal{C}_D$ and $\mathcal{C}_A$ implies that demand in both channels is now defined differently for all three segments and hence becomes a piecewise (but still continuous) function of p. As the seller will still perfectly price-discriminate in channel D, we focus on the implications of this change for channel A. There, demand is still linear for prices in segment (a) and constant for prices in segment (c). For price levels in segment (b), however, the uniformly distributed cost of anonymization leads to quadratic demand. For the parameters of Figure 2.4, this leads to the following demand in channel A:

$$q_A(p_A) = \begin{cases} 0 & \text{if } p_A \geq 1, \\ 1 - p_A & \text{if } 0.8 < p_A < 1, \\ 1 - 0.625 - \frac{(p_A - 0.55)^2}{0.5} & \text{if } 0.55 < p_A \leq 0.8, \\ 1 - 0.625 & \text{if } p_A \leq 0.55. \end{cases} \tag{2.62}$$

Due to demand being defined piecewise a closed-form solution for p_A^* for all k is not easily found. The general derivation of demand (see Appendix 2.A.1) confirms that demand is always quadratic for $p \in (\mathbb{E}[p_A] + \underline{s}, \mathbb{E}[p_A] + \bar{s}]$. However, we show in Appendix 2.A.3 that $p_{A_k}^* > \mathbb{E}_k(p_A) + \underline{s}$ as long as $\mathcal{C}_A \neq \emptyset$, i.e. the seller's optimal price is no longer equal to the valuation at the lower bound of $\mathcal{C}_A$. This in turn implies that there are some consumers in $\mathcal{C}_A$ that do not buy the product.

Continuing the example from above, this result is illustrated in Figure 2.5. There, the left panel depicts the demand function $q_A(p_A)$ as well as the optimal price in channel A, given by $p_A^* = 0.6629$ with the chosen parameters, whereas the right panel replicates Figure 2.4 to highlight the mapping from $\mathcal{C}_A$ to $q_A(p_A)$), but replacing the example point with the optimal price p_A^*. Both panels of Figure 2.5 show that the optimal price p_A^* now exceeds the lowest valuation in $\mathcal{C}_A$. Now, consumers the white area between $q_A(p_A^*)$ and $q_A(p_A)$ are not buying the product, despite having a valuation of at least p_M (the monopoly price without price-discrimination). Contrary to our main specification, the seller is now willing to forgo profits from some consumers, because the density of consumers in $\mathcal{C}_A$ across valuations is not uniform in the neighborhood of the lower bound of $\mathcal{C}_A$ anymore.

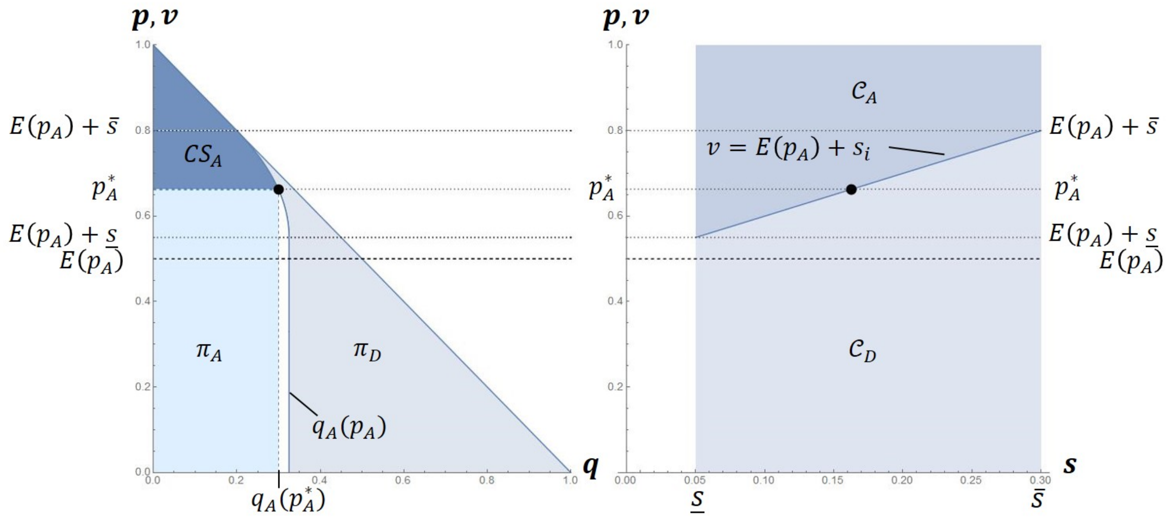


Figure 2.5: Consumers' Anonymization Choice as a Function of v and s_i with Parameters $v \sim \mathcal{U}[0,1]$, $s_i \sim \mathcal{U}[0.05, 0.3]$, $c = 0$, $k = 0$

If we increase consumers' strategic sophistication from $k = 0$ to $k = 1$, consumers form the expectation $\mathbb{E}_1(p_A) = 0.6629$ and make their anonymization choice accordingly. The difficulty of finding an analytical closed-form solution for p_A^* for any k transmits to finding the level $\bar{k}$ from which onwards channel A remains unused. Figure 2.6 illustrates that the profit-maximizing prices p_A^* in our example do not increase linearly in k as was the case in our main model specification (cf. Equation (2.13)).

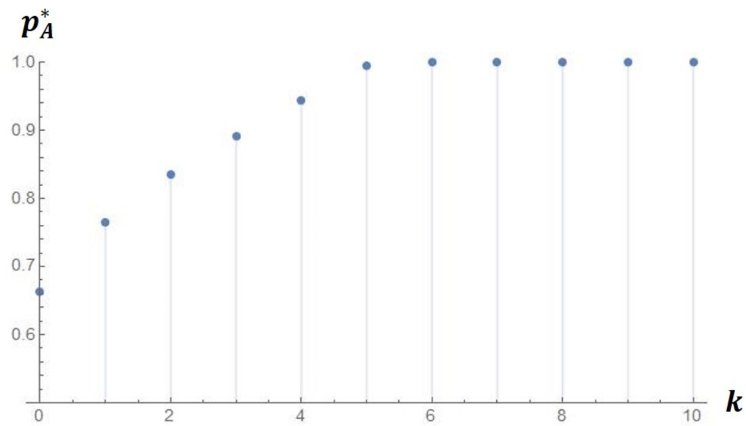


Figure 2.6: Optimal Price in Channel A p_A^* for $k = 0, 1, 2, \ldots, 10$ with Parameters $v \sim \mathcal{U}[0,1]$, $s_i \sim \mathcal{U}[0.05, 0.3]$, $c = 0$

The qualitative result, that channel A is used less for higher levels of consumers' strategic sophistication and eventually remains unused, however, is replicated: Once the optimal price falls within the interval $(0.95, 1]$, channel A remains unused at the next higher level of sophistication. Note, though, that this only holds as long as anonymization is at least somewhat costly for all consumers, i.e. if $\underline{s} > 0$. If anonymization comes for free to some consumers, i.e. if $\underline{s} = 0$, we show in Appendix 2.A.4 that $p^*_{A_k} < 1$ for all finite k, i.e. that a full breakdown of channel A is not achieved for a finite k anymore.[22]

Summing up, we conclude that despite some quantitative changes, the general pattern of a gradual breakdown of the anonymous channel and the corresponding effects are not altered qualitatively by assuming heterogeneous anonymization cost.

2.5.3. Increasing Competition

Many markets where sellers have access to large amounts of data on consumers' preferences or characteristics, a prerequisite for perfect price discrimination, are dominated by one firm.[23] But to which extent would such a dominant firm, or a monopolist in a market niche, adapt behavior if consumers had access to an (imperfect) substitute product, thereby increasing competition? Assume a rival R offers a product competing with the monopolist's product. Consumer i's net value of the rival's product is

$$v^R \equiv \sigma v - p_R, \tag{2.63}$$

where $\sigma \in [0, 1)$ denotes the degree of substitutability of the products and p_R denotes the price of the rival's product. Alternatively, σ proxies the intensity of competition. As any $p_R > 0$ can be reflected by a lower intensity of competition σ, we assume $p_R = 0$ and focus on changes in σ to study the effects of increasing competition. In this scenario, consumers buy from the "monopolist" M if, and only if, $v - p \geq v^R$, i.e. if

$$v - v^R = (1 - \sigma)v \geq p, \tag{2.64}$$

where $p \in \{p_i, p_A\}$. Knowing v precisely for all consumers in $\mathcal{C}_D$, M will set

$$p^*_i = \max\{(1 - \sigma)v, c\} \text{ for all } i \in \mathcal{C}_D, \tag{2.65}$$

thus guaranteeing that no consumer in channel D will buy the rival's product as long as M can earn a profit from this consumer. Consumers anonymize if $\mathbb{E}(u(A)) > \mathbb{E}(u(D))$, i.e. if, and only if,

$$v - \mathbb{E}_k(p_A) - s > v - p^*_i \Leftrightarrow \tag{2.66}$$

$$v > \frac{\mathbb{E}_k(p_A) + s}{1 - \sigma} \equiv \tilde{v}_k, \tag{2.67}$$

[22] This in turn would reinstate unlimited sophistication as a *necessary* condition for a complete breakdown of channel A in the particular case of $\underline{s} = 0$.

[23] Prüfer and Schottmüller (2017) explain this development in theoretical terms and cite empirical work to support the statement.

where $\tilde{v}_k$ denotes the cutoff valuation in this modified version of our model (corresponding to $\hat{v}_k$ in the baseline model). Recalling that $\hat{v}_k = \mathbb{E}_k(p_A) + s$ leads to the following lemma.

Lemma 2.8 (Anonymization and Monopolistic Competition)**.** *For any given price expectation of consumers for channel A, $\mathbb{E}_k(p_A)$, the presence of a rival selling a product of substitutability $\sigma \in [0, 1)$ raises the anonymization threshold from $\hat{v}_k$ to $\tilde{v}_k = \frac{\hat{v}_k}{1-\sigma}$.*

Understanding this increase in the anonymization threshold, M might consider to also increase his price, as in the baseline model, and set $p^*_{A_k} = \tilde{v}_k$. Recall that also consumers in channel A might still buy from R. Thus, M faces the same participation constraint in channel A as in channel D: to leave every consumer with at least a net surplus of σv. It follows that pricing at $\tilde{v}_k$ is infeasible. M has to decrease the price to fulfill:

$$(1-\sigma)\tilde{v}_k \geq p^*_{A_k} \tag{2.68}$$

which yields, in combination with Equation (2.67),

$$p^*_{A_k} = (1-\sigma)\frac{\hat{v}_k}{1-\sigma} = \hat{v} = \mathbb{E}_k(p_A) + s. \tag{2.69}$$

Thus, the seller cannot capitalize on the increased anonymization threshold as a consequence of increased competition. Even though every consumer in channel A will now end up with a positive net surplus from the entire transaction (depending on the size of s compared to the guaranteed benefit of σv), some consumers will still be worse off, having chosen channel A instead of channel D. Consumers with $k = 1$, however, will not adjust their expectation based on forgone surplus but simply update their price expectation to the price that would be M's best response to consumers with $k = 0$; just as in the baseline model. Summarizing, consumer surplus increases with competitive pressure. Therefore, there is less anonymization for any given price expectation. The prices in channel D decrease to account for consumers' improved outside option but the price in channel A is unaffected by competition.

2.6. Discussion and Conclusion

This paper started from the empirical observation that the technological developments that are alluded to as the "rise of big data" or "datafication" have led to asymmetries on markets for consumer goods (Mayer-Schönberger and Cukier 2013). Sellers making use of huge datasets with information on choices of large masses of consumers can tailor prices to individual characteristics and thereby appropriate a large share of the surplus created by market transactions. On top of the sheer amount of information that is available to sellers, consumers are at a second disadvantage. They face cognitive constraints regarding strategic sophistication (Acquisti and Grossklags 2007), while the seller's data processing capabilities enable him to find best responses to consumers' behavior.

In this paper, we have taken these developments seriously and constructed a model to study their implications on prices, consumption choices, and consumers' incentives to use anonymization technologies protecting their privacy. We have shown that under certain conditions, most notably under the assumption of imperfect strategic sophistication of consumers, a costly privacy-protective sales channel is used even if consumers do not have an explicit taste for privacy. In our model, consumers want to restore their privacy (i.e. choose channel A) solely based on their valuation of the good and their price expectation. We thereby provide a micro-foundation for consumers' privacy choices, to which the existence of a privacy-protective sales channel can cater.

Our model showed that unlimited strategic sophistication is a sufficient but not a necessary condition for the breakdown of the anonymous sales channel if anonymization is equally costly to all consumers. Allowing for heterogeneity in anonymization cost, sources of which can be different technological savviness but also differing preferences for privacy, would reinstate the necessity of unlimited strategic sophistication for a complete breakdown of the anonymous channel, though.

In general, the use of big data technologies by sellers improves total welfare by avoiding the dead weight loss usually associated with a monopoly: In contrast to uniform monopoly pricing, consumers with low valuations, $v < p_M$, can purchase the product now. This increases efficiency but not consumer surplus as the seller appropriates the entire surplus from these additional transactions. We have further shown that using the anonymous channel backfires and leads to a net loss for at least some (and under certain conditions all) anonymized consumers (forming the set $\mathcal{C}_A^-$).

We have shown that increasing consumer sophistication leads to a reduction in consumer surplus but to an increase in profits and total welfare. Given that the level of strategic sophistication, k, may be regarded as exogenous for policy makers, our results concerning the cost of anonymization, s, might be more policy relevant. Consumer surplus is largest in the extreme case of costless anonymization, $s = 0$, and profits are maximal in the extreme case of anonymization being prohibitively costly, $s = \bar{s}$.

Total welfare, however, is maximal at either extreme, $s = 0$ or $s = \bar{s}$. The two fully efficient cases differ, though, in the way in which they ensure a first-best result. If $s = 0$, consumers with high valuation, $v > p_M$, anonymize for free and receive positive surplus, whereas all others choose the direct channel, where they get perfectly discriminated against and are left with zero surplus. If $s = \bar{s}$, consumers choose the direct channel irrespective of their valuation for the good. Hence in this case, efficiency is brought about by the fact that anonymization is simply too costly and consumers rather leave the entire surplus to the seller than to incur a net loss from anonymizing.

Which of these two extremes should be favored therefore crucially depends on the objective function of a possibly intervening authority. A consumer-oriented welfare maximizer will try to eliminate anonymization cost, whereas a seller-oriented welfare maximizer will seek to increase the cost of anonymization to a prohibitive level.

A policy maker with a preference for consumer surplus could, for example, require marketers and online platforms serving as matchmakers between sellers and buyers of consumer goods to set anonymous shopping technologies as *default*. This would then require consumers to *opt in* to non-anonymous shopping instead of today's standard, where full tracking of consumers' choices is the default and a few providers offer *opt out* technologies. This proposal is also discussed by Acquisti et al. (2016). Thereby, those consumers who find it in their interest to reveal their true characteristics to sellers (those with low valuations for a given product) would log in to some service and receive the product for a price equalizing their willingness-to-pay. Consumers with higher valuation would stay with the (now default) anonymous channel and pay a higher price for the product, but still retain some surplus.

On the theory side, future research could shed light on the effects of heterogeneity in the level of strategic sophistication amongst consumers, relying on a more elaborate cognitive hierarchy model than this first attempt we undertook here. This is a complex undertaking, however, because it is not only necessary to specify a distribution of k-levels across the population of consumers (and how it might be related to their willingness-to-pay). It also requires to specify anew every consumer's belief about other consumers' level(s) of sophistication and the seller's response to them as a function of that consumer's own sophistication.

As regards empirical testing of our theory, we consider it most promising to conduct laboratory experiments because subjects can be assigned certain valuations which are then known and a discriminatory pricing algorithm can actually be implemented. Moreover, laboratory experiments are less susceptible to noise in the elicitation of subjects' levels of strategic sophistication than field experiments. Subjects can indicate their respective anonymization choice given a valuation and a known cost of anonymization. The implied thresholds for anonymization correspond to a certain level of strategic sophistication according to our model, which in turn could be compared to other measures of strategic sophistication spawned from the level-k literature. Using measures that capture differences between belief interactions of subjects could inform whether the current model, which neglects more complex cognitive hierarchies, is a fair representation of subjects' approach to such a market or whether efforts to generalize our theory are needed.

Appendix 2.A Further Results for Heterogenous Cost of Anonymization

2.A.1 General Form of Demand in Channel A with $s_i \sim \mathcal{U}[\underline{s}, \bar{s}]$

Assume a monopolistic seller producing at constant marginal cost $c \geq 0$ and consumers valuing the good heterogeneously with v, where $v \sim \mathcal{U}[0, \bar{v}]$. We further assume that consumers face heterogeneous anonymization cost $s_i \sim \mathcal{U}[\underline{s}, \bar{s}]$, where $0 \leq \underline{s} < \bar{s} \leq \bar{v}$, if they choose channel A, and zero cost, if they choose channel D. Consumers choose channel A if and only if $v > \mathbb{E}_k(p_A) + s_i$. As long as $\mathcal{C}_A \neq \emptyset$ (i.e. $\bar{v} > \mathbb{E}_k(p_A) + \underline{s}$), three cases can be distinguished, depending on the relation of the maximum valuation $\bar{v}$ to $\mathbb{E}_k(p_A)$, $\underline{s}$, and $\bar{s}$:

(1) $\bar{v} > \mathbb{E}_k(p_A) + \bar{s}$,

(2) $\bar{v} = \mathbb{E}_k(p_A) + \bar{s}$,

(3) $\mathbb{E}_k(p_A) + \bar{s} > \bar{v} > \mathbb{E}_k(p_A) + \underline{s}$.

These cases are depicted in Figure 2.7. The area of $\mathcal{C}_A$, i.e. maximal demand in channel A, for all three cases is given by:

$$q_A^{max} = \frac{(\bar{x} + \bar{y})}{2} \cdot \frac{\bar{z}}{\bar{s} - \underline{s}}, \tag{2.70}$$

with

$$\bar{x} = \max\{\bar{v} - [\mathbb{E}_k(p_A) + \underline{s}], 0\}, \tag{2.71}$$

$$\bar{y} = \max\{\bar{v} - [\mathbb{E}_k(p_A) + \bar{s}], 0\}, \tag{2.72}$$

$$\bar{z} = \max\{\min\{\bar{v}, [\mathbb{E}_k(p_A) + \bar{s}]\} - [\mathbb{E}_k(p_A) + \underline{s}], 0\}. \tag{2.73}$$

While $\bar{x}$ and $\bar{y}$ measure the range of valuations, $\bar{z}$ measures the share of consumers present in channel A at any given valuation, which requires the normalization by the factor $\frac{1}{\bar{s}-\underline{s}}$. Maximal demand is achieved for any $p_A \leq \mathbb{E}_k(p_A) + \underline{s}$. However, any $p_A > \mathbb{E}_k(p_A) + \underline{s}$ reduces demand in channel A such that q_A^{max} is reduced by q_A^{sub}, where

$$q^{sub}(p_A) = \frac{(x(p_A) + y(p_A))}{2} \cdot \frac{z(p_A)}{\bar{s} - \underline{s}}, \tag{2.74}$$

with

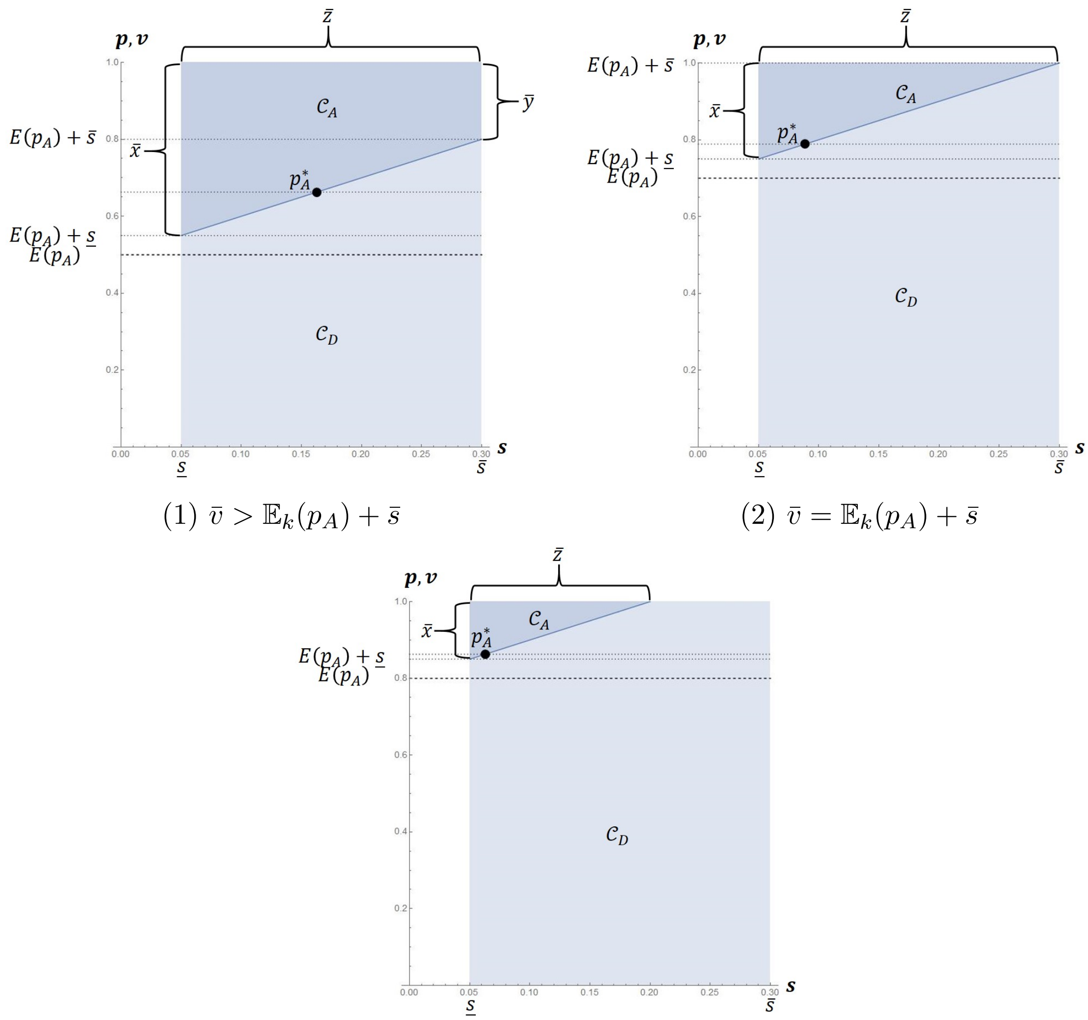
$$x(p_A) = \max\{p_A - [\mathbb{E}_k(p_A) + \underline{s}], 0\}, \tag{2.75}$$

$$y(p_A) = \max\{p_A - [\mathbb{E}_k(p_A) + \bar{s}], 0\}, \tag{2.76}$$

$$z(p_A) = \max\{\min\{p_A - [\mathbb{E}_k(p_A) + \underline{s}], \max\{\min\{\bar{v}, \mathbb{E}_k(p_A) + \bar{s}\} - [\mathbb{E}_k(p_A) + \underline{s}], 0\}\}, 0\}. \tag{2.77}$$

Demand in channel A is then given by subtracting (2.74) from (2.70) and takes the following general form (matched to the segments in Figure 2.4 in the main manuscript):

$$q(p_A) = q_A^{max} - q^{sub}(p_A) = \begin{cases} 0 & \text{for } p_A \geq \bar{v}, \\ \bar{v} - p_A & \text{for } p_A \text{ in (a)}, \\ \bar{v} - [\mathbb{E}_k(p_A) + \frac{\bar{s}+\underline{s}}{2}] - \frac{(p_A - [\mathbb{E}_k(p_A)+\underline{s}])^2}{2(\bar{s}-\underline{s})} & \text{for } p_A \text{ in (b)}, \\ \bar{v} - [\mathbb{E}_k(p_A) + \frac{\bar{s}+\underline{s}}{2}] & \text{for } p_A \text{ in (c)}. \end{cases} \tag{2.78}$$



(1) $\bar{v} > \mathbb{E}_k(p_A) + \bar{s}$

(2) $\bar{v} = \mathbb{E}_k(p_A) + \bar{s}$

(3) $\mathbb{E}_k(p_A) + \bar{s} > \bar{v} > \mathbb{E}_k(p_A) + \underline{s}$

Figure 2.7: Composition of Sets $\mathcal{C}_D$ and $\mathcal{C}_A$ Depending on v and s_i with Parameters $v \sim \mathcal{U}[0,1]$, $s_i \sim \mathcal{U}[0.05, 0.3]$

This results in the following piecewise demand functions for the three different cases:

Case 1: $\bar{v} > \mathbb{E}_k(p_A) + \bar{s}$

$$q(p_A) = \begin{cases} 0 & \text{if } p_A \geq \bar{v}, \\ \bar{v} - p_A & \text{if } \mathbb{E}_k(p_A) + \bar{s} \leq p_A < \bar{v}, \\ \bar{v} - [\mathbb{E}_k(p_A) + \frac{\bar{s}+\underline{s}}{2}] - \frac{(p_A - [\mathbb{E}_k(p_A)+\underline{s}])^2}{2(\bar{s}-\underline{s})} & \text{if } \mathbb{E}_k(p_A) + \underline{s} < p_A < \mathbb{E}_k(p_A) + \bar{s}, \\ \bar{v} - [\mathbb{E}_k(p_A) + \frac{\bar{s}+\underline{s}}{2}] & \text{if } p_A \leq \mathbb{E}_k(p_A) + \underline{s}. \end{cases} \tag{2.79}$$

Case 2: $\bar{v} = \mathbb{E}_k(p_A) + \bar{s}$

$$q(p_A) = \begin{cases} 0 & \text{if } p_A \geq \bar{v}, \\ \frac{\bar{s}-\underline{s}}{2} - \frac{(p_A - [\mathbb{E}_k(p_A)+\underline{s}])^2}{2(\bar{s}-\underline{s})} & \text{if } \mathbb{E}_k(p_A) + \underline{s} < p_A < \mathbb{E}_k(p_A) + \bar{s}, \\ \frac{\bar{s}-\underline{s}}{2} & \text{if } p_A \leq \mathbb{E}_k(p_A) + \underline{s}. \end{cases} \tag{2.80}$$

Case 3: $\mathbb{E}_k(p_A) + \bar{s} > \bar{v} > \mathbb{E}_k(p_A) + \underline{s}$

$$q(p_A) = \begin{cases} 0 & \text{if } p_A \geq \bar{v}, \\ \frac{(\bar{v} - [\mathbb{E}_k(p_A)+\underline{s}])^2}{2(\bar{s}-\underline{s})} - \frac{(p_A - [\mathbb{E}_k(p_A)+\underline{s}])^2}{2(\bar{s}-\underline{s})} & \text{if } \mathbb{E}_k(p_A) + \underline{s} < p_A < \bar{v}, \\ \frac{(\bar{v} - [\mathbb{E}_k(p_A)+\underline{s}])^2}{2(\bar{s}-\underline{s})} & \text{if } p_A \leq \mathbb{E}_k(p_A) + \underline{s}. \end{cases} \tag{2.81}$$

Section 2.A.2 shows the construction of the demand function for each case separately.

2.A.2 Demand in Channel A with $s_i \sim \mathcal{U}[\underline{s}, \bar{s}]$ - case by case

Case 1: $\bar{v} > \mathbb{E}_k(p_A) + \bar{s}$

In this case

$$\bar{x} = \bar{v} - [\mathbb{E}_k(p_A) + \underline{s}], \tag{2.82}$$

$$\bar{y} = \bar{v} - [\mathbb{E}_k(p_A) + \bar{s}], \tag{2.83}$$

$$\bar{z} = \bar{s} - \underline{s} \tag{2.84}$$

yield in combination with Equation (2.70)

$$q^{max} = \bar{v} - [\mathbb{E}_k(p_A) + \frac{\bar{s} + \underline{s}}{2}]. \tag{2.85}$$

Further

$$x(p_A) = \max\{p_A - [\mathbb{E}_k(p_A) + \underline{s}], 0\}, \tag{2.86}$$

$$y(p_A) = \max\{p_A - [\mathbb{E}_k(p_A) + \bar{s}], 0\}, \tag{2.87}$$

$$z(p_A) = \max\{\min\{p_A - \mathbb{E}_k(p_A), \bar{s}\} - \underline{s}, 0\} \tag{2.88}$$

yield in combination with Equation (2.74)

$$q^{sub}(p_A) = \begin{cases} q^{max} & \text{if } p_A \geq \bar{v}, \\ p_A - [\mathbb{E}_k(p_A) + \frac{\bar{s}+\underline{s}}{2}] & \text{if } \mathbb{E}_k(p_A) + \bar{s} \leq p_A < \bar{v}, \\ \frac{(p_A - [\mathbb{E}_k(p_A) + \underline{s}])^2}{2(\bar{s} - \underline{s})} & \text{if } \mathbb{E}_k(p_A) + \underline{s} < p_A < \mathbb{E}_k(p_A) + \bar{s}, \\ 0 & \text{if } p_A \leq \mathbb{E}_k(p_A) + \underline{s}. \end{cases} \tag{2.89}$$

Subtracting (2.89) from (2.85) then yields

$$q(p_A) = \begin{cases} 0 & \text{if } p_A \geq \bar{v}, \\ \bar{v} - p_A & \text{if } \mathbb{E}_k(p_A) + \bar{s} \leq p_A < \bar{v}, \\ \bar{v} - [\mathbb{E}_k(p_A) + \frac{\bar{s}+\underline{s}}{2}] - \frac{(p_A - [\mathbb{E}_k(p_A) + \underline{s}])^2}{2(\bar{s} - \underline{s})} & \text{if } \mathbb{E}_k(p_A) + \underline{s} < p_A < \mathbb{E}_k(p_A) + \bar{s}, \\ \bar{v} - [\mathbb{E}_k(p_A) + \frac{\bar{s}+\underline{s}}{2}], & \text{if } p_A \leq \mathbb{E}_k(p_A) + \underline{s}. \end{cases} \tag{2.90}$$

Case 2: $\overline{v} = \mathbb{E}_k(p_A) + \overline{s}$

In this case

$$\bar{x} = \max\{\bar{v} - [\mathbb{E}_k(p_A) + \underline{s}], 0\}, \tag{2.91}$$

$$\bar{y} = \max\{\bar{v} - [\mathbb{E}_k(p_A) + \bar{s}], 0\} = 0, \tag{2.92}$$

$$\bar{z} = \max\{\min\{\bar{v}, [\mathbb{E}_k(p_A) + \bar{s}]\} - [\mathbb{E}_k(p_A) + \underline{s}], 0\} = \bar{s} - \underline{s} \tag{2.93}$$

yield in combination with Equation (2.70)

$$q^{max} = \frac{\bar{v} - [\mathbb{E}_k(p_A) + \underline{s}]}{2}. \tag{2.94}$$

$$\tag{2.95}$$

Further

$$x(p_A) = \max\{p_A - [\mathbb{E}_k(p_A) + \underline{s}], 0\}, \tag{2.96}$$

$$y(p_A) = \max\{p_A - [\mathbb{E}_k(p_A) + \bar{s}], 0\} = 0, \tag{2.97}$$

$$z(p_A) = \max\{\min\{p_A - \mathbb{E}_k(p_A), \bar{s}\} - \underline{s}, 0\} \tag{2.98}$$

yield in combination with Equation (2.74)

$$q^{sub}(p_A) = \begin{cases} q^{max} & \text{if } p_A \geq \bar{v}, \\ \frac{(p_A - [\mathbb{E}_k(p_A) + \underline{s}])^2}{2(\bar{s} - \underline{s})} & \text{if } \mathbb{E}_k(p_A) + \underline{s} < p_A < \mathbb{E}_k(p_A) + \bar{s} = \bar{v}, \\ 0 & \text{if } p_A \leq \mathbb{E}_k(p_A) + \underline{s}. \end{cases} \tag{2.99}$$

Subtracting (2.99) from (2.95) then yields

$$q(p_A) = \begin{cases} 0 & \text{if } p_A \geq \bar{v}, \\ \frac{\bar{s} - \underline{s}}{2} - \frac{(p_A - [\mathbb{E}_k(p_A) + \underline{s}])^2}{2(\bar{s} - \underline{s})} & \text{if } \mathbb{E}_k(p_A) + \underline{s} < p_A < \mathbb{E}_k(p_A) + \bar{s}, \\ \frac{\bar{s} - \underline{s}}{2} & \text{if } p_A \leq \mathbb{E}_k(p_A) + \underline{s}. \end{cases} \tag{2.100}$$

Case 3: $\mathbb{E}_k(p_A) + \bar{s} > \bar{v} > \mathbb{E}_k(p_A) + \underline{s}$

In this case

$$\bar{x} = \max\{\bar{v} - [\mathbb{E}_k(p_A) + \underline{s}], 0\} = \bar{v} - [\mathbb{E}_k(p_A) + \underline{s}], \tag{2.101}$$
$$\bar{y} = \max\{\bar{v} - [\mathbb{E}_k(p_A) + \bar{s}], 0\} = 0, \tag{2.102}$$
$$\bar{z} = \max\{\min\{\bar{v}, [\mathbb{E}_k(p_A) + \bar{s}]\} - [\mathbb{E}_k(p_A) + \underline{s}], 0\} = \bar{v} - [\mathbb{E}_k(p_A) + \underline{s}] \tag{2.103}$$

yield in combination with Equation (2.70)

$$q^{max} = \frac{(\bar{v} - [\mathbb{E}_k(p_A) + \underline{s}])^2}{2(\bar{s} - \underline{s})}. \tag{2.104}$$

Further

$$x(p_A) = \max\{p_A - [\mathbb{E}_k(p_A) + \underline{s}], 0\}, \tag{2.105}$$
$$y(p_A) = \max\{p_A - [\mathbb{E}_k(p_A) + \bar{s}], 0\} = 0, \tag{2.106}$$
$$z(p_A) = \max\{p_A - [\mathbb{E}_k(p_A)] + \underline{s}], 0\} \tag{2.107}$$

yield in combination with Equation (2.74)

$$q(p_A) = \begin{cases} 0 & \text{if } p_A \leq \mathbb{E}_k(p_A) + \underline{s}, \\ \frac{(p_A - [\mathbb{E}_k(p_A) + \underline{s}])^2}{2(\bar{s} - \underline{s})} & \text{if } \mathbb{E}_k(p_A) + \underline{s} < p_A < \bar{v}, \\ q^{max} & \text{if } p_A \geq \bar{v}. \end{cases} \tag{2.108}$$

Subtracting (2.108) from (2.104) then yields

$$q(p_A) = \begin{cases} \frac{(\bar{v} - [\mathbb{E}_k(p_A) + \underline{s}])^2}{2(\bar{s} - \underline{s})} & \text{if } p_A \leq \mathbb{E}_k(p_A) + \underline{s}, \\ \frac{(\bar{v} - [\mathbb{E}_k(p_A) + \underline{s}])^2}{2(\bar{s} - \underline{s})} - \frac{(p_A - [\mathbb{E}_k(p_A) + \underline{s}])^2}{2(\bar{s} - \underline{s})} & \text{if } \mathbb{E}_k(p_A) + \underline{s} < p_A < \bar{v}, \\ 0 & \text{if } p_A \geq \bar{v}. \end{cases} \tag{2.109}$$

2.A.3 Proof that $p^*_{A_k} > \mathbb{E}_k(p_A) + \underline{s}$

Claim. If consumers' anonymization cost are $s_i \sim \mathcal{U}[\underline{s}, \bar{s}]$ with $0 \leq \underline{s} < \bar{s} \leq \bar{v}$, then $p^*_{A_k} > \mathbb{E}_k(p_A) + \underline{s}$ as long as $\mathcal{C}_A \neq \emptyset$. To this end it suffices to show that $\pi_{A_k}(\mathbb{E}_k(p_A) + \underline{s} + \varepsilon) > \pi_{A_k}(\mathbb{E}_k(p_A) + \underline{s})$, for some $\varepsilon > 0$ without determining $p^*_{A_k}$ exactly.

Proof. Recall that demand in segment (b) (just above $\mathbb{E}_k(p_A) + \underline{s}$) is of the form

$$q(p_A) = q_A^{max} - \frac{(p_A - [\mathbb{E}_k(p_A) + \underline{s}])^2}{2(\bar{s} - \underline{s})} \tag{2.110}$$

in all three possible cases, with $q_A^{max} = q(\mathbb{E}_k(p_A) + \underline{s})$, i.e.

$$q_A^{max} = \begin{cases} \bar{v} - [\mathbb{E}_k(p_A) + \frac{\bar{s}+\underline{s}}{2}] & \text{if } \bar{v} > \mathbb{E}_k(p_A) + \bar{s} \text{ (Case 1)}, \\ \frac{\bar{s}-\underline{s}}{2} & \text{if } \bar{v} = \mathbb{E}_k(p_A) + \bar{s} \text{ (Case 2)}, \\ \frac{(\bar{v}-[\mathbb{E}_k(p_A)+\underline{s}])^2}{2(\bar{s}-\underline{s})} & \text{if } \mathbb{E}_k(p_A) + \bar{s} > \bar{v} > \mathbb{E}_k(p_A) + \underline{s} \text{ (Case 3)}. \end{cases} \tag{2.111}$$

Thus, profits in all three cases can be written as

$$\pi_{A_k}(p_A) = \left(q_A^{max} - \frac{(p_A - [\mathbb{E}_k(p_A) + \underline{s}])^2}{2(\bar{s} - \underline{s})} \right) (p_A - c). \tag{2.112}$$

The corresponding first order derivative in p_A is given by

$$\frac{\partial \pi_{A_k}}{\partial p_A} = q_A^{max} - \frac{3p_A^2 - p_A(2c + 4[\mathbb{E}_k(p_A) + \underline{s}]) + [\mathbb{E}_k(p_A) + \underline{s}](2c + [\mathbb{E}_k(p_A) + \underline{s}])}{2(\bar{s} - \underline{s})}. \tag{2.113}$$

Evaluating at $p_A = \mathbb{E}_k(p_A) + \underline{s} + \varepsilon$ with $\varepsilon > 0$ gives

$$\pi'_{A_k}(\mathbb{E}_k(p_A) + \underline{s} + \varepsilon) = q_A^{max} - \varepsilon \cdot \frac{[\mathbb{E}_k(p_A) + \underline{s}] + \frac{3}{2}\varepsilon - c}{\bar{s} - \underline{s}}. \tag{2.114}$$

Approaching $p_A = \mathbb{E}_k(p_A) + \underline{s}$ from above, the limit is given by

$$\lim_{\varepsilon \to 0} q_A^{max} - \varepsilon \cdot \frac{[\mathbb{E}_k(p_A) + \underline{s}] + \frac{3}{2}\varepsilon - c}{\bar{s} - \underline{s}} = q_A^{max}. \tag{2.115}$$

From $q_A^{max} > 0$ in all three possible cases it follows that

$$\frac{\partial \pi_{A_k}}{\partial p_A} > 0 \text{ for at least some } p_A = \mathbb{E}_k(p_A) + \underline{s} + \varepsilon \text{ with } \varepsilon > 0 \tag{2.116}$$

Hence, $\pi_{A_k}(\mathbb{E}_k(p_A) + \underline{s} + \varepsilon) > \pi_{A_k}(\mathbb{E}_k(p_A) + \underline{s})$ for some $\varepsilon > 0$ and $p^*_{A_k} > \mathbb{E}_k(p_A) + \underline{s}$.
Q.E.D.

2.A.4 Proof that $p^*_{A_k} < 1$ for all finite k if $\underline{s} = 0$

Claim. If consumers' anonymization cost are $s_i \sim \mathcal{U}[0, \bar{s}]$ with $0 < \bar{s} \leq \bar{v}$, then $p^*_{A_k} < \bar{v}$ for all finite k.

Proof. Suppose not. Then, there must be a finite $k = \tilde{k}$ at which the seller sets $p^*_{A_{\tilde{k}}} = \bar{v}$ for the first time (when k increases). Because consumers indifferent between channel A and channel D are assumed to choose channel D, it follows that $\mathcal{C}_{A_{\tilde{k}}} = \emptyset$. Otherwise, the seller forgoes profits from consumers in $\mathcal{C}_{A_{\tilde{k}}}$ with $v < \bar{v}$ by setting $p^*_{A_{\tilde{k}}} = \bar{v}$.

Recall further that, with level-k thinking, for any $k > 0$ it holds that

$$\mathbb{E}_k(p_A) = p^*_{A_{k-1}}. \tag{2.117}$$

and that consumers choose channel A if, and only if $v > \mathbb{E}(p_A) + s_i$. As $\underline{s} = 0$ by assumption, it follows that the condition that $\mathcal{C}_{A_{\tilde{k}}} = \emptyset$ requires

$$\mathbb{E}_{\tilde{k}}(p_A) + \underline{s} = \mathbb{E}_{\tilde{k}}(p_A) \geq v \text{ for all } v. \tag{2.118}$$

Combining Equations (2.117) and (2.118), gives

$$\mathbb{E}_{\tilde{k}}(p_A) = p^*_{A_{\tilde{k}-1}} \geq v \text{ for all } v, \tag{2.119}$$

implying that already at $\tilde{k} - 1$ the seller sets $p^*_{A_{\tilde{k}-1}} = \bar{v}$ (as the seller is restricted to set prices within the support of the demand), leading to a contradiction with the assumption that the seller sets $p^*_{A_{\tilde{k}}} = \bar{v}$ at $k = \tilde{k}$ for the first time (when k increases).

By transitivity, this further implies that also $p^*_{A_0} = \bar{v}$ which is only possible if $p_M = \frac{\bar{v}+c}{2} = \bar{v}$, which itself is ruled out by the assumption underlying the model that c is not prohibitively costly.

Therefore, there can be no finite $k = \tilde{k}$ at which the seller sets $p^*_{A_{\tilde{k}}} = \bar{v}$.

Thus, it has to hold that $p^*_{A_k} < 1$ for all finite k if $\underline{s} = 0$. *Q.E.D.*

Chapter 3

Predictive Algorithms and Consumer Behavior

This chapter is single-authored

Abstract

I analyze consumer behavior in a laboratory experiment where subjects face a computerized seller. Unless consumers make use of a costly hiding technique to hide (some) of their valuations for the offered good, the seller's algorithm implements perfect price discrimination. Depending on the hiding cost treatment, Nash equilibrium behavior is characterized by either complete or partial hiding by consumers. Addressing cognitive constraints often present in privacy choices, I consider two alternative explanations: level-k thinking and reinforcement learning. I find substantial deviations from the Nash predictions. Further, there is some evidence that level-k thinking measures can explain subjects' hiding choices in the first period. Additionally, I find evidence for the occurence of reinforcement learning suggesting that subjects react significantly to having realized losses in the previous period, but not to gains. Finally, a within-subject policy treatment resembling privacy-by-default mechanisms leads to a strong increase in hiding behavior: among various Nash equilibria, the majority of subjects chooses the one with maximal hiding behavior.

3.1. Introduction

Intelligent personal assistants (such as Google Now (Google), Siri (Apple), Cortana (Microsoft), and Alexa (Amazon)) are becoming vital parts of many people's lives. Some of them can already provide information before they have been asked for it. While data stored and used by the algorithms behind such services improves user experience via more accurate predictions of a user's interests and needs, this data can be of secondary use and additionally improve the tailoring capabilities of a seller. Such tailoring is not limited to consumers' preferences alone, but can manifest as so-called "mobile targeting" (Luo et al. 2014) pitching products in real time based on location data and potentially at an individualized price that increases the likelihood of a purchase.

An exemplary case of this trend is the shopping app Shopkick that rewards its users not only for visiting particular stores, but even guides them to interact with specific products (e.g., by scanning bar codes, taking them to the dressing room, etc.) to make purchases more likely. According to a 2012 study by Nielsen, Shopkick (founded in 2009) already had a reach of 6 million users spending more than three hours per month using the app rendering it the most-used shopping app (Nielsen 2012). Shopkick itself reported that it had created a total of more than USD 1 billion in revenues for partner stores by generating more than 90 million walk-ins and over 150 million product scans since its foundation (Shopkick 2015).

The more data is generated by consumers and provided to sellers, the closer they will get to first-degree (or perfect) price discrimination (Odlyzko 2003), where the seller has complete information about every specific consumer's willingness to pay for any given product (and at any given time and/or location) (Pigou 1920). Because first-degree price discrimination can deprive consumers of all surplus from the transaction, they may want to protect their privacy and hide their willingness-to-pay from a seller capable of such price discrimination.

Regarding privacy choices, however, Acquisti and Grossklags (2007, p. 369) note that "consumers will often be overwhelmed with the task of identifying possible outcomes related to privacy threats and means of protection. [...] Especially in the presence of complex, ramified consequences associated with the protection or release of personal information, our innate bounded rationality limits our ability to acquire, memorize and process all relevant information, and it makes us rely on simplified mental models, approximate strategies, and heuristics".

In the context of predictive algorithms, consumers might therefore fail to see the secondary use of data for increasingly personalized pricing despite their general awareness of the capabilities of predictive algorithms as personal assistants. Secondly, even if consumers were aware of the potential for personalized pricing, they might not be able to figure out how to react optimally to maximize their own surplus due to cognitive constraints.

As Goldfarb and Tucker (2012) document that people are becoming increasingly aware of privacy-related contexts, I focus on the latter issue of the optimality of choices once it has become apparent that such choices ought to be made. In this paper, I present a laboratory experiment inspired by the theoretical model in Dengler and Prüfer (2017). Both, the model and the experiment, mimick a situation in which consumers may want to restore their privacy after a seller has already collected the necessary data to implement perfect price discrimination. Do subjects in the experiment make optimal privacy choices when they are matched against a pricing algorithm implementing perfect price discrimination?

The reduced complexity of a laboratory experiment as compared to everyday privacy choices notwithstanding, cognitive constraints might still play a role in the decision-making process of subjects. Therefore, I do not restrict the analysis to Nash equilibrium predictions alone and investigate two possible alternative explanations. The selection of the first concept, *level-k thinking*, is motivated by the model in Dengler and Prüfer (2017), whereas the second concept, *reinforcement learning*, stems from earlier experimental literature.

Level-k thinking is an iterative reasoning model introduced by Stahl and Wilson (1994; 1995) and Nagel (1995) where k denotes the number of strategic iterations a player can perform. While there is considerable experimental literature on level-k thinking itself, with respect to privacy choices only Benndorf et al. (2015) conducted a laboratory experiment concerned with level-k thinking. They found that voluntary, but costly, disclosure choices in a labor market setting are consistent with level-k thinking. In contrast to their experiment, I focus on a consumer market and subjects have to incur costs to restore their privacy rather than enjoying privacy by default which can be voluntarily surrendered. Because level-k thinking describes iterative reasoning in static frameworks, this concept cannot be used to describe a learning trajectory. Therefore, the level-k thinking analysis focuses on subjects' behavior in the first period.

Reinforcement learning as introduced by Erev and Roth (1998), in contrast, is applicable to the dynamic framework of repeated play in the experiment. It captures the idea that adjustments in behavior are a response to payoffs received in previous periods. Gains are positively reinforcing the chosen strategy (i.e., increase the likelihood of it being played) whereas losses are negatively reinforcing the chosen strategy (i.e., decrease the likelihood of it being played). Perfect price discrimination leaves consumers with zero surplus. Thus, gains and losses in such a setting can only occur if subjects engage in privacy-seeking behavior. Realized gains are then expected to increase privacy-seeking, realized losses to decrease privacy-seeking.[24]

While understanding how cognitive constraints affect behavior can help to improve economic models, it may be difficult to change. The cost of privacy protection, however, seems more actionable from a policy perspective as exemplified by the inclusion of *data protection by default* in Article 25 of the EU Data Protection Regulation.[25] Therefore, this experiment includes a policy treatment where privacy protection is available for free.

This experiment thus addresses the following research questions: Do subjects in the role of consumers make optimal use of costly privacy protective options when facing a perfect price discrimination algorithm? If not, do the behavioral models of level-k thinking and reinforcement learning provide a good alternative explanation for the observed behavior? Does a policy change to privacy protection being available for free lead to substantially different behavior?

Despite some differences between the experimental implementation and the model which inspired it, the model's central proposition that a costly privacy-protective option is used, even when this implies an expected loss for consumers, receives empirical backing. I find some evidence that level-k thinking measures can explain subjects' hiding choices in the first period. Additionally, I find evidence for the occurence of reinforcement learning suggesting that subjects react significantly to having realized losses in the previous period, but not to gains. The within-subject policy treatment which makes hiding available for free strongly increases hiding behavior: among various Nash equilibria, the majority of subjects chooses the one with maximal hiding behavior.

The remainder of this paper is organized as follows. Section 3.2 introduces the Market Game. Section 3.3 introduces three level-k elicitation games. Section 3.4 describes the experimental procedures. Section 3.6 presents results from the experiment. Section 3.7 concludes.

[24] *Learning direction theory* (Selten et al. 2005) is a related but different concept of learning in which subjects adjust their behavior, not based on realized payoffs of the last chosen action, but rather on the additional payoffs that another action would have yielded. However, subjects in this experiment are not informed about the payoffs of all other actions they could have chosen. Hence, "the counterfactual causal reasoning about the past [which] is a crucial feature of learning direction theory" (Selten et al. 2005, p. 7) is not fully feasible. This notwithstanding, the regression analysis also includes an analysis based on only the sign of the payoff (loss or gain), which resembles the idea of directional learning theory to some extent.

[25] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

3.2. The Market Game

3.2.1. Model Outline

My experimental design is inspired by the model developed in Dengler and Prüfer (2017). I consider an economy where a monopolistic seller faces consumers who can buy at most one unit of a good. I assume that the monopolist can produce the good at constant marginal cost $c = 0$ and consumers have heterogeneous valuations v following a discrete uniform distribution ranging from 0.00 € to 9.50 € with increments of 0.50 €, i.e. $v \sim U\{0.00\text{ €}, 9.50\text{ €}\}$. in the experiment.[26] Consumers face the decision whether to engage in privacy protective behavior, or not. If consumers do not hide their valuation from the seller, the seller can engage in perfect price discrimination and charge targeted prices p_i. If consumers choose to hide their valuation, which comes at cost s, the seller has to set a uniform price p_A for all hidden valuations of a particular consumer. The seller can neither directly influence consumers' hiding choices nor decide to not sell to consumers with a hidden valuation. The distribution of v, the hiding cost s, the monopolist's marginal cost $c = 0$ as well as the timing of the game are common knowledge among all players. Further, the model assumes that all players are solely interested in their own material payoff (i.e., net monetary profit or consumer surplus). Specifically, consumers do not have any exogenous taste for privacy. The timing is as follows:

- Stage 1 (*Hiding*): Consumers choose whether to hide their valuation at cost s for each possible valuation v.

- Stage 2 (*Pricing*): The seller sets prices $p = \{p_i(v), p_A\}$, where $p_i(v)$ are targeted prices for the non-hidden valuations of a consumer, and p_A is the uniform price p_A the seller has to set for all hidden valuations of a consumer.

- Stage 3 (*Buying*): Consumers decide whether to buy the good for the offered price.

3.2.2. Experimental Implementation

To reflect the idea of predictive algorithms and to ensure optimal play, the seller's role was taken on by a computer rather than a human subject. This also reduces the potential for other-regarding preferences which might play a role on either side of the interaction if the seller role were instead taken on by a human subject. The experimental setup deviates from the theoretical model of Dengler and Prüfer (2017) in several aspects.

Firstly, the experiment is an individual choice experiment, which rules out dependence between subjects' choices as there is no interaction between them, in turn increasing the

[26] In the model, though, consumers' valuations follow a continuous uniform distribution.

number of independent observations. Thus, each subject faces the seller alone and the seller reacts to each consumer separately rather than to a mass of atomistic consumers forming an entire consumer population.

Secondly, subjects do not know their exact valuation of the good when making their hiding choice, but are only assigned one after submitting their hiding choice for all possible valuations. This procedure allows to analyze subjects' complete strategies (thus called "strategy method") rather than only one choice for a randomly assigned valuation. To ensure that at least a minimal rationality requirement is met subjects' choices are limited to transitive strategies: If a consumer wants to hide a particular valuation v, then the same consumer should also want to hide any valuation $v' > v$.[27] Therefore, subjects in the experiment are only asked to submit a hiding threshold $\hat{v}$ which is equivalent to having them submit a strategy for all possible valuations with the aforementioned transitivity requirement imposed. If the subsequently assigned valuation is lower than the hiding threshold ($v < \hat{v}$), the valuation is not hidden and the seller is informed exactly about v. If the subsequently assigned valuation v is equal to or higher than the hiding threshold ($v \geq \hat{v}$), the exact valuation is hidden from the seller and the subject incurs the hiding cost s.

Thirdly, in the model it is assumed that the seller forms a correct belief about the level of sophistication of consumers, which is assumed to be uniform within the consumer population. From this, the seller can infer the hiding threshold in the population and respond optimally. In the experiment, though, the level of sophistication of subjects is unknown ex ante and cannot be ruled out to differ between subjects. Nonetheless, the seller could still infer which valuations were hidden by combining his observation of non-hidden valuations and his knowledge about the demand function. Due to the strategy method, however, only the decision for *one* valuation is implemented in the interaction with any given subject. Hence, the seller does not face an entire consumer population from which to infer the residual demand of the hidden valuations. Therefore, the seller in the experiment does not infer but is informed about the exact value of the hiding threshold $\hat{v}$ and then infers only that $v \geq \hat{v}$ whenever he is approached by a consumer with a hidden valuation.[28]

[27] In an earlier (and shorter) version of the experiment where transitivity was not imposed, 48% of subjects submitted transitive strategies throughout all seven periods of the earlier experiment. The other 52% submitted a non-transitive strategy at least once during the experiment. Overall, 77% of the submitted strategies were transitive.

[28] This choice results from prioritizing independence of observations on the side of the consumers (whose behavior is the main focus of this study) over an exact replication of the seller's superior sophistication in the model. However, if such a seller outside the laboratory were considered incapable of inferring either the level of sophistication in the consumer population or the residual demand by observing the non-hidden valuations, the findings of this experiment may have limited applicability.

Thus, the hiding choice in Stage 1 is implemented by subjects submitting any of the 20 possible valuations (from 0.00 € to 9.50 €) as their hiding threshold $\hat{v}$. Then, valuations are hidden from the seller accordingly. If a subject does not want to hide any valuation at all, a hiding threshold of 10.00 € has to be entered (continuing the increments of 0.50 € above the maximum possible valuation). Subjects were explicitly made aware of the fact that their hiding threshold, but not their assigned valuation, would be communicated to the seller if a hidden valuation was subsequently assigned to them (see also the experimental instructions in Appendix 3.C). Table 3.1 was used to illustrate the consequences of their choice of a hiding threshold.

If your hiding threshold is	**and if you are assigned**	**then the seller knows ...**
0.00 €	... any valuation ...	... that your valuation is 0.00 € or higher.
3.00 €	... a valuation from 0.00 € to 2.50 € ...	... what your exact valuation is.
3.00 €	... a valuation from 3.00 € to 9.50 € ...	... that your valuation is 3.00 € or higher.
7.00 €	... a valuation from 0.00 € to 6.50 € ...	... what your exact valuation is.
7.00 €	... a valuation from 7.00 € to 9.50 € ...	... that your valuation is 7.00 € or higher.
10.00 €	... any valuation ...	... what your exact valuation is.

Table 3.1: Examples for Hiding Threshold Choices and Resulting Implementations

The seller's optimal pricing strategy in Stage 2 is unaffected by the differences to the model. Given the ability to perfectly price discriminate, the seller sets the price equal to the valuation whenever possible.[29] Where perfect price discrimination is impossible due to the valuation being hidden, the price is set to the globally profit-maximizing price without price discrimination p_M or the hiding threshold $\hat{v}$, whichever is higher.[30] Thus, the optimal pricing strategy for the seller is given by

$$p^* = \begin{cases} p_A^* = \max\{p_M, \hat{v}\} & \textit{if } v \geq \hat{v} \text{ (valuation hidden)}, \\ p_i^* = v & \textit{if } v < \hat{v} \text{ (valuation not hidden)}. \end{cases} \tag{3.1}$$

The range of valuations from 0.00 € to 9.50 € and their discrete distribution further imply that the globally profit-maximizing monopoly price without price discrimination is given by $p_M = 5.00$ €, irrespective of the treatment.[31] Note, that if a subject selected a hiding threshold of 10.00 €, the seller set $p_A = 9.50$ to account for the fact that the

[29] More precisely, the optimal price when price discrimination is possible, is given by $p_i^* = \max\{v, c\}$. Because $c = 0$, which is equal to the lowest possible valuation, this constraint is immaterial here.

[30] For a more elaborate discussion of this pricing strategy see Dengler and Prüfer (2017, p.9 ff.).

[31] Table 3.11 in Appendix 3.A shows the expected profits associated with pricing at any of the 20 valuations.

price should not exceed the highest possible valuation of 9.50 € as the 10.00 € threshold is merely a logical prolongation of the 0.50 € for the subjects' convenience.

While subjects were informed about the consequences of their choice of $\hat{v}_i$ for the information the seller obtains about v, the optimal pricing strategy of the seller was not disclosed. However, it was announced that the seller maximizes revenues, which in theory enables subjects to deduce the seller's optimal strategy.[32] This admittedly minimal disclosure was deliberate as the question to be examined is how consumers react to an algorithm that they might not be able to fully understand in terms of the best strategy.

Lastly, the buying decision in Stage 3 is implicitly implemented optimal in the payoff calculation. Subjects are *not* given an explicit choice to buy the product or not. Rather, they directly receive the difference between the assigned valuation and the price as payoff, if this difference is positive, and a zero payoff otherwise. Subjects were explicitly made aware of the buying choice being implemented in this way and the results were referred to as "buying the product" and "(not) buying the product".[33] In the experiment, subjects thus knew that their earnings, denoted here by y, are determined in the following way:

$$y = \begin{cases} \max\{v - p_A, 0\} - s & \textit{if } v \geq \hat{v} \text{ (valuation hidden)}, \\ \max\{v - p_i(v), 0\} & \textit{if } v < \hat{v} \text{ (valuation not hidden)}. \end{cases} \tag{3.2}$$

Using the seller's optimal pricing strategy (which remained undisclosed to subjects) from Equation (3.1) in Equation (3.2) yields the following determination of subjects' earnings:

$$y = \begin{cases} \max\{v - \max\{p_M, \hat{v}\}, 0\} - s & \textit{if } v \geq \hat{v} \text{ (valuation hidden)}, \\ 0 & \textit{if } v < \hat{v} \text{ (valuation not hidden)}. \end{cases} \tag{3.3}$$

Recall that subjects do not know their valuation when making their hiding choice and play alone with the seller. Thus, maximizing their expected earnings $\mathbb{E}(y|\hat{v})$ from the market game is isomorph to maximizing aggregated consumer surplus in the theoretical model rather than mimicking an atomistic consumer's choice. In contrast to the seller, consumers' optimal strategy depends on the treatment and is therefore explained in context of the respective treatments.

[32] With $c = 0$, this is identical to profit-maximization but avoids mentioning production costs.

[33] To avoid misunderstandings, it was additionally pointed out at the beginning of the experiment that no real but only a hypothetical good is "bought" (cf. Appendix 3.C).

3.2.3. Treatments

In the experiment, the Market Game is implemented with three different levels of hiding cost s. In two of the treatments the hiding cost is strictly positive at 1.00 € or 2.50 €, respectively, whereas in the third treatment hiding is available for free, i.e., the hiding cost is set to 0.00 €.

Table 3.2a shows an example of a subject's *actual* payoffs of choosing a hiding threshold of $\hat{v} = 7.00$ € for any assigned valuation v and across all three hiding cost treatments. Note that the valuations below the line at $\hat{v} = 7.00$ € are not hidden from the seller and hence perfect price discrimination takes place, whereas for all valuations above the line at $\hat{v} = 7.00$ € the seller sets a uniform price.

Table 3.2b then shows a subject's *expected* earnings of choosing any particular hiding threshold $\hat{v}$ not yet knowing which valuation will be assigned afterwards. Note that the last row of Table 3.2a is repeated in the line $\hat{v} = 7.00$ € of Table 3.2b. Likewise, all other rows in Table 3.2b stem from a table similar to Table 3.2a for another hiding threshold $\hat{v}$.

It can then be seen from Table 3.2b, that subjects maximize their expected earnings in the 1.00 € treatment by choosing a hiding threshold of 5.00 €. In the 2.50 € treatment, however, they maximize their expected earnings by choosing not to hide any valuation at all (i.e., submitting a hiding threshold of 10.00 €). Any lower hiding threshold leads to losses in expectation, although for some $\hat{v} < 10.00$ € strictly positive earnings are possible.

In contrast, in the 0.00 € treatment, every hiding threshold $\hat{v} \leq 5.00$ € maximizes a subject's expected earnings. This stems from the fact that for all these hiding thresholds the seller sets $p_A^* = p_M = 5.00$ €. Thus, all hiding thresholds $\hat{v} \leq 5.00$ € have the same potential for gains while they do not differ in incurred cost because $s = 0.00$ €.

v	$\mathbf{p}^*$	**y**		
		s = 1.00	s = 2.50	s = 0.00
9.50	7.00	1.50	0.00	2.50
9.00	7.00	1.00	-0.50	2.00
8.50	7.00	0.50	-1.00	1.50
8.00	7.00	0.00	-1.50	1.00
7.50	7.00	-0.50	-2.00	0.50
7.00	7.00	-1.00	-2.50	0.00
6.50	6.50	0.00	0.00	0.00
6.00	6.00	0.00	0.00	0.00
5.50	5.50	0.00	0.00	0.00
5.00	5.00	0.00	0.00	0.00
4.50	4.50	0.00	0.00	0.00
4.00	4.00	0.00	0.00	0.00
3.50	3.50	0.00	0.00	0.00
3.00	3.00	0.00	0.00	0.00
2.50	2.50	0.00	0.00	0.00
2.00	2.00	0.00	0.00	0.00
1.50	1.50	0.00	0.00	0.00
1.00	1.00	0.00	0.00	0.00
0.50	0.50	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00
$\mathbb{E}(y\|\hat{v} = 7.00)$		0.075	-0.375	0.375

(a) Actual Earnings y
(for the exemplary case $\hat{v} = 7.00$)

$\hat{\mathbf{v}}$	$\mathbb{E}(\mathbf{y})$		
	s = 1.00	s = 2.50	s = 0.00
10.00	0.000	**0.000**	0.000
9.50	-0.050	-0.125	0.000
9.00	-0.075	-0.225	0.250
8.50	-0.075	-0.300	0.075
8.00	-0.050	-0.350	0.150
7.50	0.000	-0.375	0.250
7.00	0.075	-0.375	0.375
6.50	0.175	-0.350	0.525
6.00	0.300	-0.300	0.700
5.50	0.450	-0.225	0.900
5.00	**0.625**	-0.125	**1.125**
4.50	0.575	-0.250	**1.125**
4.00	0.525	-0.375	**1.125**
3.50	0.475	-0.500	**1.125**
3.00	0.425	-0.625	**1.125**
2.50	0.375	-0.750	**1.125**
2.00	0.325	-0.875	**1.125**
1.50	0.275	-1.000	**1.125**
1.00	0.225	-1.125	**1.125**
0.50	0.175	-1.250	**1.125**
0.00	0.125	-1.375	**1.125**

(b) Expected Earnings $\mathbb{E}(y)$
(for all $\hat{v}$)

Table 3.2: Actual (y) and Expected ($\mathbb{E}(y)$) Earnings Across Hiding Cost Treatments (optimal strategies are highlighted)

3.3. Level-k Elicitation Games

The level-k thinking literature focuses on players' potentially limited capacities to reason iteratively in order to explain deviations from Nash equilibrium play. A level of strategic sophistication k is then given by the number of iterative reasoning steps a player performs. The underlying idea is that players with a higher level are able to anticipate actions by players with a lower level and respond optimally.[34]

The model by which this experiment is inspired predicts that an increase in the level of sophistication in the consumer population leads to less anonymization, i.e., a higher hiding threshold. The mechanism underlying this result stems from the following level-k analysis: If all consumers are of a naïve type ($k = 0$) they do not reason strategically and expect the monopolistic seller to engage in monopoly pricing. The more sophisticated seller ($k = 1$) infers consumers' resulting hiding behavior and responds optimally. In his price setting the seller exploits the fact that at the buying stage, the hiding cost is sunk for consumers. Thus, the price the seller sets for hidden consumers exceeds the price they expected. If consumers had a level of sophistication of $k = 1$, however, they would anticipate this first strategic iteration and raise their price expectation accordingly, leading to fewer consumers hiding their valuation. Again, the more sophisticated seller ($k = 2$) anticipates this and raises the price further, which in turn is expected if consumers have a level of sophistication of $k = 2$, and so on.

This prediction can be tested here as every subject decides on a hiding threshold, which corresponds to a particular level of sophistication. However, assigning levels of sophistication to given observed hiding thresholds according to the theory would not allow to test the prediction. Therefore, my experimental design features additional games from the level-k thinking literature to elicit subjects' iterative reasoning capabilities: the Adding Game (based on Dufwenberg, et al. (2010)), the Money Request Game (based on Arad and Rubinstein (2012)), and the Beauty Contest (based on Nagel, (1995)). Although all games elicit iterative thinking, they all differ to some extent from the interaction in the Market Game. The Adding Game, in the version I implemented here, shares the fact that subjects are matched to a computer player with the Market Game, but does not feature strategic uncertainty because the computer's strategy is known to subjects. The Money Request Gamen and the Beauty Contest, on the other hand, feature strategic uncertainty, but are played with one of the other subjects (Money Request Game) or all other subjects (Beauty Contest) in the same session.[35]

[34] For an overview of the theoretical and empirical literature on level-k thinking and related concepts see for example Crawford, et al. (2013).

[35] To the best of my knowledge, no typical level-k game matches the Market Game in both aspects. To not declare one of these features irrelevant ex ante, I implemented this multiplicity of elicitations.

The following subsections introduce the original games and explain the adjustments made for my experimental implementation. Further, the procedures by which levels of iterative reasoning were assigned given the choices subjects made in the respective games are presented.

3.3.1. Level-k Elicitation Game 1: Adding Game

> "Two players, call them *White* and *Green*, take turns. *White* begins. To start off, he can choose either 1 or 2. *Green* observes this choice, then increments the "count" by adding one or two. That is, if *White* chooses 1 *Green* can follow up with 2 or 3; if *White* chooses 2 *Green* can follow up with 3 or 4. *White* then observes *Green*'s choice, and again increments the count by adding one or two. The game continues with the players taking turns, each player incrementing the count by one or two. The player who reaches 21 wins."
>
> (Dufwenberg, et al. 2010, p. 132–133, *original emphasis*)

In this specification of the game, the second-mover (here *Green*) can force a win by always adding up to a multiple of 3. This can be uncovered by a backward reasoning process. First note, that adding up to 21 wins. Secondly, as *White* can only add 1 or 2 to the current total, *Green* can ensure that *White* cannot win the game if *White* can maximally add up to a "count" of 20. *Green* can do so by adding up to 18 before, leaving *White* with the two options to add to a "count" of 19 or 20. Both, however, allow *Green* to add to 21 on the next turn and win the game. Applying the same logic to understand how *Green* can ensure to be able to add up to a "count" of 18, and so on the (backward iterating) series of win-ensuring "counts" is uncovered to be 21, 18, 15, 12, 9, 6, and 3. As *White* has to start the game by adding to either 1 or 2, *Green* can always reach 3 and thus can always guarantee to win the game. As can be seen, several steps of reasoning are necessary to uncover the full series. Because the game is a game with complete information, the optimal strategy is completely independent of beliefs about the other player. Therefore, this game can serve as a means to elicit the pure iterative reasoning capabilities of a subject by the number of optimal plays counted backwards from the end.

To not lose observations by forming pairs of which only the winning subject's iterative reasoning capabilities can be measured, subjects played this game against the computer. As the strategic advantage moves from one player to the other after any non-optimal choice, the computer player was picking numbers at random. Otherwise, all variation in subjects' iterative reasoning capacities (aside of perfect play from the start) would be lost, because the computer player would be able to ensure a win after the first mistake of each subject and not deviate anymore.

To reduce measurement error due to "accidental" perfect play by subjects as well as to reduce the chance of the computer playing multiple perfect rounds, a modified version of the game was played. Instead of adding numbers 1 or 2 to a target value of 21, each player in the implemented version can choose to add 1, 2, 3, 4, 5, or 6 and the target is to reach 50 or more (the inequality being necessary due to the random computer play).

In this specification, the *first-mover* can force a win by adding up to 1, 8, 15, 22, 29, 36, 43, and finally 50 (or more). This switch to a first-mover advantage allowed to let subjects play as the first-movers eliminating the need for a first move by the computer. Despite them being told that the computer plays randomly, subjects might otherwise have taken the computer's first choice as an indication as to how to play the game nonetheless. Any such anchoring effects are thus eliminated. Hence, in the implemented version, subjects begin the game. If they won the game, they earned 10 €, if they lost the game they earned 0 €.

In the spirit of the paper by Dufwenberg et al. (2010), levels of iterative strategic sophistication were assigned by counting how many continuing choices a subject made to guarantee a win in the game (counting from the end backwards). Table 3.3 shows how different k-levels were assigned to different plays of the game and reports the frequency as well as the proportion of their occurence in the subject population. To explain the meaning of the series of numbers in the second column of this table and their relation to the assigned levels of sophistication, consider level $k = 1$ as an example.

Level $k = 1$ was assigned to all subjects that managed to win the game (by adding to 50 or more, which itself is denoted by 50_{+}) *and* forced this one step ahead by adding up to a current total of 43. Note that a struckout number does *not* mean that a subject did never add the current total up to the respective number. A struckout number only implies that this number was not the current total after the subject's choice immediately preceeding the next indicated number in the respective series of column 2.

This means for the example of subjects with a level of $k = 1$, that while they forced their win one step ahead, they may not have forced their win two steps ahead. Had they added to 36 in their previous turn it would set their level to $k = 2$ as indicated by the row above. Subjects with a level of $k = 1$ may have arrived at a current total of 36 at some point during the game, though, e.g., by a series such as 36, (38), 40, (42), 43, (47), 50, (numbers in parentheses are the current total after the computer's turn). While a subject with such a series played 36 at *some* point during the game, this is not a two-step forcing play, because of the choice to add to a current total of 40 instead of immediately adding to the forcing winning choice of 43. Thus, the highest possible level is $k = 7$ which was assigned for forcing the win right from the start of the game.

The group that was assigned a level of $k = 0$ is somewhat particular as it groups two types of subjects together: those that lost the game (explaining the absence of the 50_+ element in the respective series) and those that won the game but did not force it one step ahead. As both groups have in common that their choices imply that they did not engage in iterative reasoning, both were assigned a level of $k = 0$.

Adding-k	**Subject series ends on...**
7	$01, 08, 15, 22, 29, 36, 43, 50_+$
6	$..., \cancel{01}, 08, 15, 22, 29, 36, 43, 50_+$
5	$..., \cancel{08}, 15, 22, 29, 36, 43, 50_+$
4	$..., \cancel{15}, 22, 29, 36, 43, 50_+$
3	$..., \cancel{22}, 29, 36, 43, 50_+$
2	$..., \cancel{29}, 36, 43, 50_+$
1	$..., \cancel{36}, 43, 50_+$
0	$..., \cancel{43}, ...$

Table 3.3: Level-k Assignment - Adding Game

3.3.2. Level-k Elicitation Game 2: Money Request Game

> "You and another player are playing a game in which each player requests an amount of money. The amount must be (an integer) between 11 and 20 shekels. Each player will receive the amount he requests. A player will receive an additional amount of 20 shekels if he asks for exactly one shekel less than the other player."
>
> (Arad and Rubinstein 2012, p. 3562)

The 11-20 Money Request Game was specifically designed to "naturally [trigger] level-k reasoning and is not likely to induce other types of decision rules" (Arad and Rubinstein 2012, p. 3562) as it is essentially ingrained in the description of the game. In contrast to the Adding Game (and the Beauty Contest), a pure-strategy Nash equilibrium does not exist in the Money Request game (Arad and Rubinstein, 2012). However, best-responding to any belief about the other player is straightforward. Suppose, player i believes that the matched player j requests the amount of 20 shekels (the supposed level-0 behavior). Clearly, the response that maximizes i's payoff is to request exactly one shekel less than player j's request, i.e., 19. The same holds true for any belief about j's choice $x_j \geq 12$. If player i expects that player j requests $x_j = 11$, though, the payoff-maximizing choice is $x_i = 20$ as it is impossible to request exactly one shekel less than the other player and hence requesting the maximum possible amount becomes optimal.

In my experiment the denomination of earnings was € rather than shekels and subjects could choose numbers from 1 € to 10 € with a bonus payment of 10 € for requesting exactly 1 € less than the other player. This was done to avoid that stakes in the Money Request Game became much higher than in the other level-k elicitation tasks.[36] Shifting the choice set does not change the fact that there is no Nash equilibrium in pure strategies and keeps the same iterative reasoning structure in place as the original version. Hence, the only adjustment to assign levels is to use the formula *Request-k* $= 10 - x_i$ (instead of *Request-k* $= 20 - x_i$), leading to the assignment shown in Table 3.4.

Request-k	0	1	2	3	4	5	6	7	8	9
Request x_i	10	9	8	7	6	5	4	3	2	1

Table 3.4: Level-k Assignment - Money Request Game

3.3.3. Level-k Elicitation Game 3: Beauty Contest

> "A large number of players have to state simultaneously a number in the closed interval $[0, 100]$. The winner is the person whose chosen number is closest to the mean of all chosen numbers multiplied by a parameter p, where p is a predetermined positive parameter of the game; p is common knowledge. The payoff to the winner is a fixed amount, which is independent of the stated number and p. If there is a tie, the prize is divided equally among the winners. The other players whose chosen numbers are further away receive nothing."
>
> (Nagel 1995, p. 1313)

If the parameter p is chosen such that $0 < p < 1$, the only Nash equilibrium is for everyone to choose 0.[37] This only holds true, though, if all players are unlimitedly strategically sophisticated. If there are players who do not achieve this level of strategic reasoning, the optimal response changes upwards.

The implemented version aims to avoid too much overlap with the usual presentation of the game by setting the allowed range from 1 to 200 (in increments of 0.01) rather than 0 to 100, and the parameter p to ¾ rather than the almost colloquial "guess ⅔ of the average". The unique Nash equilibrium then is for everyone to choose 1.[38]

[36] The game could have been played on the original interval by introducing an experimental currency that would later be translated into €. This, however, would have required an explanation of this currency, changing the immediate understanding of which choice leads to which payoff and potentially distracting from the rules of the game.

[37] If there is discreteness in the implementation of the interval from which numbers can be chosen (e.g., only integers) multiple Nash equilibria can exist. In the case of $p =⅔$, there are two Nash equilibria. Aside from "everyone chooses 0", also "everyone chooses 1" is a Nash equilibrium. Lopez (2001) provides a full characterization of the pure-strategy Nash equilibria if players' choices are restricted to integers.

[38] Despite the discreteness of choices (multiples of 0.01 instead of all real numbers), there is a unique Nash equilibrium, because the lower bound of the choice set is 1 and any $p \in (0, 1)$ ensures uniqueness.

As the Beauty Contest is most widely implemented and analyzed game in the level-k literature, a variety of ways to translate choices in the Beauty Contest game into levels of strategic sophistication have been developed and discussed. This especially concerns the starting point level-0 behavior. For my analysis, I chose to assign levels according to iterative elimination of dominated strategies which is more closely related to the iterative reasoning expected in the Market Game according to the underlying theoretical model.[39] For the iterative elimination of dominated strategies, note first that the target value of "¾ of the average" can never exceed 150 - even if everyone chose the upper bound of the permitted range 200. Thus, any choice above 150 is dominated by any other choice equal to or below 150. As choosing any number above 150 can only be explained by not performing this first step of elimination of dominated strategies, subjects submitting such a choice were assigned a level of $k = 0$. Continuing the iterative process by multiplication with ¾ and rounding to the nearest multiple of 0.01 leads to the assignment of higher levels. To prevent that intervals become too small and k-levels too large compared to the other two elicitation games, a maximum level of $k = 11$ was imposed. To achieve this the assignment procedure was changed for $k \geq 10$, such that the remaining interval of the choice set, the range from 1.00 to 11.27 was split in half to assign levels $k = 10$ and $k = 11$. The complete assignment is shown in Table 3.5.

Beauty-k	Choice in Interval
0	(150.00, 200.00]
1	(112.50, 150.00]
2	(84.38, 112.50]
3	(63.29, 84.38]
4	(47.47, 63.29]
5	(35.60, 47.47]
6	(26.70, 35.60]
7	(20.03, 26.70]
8	(15.02, 20.03]
9	(11.27, 15.02]
10	(6.14, 11.27]
11	[1.00, 6.14]

Table 3.5: Level-k Assignment - Beauty Contest

[39]Other assignment procedures, e.g., the one used in Nagel (1995), have been used as well. Results are mostly statistically insignificant and can be provided upon request.

3.4. Experimental Procedures

Experimental sessions were conducted at the CentERLab of Tilburg University with student subjects, using the z-Tree software package (Fischbacher 2007). Each session lasted about 80 minutes in total including time for individually recording payment information of subjects. On average subjects earned 11.06 € (including the show-up fee). Within each session, subjects were randomly allocated to one of the workstations in the laboratory and completed a series of three parts for a total of 23 periods.

Part 1 consisted of 15 identical periods of the Market Game with all subjects in one session assigned to the same of two treatments (hiding cost of 1.00 € or 2.50 €). Part 2 consisted of 5 identical periods of the Market Game with all subjects assigned to the third treatment (hiding cost of 0.00 €) irrespective of which treatment they were assigned in Part 1. So, the two positive hiding cost treatments (1.00 € v. 2.50 €) were varied between subjects in Part 1, while there was within subjects variation between positive hiding cost and zero cost (1.00 € v. 0.00 € and 2.50 € v. 0.00 €, respectively) from Part 1 to Part 2. Part 3 consisted of 3 non-identical periods implementing the Adding Game, the Money Request Game, and the Beauty Contest each played for one period and in this order. After Part 3, subjects filled in a questionnaire about individual characteristics, such as gender and age, as well as their banking information which was deleted after all payments were processed.

In the Market Game, subjects were made aware that they played the game with the computer alone. They were further told that the seller would act in a way "which maximizes his revenues",[40] but not informed about the exact pricing algorithm of the seller. However, subjects were given the information that the computer has to set one uniform price for all their hidden valuations, while he can adjust his price for all non-hidden valuations. Finally, they were told that the seller knows the range of possible valuations and will be informed about the submitted hiding threshold, but not which exact valuation was assigned if that valuation was equal to or higher than the submitted hiding threshold (cf. Table 3.1).

In the Adding Game, subjects were informed that they played the game with the computer alone and that the computer's decisions are "determined at random and hence are not necessarily the best possible response to your choice". In the Money Request Game, subjects were made aware that they played the game "with a randomly selected other participant in this session", and in the Beauty Contest, they were informed that they played "with all other participants in this session".

[40] In the case of zero marginal cost, this is identical to profit-maximization and avoids having to mention production costs.

All subjects in a session were assigned to the same treatment in Part 1 (1.00 € or 2.50 € hiding cost) and received the same show-up fee. However, the show-up fee differed for the two treatments in Part 1: in sessions with the 1.00 € treatment in Part 1 the show-up fee was set to 4.00 €, in sessions with the 2.50 € treatment in Part 1 the show-up fee was set to 6.00 €.[41]

At the end of each period in Part 1 and Part 2 (i.e., all Market Game periods), subjects were reminded of the hiding threshold they entered, informed about the valuation randomly assigned to them, whether they had decided to hide that valuation, the price the seller set for their valuation, their resulting payoff, whether they paid the hiding cost, and their resulting earnings in this period. Also, they were always informed about the uniform price the seller had set for all their hidden valuations, even when they did not hide the valuation assigned to them in this period. At the end of each period in Part 3, subjects were informed about the relevant choices and resulting earnings in a similar fashion. More details on the experimental instructions can be found in Appendix 3.C.

To avoid income effects within the parts, one period per part was selected at random using the following procedure. Before the start of each part the experimenter asked one subject to pick an envelope from a pile of sealed envelopes, sign it, and place it in a corner to be retrieved at the end of that part. After this part was finished, the same subject verified the signature, opened the envelope, announced the selected period to all other subjects and verified the correct entry by the experimenter into the computer program. The subject returned to the seat and another subject was asked to randomly select an envelope for the next part. At the end of the experiment, subjects were called individually to confirm their banking information and sign a receipt.[42] Their earnings were transferred to their bank accounts within 24 hours.[43]

In total, 10 experimental sessions were conducted (5 for each treatment in Part 1) and a total of 188 subjects participated. One subject participated in two different sessions, which lead to the exclusion of the data record from this subject's second participation, resulting in 187 unique subject observations. Due to a bug in the code[44] which affected 22 subjects, the analysis was further restricted to the 165 unaffected subjects.

[41] On the one side, keeping the hiding cost at 25% of the show-up fee would have required to raise the show-up fee to 10.00 €, larger than any valuation in the Market Game and potentially crowding out subjects' efforts. On the other side, simply compensating for the 1.50 € difference in hiding cost would have asked subjects to invest more than half of their show-up fee to hide. The chosen values set the hiding cost at 41.6% of the show-up fee and the show-up fees for both treatments reasonably close.

[42] All envelopes were opened and subjects could verify that all periods could have been selected.

[43] Due to banking holidays, however, it took up to 7 calendar days for some subjects.

[44] It was discovered after the experiment that subjects that submitted a hiding threshold of 9.50 € received a price of 0.00 € rather than 5.00 € for all hidden valuations. Every subject that submitted a hiding threshold of 9.50 € at least once was excluded to eliminate potentially confounded observations and keep a balanced sample throughout.

3.5. Hypotheses

Given this design different hypotheses can be formulated to address the research questions posed earlier. To assess whether subjects make optimal use of the costly privacy protective option, their observed behavior in the Market Game is tested against the Nash equilibria in the 1.00 € and 2.50 € treatment in Part 1 of the Market Game.

Hypothesis 1 (Nash Equilibrium Predictions)**.**
Subjects' chosen hiding thresholds are equal to 5.00 € in the 1.00 € treatment and to 10.00 € in the 2.50 € treatment.

Subjects might not be able to fully anticipate the algorithm employed by the seller and cognitive constraints might play a role in their decision-making. Driven by the results of theoretical model inspiring this experiment, I first investigate level-k thinking as an alternative explanation for subjects' observed behavior in the Market Game and hypothesize that higher level-k thinking capacities, as measured by the three level-k elicitation games in Part 3, correspond to higher hiding thresholds in the Market Game.

Because level-k thinking describes iterative reasoning in static frameworks, the elicited levels cannot be used to describe a full learning trajectory across all periods. Thus, this analysis focuses on subjects' initial response in the first period of the Market Game akin to their initial reaction in the one-shot level-k elicitation games. Additionally, I hypothesize that level-k thinking is a better explanatory model for behavior in the 2.50 € treatment as compared to the 1.00 € treatment. This stems from the fact that the 2.50 € treatment resembles the equilibrium in the theoretical model more closely, because the Nash equilibrium is at the upper limit of the choice set.

Recall that the different decision environments in the level-k elicitations lead to the assigned levels carrying slightly different notions. As subjects play the Adding Game alone against a computer whose strategy is announced to them, there is no strategic uncertainty, and the assigned level represents subjects pure iterative reasoning capabilities. In the Market Game, though, there is strategic uncertainty because subjects did not exactly know the seller's algorithm (especially so in the first period).

The Money Request Game (played with one other subject) and the Beauty Contest (played with all subjects in the same session), on the contrary, feature strategic uncertainty. However, if cognitive constraints play a role, the presence of strategic uncertainty implies in turn that subjects' choices do not only depend on their own iterative reasoning capabilities but also on their beliefs about the capabilities of others. While potentially capturing the strategic uncertainty of the Market Game better, the two measures should be expected to be noisier, thus less likely to reach statistical significance.

Hypothesis 2 (Level-k Thinking).

(a) Subjects' with higher level-k thinking capacities in the Adding Game, the Money Request Game, and the Beauty Contest choose higher hiding thresholds than subjects with lower level-k thinking capacities in Period 1 of the Market Game.

(b) Effects of level-k thinking on chosen hiding thresholds are larger in the 2.50 € treatment than in the 1.00 € treatment.

(c) The levels elicited from the Adding Game provide a better explanation for behavior in the Market Game than the levels elicited from the Money Request Game and the Beauty Contest.

In contrast to level-k thinking, the alternative behavioral model of reinforcement learning is applicable to the dynamic framework of repeated play in the experiment. As subjects interact with the seller multiple times and observe the result of each interaction, they can adjust their behavior in subsequent rounds based on experiences made earlier. More specifically, subjects are expected to respond to realized payoffs. Gains positively reinforce past behavior, whereas losses negatively reinforce past behavior. As the sellers' perfect price discrimination algorithm leaves subjects with a zero payoff gains and losses can only result from subjects' hidden valuations. Thus, I hypothesize that in response to realizing a gain from a hidden valuation being selected, subjects decrease their hiding threshold (i.e., increase the number of hidden valuations). On the other hand, realizing a loss from a hidden valuation should lead to subjects increasing their hiding threshold (i.e., decreasing the number of hidden valuations). In light of the theoretical and experimental literature on prospect theory and loss aversion (Kahneman and Tversky (1979), Kahneman, et al. (1990)), I hypothesize further that the negative reinforcement by losses will be stronger than the positive reinforcement by gains.

Hypothesis 3 (Reinforcement Learning).

(a) Subjects choose a higher hiding threshold after having realized a loss.

(b) Subjects choose a lower hiding threshold after having realized a gain.

(c) Reinforcement learning effects are expected to be larger for losses than for gains.

Lastly, I analyze the consequences of making privacy protection available for free. Lowering hiding cost is in general expected to lead to more hiding as it weakly increases payoffs from hidden valuations (cf. Table 3.2). This within-subject treatment is tested comparing the immediate difference between the last period under the old regime (Period 15) and the first one following the exogenous change (Period 16) as well as the overall differences between Part 1 and Part 2. Recall moreover, that the 0.00 € treatment is particular due to the multiplicity of optimal choices. From the consumers' perspective all hiding thresholds $\hat{v} \leq 5.00$ € lead to the same payoff. Note, though, that among consumers' optimal hiding thresholds only $\hat{v} = 5.00$ € is welfare-maximizing in the theoretical model framework[45] as not hiding valuations below the optimal monopoly price p_M allows the seller to perfectly price discriminate and recover the otherwise resulting deadweight loss from monopoly pricing. In theory, this equilibrium can therefore be reached if the seller offers a marginal benefit for valuations below the monopoly price to be revealed. However, in the experiment no such marginal benefit was offered. Hence, subjects are not expected to prefer any particular optimal hiding threshold by standard Nash equilibrium predictions. However, if subjects understand that in the 0.00 € treatment hiding a valuation never leads to a lower payoff than not hiding it (but can create the possibility to gain), choosing to hide all valuations (i.e., $\hat{v} = 0.00$ €) is particularly appealing. I therefore hypothesize that among the multiple Nash equilibrium choices subjects have a strong tendency to choose to hide all valuations ($\hat{v} = 0.00$ €).

Hypothesis 4 (No Hiding Cost)**.**

(a) Compared to Period 15 (1.00 € or 2.50 €) subjects' chosen hiding thresholds are lower in Period 16 (0.00 €).

(b) Compared to Part 1 (1.00 € or 2.50 €) subjects' chosen hiding thresholds are lower in Part 2 (0.00 €).

(c) In the 0.00 € treatment, subjects' chosen hiding thresholds are not larger than 5.00 € (the upper bound for all Nash equilibria).

(d) In the 0.00 € treatment, the hiding threshold $\hat{v} = 0.00$ € *is chosen most often.*

3.6. Results

In this section, I first describe subjects' observed behavior in the Market Game in Part 1 in general in light of Hypothesis 1 and the two between-subjects treatments (1.00 € and 2.50 €). This is followed by a description of subjects' behavior in the three level-k elicitation games and regression results with respect to Hypothesis 2. Afterwards, regression results concerning Hypothesis 3 on reinforcement learning are presented. An analysis of the within-subjects hiding cost reduction (Hypothesis 4) concludes this section.

[45] For an explicit welfare analysis within the experiment the inclusion of a computer player, programmed by the experimenter, creates an ambiguity: a sharp distinction between the seller's profits are payoffs "inside the game" and the experimenter's budget "outside the game" is hard to draw. To avoid this ambiguity the welfare claim refers back to the theoretical model framework, where the seller's profits are included in total welfare.

3.6.1. Market Game - Between Subjects Treatments

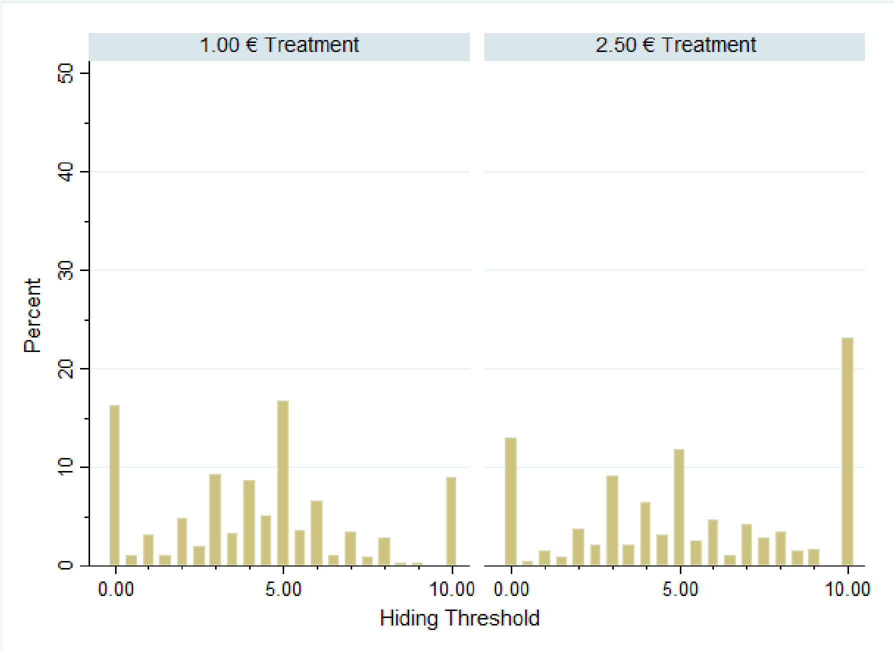


Figure 3.1: Hiding Threshold Frequency in Part 1 in Percent

Figure 3.1 shows the frequency of chosen hiding thresholds in percent for both treatments aggregated across all 15 periods of Part 1.[46] Note first, that in both treatments the respective Nash equilibrium is the modal choice. In the 1.00 € treatment, 16.78% of the chosen hiding thresholds (i.e., 214 out of 1275) are equal to Nash equilibrium of 5.00 €. In the 2.50 € treatment, 23.08% of the chosen hiding thresholds (i.e., 277 out of 1200) are equal to the Nash equilibrium of 10.00 €. In both treatments, the second most frequent choice of 0.00 € as hiding threshold is a close second with 16.31% (i.e., 208/1275) in the 1.00 € treatment and 13.08% (i.e., 157/1200) in the 2.50 € treatment, indicating that subjects do not necessarily understand that hiding all valuations is not optimal in either treatment (and actually leads to the lowest possible expected payoff in the 2.50 € treatment).

The mean of the chosen hiding thresholds in Part 1, $\bar{v}_1$, is equal to 4.18 € in the 1.00 € treatment and equal to 5.37 € in the 2.50 € treatment. In both treatments $\bar{v}_1$ is found to be statistically highly significantly different ($p < 0.001$) from their respective Nash equilbria according to a one-sample t-test (against the respective alternatives H_1: $\bar{v}_1^{1.00\,€} \neq 5.00$ €, and H_1: $\bar{v}_1^{2.50\,€} < 10.00$ €, respectively). Thus, despite the respective equilibria being the modal choice in both treatments, there is substantial deviation from Nash equilibrium play, rejecting the Nash equilibrium predictions in Hypothesis 1.

[46]Recall that subjects having chosen 9.50 € were excluded from the analysis.

Additionally, the difference $\bar{v}_1^{2.50\,€} - \bar{v}_1^{1.00\,€} = 1.19$ € is statistically significant at the 0.1% level based on a two-sample Mann-Whitney-U test, indicating that there is a between-subjects treatment effect from varying the hiding cost. Figure 3.2 displays the means for both treatments for each of the 15 periods.

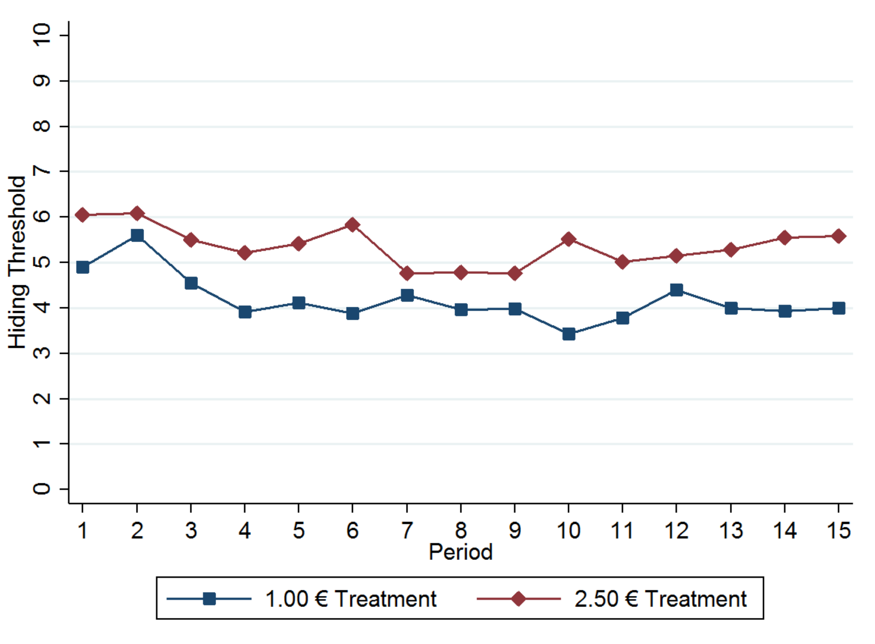
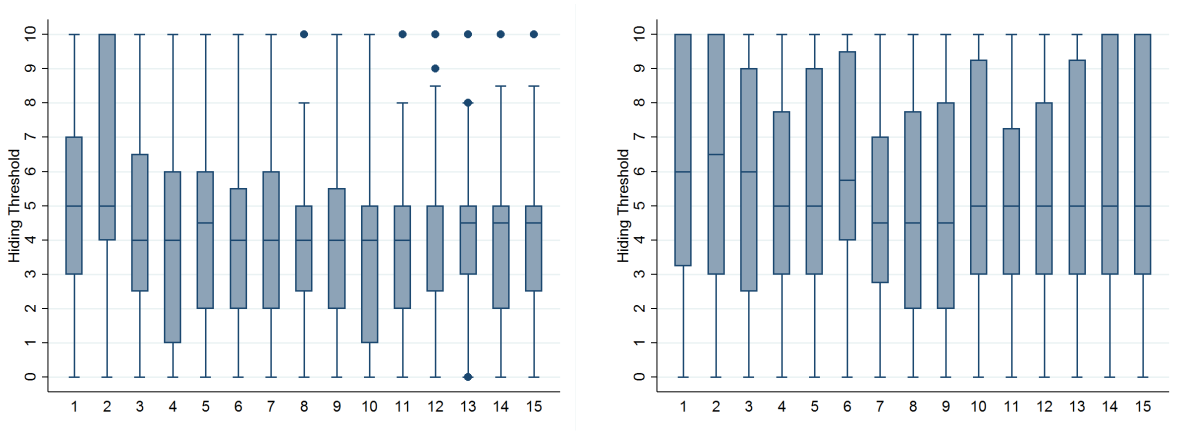


Figure 3.2: Hiding Thresholds in Part 1 by Treatment (Means)

In the 1.00 € treatment, $\bar{v}_1$ is statistically significantly different (at the 5% level) from the Nash equilibrium prediction in all but the first 3 periods (see Table 3.12 in Appendix 3.A for statistics by period). However, the equilibrium prediction of 5.00 € in the 1.00 € treatment is at the center of the permitted action space, which might have served as a focal point to start with. Thus, the non-detectable difference in the first periods is not necessarily indicative of intentional equilibrium play.

In the 2.50 € treatment, $\bar{v}_1$ is statistically different from the Nash equilibrium prediction at the 0.1% level in every period. However, as Figure 3.2 shows, there are fluctuations across periods, from a mean hiding threshold above 6.00 € in the initial periods down to 4.77 € and 4.79 € in periods 7-9 and subsequently increasing again throughout the last periods. The initially decreasing hiding thresholds are difficult to explain by anything other than subjects trying to get a better understanding of the seller's pricing mechanism. The slight upwards trend back in later periods, however, could result from subjects gaining experience.

Analyzing the development of hiding thresholds further, Figure 3.3 shows distribution box plots for all 15 periods in Part 1, separated by treatment. The boxes visualize the lower and upper quartiles (25th and 75th percentiles) around the median of the chosen hiding thresholds, which itself is indicated by the horizontal bar inside each box. Outliers are defined as values further away from the respective quartile than 3/2 the range between the two central quartiles.



(a) 1.00 € Treatment

(b) 2.50 € Treatment

Figure 3.3: Hiding Tresholds in Part 1 by Treatment (Median and Quartiles)

In the 1.00 € treatment, chosen hiding thresholds concentrate closer to the median as subjects gain experience. Observations at 10.00 € become less frequent and are considered outliers in the last 5 periods as the range between the 25th and the 75th percentile range. The range decreases especially away from the upper bound of the choice set, indicating that subjects do learn to avoid the range from $\hat{v} = 8.00$ € to $\hat{v} = 9.50$ € where expected profits are negative in the 1.00 € treatment. However, such narrowing is not as strongly observed on the lower bound of the choice set despite the optimal choice being located in the center of the choice set. This might be due to the fact that subjects choosing hiding thresholds $\hat{v} \leq 5.00$ € always observe the uniform price for all hidden valuations to be set to 5.00 € by the monopolist. At this price these subjects realize positive payoffs in 35% of the cases (for any assigned valuation $v \geq 6.50$ €) irrespective of their exact hiding threshold. The chance for negative payoffs, on the other hand, *does* depend on their exact hiding threshold and ranges from 10% (if $\hat{v} = 5.00$ €) to 60% (if $\hat{v} = 0.00$ €), but their expected payoffs remain positive nonetheless. The analysis on reinforcement learning in Section 3.6.3 sheds additional light on this potential explanation.

In the 2.50 € treatment, the range of chosen hiding thresholds does not exhibit such an overall trend to more concentration in later periods. While subjects' choices become more concentrated initially, the range expands again in later periods. It becomes apparent that also in this treatment choices in the upper half of the choice set seem to vary stronger than in the lower half of the choice set. Recalling that the optimal choice in this treatment is to not hide any valuation at all (i.e., $\hat{v} = 10.00$ €), this is to be expected. However, the direction of the variation is unexpected as subjects choices move away from the Nash equilibrium prediction. Focusing on the aforementioned drop in mean hiding thresholds in periods 7-9 and the subsequent slight upwards trend, this is noticeable, though. While from period 6 to 7 the 75th percentile drops by 2.50 €, the 25th percentile drops only by 1.25 € and hardly changes afterwards. Thus, it seems that subjects that are initially not hiding their valuations are starting to hide more rather than those who already hide a many of their valuations to engage in even more hiding. Following the same reasoning as for the 1.00 € treatment, note that subjects with a hiding threshold $\hat{v} \leq 5.00$ € have only a 20% chance of realizing a positive payoff in the 2.50 € treatment. Further, their chance to realize a negative payoff ranges from 25% (if $\hat{v} = 5.00$ €) to 75% (if $\hat{v} = 0.00$ €). Thus, despite the respective Nash equilibria being the modal choices across all periods, substantial deviations from optimal play are observed, giving room for the analysis of the alternative explanations of level-k thinking and reinforcement learning.

3.6.2. Level-k Thinking

Tables 3.6, 3.7, and 3.8 present the elicited k-levels in the Adding Game, the Money Request Game, and the Beauty Contest, respectively. Each table shows the absolute frequencies (Freq.) of each assigned level as well as their proportions (Prop.) within each initial treatment group (1.00 € and 2.50 €) and the entire subject population (Total).

First note that in the Adding Game the vast majority of subjects is assigned a level of either $k = 0$ or $k = 1$ (91.52%)[47], and that levels assigned in the Money Request Game are slightly higher (75.76% for $k \leq 2$). The levels assigned in the Beauty Contest, though, span a much larger set of levels (78.18% for $k \leq 4$).[48] Across the two Market Game treatments, however, there are no statistically significant differences in the distribution of levels within each elicitation game.[49] However, pairwise correlations across the elicitation methods are weak and, if at all different from zero, imply a negative relationship between the assigned levels from the three games.[50]

[47]Of the 96 subjects that were assigned a level of $k = 0$, 78 (overall proportion 0.47) won the game without having forced it, whereas 18 (overall proportion 0.11) did not win the game

[48]Such high levels are not usually observed in Beauty Contest games. The larger range of possible choices from 1 to 200 in combination with the higher value of $p = 0.75$ might imply a too slow convergence in the interval bounds to the Nash equilibrium at 1, in turn resulting in levels being inflated. For the entire distribution of the submitted choices, see Figure 3.6 in Appendix 3.B.

[49]The two sample Mann-Whitney U test statistics and associated probabilities are: in the Adding Game $z = 0.069$ (Prob $> |z| = 0.9450$), in the Money Request Game $z = -0.416$ (Prob $> |z| = 0.6774$), and in the Beauty Contest $z = -0.597$ (Prob $> |z| = 0.5507$).

[50]Pearson's correlation coefficient between Adding-k and Request-k is equal to -0.1980, between Adding-k and Beauty-k is equal to -0.0206, and between Request-k and Beauty-k is equal to -0.0456.

Adding-k		0	1	2	3	4	5	6	7	N
1.00 €	Freq.	50	26	6	1	1	1	0	0	85
	Prop.	0.59	0.31	0.07	0.01	0.01	0.01	0.00	0.00	
2.50 €	Freq.	46	29	3	1	0	0	0	1	80
	Prop.	0.58	0.36	0.04	0.01	0.00	0.00	0.00	0.01	
Total	Freq.	96	55	9	2	1	1	0	1	165
	Prop.	0.58	0.33	0.05	0.01	0.01	0.01	0.00	0.01	

Table 3.6: Level-k Elicitation - Adding Game

Request-k		0	1	2	3	4	5	6	7	8	9	N
1.00 €	Freq.	11	22	32	10	3	4	0	1	1	1	85
	Prop.	0.13	0.26	0.38	0.12	0.04	0.05	0.00	0.01	0.01	0.01	
2.50 €	Freq.	11	16	33	12	2	0	2	1	0	3	80
	Prop.	0.14	0.2	0.41	0.15	0.03	0.00	0.03	0.01	0.00	0.04	
Total	Freq.	22	38	65	22	5	4	2	2	1	4	165
	Prop.	0.13	0.23	0.39	0.13	0.03	0.02	0.01	0.01	0.01	0.02	

Table 3.7: Level-k Elicitation - Money Request Game

Beauty-k		0	1	2	3	4	5	6	7	8	9	10	11	N
1.00 €	Freq.	9	8	18	18	12	3	4	3	4	0	2	4	85
	Prop.	0.11	0.09	0.21	0.21	0.14	0.04	0.05	0.04	0.05	0.00	0.02	0.05	
2.50 €	Freq.	5	8	12	22	17	6	2	0	2	2	4	0	80
	Prop.	0.06	0.10	0.15	0.28	0.21	0.08	0.03	0.00	0.03	0.03	0.05	0.00	
Total	Freq.	14	16	30	40	29	9	6	3	6	2	6	4	165
	Prop.	0.08	0.10	0.18	0.24	0.18	0.05	0.04	0.02	0.04	0.01	0.04	0.02	

Table 3.8: Level-k Elicitation - Beauty Contest

The model estimating the relationship of subjects hiding thresholds in the first period and the different level-k elicitation is given by Equation (3.4) (with a slight abuse of notation due to the definition of the chosen hiding threshold as $\hat{v}$):

$$\hat{v}_i = \beta_0 + \beta_1 k_i^{\text{Add}} + \beta_2 k_i^{\text{Req}} + \beta_3 k_i^{\text{Beau}} + \delta X_i + \zeta S_i + \varepsilon_{it}, \tag{3.4}$$

where $\hat{v}_i$ denotes the hiding threshold chosen by subject i, k_i^{Add}, k_i^{Req}, and k_i^{Beau} are the k-levels assigned to subject i in the Adding Game, the Money Request Game, and the Beauty Contest, respectively, X_i is a set of subject characteristics (age and gender), S_i is the set of session dummies, and ε_i denotes the error term.

Table 3.9 presents the results from estimating different specifications of the model in Equation (3.4) with the sample being separated by the two treatments (1.00 € treatment in columns (1)-(4), 2.50 € treatment in columns (5)-(8)). Columns (1)-(3) and (5)-(7) show results from regressions on the levels from each level-k elicitation method separately, while columns (4) and (8) report the results regressing on them jointly. Robust standard errors are reported in parentheses.

Hid. Thresh.	**1.00 €**				**2.50 €**			
Period 1	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Adding-k	-0.62^{+}			-0.52	0.67*			0.62^{+}
	(0.340)			(0.335)	(0.310)			(0.332)
Request-k		-0.19		-0.21		-0.24^{+}		-0.18
		(0.192)		(0.174)		(0.133)		(0.137)
Beauty-k			0.34**	0.31**			-0.03	-0.07
			(0.108)	(0.106)			(0.159)	(0.133)
Age	0.23***	0.23***	0.25***	0.25***	-0.27*	-0.25*	-0.25*	-0.27*
	(0.058)	(0.060)	(0.062)	(0.061)	(0.111)	(0.109)	(0.112)	(0.108)
Female	0.74	0.88	1.17^{+}	1.06^{+}	0.32	0.12	0.02	0.32
	(0.648)	(0.648)	(0.598)	(0.605)	(0.733)	(0.741)	(0.740)	(0.737)
Session FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	85	85	85	85	80	80	80	80

Notes: $^{+}$, *, **, *** indicate significance at the 10%, 5%, 1% and 0.1% level, respectively. Coefficients are based on OLS regressions of subjects' hiding thresholds in period 1 on the indicated variables. Robust standard errors are reported in parentheses.

Table 3.9: Level-k Thinking in the Market Game (Period 1)

In the 1.00 € treatment, there is mild statistical evidence that a higher level in the Adding Game corresponds to a 0.62 € decrease in a subject's hiding threshold in the first period of the Market Game. While there is no statistical evidence for higher levels in the Request Game, there is strong statistical evidence for a higher level in the Beauty Contest to correspond to a 0.34 € increase in a subject's hiding threshold in the first period of the Market Game. In the joint regression reported in column (4), however, the effect of the levels elicited from the Adding Game is not statistically significant, while the effect of the levels from the Beauty Contest is still detected.[51] Both effect sizes, for the Adding Game and the Beauty Contest, are roughly of the magnitude of one increment (0.50 €) in the Market Game setting. While the theoretical model would predict an increase of the size of the hiding cost for a one level increase in sophistication, it needs to be taken into account that in the 1.00 € treatment the Nash equilibrium is given by a hiding threshold 5.00 € rather than at the permitted maximum. This might also explain that the sign for the Adding Game level is negative as in the 1.00 € treatment it is not always the case that choosing a higher hiding threshold is necessarily more beneficial.

In the 2.50 € treatment, the results indicate that there is statistical evidence that a higher level in the Adding Game corresponds to a 0.67 € increase in a subject's hiding threshold in the first period of the Market Game. There is also mild evidence that a higher level in the Request Game corresponds to a 0.24 € decrease in a subject's hiding threshold. No evidence is found for the levels from the Beauty Contest to have statistically significant effects on behavior in this treatment. In the joint regression (column (8)), the effect of the levels from the Request Game is no longer statistically significant, while the effect of the levels from the Adding Game persists. Similar to the 1.00 € treatment the effect size for the Adding Game levels is estimated to be roughly of the magnitude of one increment (0.50 €) in the Market Game setting, whereas the effect size for the Request Game levels is estimated at about half the magnitude of one such increment. Contrary, to the 1.00 € treatment, a decrease of the hiding threshold is always reducing subjects' payoffs. Hence the effect of the Request Game levels is contradictory to representing "more" sophistication in this setting.

In light of Hypothesis 2, it has to be concluded that there is some, but no general, statistical evidence for higher levels in all level-k elicitations to result in higher hiding thresholds in the Market Game.[52] While the effects of the levels elicited from the Money Request Game and the Beauty Contest are only statistically significant in one of the treatments, the effects of the levels from the Adding Game are statistically significant in both treatments. At first glance, this provides support for the prediction posed in

[51] The effect from the Beauty Contest, however, is not completely robust to other level-k assignment procedures.

[52] This also holds if the average hiding threshold across all 15 periods in Part 1 is the dependent variable. Regression results on the average hiding threshold can be found in Table 3.14 in Appendix 3.A.

Hypothesis 2 that the Adding Game might outperform the two other level-k elicitations. However, as the Adding Game effect is not of the predicted sign in the 1.00 € treatment, and as the Beauty Contest levels are unusually high, these results should not be overstated.

3.6.3. Reinforcement Learning

With respect to the analysis of the effects of reinforcement learning, I estimated two different models for each treatment. The first model considers only whether a subject experienced any gain or loss in the previous period to assess the overall effect, while the second model takes into account the absolute size of the realized positive or negative payoff in the previous period. As the analysis requires lagged terms, effects are only identified for periods 2 to 15. The two models are given by Equations (3.5) and (3.6) (with a slight abuse of notation due to the definition of the chosen hiding threshold as $\hat{v}$):

$$\hat{v}_{it} = \beta_0 + \beta_1 \mathbf{I}(y_{it-1} < 0) + \beta_2 \mathbf{I}(y_{it-1} > 0) + \gamma \hat{v}_{it-1} + \delta X_i + \zeta S_i + \theta_t + \varepsilon_{it}, \tag{3.5}$$

$$\hat{v}_{it} = \beta_0 + \beta_1 \mathbf{I}(y_{it-1} < 0)|y_{it-1}| + \beta_2 \mathbf{I}(y_{it-1} > 0) y_{it-1} + \gamma \hat{v}_{it-1} + \delta X_i + \zeta S_i + \theta_t + \varepsilon_{it}, \tag{3.6}$$

where $\hat{v}_{it}$ denotes the hiding threshold chosen by subject i in period t, $\mathbf{I}$ is the indicator function operator, y_{it-1} denotes the payoff of subject i in period $t-1$, $\hat{v}_{it-1}$ is the hiding threshold chosen by subject i in period $t-1$, X_i is a set of period-invariant subject characteristics (age and gender), S_i is the set of session dummies, θ_t is the set of period dummies, and ε_{it} denotes the error term.

Table 3.10 presents the results from estimating these models with the sample being separated by the two treatments (1.00 € treatment in columns (1)-(2), 2.50 € treatment in columns (3)-(4)). Columns (1) and (3) show the results from the model in Equation (3.5), whereas columns (2) and (4) show the results from the model in Equation (3.6). Standard errors clustered at the subject level are reported in parentheses.

Note first, that in neither specification in either treatment the coefficients of the gain variables are statistically significant, whereas in both treatments there is statistical evidence for an effect of having realized a loss in the previous period, controlled for the previously chosen hiding treshold (and thus implicitly for the chance of incurring a loss).

In the 1.00 € treatment, there is significant statistical evidence that a negative payoff of any size leads to an increase of the hiding threshold in the subsequent period by 0.65 €. When regressing on the exact size of the loss in the previous period instead, the effect size decreases slightly to an estimated increase of 0.56 € per experienced loss of 1.00 €.

	1.00 €		2.50 €	
Hiding Threshold	(1)	(2)	(3)	(4)
Loss-Before	0.65*		1.00***	
	(0.279)		(0.271)	
Gain-Before	-0.00		-0.03	
	(0.176)		(0.312)	
LossSize-Before		0.56+		0.43***
		(0.291)		(0.122)
GainSize-Before		-0.07		-0.22
		(0.070)		(0.223)
Age	0.02	0.02	-0.07+	-0.07+
	(0.029)	(0.029)	(0.040)	(0.039)
Female	-0.41*	-0.40*	-0.56*	-0.55*
	(0.196)	(0.196)	(0.280)	(0.277)
HT-Before	0.27***	0.25***	0.40***	0.38***
	(0.044)	(0.042)	(0.056)	(0.058)
Period FE	Yes	Yes	Yes	Yes
Session FE	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes
Observations	1190	1190	1120	1120

Notes: +, *, **, *** indicate significance at the 10%, 5%, 1% and 0.1% level, respectively. Coefficients are based on OLS regressions of subjects' hiding thresholds from period 2 to period 15 on the indicated variables. Standard errors clustered at the subject level are reported in parentheses. The dummy variables Loss-Before and Gain-Before equal 1 if the one-period lagged payoff was negative (Loss-Before) or positive (Gain-Before). LossSize-Before and GainSize-Before denote the absolute value of the one-period lagged payoff if this payoff was negative (LossSize-Before) or positive (GainSize-Before). HT-Before denotes the one-period lagged hiding threshold.

Table 3.10: Reinforcement Learning in the Market Game in Part 1

Taking into account that the specification in column (2) differentiates between the two possible negative payoffs (-0.50 € and -1.00 €) this decrease is not unexpected.

In the 2.50 € treatment, the reinforcement learning effect of a negative payoff of any size is estimated to lead to an increase of the hiding threshold in the subsequent period by 1.00 €. For the exact size of the loss in the previous period the effect size decreases to an average increase of 0.43 € for a marginal experienced loss of 1.00 €. The difference between the two estimates is larger in this treatment due to larger number of possible negative payoffs (-0.50 €, -1.00 €, -1.50 €, -2.00 € and -2.50 €) that the binary variable Loss-Before pools together, which also explains the magnitude of the estimated effect being equal to two increments in the Market Game. The size of the effect controlling for the size of the loss is again similar in magnitude to one increment in the Market Game.

Thus, with regard to Hypothesis 3 it can be concluded that there is substantial statistical evidence that subjects' behavior is consistent with reinforcement learning. However, this only holds as far as losses are concerned. Realizing gains does not seem to change subjects' behavior in either direction.

3.6.4. Market Game - Within Subjects Treatment

Figure 3.4 shows the frequency of chosen hiding thresholds in percent for both treatments aggregated across all 5 periods of Part 2 where hiding cost was set to 0.00 €. Note first, that in both treatments the hiding threshold of 0.00 € (i.e., hiding *all* valuations) is chosen more frequently than all other hiding thresholds combined in line with Hypothesis 4. In the 1.00 € treamtent, 54.35% of the chosen hiding thresholds are equal to 0.00 €. In the 2.50 € treatment, 58.50% of the chosen hiding thresholds are equal to 0.00 €. Furthermore, 92.71% of the hiding thresholds chosen in the 1.00 € treatment and 93.50% in the 2.50 € treatment are in line with the Nash equilibrium prediction $\hat{v} \leq 5.00$ €.

Figure 3.5 shows the immediate reaction of subjects to the reduction of hiding cost to 0.00 € by depicting the distributions of hiding threshold choices for the last period in Part 1 (Period 15) and the first period in Part 2 (Period 16), separated by the two initial treatments (1.00 € and 2.50 €). In line with the prediction in Hypothesis 4 the reduced hiding cost leads to an immediate and statistically highly significant decrease in hiding thresholds. The Wilcoxon signed rank test for within-subject treatment effects confirms this for the immediate effect (Period 15 v Period 16: $p < 0.001$ in both treatments), as well as for the overall means in both parts (Part 1 v Part 2: $p < 0.001$ in both treatments). Thus, the effect of the policy treatment does not only appear on an aggregate but also on the individual level.

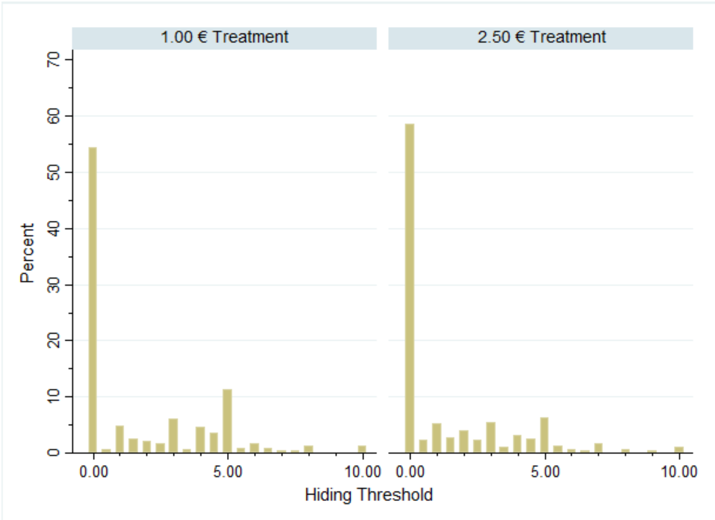
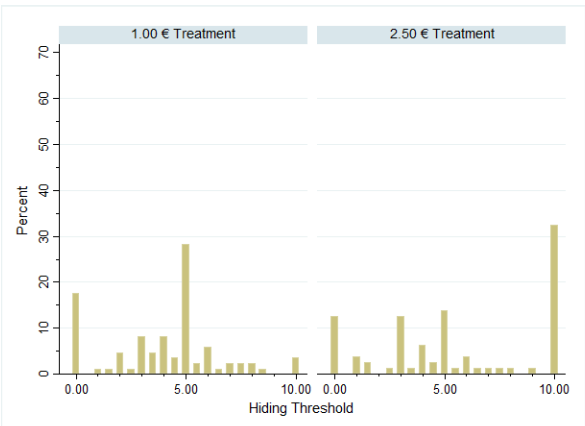
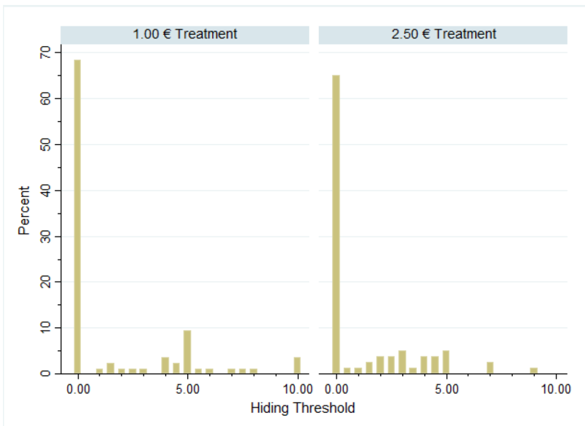


Figure 3.4: Hiding Threshold Distribution in Part 2 in Percent
Treatment 1 - 1.00 €, Treatment 2 - 2.50 €



(a) Period 15



(b) Period 16

Figure 3.5: Hiding Threshold Frequency in Period 15 and Period 16 in Percent
Period 15: Positive Hiding Cost; Period 16: Zero Hiding Cost

3.7. Discussion

In this paper, I analyzed consumer behavior in a laboratory experiment where subjects faced a computerized seller. Unless consumers made use of a costly hiding technique to hide (some) of their valuations for the offered good, the seller's algorithm implemented perfect price discrimination. Despite some differences between the experimental implementation and the model which inspired it, the model's central proposition that a costly privacy-protective option is used, despite an expected loss for consumers, receives empirical backing.

Although the respective Nash equilibrium is the modal choice in each treatment, I found that subjects, by and large, do not make optimal hiding choices when doing so comes at a cost. The results underline that consumers do not fully grasp the complex ramifications of privacy-related choices when they are matched with an optimizing computerized seller.

Considering alternative explanations that address such cognitive constraints, I found some evidence for level-k thinking measures to explain the observed behavior in the first period of each treatment. First, the iterative reasoning levels from the Adding Game are estimated to decrease the hiding threshold by 0.62 € in the 1.00 € treatment. In the 2.50 € treatment, they are estimated to increase the first period hiding threshold by 0.67 €. Second, the k-levels assigned based on the Beauty Contest are only found to be statistically significant in the 1.00 € treatment, where they are estimated to increase the hiding threshold by 0.34 €. Third, the Money Request Game levels are only found to be mildly statistically significant in the 2.50 € treatment and estimated to decrease the hiding threshold by 0.24 €. In general, the size of statistically significant effects is about the size of the increment between any two of the possibly assigned discrete valuations in the Market Game, which is set at 0.50 €.

The experiment provides stronger evidence for the second alternative explanation of reinforcement learning, implying that subjects adjust their behavior in response to previously realized payoffs. Although there is no evidence that subjects react to realized gains, there is substantial evidence that incurring losses leads to increased hiding thresholds in subsequent periods. In both treatments, subjects increased their hiding threshold by about 0.50 € for every 1.00 € loss incurred in the preceeding period.

Mapping these results back to the theoretical model in Dengler and Prüfer (2017) suggests, that in a repeated setting not all consumers, but only those that incur a loss, would take the outcome into account in a potential belief updating process. While the proposed level-k model itself is a static game and does not allow for this type of updating, this could be incorporated by combining multi-period behavior-based price discrimination models with level-k distributions.

Lastly, I found strong evidence for the efficacy of reducing hiding cost to 0.00 € resembling a policy change towards a privacy-by-default regime. In a within-subject treatment design, this resulted in an immediate and drastic decrease in chosen hiding thresholds. An absolute majority of subjects opted for the hiding threshold at which all information about their valuations (except for the general distribution) is hidden from the seller.

Noting that subjects exhibit such a strong tendency to not opt for information sharing in this setting, provides an important caveat to policy recommendations. If consumers understand that available information can be used by predictive algorithms to price discriminate against them (despite not exactly understanding how or why), they might simply decide to hide everything as soon as doing so comes for free. Thus, while reducing hiding cost increases consumer surplus, it may lead to too little information being disclosed and efficiency gains made possible by predictive algorithms might get lost in the process.

Appendix 3.A Additional Tables

Table 3.11: Expected Profits $\mathbb{E}(\pi)$ for all Prices p_A for Hidden Valuations

p_A	$Pr(v \geq p_M)$	$\mathbb{E}(\pi)$
9.50	0.05	0.475
9.00	0.10	0.900
8.50	0.15	1.275
8.00	0.20	1.600
7.50	0.25	1.875
7.00	0.30	2.100
6.50	0.35	2.275
6.00	0.40	2.400
5.50	0.45	2.475
5.00	**0.50**	**2.500**
4.50	0.55	2.475
4.00	0.60	2.400
3.50	0.65	2.275
3.00	0.70	2.100
2.50	0.75	1.875
2.00	0.80	1.600
1.50	0.85	1.275
1.00	0.90	0.900
0.50	0.95	0.475
0.00	1.00	0.000

Table 3.12: Hiding Thresholds in Part 1

	1.00 € (N=85)		**2.50 €** (N=80)		**1.00 € v. 2.50 €**	
Period	**Mean**	**St.Err.**	**Mean**	**St.Err.**	**MW**	**KS**
1	4.91	0.344	6.06***	0.385	−1.16*	0.20*
2	5.60	0.390	6.09***	0.380	−0.49	0.12
3	4.55	0.346	5.51***	0.406	-0.95^{+}	0.24**
4	3.91**	0.342	5.21***	0.364	−1.30*	0.20*
5	4.12**	0.318	5.43***	0.366	−1.30*	0.21*
6	3.89***	0.300	5.83***	0.360	−1.94***	0.27**
7	4.28*	0.328	4.77***	0.347	−0.49	0.10
8	3.96***	0.284	4.79***	0.391	−0.83	0.21*
9	3.98***	0.291	4.77***	0.405	−0.79	0.20*
10	3.44***	0.279	5.52***	0.377	−2.08***	0.32***
11	3.79***	0.269	5.02***	0.369	−1.23*	0.20*
12	4.39*	0.303	5.16***	0.368	−0.76	0.16
13	4.01***	0.260	5.29***	0.378	−1.28*	0.23*
14	3.94***	0.276	5.56***	0.380	−1.63**	0.29***
15	4.01***	0.275	5.59***	0.405	−1.59*	0.30***
Total	4.18***	0.081	5.37***	0.098	−1.19***	0.20***

Notes: $^{+}$, *, **, *** indicate significance at the 10%, 5%, 1% and 0.1% level, respectively. In the 1.00 € treatment, means are tested using a two-sided one-sample t-Test against the null hypothesis that $\hat{v} = 5.00$ €. In the 2.50 € treatment, means are tested using a one-sided one-sample t-test against the null hypothesis that $\hat{v} = 10.00$ €. The difference in means between the 1.00 € treatment and the 2.50 € treatment is reported in column MW, where significance is based on a two-sample Mann-Whitney-U test under the null hypothesis that the treatments are equal. The maximum difference between the cumulative distribution functions of hiding thresholds in the two treatments is reported in column KS where significance is based on a Kolmogorov-Smirnov test under the null hypothesis that treatments are equal against the alternative that $F(\hat{v}^{1.00\,€}) < F(\hat{v}^{2.50\,€})$.

Table 3.13: Hiding Thresholds in Part 2

	1.00 € (N=85)	**2.50 €** (N=80)	**1.00 € v. 2.50 €**		**1.00 € & 2.50 €** (N=165)		
Period	**Mean**	**Mean**	**MW**	**KS**	**Mean**	$\hat{v} = 0.00$	$\hat{v} \leq 5.00$
16	1.61 (0.297)	1.28 (0.233)	0.33	0.15	1.45 (0.190)	0.67	0.93
17	1.73 (0.250)	1.69 (0.265)	0.04	0.14	1.71 (0.181)	0.52	0.93
18	1.89 (0.261)	1.55 (0.245)	0.34	0.11	1.72 (0.179)	0.51	0.92
19	1.85 (0.243)	1.36 (0.251)	0.49^{+}	0.09	1.61 (0.175)	0.55	0.95
20	2.06 (0.276)	1.48 (0.269)	0.58	0.10	1.78 (0.194)	0.57	0.92
Total	1.83 (0.119)	1.47 (0.113)	0.36^{+}	0.09^{+}	1.65 (0.082)	0.56	0.93

Notes: $^{+}$, *, **, *** indicate significance at the 10%, 5%, 1% and 0.1% level, respectively. Means are tested against the Nash equilibrium prediction that $\hat{v} \leq 5.00$ € using a one-sided one-sample t-test. The difference in means between the 1.00 € and 2.50 € treatment is reported in column MW, where significance is based on a two-sample Mann-Whitney-U test under the null hypothesis that the 1.00 € and 2.50 € treatments are equal. The maximum difference between the cumulative distribution functions of hiding thresholds in the 1.00 € and 2.50 € treatment is reported in column KS where significance is based on a Kolmogorov-Smirnov test for the larger difference when comparing distribution functions (i.e., without *a priori* judgment about the two distributions).

Table 3.14: Level-k Thinking in the Market Game (Average Hiding Threshold)

Hid. Thresh.	**1.00 €**				**2.50 €**			
Average	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Adding-k	-0.08			-0.06	0.06			0.01
	(0.156)			(0.162)	(0.153)			(0.153)
Request-k		0.10		0.10		-0.07		-0.07
		(0.071)		(0.072)		(0.070)		(0.068)
Beauty-k			-0.00	-0.00			0.06	0.06
			(0.044)	(0.046)			(0.084)	(0.086)
Age	0.05	0.05	0.05	0.05	-0.12*	-0.12*	-0.12*	-0.12*
	(0.037)	(0.037)	(0.038)	(0.037)	(0.054)	(0.052)	(0.050)	(0.054)
Female	-0.38	-0.38	-0.37	-0.39	-0.77^{+}	-0.77^{+}	-0.76*	-0.74^{+}
	(0.248)	(0.248)	(0.254)	(0.262)	(0.409)	(0.391)	(0.380)	(0.402)
Session FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	85	85	85	85	80	80	80	80

Notes: $^{+}$, *, **, *** indicate significance at the 10%, 5%, 1% and 0.1% level, respectively. Coefficients are based on OLS regressions of subjects' mean hiding treshold across all 15 periods of Part 1 on the indicated variables. Robust standard errors are reported in parentheses.

Appendix 3.B Additional Graphs

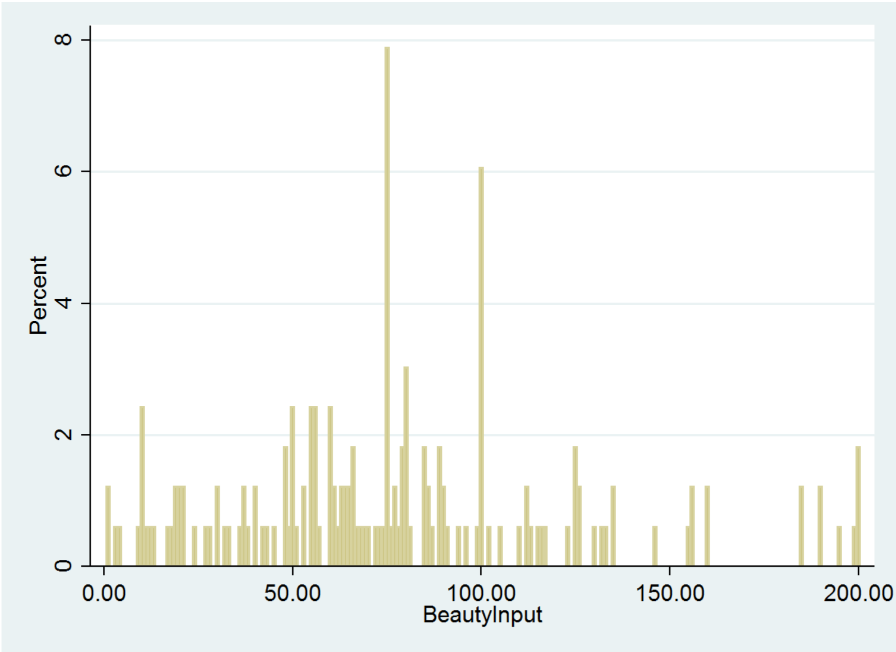


Figure 3.6: Beauty Contest Choices
(rounded to the next integer)

Appendix 3.C Experimental Instructions

The following instructions have been distributed and read out aloud to all subjects. In the original layout used in the experiment all instruction sections were printed on separate sheets of paper. The instructions here are from the treatment with hiding cost of 1.00 € and a show-up fee of 4.00 €. The instructions for the treatment with hiding cost of 2.50 € and a show-up fee of 6.00 € were changed accordingly.

General instructions for part 2 and 3 were not distributed, but only read out aloud and are marked by (*).

The table on page 2 of the Market Game instructions, on the contrary, was not read out aloud and only meant as a reference for subjects.

The subject selected for selection of an envelope was a different subject for each part. They were always the three subjects sitting closest to the front of the laboratory experiment.

General Instructions

Welcome to this experiment. During the experiment:

- please refrain from talking
- please turn off your cell phone
- please do NOT write on any of the sheets you are handed, nor use anything else to take notes.

This experiment consists of 3 parts. Part 1 concerns a game you will play repeatedly for 15 rounds. Information about part 2 will be given after part 1 is finished. Information about part 3 will be given after part 2 is finished.

For now, it is important to know that in all parts, exactly 1 round will be randomly selected to be paid for real at the end of the experiment. For this purpose, the experimenter will ask one of you to select and sign an envelope from a pile of sealed envelopes. As part 1 consists of 15 rounds the envelopes for part 1 contain cards with the numbers 1-15 on them. (You will be able to check this at the end of the experiment.)

At the end of part 1, the signed envelope will be opened, the numbered card will be shown to you, and your earnings in the task that corresponds to the number on the card will be added to your earnings. Suppose for example that the number on the card is 9. Then you will be paid your earnings of round 9.

At the beginning of part 2 and part 3 another envelope will be selected. At the end of these parts, the respective envelope will be opened, the numbered card will be shown to you, and your earnings in the round that corresponds to the number on the card will be added to your earnings.

Further, you receive a show-up fee of 4 €. Your total earnings in the experiment are determined as follows:

total earnings =
earnings in one round in part 1 +
earnings in one round in part 2 +
earnings in one round in part 3 +
show-up fee of 4 €.

Information about your earnings will be private and your earnings will be transferred to your bank account.

<The experimenter will ask one participant to pick and sign an envelope and put the envelope in a corner.>
<The experimenter will now hand out the instructions for part 1 and read them aloud.>

Good luck!

Instructions Part 1 (page 1)

Part 1 consists of 15 identical rounds. You will be in the role of a consumer buying a hypothetical good from a seller who is played by the computer. Note that you will play this game with the computer alone. This means that your outcomes are only affected by your choices, the computer's choices and randomness, but not by choices made by other participants. Also, the computer's decisions can therefore differ between you and other participants.

Each round, you will be randomly assigned one of 20 valuations for the good, which range from 0.00 € to 9.50 € with increments of 0.50 € (i.e. 0.00 €, 0.50 €, 1.00 €, ..., 8.50 €, 9.00 €, 9.50 €). It is equally likely that you will be assigned any of the 20 valuations and there is a new random draw in every round.

The computer will act as a seller and set a price for the good which maximizes his revenues from interacting with you. If the price the seller sets for you is equal to or lower than your valuation, you receive the difference between your valuation and the price as payoff (i.e. you "buy" the product as it is cheaper than or equal to your valuation). If the price is higher than your valuation, you receive no payoff (i.e. you "do not buy" the product as it is more expensive than your valuation). Thus, your payoff is determined as follows:

payoff = valuation - price ***if your valuation is larger than or equal to the price***

payoff = 0 ***if your valuation is lower than the price***

Suppose your valuation is 7.00 € and the price is 5.00 €, then your payoff is 2.00 €.
Suppose your valuation is 3.00 € and the price is 5.00 €, then your payoff is 0.00 € (not -2.00 €).

Your decision in every round will be to either hide or not hide your valuation from the seller. If you do not hide it, the seller will know your exact valuation before setting his price. If you hide it, the seller will not know your exact valuation before setting his price. This implies that the seller has to set one uniform price for all your hidden valuations, while he can adjust his price for all non-hidden valuations. Note, that even if you hide your valuation, the seller always knows that valuations range from 0.00 € to 9.50 € with increments of 0.50 € and that every valuation is equally likely to be assigned to you.

Hiding your valuation from the seller costs you 1.00 €, which will be subtracted from your surplus. Not hiding your valuation costs you nothing. Thus, your earnings in each round are determined as follows:

round earnings = payoff ***if your valuation is not hidden***

round earnings = payoff − 1.00 € ***if your valuation is hidden***

This means, that even though your ***payoff*** cannot be negative, your ***round earnings*** can be negative (but this is limited to a loss of 1.00 €). If a round in which that happens is selected to be paid (out of the 15 identical rounds), this will be deducted from your total earnings (i.e. your earnings from part 2, part 3, or your show-up fee).

Instructions Part 1 (page 2)

In each round, you will have to make a decision for all possible valuations before a valuation will be assigned to you. You will do so by entering a hiding threshold. For all valuations equal to or higher than this hiding threshold, your valuation will be hidden from the seller. For all valuations lower than this hiding threshold, your valuation will not be hidden from the seller. If you do not want to hide any valuation at all, you must enter a hiding threshold of 10.00 €.

Suppose you enter a hiding threshold of 0.00 €: You hide all valuations.
Suppose you enter a hiding threshold of 6.00 €: You hide the valuations from 6.00 € to 9.50 €.
Suppose you enter a hiding threshold of 10.00 €: You hide no valuation.

After you made your decision, one of the valuations will be randomly assigned to you and the respective choice will be implemented: If you are assigned a valuation for which you decided not to hide (i.e. a valuation lower than your cutoff), the seller will know your exact valuation before setting his price. If you are assigned a valuation for which you decided to hide (i.e. a valuation equal to or higher than your cutoff), the seller will only know that your valuation is equal to or higher than your hiding threshold as well as what your hiding threshold is. The table below illustrates this further and includes the special cases when you choose to hide all valuations (hiding threshold of 0.00 €) or no valuation (hiding threshold of 10.00 €).

If your hiding threshold is	**and if you are assigned**	**then the seller knows …**
0.00 €	… any valuation …	… that your valuation is 0.00 € or higher.
3.00 €	… a valuation from 0.00 € to 2.50 € …	… what your exact valuation is.
3.00 €	… a valuation from 3.00 € to 9.50 € …	… that your valuation is 3.00 € or higher.
7.00 €	… a valuation from 0.00 € to 6.50 € …	… what your exact valuation is.
7.00 €	… a valuation from 7.00 € to 9.50 € …	… that your valuation is 7.00 € or higher.
10.00 €	… any valuation …	… what your exact valuation is.

At the end of each round you will be reminded of the hiding threshold you entered, informed about the valuation randomly assigned to you and whether you decided to hide that valuation, the price the seller has set for you, your resulting payoff, whether you paid the hiding cost, and your resulting earnings in this round. Additionally, you will always be informed about the uniform price the seller has set for all your hidden valuations, even when you did not hide the valuation randomly assigned to you. The result screen after each round will therefore always show:

- Your hiding threshold and the range of hidden valuations
- Your assigned valuation and your hiding decision for this valuation
- The seller's price for all your hidden valuations
- The seller's price for your assigned valuation
- Your payoff from the interaction with the seller
- Your hiding cost
- Your round earnings

Note again, that for this part you will play 15 identical rounds with the computer alone. This means that your outcomes are only affected by your choices, the computer's choices and randomness, but not by choices made by other participants. Also, the computer's decisions can therefore differ between you and other participants.

Please raise your hand if you need further explanation from the experimenter.
If you have no questions, the experimenter will soon start the program for part 1.

Instructions Part 2 - General (*)

Part 2 of this experiment consists of 5 identical rounds. As part 2 consists of 5 rounds the envelopes for part 2 contain cards with the numbers 1, 2, 3, 4, and 5 on them. (You will be able to check this at the end of the experiment.) At the end of part 2, this envelope will be opened, the numbered card will be shown to you, and your earnings in the round that corresponds to the number on the card will be added to your earnings.

<The experimenter will ask one participant to pick and sign an envelope and put the envelope in a corner.>
<The experimenter will now hand out the instructions for part 2 and read them aloud.>

Instructions Part 2

Part 2 consists of 5 identical rounds. You will again be in the role of a consumer buying a hypothetical good from a seller who is played by the computer. Again, your outcomes are only affected by your choices, the computer's choices and randomness, but not by choices made by other participants.

You will again be randomly assigned one of 20 valuations for the good, which still range from 0.00 € to 9.50 € with increments of 0.50 €. It is again equally likely that you will be assigned any of the 20 valuations and there is a new random draw in every round.

The computer will still act as a seller and set a price for the good which maximizes his revenues from interacting with you and your payoff is again determined as follows:

payoff* = *valuation - price ***if your valuation is larger than or equal to the price***
payoff* = *0 ***if your valuation is lower than the price***

Again, your decision in every round will be to either hide or not hide your valuation from the seller in the same manner as before, i.e. you will enter a hiding threshold and for all valuations equal to or higher than this threshold your valuation will be hidden from the seller.

The only difference to part 1 is: Hiding your valuation from the seller costs you nothing in this part.

Not hiding your valuation also still costs you nothing. Thus, your earnings in each round are now determined as follows:

round earnings* = *payoff ***if your valuation is not hidden***
round earnings* = *payoff ***if your valuation is hidden***

The result screen after each round will therefore always show:

- Your hiding threshold and the range of hidden valuations
- Your assigned valuation and your hiding decision for this valuation
- The seller's price for all your hidden valuations
- The seller's price for your assigned valuation
- Your payoff from the interaction with the seller
- Your hiding cost
- Your round earnings

Note again, that for this part you will play 15 identical rounds with the computer alone. This means that your outcomes are only affected by your choices, the computer's choices and randomness, but not by choices made by other participants. Also, the computer's decisions can therefore differ between you and other participants.

Please raise your hand if you need further explanation from the experimenter.
If you have no questions, the experimenter will soon start the program for part 2.

Instructions Part 3 - General (*)

Part 3 of this experiment consists of 3 tasks, each played for one round only. As part 3 therefore consists of 3 rounds the envelopes for part 3 contain cards with the numbers 1, 2, and 3 on them. (You will be able to check this at the end of the experiment.)

For part 3, again, an envelope will be selected. At the end of part 3, this envelope will be opened, the numbered card will be shown to you, and your earnings in the round that corresponds to the number on the card will be added to your earnings.

<The experimenter will ask one participant to pick and sign an envelope and put the envelope in a corner.>
<The experimenter will now hand out the instructions for part 3 and read them aloud.>

Instructions Part 3 - Task 1

In this task you are matched with the computer. This means that your outcomes are only affected by your choices, the computer's choices and randomness, but not by choices made by other participants. Also, the computer's decisions can therefore differ between you and other participants.

You will take turns in alternate order. You begin the game, i.e. the order is you, computer, you, computer, . . . until the game ends.

Before your first turn the "current total" is 0.

The player whose turn it is can pick 1, 2, 3, 4, 5, or 6 to be added to the "current total".

The player who makes the "current total" reach 50 (or more) wins the game.

If you win, you earn 10 €. If the computer wins, you earn nothing.

The decisions by the computer are going to be determined at random and hence are not necessarily the best possible response to your choices.

Please raise your hand if you need further explanation from the experimenter.
If you have no questions, the experimenter will soon start the program for this task.

Instructions Part 3 - Task 2

In this task you are matched with a randomly selected other participant in this session.

In this game each participant requests an amount of money.

The amount must be an integer between 1 and 10 € (i.e. 1, 2, 3, ..., 9, 10).

Each participant will receive the amount (s)he requests as earnings.

A participant will receive an additional amount of 10 € if (s)he asks for exactly 1 € less than the participant (s)he is matched with.

Please raise your hand if you need further explanation from the experimenter.
If you have no questions, the experimenter will soon start the program for this task.

Instructions Part 3 - Task 3

In this task you are matched with all other participants in this session.

All of you have to choose a number between 1.00 and 200.00 with increments of 0.01 (i.e. 1.00, 1.01, 1.02, . . . , 199.98, 199.99, 200.00).

After all participants have made their choice, the average of all chosen numbers is calculated by adding up all chosen numbers and dividing it by the number of participants.

The person whose chosen number is closest to ¾ times the average of all chosen numbers wins this game and earns 20 €. In case of a tie, the prize is divided equally among the winners. The other players whose chosen numbers are further away receive nothing.

Be reminded that you are not aiming to be close to the average, but to ¾ of the average.

Please raise your hand if you need further explanation from the experimenter.
If you have no questions, the experimenter will soon start the program for this task.

Chapter 4

Climate Policy Commitment Devices

This chapter is based on the identically entitled working paper co-authored with Reyer Gerlagh, Gijs van de Kuilen, and Stefan Trautmann

Abstract

We develop a dynamic resource extraction game that mimics the global multi-generation planning problem for climate change and fossil fuel extraction. We implement the game under different conditions in the laboratory. Compared to a 'libertarian' baseline condition, we find that policy interventions that provide a costly commitment device or reduce climate threshold uncertainty reduce resource extraction. We also study two conditions to assess the underlying social preferences and the viability of ecological dictatorship. Our results suggest that climate change policies that focus on investments that lock the economy into carbon-free energy sources provide an important commitment device in the intertemporal cooperation problem.

4.1. Introduction

Reducing fossil fuel use is a major component of climate policy. Yet the economic mechanisms for exhaustible resources amplify the coordination difficulties for the public good. Effective climate policy requires coordination both between countries, and over time between generations, because the exhaustible resource characteristics of fossil fuels tend to annul unilateral action (cf. Karp 2015 and references therein). A decrease of fuel demand induces increasing demand by others, both by other countries as emphasized in the carbon leakage literature (Michielsen 2014), and over time as reported in the green paradox literature (Sinn 2008, Gerlagh 2011, van der Ploeg 2016). Policy makers who are unaware of these challenges tend to develop too optimistic plans and early climate targets are subsequently relaxed. Such lack of commitment is especially problematic if climate change is uncertain, and the threshold for a 'catastrophe' is unknown (Barrett and Dannenberg 2012, Gerlagh and Michielsen 2015, Dannenberg et al. 2015). There is an abundant literature on the *international* coordination problem, mostly presenting a pessimistic free-rider perspective (Barrett 1994, 2013), though recent papers present a more constructive contracting approach (Harstad 2015b).[53] In contrast, we focus on the *intertemporal* coordination problem. Recently, some more optimistic analyses suggest democratic rules (Hauser et al. 2014) and investments in renewable energy as a commitment device as solutions of the dynamic problem (Holtsmark and Midttomme 2015, Harstad 2015a).[54] Our contribution is threefold. We develop a dynamic threshold public good game where players choose their actions sequentially, focusing on intergenerational trade-offs rather than international negotiations.[55] Second, in the context of the intertemporal resource extraction dilemma, we show that a commitment device that reduces future resource demand can help to implement resource conservation. This holds even though the commitment device is costly, meaning that its use is inefficient – a waste of welfare – from a first-best planner's perspective. Yet, in the context of a strategic interaction between generations, it helps to improve the outcome compared to a context

[53]The free-rider incentive in Barrett (1994) derives from the concept of a group of countries who set up an agreement, while the outsiders to the agreement play a Nash strategy. The free entry-exit condition for the agreement then constrains the maximum effort inside the agreement. Harstad (2015b) invokes a contract mechanism that involves all players.

[54]In Harstad (2015a) players have a positive incentive to invest in clean technologies as these serve as commitment over time that reduce emissions. In Harstad (2015b) countries reduce investment in clean technologies as these enable other countries to exploit the commitment that comes with clean energy. This paper is more closely related to the first mechanism; the insights complement those of the second paper.

[55]Compare with Calzolari et al. (2016). In their dynamic climate change model, players represent countries that are active over all rounds. In our model players represent generations who, when time passes, enter and leave the game.

where this commitment device is not available. The third major contribution is that we connect subjects' behavior *across* conditions (within-rule choice) with their votes *for* a game condition (rule choice). The findings suggest that successful cooperation not only needs to overcome a gap between individual incentives and public interests, but also a fundamental heterogeneity between subjects with respect to beliefs and preferences about the way in which this should be achieved.

We develop a simple 3-player 3-period sequential resource extraction game in the spirit of Erev and Rapoport (1990) and Budescu et al. (1992). The game mimics the essential characteristics of the climate change, fossil fuel resource-planning problem through four key features. First, players in the game can exploit or conserve a resource, but conservation by one player does not prevent exhaustion by others. Second, resource conservation is a public good. Each generation values its own consumption, but also derives utility from contributing to a stable climate. We model the public good feature through a payoff for all players that depends on both their own resource extraction and the end-of-game resource conservation as in Schmidt et al. (2011), Neubersch et al. (2014) and Gerlagh and Michielsen (2015). Third, the public good is uncertain, so that the benefits from resource conservation are not precisely known (akin to the threshold for climate catastrophe being uncertain). This important difference with Hauser et al. (2014), who have a perfectly known sustainability threshold will be a policy variable (akin to Tavoni 2014) in the current study, because the certainty of the climate threshold has been found to have profound effects on cooperation (Barrett and Dannenberg 2012, Dannenberg et al. 2015). Fourth, the game is played sequentially, so that strategies are asymmetric. Players in earlier positions must base their decisions on expectations regarding future strategies by other players and the consequential conservation outcomes. Players in later positions can condition their actions on outcomes of choices by others.

We study the game's outcomes in an incentivized laboratory experiment under five different conditions. A benchmark condition called *Libertarian* reflects a democratic business-as-usual scenario in which submitted choices are simply implemented. Two conditions refer to potential policy interventions, two others investigate ethical aspects of the public good dilemma in resource conservation. In addition, our experiment includes incentivized voting for the most preferred among the five conditions.[56]

The first policy condition, *Certainty*, eliminates uncertainty about the catastrophe threshold, for example through an increased funding of research to improve climate change predictions. This condition resembles the setting studied in more detail by Barrett and Dannenberg (2012). Importantly, in our design we impose a conservative as-

[56]The experimental setting is especially useful to analyze behavior in the two ethical conditions which could hardly be observed otherwise due to their hypothetical "thought experiment"-like nature.

sessment, with any non-zero degree of exploitation of the resource leading to climate catastrophe resulting, i.e., the Certainty condition provides a physically strictly worse environment.[57] However, in equilibrium, the threat of climate catastrophe leads to full resource conservation. The second policy condition, *Solar*, introduces a costly commitment device resembling the development of renewable energy, based on a recently developing literature emphasizing the commitment problem as one of the keyelements of effective climate policy (Gerlagh and Michielsen 2015, Harstad 2015a, Holtsmark and Midttomme 2015). In the Solar condition, the first generation is presented with the choice of using a costly commitment device. The commitment is intertemporal; we do not study intratemporal commitment mechanisms, e.g. public choices enforced through voting (Hauser et al. 2014).

The first ethical condition, *Dictator*, introduces commitment through dictatorship of the first generation. The first generation has the possibility to become a benevolent ecological dictator installing full resource conservation to the benefit of future generations. However, it is also possible that the first generation exploits the resource for own benefits, while restricting only future generations. This trade-off reflects the discussion on the desirability of ecological dictatorship (Stehr 2015). The second ethical condition, *Rawls*, introduces a ‘veil of ignorance:’ subjects propose an entire plan for all three generations in the game, not yet knowing which generational position they will hold. It aims to move the perspective of the decision maker away from individual to group benefits, i.e., considering the outcome of all generations jointly. These two conditions provide empirical insight into the trade-off between own benefits and group benefits, which is inherent to the intergenerational resource extraction dilemma. Note that meaningful labels are used for convenience here, but were not part of the experiment.

The experiment has three stages. In Stage 1, subjects are matched in groups of 3 players and play each condition exactly *once* without receiving feedback on outcomes. This allows for within-person comparison of behavior across the different institutions. In Stage 2, subjects vote in their groups for their preferred condition under a simple majority rule with tie breaking, and play the selected game *once* without receiving feedback on outcomes. Here we aim to understand the perceived legitimacy of the different institutions and how it affects endogenous institutional choice (Sutter et al. 2010, Barrett and Dannenberg 2017). For example, even if dictatorship yields better social outcomes

[57]New research on climate dynamics may result in both, more optimistic or more pessimistic, views. To keep the number of treatments limited, we study only the more pessimistic research outcome. This choice is theoretically appealing, because in equilibrium a more pessimistic view, counterintuitively, increases the expected payoffs. Note additionally, that a certainty treatment with a more optimistic outcome (a stable climate whenever the resource is not completely exhausted) leads to the same equilibrium play as the Dictator treatment, albeit with higher payoffs for everyone. The climate stability condition would then serve as Player 1’s commitment device.

than the libertarian condition, it may not be acceptable as an institution. In Stage 3, players are randomly regrouped and each group plays one of the conditions *repeatedly* for six rounds *with* feedback on outcomes, allowing players to accumulate experience. We thus observe whether a condition directly implements a certain level of resource conservation, or whether learning and experience of realized outcomes are crucial for players to understand the underlying mechanisms.

4.2. The Resource Extraction Game

The resource extraction game models the behavior of three players in a sequential resource extraction setting. We assume that the resource stock S_t develops according to the dynamic equation

$$S_{t+1} = S_t - R_t, \tag{4.1}$$

where $R_t \in \{0, 1\}$ denotes exploitation by the player living in period t, and an initial resource stock of $S_1 = 2$. S_4 denotes the final stock, which determines whether a stable climate can be attained. Each player receives direct benefits from resource exploitation, but also from resource conservation through climate protection.[58] Payoffs are given by

$$V_t = 6R_t + 8C, \tag{4.2}$$

where V_t denotes the payoffs to the player living in period t, and $C \in \{0, 1\}$ is an indicator for a stable climate.

The payoff structure presents a subtle difference in preferences between generations in the climate change context. Each generation gives a higher weight to its own consumption vis-à-vis the consumption of next generations; such myopic preferences are typical for descriptive models of intertemporal allocation problems. The first term in Equation (4.2) represents a simplified version of these: it associates zero weight to the benefits of resource extraction by next generations. Note that our empirical data will inform about the importance of any direct (and unspecified) altruism towards the other parties in the context of the current dilemma. Such attitudes would be relevant if the current generation forgoes benefits of extraction in favor of the next generation at its own expense after accounting for long-term benefits. The payoff function specifies that generations are not

[58] At first glance, it may seem impossible that welfare of the present generations depends on the future state of the climate, as the present cannot observe the state of climate after its passing. The interpretation is that present utility depends on the expected state of conservation. For risk-neutral players, letting the payoff depend on the actual state of conservation at the end of the game is consistent with a payoff that is dependent on the expected state at the end of the game.

fully myopic. They are concerned about far-future outcomes of their decisions. In climate economy models, the positive weight for long-term outcomes can be modelled through quasi-hyperbolic time discounting (Gerlagh and Liski 2016), or through a positive weight for the long-term future through an additional payoff (Chichilnisky 1996, Gerlagh and Michielsen 2015). The second term in Equation (4.2) captures such a dislike of each generation to add risk to the climate system, representing decision making under scientific uncertainty. The altruism towards far-future generations is explicitly modelled in the payoffs through the stable climate indicator. That is, we do not measure home-grown altruism regarding far off generations, but aim to study intertemporal cooperation *conditional* on the presumed empirical relevance of such altruism, using induced preferences (Smith 1976).

The payoff structure constructs a paternalistic view in which the first generation prefers a stable climate, but also likes to reap the gains from fossil fuel use and to shift costs of achieving a climate target on to the second and third generation. The second generation also appreciates a stable climate, but as well the own gains from fossil fuel use. This setup constructs an intertemporal dilemma. Each generation would like to commit the next generations to abandon fossil fuel use, but without commitment device, the accumulation of short-term exploitation gains results in long-term climate damages.

4.2.1. The Benchmark Game *(Libertarian)*

In the benchmark, or *Libertarian*, condition, there are no restrictions on the players' exploitation choices. If both resource units of the initial stock are conserved, a stable climate is ensured. If only one of the two resource units is conserved, the probability of a stable climate is 50%. If the resource is fully exploited, a stable climate is impossible:

$$S_4 = 2 \Rightarrow \quad C = 1, \tag{4.3}$$

$$S_4 = 1 \Rightarrow P(C = 1) = \frac{1}{2}, \tag{4.4}$$

$$S_4 = 0 \Rightarrow \quad C = 0. \tag{4.5}$$

Thus, the expected climate variable is linear in the final resource stock: $\mathbb{E}(C) = \frac{1}{2}S_4$. It follows that for the third player, expected payoffs are maximized by extracting the resource:

$$\mathbb{E}(V_3) = 6R_3 + 4S_4 = 4S_3 + 2R_3 \tag{4.6}$$

Through backward induction, it is clear that for each player it is optimal to extract the resource if given the opportunity. In contrast, the expected group payoff is maximized by full resource conservation:

$$\mathbb{E}(V_1 + V_2 + V_3) = 6R_1 + 6R_2 + 6R_3 + 24\mathbb{E}(C) = 12 - 6S_4 + 12S_4 = 12 + 6S_4 \quad (4.7)$$

The expected group payoff increases in conservation S_4, but resource extraction is individually optimal for players who are only concerned about their own payoffs as given by Equation (4.2). For the earlier players resource conservation is particularly risky as it leaves the opportunity for the subsequent players to extract the resource, not leaving any reserve at the end, so that good deeds might not be paid back by gains of enjoying a long-term stable climate.[59]

4.2.2. Two Policy Conditions (*Certainty* and *Solar*)

We study two policy conditions that aim to overcome coordination failure. First, the *Certainty* condition assumes that scientific research has sufficiently progressed to pinpoint the precise catastrophe threshold. As a conservative assessment, we assume that the threshold is found to be at the lower end. That is, a catastrophe is certain whenever any part of the resource is exhausted:

$$S_4 = 2 \Rightarrow C = 1, \quad (4.8)$$

$$S_4 = 1 \Rightarrow C = 0, \quad (4.9)$$

$$S_4 = 0 \Rightarrow C = 0. \quad (4.10)$$

In this case, exploiting the first unit is more harmful than in the above setting with uncertainty, while there is no additional harm from exploiting the second unit. For the third player, the individually rational conservation decision depends on the inherited resource stock. If two resource units are inherited, conservation leads to payoff of 8 units, while extraction pays 6 units. Thus, conservation is individually rational. If one resource unit is inherited, a stable climate is unattainable and resource extraction is the superior strategy. By backward induction, one can see that full resource conservation is the unique Nash equilibrium, but it requires a supporting belief structure of the first and second player in the conservationist strategy by the subsequent player(s). Empirically, it may be easier to maintain cooperation in the absence of uncertainty (cf. Barrett and Dannenberg, 2012).

[59] The game resembles a centipede game in the sense that everyone benefits if the resource is conserved period after period, while backwards induction leads to equilibrium strategies with the quickest possible resource extraction. Differently from the centipede game, in each period a new player enters; which makes coordination between the players more difficult.

The second policy condition, *Solar*, concerns an investment $I_1 \in \{0, 1\}$ of the first player into a technology (e.g., renewables) that makes future resource exploitation redundant. As in the Libertarian game, however, there exists uncertainty about the threshold for climate catastrophe. We model the treatment such that an up-front cost can be incurred by Player 1, which fixes future exploitation at zero for Player 2 and 3 ($R_2 = R_3 = 0$); Player 1 can still exploit the resource though. The investment is costly for Player 1, reducing payoffs by 1 unit:

$$V_1 = 6R_1 - I_1 + 8C \tag{4.11}$$

In the *Solar* condition Player 1's expected payoffs are maximized by exploiting the resource and simultaneously investing in the renewable, with a probability of a climate catastrophe of 50%. We emphasize that both policy conditions (at least potentially) restrict the opportunity set. In the Certainty condition the opportunity set is restricted as payoffs in the case of exactly one resource unit being extracted are now deterministic. In the Solar condition, using the commitment device restricts actions by the other players and reduces the maximal group payoff due to its cost. The important empirical question we aim to answer in this paper is whether in the presence of a conflict between individual and social rationality, these policies can implement better outcomes despite the restricted opportunity sets.

4.2.3. Two Ethical Conditions (*Dictator* and *Rawls*)

Compared to the benchmark *Libertarian* condition, the policy conditions *Certainty* and *Solar* change the technology of the game described in Equations (4.1)–(4.7). We now define two ethical conditions that do not change technology compared to the benchmark, but do change the mapping from players and their decisions to payoffs. In both conditions, players need to propose a complete plan of action for the entire group. Again, as in the benchmark condition, there is uncertainty about the climate threshold.

The *Dictator* condition requires players to propose a plan for all three positions in the game, knowing that they will be in position 1, if their proposal is implemented. The proposal that is implemented and the positions of the players whose proposal is not implemented are determined randomly. This treatment resembles a 'perfect commitment' equilibrium, with exploitation power for the first player and no exploitation power for the second and third player. The condition relates to the discussion on ecological dictatorship (Stehr 2015); because there is no cooperation problem, the dictator has the possibility to implement a benevolent dictatorship with full resource conservation and maximum social welfare. However, selfish motives may lead the dictator to forgo social welfare maximization. The treatment thus informs on how much of the failure to achieve the socially optimal level of conservation is accounted for by cooperation failure compared to selfish preferences.

The *Rawls* treatment requires players to propose a plan for all three positions in the game, not yet knowing which position they will hold. The proposal that is implemented and the positions in the game are determined randomly. Thus, this treatment implements Rawls' 'veil of ignorance:' the players do not know in advance their role in the game, but need to submit a full plan of action for the whole society of roles. Thus, their expected payoff is maximized by maximizing the group payoff (full conservation). This condition should therefore shift the players' focus towards social welfare maximization.

4.2.4. Predictions

On the basis of the different policy and ethical conditions for the cooperation dilemma we can summarize the theoretical predictions. Table 4.1 shows the expected equilibrium payoffs for each player ex-ante, before player roles are allocated, as well as the overall expected equilibrium payoff. The table also shows the equilibrium final resource conservation. Certainty and Rawls provide the highest expected payoffs in equilibrium, followed by Dictator, Solar and Libertarian. Note that equilibrium payoffs differ in their distributional riskiness (variation of payoffs over generations), and also in terms of their strategic uncertainty, caused by the possibility of non-equilibrium play of later players. Importantly, we observe that conditions Solar and Dictator provide insurance against exploitation by later players.

Condition	**Equilibrium Play**	**Final Stock** S_4	**Expected Payoffs(€)** P1	P2	P3	Avg.
Libertarian	Always exploit if possible	0	6	6	0	4
Certainty	Never exploit	2	8	8	8	8
Solar	Player 1 exploits and invests; Players 2 and 3 cannot exploit	1	9	4	4	5.67
Dictator	Player 1 exploits and forces Players 2 and 3 to conserve	1	10	4	4	6
Rawls	Never exploit	2	8	8	8	8

Notes: The table shows game-theoretic Nash equilibrium predictions of the five conditions, including the final resource conservation S_4 and the expected payoffs in € of the three players.

Table 4.1: Predictions of Expected Payoffs in Equilibrium

Our benchmark prediction is risk-neutral equilibrium play in each condition. Thus, when institutional choice becomes relevant in the voting stage, we predict that Certainty and Rawls will be selected. Behavioral patterns may deviate from the current predictions for various reasons. Riskiness of the different conditions may affect behavior, and so may beliefs about other players. More subtle effects may be due to the degree of control over outcomes. In Libertarian and Certainty all subjects' decisions are part of the outcome. In the other three conditions players may experience that decisions are explicitly imposed on them. In contrast to Dictator, Solar requires the first-generation player to invest some of her own funds to obtain such commitment power. A central question that the explicit voting addresses is how people perceive the value and the legitimacy of the different conditions and how these aspects may affect realized outcomes in each condition. From a practical viewpoint this is especially relevant with respect to the policy interventions.

4.3. Empirical Methods

Our computerized experiment (zTree; Fischbacher 2007) involved 120 student participants from Tilburg University.[60] The games and their payoffs were translated into Euro values by a 1-for-1 mapping of Equation (4.2) (resp. Equation (4.11)). Participants played multiple games, one of which was randomly selected for monetary payments according to participants' actual choices in this game at the end of the experiment (paying one game prevents income and portfolio effects across games).

The experiment consisted of three stages. In Stage 1, groups of three players played each of the five games exactly once (i.e., no repetition). Participants did not know the identity of their group members, and no feedback on choices or payoffs was given between the games. Importantly, each participant made decisions for all three positions in the game. For positions 2 and 3 these decisions were conditional on the potential stocks at the respective position. Thus, when making decisions, subjects did not know the exact amount of resources that were taken from the common pool as in Budescu et al. (1995). In condition Rawls this elicitation of full strategies was necessary to implement payoffs (because one person's decisions determine the full vector of choices). In the other conditions, the procedure allows us to observe strategies also for events that rarely obtain in sequential play (e.g., Brandts and Charness 2011), and maintains comparability of

[60]While the external validity of laboratory experiments conducted on student populations can (and is) debated, there are two factors in favor of using a student sample. First, being of younger age than the general population, students are more exposed to challenges posed by a changing climate. Second, the acceptability of different institutions by those potentially subjected to them is crucial for their successful implementation (at least in democratic societies).

structure to the Rawls condition. For each game, after all strategies had been submitted, the position of the three players was randomly determined and choices were implemented according to the rules of the specific condition (but no feedback was given until the end of the experiment). To control for order effects when making choices in the 5 games in Stage 1, each group was assigned one of 40 pre-selected orders (out of the 120 possible orders) of the 5 games. We pre-selected these orders such that each game was played equally often in each position (i.e. 8 times as the first, second, . . . , last game) while also being played equally often earlier or later than any other game in pair-wise comparison (i.e. Libertarian appeared earlier than Rawls in 20 of the orders and later than Rawls in the other 20, etc.).

Stage 2 measures participants' preferences over the different institutions. Players were asked to vote for the game that they would prefer to repeat once more. The different conditions were listed to ensure that subjects knew what they vote for. We used a simple majority rule to determine for each group of 3 players which game was played again. In case of a tie, each of the three treatments that received a vote had a chance of 1/3 to be selected. While subjects were informed about which game was repeated once more, they were not informed about votes by other subjects. That is, subjects did not know whether the vote was consensual, a simple majority, or tied. The instructions of the selected game were repeated once more, subjects made their choices and were again not informed about choices by others or the outcome of the game.

Before Stage 3, participants were re-matched in new groups of three players which had not interacted with each other before (stranger matching). Then, each of the groups was randomly assigned to one of the five conditions, such that each condition was played by the same number of groups over the course of the experiment. These groups then played the selected condition 6 times repeatedly with feedback after each round. That is, at the end of each round, players were informed about their assigned position, the remaining resource units at their position and the implemented decision, as well as the final resource conservation and the resulting payoffs. Individual actions by others were not identifiable by the participants. This allowed the participants to learn about the behavior of their group members (which was impossible in Stages 1 and 2), and to adjust their behavior accordingly, without providing the opportunity to react to actions by specific other individuals.

After Stage 3 of the experiment, participants filled in a questionnaire eliciting individual characteristics such as attitudes towards risk (using an incentivized elicitation task), political orientation, views on climate change, numeracy, gender, age, study, and year of study. [61]

[61]More details on the experimental implementation, as well as printed and on-screen instructions can

4.4. Results

We first discuss results at the group level. In the following subsections we then analyze individual strategies in Stages 1 and 3 and discuss Stage 2 voting behavior. When discussing Stage 3 results, we always report results from the sixth iteration of the game in Stage 3, that is, for behavior of experienced players.

4.4.1. Outcomes at Group Level

Figure 4.1 shows resource conservation in the five conditions. For each condition we show the average conservation level in Stage 1 (one-shot interaction), Stage 2 (self-selected conditions after voting), and Stage 3 (last iteration). In the calculation of conservation outcomes the figure accounts for within-group interaction by averaging game outcomes over all possible permutations of allocating subjects to generational positions (and hypothetically implementing their respective decisions). Results for the Libertarian condition demonstrate the essence of the intertemporal resource extraction dilemma, with low levels of resource conservation in all three stages. This allows for sensible assessments of the effects of the different policy interventions and the ethical conditions.

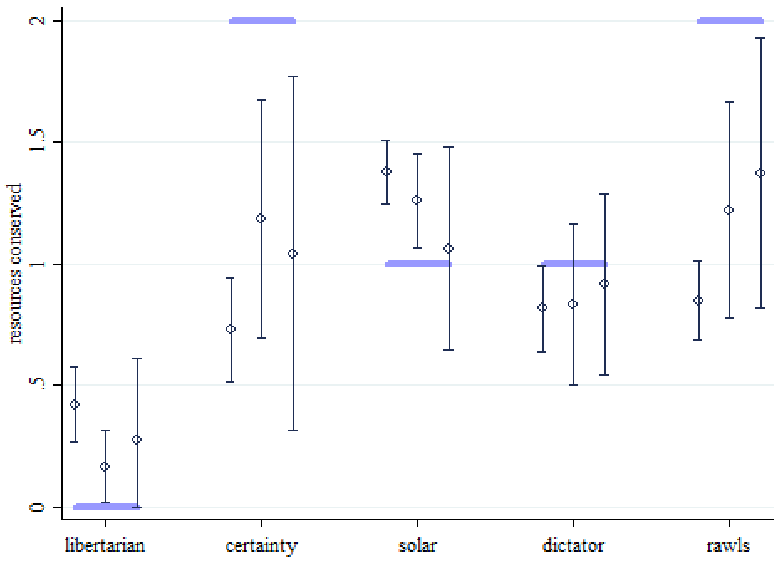


Figure 4.1: Resource Conservation Averages Dependending on Conditions and Stages

Notes: The figure shows the mean and 95% confidence intervals for the resources conserved over all games in each condition. Within each condition, the first line presents the results for Stage 1 (one-shot), the second line shows results for Stage 2 (after voting), and the third line shows results for Stage 3 (last iteration). Thick horizontal bars across all stages show the theoretical equilibrium prediction for the respective condition.

be found in Appendix 4.B and 4.C.

Compared to the Libertarian condition, the other conditions increase conservation. However, in contrast to the predictions in Table 4.1, Certainty and Rawls do not outperform Dictator and Solar in the initial round (Stage 1); only after sufficient experience does Rawls lead to higher conservation. Clearly, even in the last round of Stage 3 the average conservation over groups is substantially lower than the predicted full conservation. Interestingly, while Certainty and Rawls lead to predominantly full or zero conservation, Solar and Dictator show more prevalence of exactly one resource unit conserved (see Figure 4.2 in Appendix 4.A). This finding is consistent with equilibrium predictions. The dynamics of behavior differ across conditions. The effect of Solar seems to be immediate and becomes more moderate with repetition; the effect of Rawls requires some experience to emerge fully.

Table 4.2 provides statistical analyses of group outcomes. It shows conservation rates in Stages 1 (columns (1)-(3)) and 3 (columns (4)-(6)), and the social welfare effects of the different conditions (columns 7 and 8), as a percentage of the maximum potential outcome. The table provides two perspectives on conservation outcomes. The first perspective measures the differences across conditions in terms of subjects' conservation strategies (columns (1) and (4)). More precisely, the variable $\bar{S}_4^O$ is defined as the resource stock left at the end of the game averaged over all players in the fictitious case when for each player the own strategy would be implemented for all 3 player roles (i.e. as if they had played against themselves). As this variable does not involve any group effects,[62] we consider it a measure for the individual resource conservation strategy. The second perspective measures the differences across conditions in terms of groups' expected conservation success (columns (2) and (5)). The variable $\mathbb{E}(S_4)$ is defined as the expected resource stock left at the end of the game on a group level as in Figure 4.1, averaging over all possible random selections of subjects in each role. Note that, for the Dictator and Rawls condition both perspectives, $\bar{S}_4^O$ and $\mathbb{E}(S_4)$, yield the same results because these conditions always implement the strategy of one player by design. For the Libertarian, Certainty, and Solar condition the gap between the two variables reveals the cost associated with incoherent strategies between the group members. That is, although a significant number of players may aim to conserve resources, groups in these conditions may still perform poorly due to a few subjects who exhaust the resource whenever given the opportunity.

Based on Table 4.2, we can make a few observations regarding the average performance of the different conditions. Both policy interventions (Certainty and Solar) perform better in terms of resource conservation than Libertarian, irrespective of whether we consider the individual strategies or the average group outcomes. This is true for

[62] Apart from potential learning effects in Stage 3 that indirectly affect individual behavior.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variable	$\bar{S}_4^O$	$\mathbb{E}(S_4)$	(1)-(2)	$\bar{S}_4^O$	$\mathbb{E}(S_4)$	(4)-(5)	$\mathbb{E}(V)$	$\mathbb{E}(V)$
Player Interaction	*No*	*Yes*		*No*	*Yes*		*Yes*	*Yes*
Stage	1	1	1	3	3	3	1	3
Libertarian	0.41	0.21	0.20***	0.17	0.14	0.03	0.21	0.14
Certainty	0.51**	0.36***	0.15***	0.63**	0.52##	0.10	0.24	0.48
Solar	0.75***	0.69***	0.06***	0.54**	0.53**	0.01	0.57***	0.41**
Dictator	0.41	0.41***		0.46**	0.46**		0.41***	0.46**
Rawls	0.43	0.43***		0.69**	0.69***		0.43***	0.69***

Notes: Resource conservation and expected social welfare expressed in percentage of potential maximum, on the scale from 0 to 1. Outcomes/payoffs are expected values over all subjects and positions. Columns (1) and (4) present (fictitious) outcomes if players played against themselves. Columns (2) and (5) present expected group outcomes and columns (3) and (6) present the effect of these within-group interactions between players. Columns (7) and (8) present expected group payoffs. *, **, *** indicate significance at the 10%, 5%, 1% level. For columns (1), (2), and (7), significance is determined by comparison with the Libertarian treatment, using Wilcoxon signed rank matched pairs test. For columns (3) and (6), significance is based on comparison with zero, using Wilcoxon signed rank matched pairs test. For columns (4), (5), and (8), significance is based on comparison with the libertarian treatment, using the Mann-Whitney two-sample tests. ## indicates that the certainty treatment has a higher probability of full resource conservation than Libertarian at the 5% significance level, though the Mann-Whitney test does not provide significance for the full resource conservation vector.

Table 4.2: Resource Conservation and Expected Social Welfare

unexperienced behavior in Stage 1 (columns (1) and (2)), as well as for experienced behavior in Stage 3 (columns (4) and (5)). For experienced interactions there is little difference between the two policy conditions. However, in the absence of experience and at the group level (arguably the most relevant conditions from a practical perspective), the Solar condition outperforms Certainty (69% vs. 36%, $p < 0.01$, Wilcoxon test). That is, although full resource conservation is the unique Nash Equilibrium in the Certainty condition, empirically subjects seem to hold pessimistic beliefs about others' actions and therefore often fail to coordinate.

We also find evidence for the "exhaustible resource curse," (the effect of conservation choices being substitutes over time) in Stage 1 (column (3)): we find that individual strategies are significantly less exploitative than group outcomes, i.e., the most exploitative players dominate the game outcomes. This is not the case in the last round of Stage 3 (column (6)). That is, subjects' behavior in a group converges over time.

Interestingly, this adaptation effect points into different directions in the two policy conditions. In Certainty, individual and group outcomes increase insignificantly (from 51% to 63%, n.s., Wilcoxon matched pairs test; from 36% to 52%, n.s., Mann-Whitney U test). In contrast, behavior in Solar exhibits a downward trend in conservation with experience (from 75% to 54%, $p < 0.01$ Wilcoxon matched pairs; from 69% to 53%, n.s., Mann-Whitney U test). In Libertarian, individual strategies become significantly more exploitative (from 41% to 17%, $p < 0.01$, Wilcoxon matched pairs test), while group outcomes do not change significantly with experience (from 14% to 21%, n.s, Mann-Whitney U test), presumably because they were low to start with.

The two ethics conditions (Dictator and Rawls) show a somewhat different pattern. These conditions perform better than Libertarian in Stage 1 only on the group outcome, but not the individual strategy level. However, with sufficient experience in Stage 3 both conditions clearly outperform the Libertarian condition in terms of conservation. This is driven by two effects. On the one hand, conservation in the Dictator and Rawls conditions shows an upward trend with experience (from 41% to 46%, n.s.; from 43% to 69%, $p < 0.05$, Wilcoxon matched pairs test). It seems that some experience is necessary to understand the mechanics of these social allocation mechanisms. On the other hand, conservation in Libertarian decreases with experience, thus widening the gap. The Dictator condition shows that even if coordination failure can be overcome, selfish preferences of some subjects still stand in the way of more substantial conservation outcomes.

Next, we take a look at the social welfare implications of the different conditions (columns (7) and (8)), starting with Stage 1 behavior (without experience). Both ethics conditions (Dictator and Rawls) outperform Libertarian in terms of welfare. Remarkably, of the two policy conditions only Solar improves upon Libertarian with respect to welfare, despite the additional costs involved (but in line with theoretical equilibrium predictions). While Certainty improves conservation compared to Libertarian (column (2)), the less favorable mapping from conserved resources to payoffs renders them indistinguishable in terms of welfare. For the behavior of experienced subjects, the same pattern emerges.[63]

Finally, we can put the performance of the policy conditions (Certainty and Solar) and the potentially selfish ethics condition (Dictator) in perspective to the Rawls condition, where social welfare maximization should be easiest to attain. We observe that, with sufficient experience, subjects indeed attained the highest level of welfare behind a Rawlsian 'veil of ignorance'. Therefore, we conclude that selfish motives and coordination failure constrain resource conservation in the other conditions. However, noting the fact that only 69% of the theoretical maximum is achieved for Rawls shows that either policy condition (Solar or Certainty) can bridge about half of the gap between the low Libertarian outcome and the highest observed level of attained welfare.

[63] Although Certainty induces a level of welfare comparable to that of Solar, its effect is not statistically significant due to higher variability (cf. Figure 4.1).

4.4.2. Individual Strategies

We analyze individual behavior with respect to two questions. First, do subjects condition their behavior on other subjects' decisions? Second, do they behave differently depending on the position in the game (conditional on identical resource endowment)?

Table 4.3 shows individual resource extraction strategies, for each position in the game and dependent on resources conserved by previous players. Note that for conditions Dictator and Rawls there were no such conditional strategies as the full strategy of one selected player was implemented for the group. We find a tendency for conditionality of individual choices in Libertarian, Certainty, and Solar. Entries in columns (3) and (5) are always larger than the corresponding entries in columns (2) and (4), although not all comparisons are significant. That is, if the resource has been exploited by at least one person, people are more likely to respond by also exploiting the resource. Note that in the Certainty condition, this behavior should follow directly from the fact that conservation provides no benefit once the first unit has been exploited. Comparisons of columns (2) and (4) to column (1) show that people are typically less likely to exploit the resource the later in the game they are called to make a decision, conditional on full resource conservation $S_t = 2$ at the moment of their decision.[64] This suggests that pessimistic beliefs about choices by "future" subjects negatively affect subjects' conservation choices. Beliefs thus seem to play an important role in the breakdown of cooperation.

4.4.3. Voting and Voting Effects

Stage 2 of the experiment offers insights into participants' preferences for the different conditions. Table 4.4 shows that preferences vary widely (cf. Figure 4.3 in Appendix 4.A for a graphical representation). Solar is the modal vote, receiving significantly more votes than the other conditions (37%, binomial proportion test, $p < 0.01$). Vote shares for Libertarian, Rawls, and Certainty are close to 20% each, whereas Dictator receives significantly fewer votes than the other conditions (10%, binomial proportion test, $p < 0.01$). Note that all players had experienced each of the five conditions exactly once in Stage 1 in randomized order and without any feedback on the behavior of others. Thus, differences in preference over the different conditions can neither be driven by random experiences due to behavior of the other players in the group, nor by order effects. Clearly, people have obtained only a modest degree of intuition about the potential payoffs of the different conditions from the one-shot decision made in each condition.

[64] The effect is less pronounced in the Solar condition. Recall, though, that choices of Player 2 and 3 may not become relevant if Player 1 forces non-extraction by others. Conditionality in choice may thus be diluted by the potential irrelevance of the choice (unknown at the moment the choice is submitted), as well as strategic considerations regarding why Player 1 did not force non-extraction (in which case when the choice is relevant).

	(1)	(2)	(3)	(4)	(5)
Variable	$\mathbb{E}(R_1)$	$\mathbb{E}(R_2)$	$\mathbb{E}(R_2)$	$\mathbb{E}(R_3)$	$\mathbb{E}(R_3)$
Conservation level		$S_1 = 2$	$S_1 = 1$	$S_2 = 2$	$S_2 = 1$
Stage 1 (one-shot)					
Libertarian	0.63	0.54	0.63	0.38^{***}	0.64^{***}
Certainty	0.49	0.35^{**}	0.73^{***}	0.15^{***}	0.72^{***}
Solar	0.41	0.51	0.59	0.35	0.58^{***}
Stage 2 (voting)					
Libertarian	0.67	0.67	0.78	0.39^{*}	0.78^{**}
Certainty	0.33	0.17^{**}	0.71^{**}	0.00^{**}	0.79^{***}
Solar	0.57	0.52	0.70	0.37^{**}	0.70^{**}
Stage 3 (experienced, i.e., in last repetition)					
Libertarian	0.88	0.58^{**}	0.83	0.38^{**}	0.88^{**}
Certainty	0.38	0.38	0.79^{**}	0.17^{*}	0.75^{**}
Solar	0.67	0.63	0.88^{**}	0.54	0.79

Notes: Entries are expected resource extraction averaged over participants in a condition. Stage 1 comparisons are based on one-sample tests of proportion on the individual level. Stage 2 and 3 comparisons are based on Wilcoxon matched-pair tests on the group level. In columns (2) and (4), *, **, *** indicate significance at the 10%, 5%, 1% level, compared to column (1); in columns (3) and (5), *, **, *** indicate significance at the 10%, 5%, 1% level, compared to columns (2) and (4), respectively.

Table 4.3: Individual Exploitation Strategies R_t

However, the middle panel of Table 4.3 above shows that there are pronounced differences in behavior between the conditions after voting, i.e., conditional on playing the game that has been elected in a group. Thus, although not all players in a group may play the game they have in fact voted for, this suggests that voting (and playing) was not random and that players took the special features of each condition into account once they reached Stage 2. We therefore analyze the potential predictors of voting behavior in terms of participants' Stage 1 behavior in more detail and show the results in Table 4.4.

	(1)	(2)	(3)	(4)	(5)
Voted for	Libertarian	Certainty	Solar	Dictator	Rawls
Observations / %	22 / 18%	23 / 19%	44/37%***	12/10%***	19/16%
Stage 1 behavior in...	*Resource Conservation* $\bar{S}_4^0$ *(percentage out of 2)*				
Libertarian	0.39	0.41	0.45	0.21	0.45
Certainty	0.45	0.70**	0.48	0.33	0.55
Solar	0.75	0.85*	0.77	0.67	0.66*
Dictator	0.48	0.39	0.35	0.29	0.55*
Rawls	0.36	0.59*	0.34	0.33	0.55
Stage 1 Average	0.49	0.59**	0.48	0.37**	0.55
% Solar Investment	0.68	0.52**	0.91***	0.75	0.63

Notes: Entries are expected resource extraction averaged over participants in a condition. Stage 1 comparisons are based on one-sample tests of proportion on the individual level. Stage 2 and 3 comparisons are based on Wilcoxon matched-pair tests on the group level. In columns (2) and (4), *, **, *** indicate significance at the 10%, 5%, 1% level, compared to column (1); in columns (3) and (5), *, **, *** indicate significance at the 10%, 5%, 1% level, compared to columns (2) and (4), respectively.

Table 4.4: Stage 2 Voting and Stage 1 Individual Strategies Conditional on Voting

Table 4.4 shows resource conservation shares out of the maximum conservation of 2 units, on the basis of individual behavior only ($\bar{S}_4^O$), i.e., as if subjects had played against their own strategy. Columns show data on subjects who voted for the respective condition, and rows present conservation results for these participants in each of the five conditions in Stage 1. For each group of voters the table also shows average resource conservation in Stage 1 over all treatments, as well as the choice of the investment option in the Solar condition. We first observe that subjects do not vote for the condition that performed best conditional on their own behavior, except for those voting for the Solar condition. In fact, all voter groups conserve most in the Solar treatment. But votes are significantly correlated with Stage 1 resource conservation choices. Those who vote for Solar almost all have made use of the commitment device in the Solar condition in Stage 1 (Table 4.4, last row). Certainty voters tend to conserve more resources than other participants for three specific conditions and overall, but invested the least in the commitment device available in the Solar condition. Potentially these players believe that the Certainty condition offers a cheaper commitment for coordination and therefore value the costly Solar commitment device less. At the other extreme, we find that Dictator voters conserve the least compared to other players. Rawls voters have a higher conservation in the Dictator condition than other voters, suggesting other-regarding preferences play a larger role for these players. The poor performance and voting outcome of Rawls confirms the above-discussed result that significant experience is necessary to understand how the Rawls mechanism allows groups to align individual and group preferences. Investment in the Solar commitment device differs vastly across voters. Clearly, the modest voting success of Solar despite its good performance can be explained by low take up of the investment opportunity for non-Solar voters. We note that there were no significant correlations of voting behavior with individual differences in the subject's risk aversion, gender, or concerns for global warming.

4.5. Discussion

We study the intertemporal resource extraction problem inherent to climate change mitigation in an experimental setting. We find that with uncertainty about the threshold for catastrophe, and in the absence of commitment devices, subjects do not succeed in cooperating to prevent climate catastrophe. There is clear evidence of an 'exhaustible resource curse': conservation choices are substitutes over time. Our game specification is simple and offers only discrete choices to subjects; yet we believe that this feature of our game is robust. It has been observed in other games with semi-continuous choices (Barrett and Dannenberg 2012), and studies repeatedly report that fossil fuels will continue their dominance in absence of drastic global policies (Covert et al. 2016).

We introduce two different mechanisms to mitigate the coordination problem. First, a reduction of uncertainty significantly improves resource conservation, despite being strictly worse in terms of the choice environment. This effect confirms earlier findings in the horizontal cooperation problem (Barrett and Dannenberg 2012). Second, a costly commitment device significantly improves resource conservation and social welfare, despite being a wasteful investment. An important insight is that subjects are willing to pay upfront costs to reduce resource exploitation in later rounds. Moreover, we find that this mechanism receives most support from the subjects, even if they do not know ex-ante whether they will be in charge of the commitment device or potentially be constrained by someone else being in charge of the commitment device (i.e., their position in the game). This suggests that assessments for investments in technology and infrastructure for renewable energy should also include the perspective of their benefits for intertemporal coordination. While present decision makers cannot commit to future carbon prices, they can invest in clean energy and commit to the availability of a competing non-carbon energy supply. The result does not suggest that standard economic reasoning is invalid: a global agreement that (explicitly or implicitly) sets sufficiently high carbon prices remains an efficient instrument available to reduce greenhouse gas emissions. While waiting for such an agreement, costly investments in clean energy could be an essential step. Importantly, this investment seems to be perceived as the most legitimate instrument by those who are exposed to it as the voting stage in our experiment shows.

Even if cooperation failure can be overcome, low weight on other peoples' welfare is still a constraint on socially optimal resource conservation as the Dictator condition of our experiment shows. Moreover, as shown in the analysis of conditionality, pessimistic beliefs about other people making exploitative choices at a later stage prevent a higher degree of conservation. Combining the insights from the Dictator and the Solar condition, we note that strategic instruments aiming at distorting future decisions also carry a danger (Goeschl et al. 2013).

Lastly, the voting stage of our experiment provides a lens to reassess the difficulty in reaching global climate change cooperation, because it shows that strategies in the game are related to votes for conditions. This suggest a within-subject consistency between strategies, beliefs, and preferences for conditions, while at the same time, there is a large between-subject divergence of such beliefs and preferences. Climate change coordination is more difficult than the classic public good view suggests. Our findings suggest that successful cooperation not only needs to overcome a gap between individual incentives and public interests, but also a fundamental heterogeneity between subjects with respect to beliefs and preferences about the way in which this should be achieved.

Appendix 4.A Additional Graphs

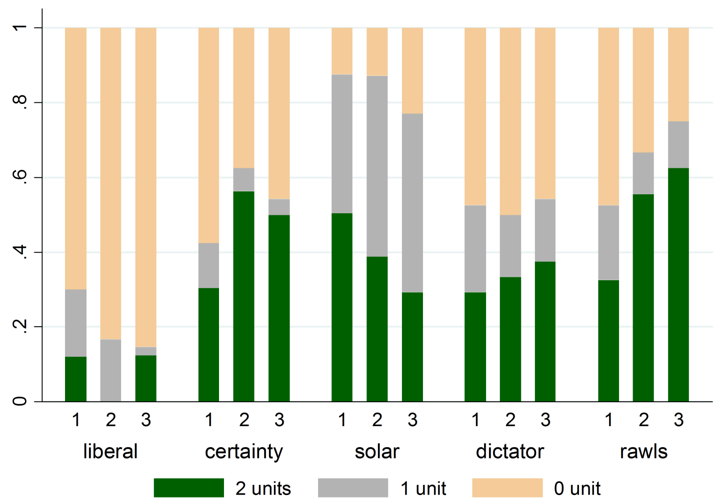


Figure 4.2: Resource Conservation Frequencies Depending on Conditions and Stages

Notes: The figure shows the frequency for conservation outcomes (0 unit, 1 unit, or 2 units) over all games in a condition. Within each condition, the first column presents the results for Stage 1, the second column shows results for Stage 2 (after voting), and the third column shows results for Stage 3 (last iteration).

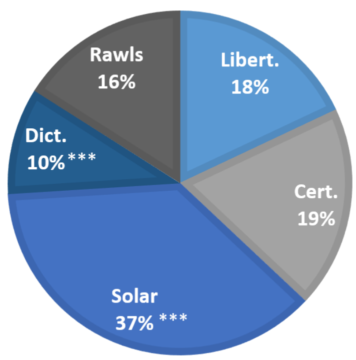


Figure 4.3: Vote Shares (Stage 2)

Notes: The figure shows frequency of votes for all conditions. Asterisks denote significant deviation from 20%, binomial proportion tests.

Appendix 4.B Additional Details on the Experimental Method

General Procedures

The experiment was conducted at Tilburg University with 120 student participants. Students were recruited via an online recruiting system. The experiment was programmed in zTree (Fischbacher 2007). The experiment consisted of 3 stages with a total of 12 tasks: Stage 1 consisting of Tasks 1–5; Stage 2 consisting of Task 6; and Stage 3 consisting of Tasks 7–12. The final Task 13 consisted of a risk attitude assessment and was followed by a questionnaire to collect background information on the participants' demographics (see "Controls").

At the start of the experiment, subjects were randomly allocated to private computers and to groups of 3 participants. At the beginning of each stage, general instructions were handed out on paper and read out aloud by the experimenter. Additionally, task-specific instructions were presented on the computer screen (see Appendix 4.C). The instructions employed a neutral frame and language; meaningful labels were removed to avoid potential confounding effects due to use of language. Any questions were answered in private. Task 1–12 concerned 5 types of games that participants played in groups of 3 (see "The Games").

The Games

In the *Libertarian* game, each group started with a common pool of 12.00 €. Subjects were informed that group members would randomly be assigned a player role (Player 1, Player 2, or Player 3), and that each group member would have the option to take 6.00 € from the common pool sequentially, in order of the player number. Every 6.00 € taken from the common pool were awarded privately. Subjects were informed that if there were 12.00 € (0.00 €) remaining in the pool at the end of the game (i.e., after the action of Player 3 was implemented), each group member would receive 8.00 € (0.00 €) privately. If there were 6.00 € remaining in the common pool, there was a 50% chance that each group member would receive 8.00 € extra, and a 50% chance that each group member would receive nothing extra.

The *Certainty* game was similar to the *Libertarian* game, except for the fact that there was no uncertainty about the amount that subjects would receive at the end of the game if there were 6.00 € remaining in the common pool; each group member would receive 0.00 € in that case.

The *Solar* game was similar to the *Libertarian* game, except for the fact that in the role of Player 1, subjects could decide to remove the possibility of the other group members to take 6.00 € from the common pool, at the cost of 1.00 €.

The *Dictator* game was similar to the *Libertarian* game, except for the fact that all actions of the group member assigned the role of Player 1 were implemented. Thus, in the *Dictator* game, the actions by Player 2 and Player 3 were in fact determined by the group member assigned the role of Player 1.

The *Rawls* game was similar to the *Libertarian* game, except for the fact that all the actions of a randomly chosen group member were implemented. That is, players were randomly assigned the position in the game, and then one person's strategy vector was implemented. For example, in case the actions of Player 2 were chosen to be implemented, the action by Player 1 and Player 3 were determined by the actions of the group member assigned the role of Player 2 in the *Rawls* game.

Specific Procedures

In the first five tasks, subjects simply played each of the five games once without receiving feedback. To measure individual preferences towards each game, subjects were asked to vote for the game that they would prefer to repeat once more with majority voting at the beginning of Task 6. The game that was thus selected was played once more in Task 6. In case of a voting tie, each of the three games that received a vote had an equal 1⁄3 chance of being implemented in Task 6. After Task 6, the groups were reshuffled such that all subjects were in a different group, i.e., each subject was assigned to a new group with 2 participants that were not in their group before. These groups then played one of the five games repeatedly with feedback six times in Task 7–12 within the same group. Thus, at the end of each game, subjects were informed about their player role, the action chosen by them, the actions chosen by the other group members, and their resulting payoff in that task.

Controls

To control for order effects when making choices in the five games, each group was assigned one of 40 pre-selected orders (out of the 120 possible orders) in Task 1–5. We pre-selected these orders such that each game was played equally often in each task while also being played equally often earlier or later than any other game in pair-wise comparison.

Individual risk attitudes were measured in Task 13 that elicits the certain monetary amount that made subjects indifferent between receiving the certain amount and between receiving a lottery yielding either 10.00 € or nothing with equal (50/50) probability, depending on the outcome of a die roll performed at the end of the experiment. In particular, respondents were asked to make a series of 21 choices between the lottery and an ascending range of certain amounts grouped together in a list. The certain amounts in the list ranged from 0.00 € in the first choice to 10.00 € in the final choice, and increased in equal steps of 0.50 €. The midpoint of the last choice in which the subject chose the lottery and the first choice in which the subject chose the certain amount was taken as the certain amount that made the subject indifferent. A certain amount lower/higher than the expected value of the lottery (5.00 €) is indicative of risk averse (seeking) preferences.

Afterwards, subjects were asked to report gender, age, study year, and type of study. In addition, we measured political orientation (left, middle, right), and attitudes towards global warming. Finally, we obtained an individual measurement of numeracy using 5 items from the numeracy scale employed by Peters et al. (2006; items 1, 2, 3, 7, and 10).

Payment

To avoid potential income effects (such as Thaler and Johnson's (1990) house money effect), 1 of the 13 tasks was randomly selected to be paid for real. For this purpose, at the start of the experiment, one participant was asked to assist the experimenter in drawing one envelope from a pile of sealed envelopes, each containing a card numbered 1-13. Participants were informed that the envelope would be opened at the end of the experiment and that the task corresponding with the number on the card inside the envelope would determine the earnings of the experiment. All non-selected envelopes were opened at the end to show that indeed all 13 tasks could have been selected for payment. Additionally to the task-contingent earnings, all subjects received a fixed show-up fee of 4.00 €. On average, subjects earned 9.32 €, while the experiment took about 1 hour and 15 minutes to complete.

Appendix 4.C Experimental Instructions

Printed Instructions

The following instructions have been distributed and read out aloud to all subjects. In the original layout used in the experiment all instruction sections were printed on separate sheets of paper.

General Instructions

Welcome to this experiment. During the experiment:

- please no talking
- please turn off your cell phone
- and raise your hand if anything is unclear, to be helped in private

This experiment consists of 13 tasks in total. Some tasks will involve a game that you will play with other participants; other tasks will involve choices between lotteries. For now, it is important to know that 1 of the 13 tasks will be randomly selected to be paid for real at the end of the experiment. For this purpose, the experimenter will now ask one of you to select an envelope from a pile of sealed envelopes containing cards with numbers 1-13 on them. In particular, the experimenter will ask one of you to draw an envelope from the pile of envelopes at random and sign it, so that you know that the envelope that will be opened at the end of the experiment indeed was the envelope selected by one of you.

<The experimenter will ask one participant to draw an envelope.>

Thus, at the end of the experiment, the randomly selected envelope will be opened, the numbered card will be shown to you, and your earnings in the task that corresponds with the number on the card will be paid for real. Suppose for example that the number on the card is 9. Then, you will be paid your earnings in the 9th task. On top of these earnings, you will receive a show-up fee of €4. Thus, your total earnings in the experiment are determined as follows:

<Total earnings = Earnings of 1 of the 13 tasks (randomly selected) + show-up fee of €4>

All earnings will be paid to you in private. Your earnings in this experiment will be transferred to your bank account. The experimenter will now hand out the instructions for the first 5 tasks and read these instructions aloud. When everybody has completed the first 5 tasks, additional instructions will be handed out. Good luck!

Instructions Task 1–5

The first 5 tasks concern a game that you will play with two other participants. For this purpose, the computer will randomly match you with 2 other participants of this experiment for the duration of the first 5 tasks. You will not learn the identity of your group members; neither will your group members learn your identity. Each group has 3 members.

The 5 games that you and your group members will play all involve 3 player roles: Player 1, Player 2, and Player 3. The computer will randomly assign a role to you and to your fellow group members after you made your decisions. Thus, you do not know yet what your role will be when you are asked to make a decision. Hence, each group member is asked to make a decision for the three possible cases; that (s)he is selected as Player 1, Player 2, or Player 3. Notice that it is equally likely that you will be assigned the role of Player 1, Player 2 or Player 3 (i.e. the chance for each role is equal to 1⁄3).

Each game is played as follows: The group starts with a common pool of €12. In each role you can take out exactly €6 from the common pool, as long as there is money in the pool. Hence, each member decides whether or not to take €6 from the common pool for the three cases of being selected as Player 1, Player 2, or Player 3. The choices will be implemented sequentially, that is, Player 1 decides first whether to take €6 from the common pool, followed by Player 2, and finally by Player 3. Therefore, the decisions made by Player 2 and Player 3 are conditional on the amount of euros remaining in the pool after the previous players have made their decisions. For example, in the role of Player 2, you will be asked separately whether you want to take €6 from the pool if there are €6 remaining, and whether you want to take €6 if there are €12 remaining in the common pool. Which of the two cases holds depends on the choice of Player 1.

In each game, each player who takes €6 from the common pool receives these €6 privately. After all 3 players made their decision, the computer checks how much euros are left in the common pool at the end: 0, 6 or 12. Then, each player receives an amount of euros on top of the private earnings depending on how many euros are left in the common pool as follows:

- If there are €0 left in the common pool, each player receives nothing.
- If there are €12 left in the common pool, each player receives €8.
- If there are €6 left in the common pool, then what happens depends on the game; game specific details will be given on your decision screen.

Please raise your hand if you need further explanation from the experimenter. If you have no questions, the experimenter will soon start the program for the first 5 tasks.

Instructions Task 6

In this task, you and your fellow group members will decide which of the first 5 tasks will be repeated once more. Thus, task 6 will be a repetition of either task 1, task 2, task 3, task 4 or task 5. To select the task that is going to be repeated once more, each group member will be asked to cast a vote on the task that (s)he prefers. The task that has the majority of the votes will be the one that will be repeated once more. If all group members vote for a different task – i.e., if each task receives 1 vote – the task will be selected at random from those that have received a vote. For example, if group member 1 votes to repeat task 2, group member 2 votes to repeat task 5 and group member 3 votes to repeat task 1, the computer will select either task 2, task 5 or task 1 with equal (1⁄3) chance.

In each task, you were asked to indicate whether you wanted to take €6 from the pool in case you are selected as Player 1, Player 2, and Player 3, conditional on how many euros were remaining in the common pool. Each player who took €6 from the common pool received these €6 privately. If after all three players made their decisions there were €0 remaining in the common pool, then there would be no payment to the players additional to their payments based on the private decisions. If there were €12 remaining in the common pool, then each player received an additional €8.

On your screen, you will find a summary of the task-specific instructions, so you can make a well-informed vote. Please raise your hand if you need further explanation from the experimenter. If you have no questions, the experimenter will soon start the program.

Instructions Task 7–12

The next 6 tasks again concern a game that you will play with two other participants. For this purpose, the computer will again randomly match you with 2 other participants of this experiment for the duration of the 6 tasks. Importantly, you will be matched with 2 **other** participants; your group members will not be same ones you were matched with in the first 6 tasks of today's experiment. Again, you will not learn the identity of your group members; neither will your group members learn your identity, and each group has 3 members.

Again, the 6 games that you and your group members will play all involve 3 player roles: Player 1, Player 2, and Player 3. The computer will randomly assign a role to you and to your fellow group members after you made your decisions. Thus, you do not know yet what your role will be when you are asked to make a decision. Hence, each group member is asked to make a decision for the three possible cases; that (s)he is selected as Player 1, Player 2, or Player 3. Notice that it is equally likely that you will be assigned the role of Player 1, Player 2 or Player 3 (i.e. the chance for each role is equal to 1⁄3).

Each game is again played as follows: The group starts with a common pool of €12. In each role you can take out exactly €6 from the common pool, as long as there is money in the pool. Hence, each member decides whether or not to take €6 from the common pool for the three cases of being selected as Player 1, Player 2, or Player 3. The choices will be implemented sequentially, that is, Player 1 decides first whether to take €6 from the common pool, followed by Player 2, and finally by Player 3. Therefore, the decisions made by Player 2 and Player 3 are conditional on the amount of euros remaining in the pool after the previous players have made their decisions. For example, in the role of Player 2, you will be asked separately whether you want to take €6 from the pool if there are €6 remaining, and whether you want to take €6 if there are €12 remaining in the common pool. Which of the two cases holds depends on the choice of Player 1.

In each game, each player who takes €6 from the common pool receives these €6 privately. After all 3 players made their decision, the computer checks how much euros are left in the common pool: 0, 6 or 12. Then, each player receives an amount of euros on top of the private earnings depending on how many euros are left in the common pool as follows:

- If there are €0 left in the common pool, each player receives nothing.
- If there are €12 left in the common pool, each player receives €8.
- If there are €6 left in the common pool, then what happens depends on the game; game specific details will be given on your decision screen.

In the next 6 tasks, you will play the same game with the same group members and you will directly receive feedback about the outcome of the game after each group member has made their decision. Thus, you know your payoffs in each task.

Please raise your hand if you need further explanation from the experimenter. If you have no questions, the experimenter will soon start the program for the next 6 tasks.

Instructions Task 13

The final task concerns several choices between two options, labelled LEFT and RIGHT, grouped together in a list. Both options will yield an amount of money, potentially depending on the throw of a standard six-sided die performed at the end of the experiment. If this task is selected to be paid for real, you payoff will depend on the option you have chosen and, potentially, on the throw of the six-sided die.

Please raise your hand if you need further explanation from the experimenter. If you have no questions, the experimenter will soon start the program for the final task.

Detailed On-Screen Instructions for each Condition

The following instructions have been shown to all subjects on their computer screens. In the original layout used in the experiment the general instructions and the respective task-specific instructions occupied the left half of the screen, whereas the right half of the screen was occupied by the respective choice screen.

4.C.0.1 Left Half of the Screen

General Instructions

On the right, you are asked to indicate whether you want to take € 6 from the pool in case you are selected as Player 1, Player 2, and Player 3, conditional on how many euros are remaining in the common pool.

Each player who takes € 6 from the common pool receives these € 6 privately.

If after all three players made their decisions there are € 0 remaining in the common pool, then there will be no payment to the players additional to their payments based on the private decisions.

If there are € 12 remaining in the common pool, then each player receives an additional € 8.

Task-Specific Instructions

Libertarian

In this task, if there are € 6 remaining in the common pool, then there is a 50% chance that each player receives nothing extra, and a 50% chance that each player receives € 8 extra.

If this task is selected to be paid, then the game will be played based on the decisions made by you and the other group members, in the respective role that has been assigned to each of you.

Please raise your hand if you need further explanation from the experimenter. If you have no questions, please make your choices for this task on the right side of the screen.

Certainty

In this task, if there are € 6 remaining in the common pool, then each player receives nothing extra.

Only if there are € 12 left in the common pool will all players receive € 8 extra on top of their private payoff.

If this task is selected to be paid, then the game will be played based on the decisions made by you and the other group members, in the role that has been assigned to you.

Please raise your hand if you need further explanation from the experimenter. If you have no questions, please make your choices for this task on the right side of the screen.

Solar

In this task, if there are € 6 remaining in the common pool, then there is a 50% chance that each player receives nothing extra, and a 50% chance that each player receives € 8 extra.

In this task, Player 1 has one extra option. You can choose to force Player 2 and Player 3 to leave the euros in the common pool.

Choosing this option costs € 1, which is subtracted from your pay if you are assigned and the role of Player 1.

Please raise your hand if you need further explanation from the experimenter. If you have no questions, please make your choices for this task on the right side of the screen.

Dictator

In this task, if there are € 6 remaining in the common pool, then there is a 50% chance that each player receives nothing extra, and a 50% chance that each player receives € 8 extra.

If this task is selected to be paid, then the game will be played based on the decisions made by the group member that has been assigned the role of Player 1.

The payments to you depend on the role that has been assigned to you. Player 2 receives the payments for Player 2 based on the decisions made by Player 1. Player 3 receives the payment for Player 3, based on the decisions made by Player 1.

Please raise your hand if you need further explanation from the experimenter. If you have no questions, please make your choices for this task on the right side of the screen.

Rawls
In this task, if there are € 6 remaining in the common pool, then there is a 50% chance that each player receives nothing extra, and a 50% chance that each player receives € 8 extra.

If this task is selected to be paid, then the game will be played based on the decisions made by a randomly chosen group member ('the representative').

The payments to you depend on the role that has been assigned to you. Player 1, 2 and 3 receive the payments for Player 1, Player 2, and Player 3, based on the decisions made by 'the representative'.

It is equally likely that you are selected as 'the representative', independent of your chances to be Player 1, 2, or 3, respectively.

Please raise your hand if you need further explanation from the experimenter. If you have no questions, please make your choices for this task on the right side of the screen.

4.C.0.2 Right Half of the Screen

Solar Choice Screen *(preceding the general choice screen in the Solar treatment)*
As explained, before making a choice in each situation, you can force Player 2 and Player 3 to leave the euros in the pool, if you are selected as Player 1.

Doing so, will cost you € 1, if you are selected as Player 1.

Do you want to force Player 2 and Player 3 to leave the euros in the pool?
BUTTON: YES BUTTON: NO

General Choice Screen *(in all treatments, in Solar this is the second decision screen)*
For each situation described below, please indicate whether you take € 6 from the pool or not.

SITUATION 1: YOU ARE Player 1
There are € 12 remaining in the pool, your decision: take ○ ○ do not take (radio buttons)

SITUATION 2: YOU ARE Player 2
If there are € 12 remaining in the pool, your decision: take ○ ○ do not take (radio buttons)
If there are € 6 remaining in the pool, your decision: take ○ ○ do not take (radio buttons)

SITUATION 3: YOU ARE Player 3
If there are € 12 remaining in the pool, your decision: take ○ ○ do not take (radio buttons)
If there are € 6 remaining in the pool, your decision: take ○ ○ do not take (radio buttons)

Chapter 5

Bibliography

Acquisti, Alessandro, and Jens Grossklags. 2007. "What Can Behavioral Economics Teach Us About Privacy" In *Digital Privacy: Theory, Technologies and Practices*, edited by Alessandro Acquisti, Stefanos Gritzalis, Costas Lambrinoudakis, and Sabrina De Capitani di Vimercati, 363–77. New York & London: Auerbach Publications.

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman. 2016. "The Economics of Privacy." *Journal of Economic Literature* 54 (2): 442–92. doi:10.1257/jel.54.2.442.

Acquisti, Alessandro, and Hal R. Varian. 2005. "Conditioning Prices on Purchase History." *Marketing Science* 24 (3): 367–81. doi:10.1287/mksc.1040.0103.

Arad, Ayala, and Ariel Rubinstein. 2012. "The 11–20 Money Request Game: A Level-K Reasoning Study." *The American Economic Review* 102 (7): 3561—73. doi:10.1257/aer.102.7.3561.

Argenton, Cédric, and Jens Prüfer. 2012. "Search Engine Competition with Network Externalities." *Journal of Competition Law and Economics* 8 (1): 73–105. doi:10.1093/joclec/nhr018

Armstrong, Mark. 2006. "Recent Developments in the Economics of Price Discrimination." In *Advances in Economics and Econometrics: Theory and Applications: Ninth World Congress: Volume II*, edited by R. Blundell, W. K. Newey, and T. Persson, 97–141. Cambridge, UK: Cambridge University Press. http://www.cambridge.org/catalogue/catalogue.asp?isbn=9780521871532.

Barrett, Scott. 1994. Self-enforcing international environmental agreements. *Oxford Economic Papers* 46, 878–894. http://www.jstor.org/stable/2663505

Barrett, Scott. 2013. Climate treaties and approaching catastrophes. *Journal of Environmental Economics and Management*, 66 (2): 235–250. https://doi.org/10.1016/j.jeem.2012.12.004

Barrett, Scott and Astrid Dannenberg. 2012. Climate negotiations under scientific uncertainty. *Proceedings of the National Academy of Sciences* 109 (43): 17372-17376. doi:10.1073/pnas.1208417109

Barrett, Scott and Astrid Dannenberg. 2016. Tipping versus cooperating to supply a public good. *Journal of the European Economic Association* 15 (4): 910–941. https://doi.org/10.1093/jeea/jvw022

Benndorf, Volker, Dorothea Kübler, and Hans-Theo Normann. 2015. "Privacy Concerns, Voluntary Disclosure of Information, and Unraveling: An Experiment." *European Economic Review* 75: 43–59. https://doi.org/10.1016/j.euroecorev.2015.01.005

Brandts, Jordi and Gary Charness. 2011. The strategy versus the direct-response method: a first survey of experimental comparisons. *Experimental Economics* 14 (3): 375–398. https://doi.org/10.1007/s10683-011-9272-x

Budescu, David V., Amnon Rapoport, and Ramzi Suleiman. 1992. Simultaneous vs. Sequential Requests in Resource Dilemmas with Incomplete Information. *Acta Psychologica* 80 (1-3): 297–310. http://dx.doi.org/10.1016/0001-6918(92)90052-F

Budescu, David V., Ramzi Suleiman, and Amnon Rapoport. 1995. Positional order and group size effects in resource dilemmas with uncertain resources. *Organizational Behavior and Human Decision Processes* 61 (3): 225–238. http://dx.doi.org/10.1006/obhd.1995.1018

Calzolari, Giacomo, Casari, Marco, and Riccardo Ghidoni. 2016. Carbon is forever: A climate change experiment on cooperation, Quaderni - *Working Paper DSE N°1065*, Universita di Bologna. http://dx.doi.org/10.2139/ssrn.2775276

Calzolari, Giacomo, and Alessandro Pavan. 2006. "On the Optimality of Privacy in Sequential Contracting." *Journal of Economic Theory* 130 (1): 168–204. doi:10.1016/j.jet.2005.04.007.

Camerer, Colin F., Ho Teck-Hua Ho, and Juin-Kuan Chong. 2004. "A Cognitive Hierarchy Model of Games." *The Quarterly Journal of Economics* 119 (3): 861–898.

Casadesus-Masanell, Ramon, and Andres Hervas-Drane. 2015. "Competing with Privacy." *Management Science* 61 (1): 229–246. doi:10.1287/mnsc.2014.2023.

Chichilnisky, Graciela. 1996. An axiomatic approach to sustainable development. *Social Choice and Welfare* 13 (2): 231–257. https://doi.org/10.1007/BF00183353

Conitzer, Vincent, Curtis R. Taylor, and Liad Wagman. 2012. "Hide and Seek: Costly Consumer Privacy in a Market with Repeat Purchases." *Marketing Science* 31 (2): 277–92. doi:10.1287/mksc.1110.0691.

Costa-Gomes, Miguel A., and Vincent P. Crawford. 2006. "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study." *The American Economic Review* 96 (5): 1737–68. doi:10.1257/aer.96.5.1737.

Costa-Gomes, Miguel A., Vincent P. Crawford, and Bruno Broseta. 2001. "Cognition and Behavior in Normal-Form Games: An Experimental Study." *Econometrica* 69 (5): 1193–1235. doi:10.1111/1468-0262.00239.

Covert Thomas, Michael Greenstone, and Christopher R. Knittel. 2016. Will we ever stop using fossil fuels? *Journal of Economic Perspectives* 30 (1): 117–138. doi:10.1257/jep.30.1.117

Crawford, Vincent P. 2003. "Lying for Strategic Advantage: Rational and Boundedly Rational Misrepresentation of Intentions." *The American Economic Review* 93 (1): 133–49. doi:10.1257/000282803321455197.

Crawford, Vincent P., and Nagore Iriberri. 2007a. "Level-K Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse and Overbidding in Private-Value Auctions?" *Econometrica* 75 (6): 1721–70. doi:10.1111/j.1468-0262.2007.00810.x.

Crawford, Vincent P., and Nagore Iriberri. 2007b. "Fatal Attraction: Salience, Naïveté, and Sophistication in Experimental 'Hide-and-Seek' Games." *The American Economic Review* 97 (5): 1731–50. doi:10.1257/aer.97.5.1731.

Crawford, Vincent P., Miguel A. Costa-Gomes, and Nagore Iriberri. 2013. "Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications." *Journal of Economic Literature* 51 (1): 5–62. doi:10.1257/jel.51.1.5

de Cornière, Alexandre, and Romain de Nijs. 2016. "Online Advertising and Privacy." *The RAND Journal of Economics* 47 (1): 48–72. doi:10.1111/1756-2171.12118.

Dannenberg, Astrid, Andreas Löschel, Gabriele Paolacci, Christiane Reif, and Alessandro Tavoni. 2015. On the provision of public goods with probabilistic and ambiguous thresholds. *Environmental and Resource Economics*, 61 (3): 365-383. https://doi.org/10.1007/s10640-014-9796-6

Dengler, Sebastian, and Jens Prüfer. 2017. "Consumers' Privacy Choices in the Era of Big Data." Working Paper.

Dufwenberg, Martin, Ramya Sundaram, and David J. Butler. 2010. "Epiphany in the Game of 21." *Journal of Economic Behavior & Organization* 75 (2): 132–43. doi:10.1016/j.jebo.2010.03.025.

Einav, Liran, and Jonathan D. Levin. 2013. "The Data Revolution and Economic Analysis." Working Paper 19035. National Bureau of Economic Research. http://www.nber.org/papers/w19035.

Erev, Ido, and Amnon Rapoport. 1990. Provision of step-level public goods: the sequential contribution mechanism. *The Journal of Conflict Resolution* 34 (3): 401–425. https://www.jstor.org/stable/174222

Erev, Ido, and Alin E. Roth. 1998. "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria." *The American Economic Review* 88 (4): 848–881. https://www.jstor.org/stable/117009

Fischbacher, Urs. 2007. "z-tree: Zurich toolbox for ready-made economic experiments" *Experimental Economics* 10 (2): 171—178. doi.org/10.1007/s10683-006-9159-4

Fudenberg, Drew, J. Miguel Villas-Boas, and T. Hendershott. 2006. "Behavior-Based Price Discrimination and Customer Recognition." *Handbook on Economics and Information Systems* 1: 377–436.

Gerlagh, Reyer. 2011. Too much oil. *CESifo Economic Studies* 57, 79–102. doi.org/10.1093/cesifo/ifq004

Gerlagh, Reyer, and Matti Liski. 2016. Consistent climate policies. *Journal of the European Economic Association*, forthcoming.

Gerlagh, Reyer, and Thomas Michielsen. 2015. Moving targets – cost-effective climate policy under scientific uncertainty. *Climatic Change* 132, 519-529.

Goeschl Timo, Daniel Heyen, and Juan Moreno-Cruz. 2013. The Intergenerational Transfer of Solar Radiation Management Capabilities and Atmospheric Carbon Stocks. *Environmental and Resource Economics* 56 (1): 85–104. doi.org/10.1007/s10640-013-9647-x

Goldfarb, Avi, and Catherine Tucker. 2012. "Shifts in Privacy Concerns." *The American Economic Review* 102 (3): 349–53. doi:10.1257/aer.102.3.349.

Goldfarb, Avi, and Botao Yang. 2009. "Are All Managers Created Equal?" *Journal of Marketing Research* 46 (5): 612–22. doi:10.1509/jmkr.46.5.612.

Harstad, Bard. 2015a. Investment policy for time-inconsistent discounters. *Working Paper.*

Harstad, Bard. 2015b. The dynamics of climate agreements. *Journal of the European Economic Association* 14 (3): 719–752. doi:10.1111/jeea.12138

Hauser Oliver P., David G. Rand, Alexander Peysakhovich, and Martin A. Nowak. 2014. Cooperating with the future. *Nature* 511, 220–223. doi:10.1038/nature13530

Hermalin, Benjamin E., and Michael L. Katz. 2006. "Privacy, Property Rights and Efficiency: The Economics of Privacy as Secrecy." *Quantitative Marketing and Economics* 4 (3): 209–239.

Ho, Teck-Hua, Colin Camerer, and Keith Weigelt. 1998. "Iterated Dominance and Iterated Best Response in Experimental 'P-Beauty Contests."' *The American Economic Review* 88 (4): 947–969. http://www.jstor.org/stable/117013

Holtsmark, Katinka, and Kristoffer Midttømme. 2015. The dynamics of linking permit markets. *Working Paper*.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1990. "Experimental Tests of the Endowment Effect and the Coase Theorem." *Journal of Political Economy* 98 (6): 1325–1348.

Kahneman, Daniel, and Amos Tversky 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, 47 (2): 263–291

Karp, Larry. 2015. Provision of a public good with multiple dynasties. *Economic Journal*, forthcoming.

López, Rafael. 2001. "On p-Beauty Contest Integer Games." Working Paper No. 608, Universitat Pompeu Fabra.

Luo, Xueming, Michelle Andrews, Zheng Fang, and Chee Wei Phang. 2014. "Mobile Targeting." *Management Science* 60 (7): 1738–56. doi:10.1287/mnsc.2013.1836.

Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work and Think*. John Murray.

Michielsen, Thomas. 2014. Brown backstops versus the green paradox. *Journal of Environmental Economics and Management* 68, 87–110

Mikians, Jakub, László Gyarmati, Vijay Erramilli, and Nikolaos Laoutaris. 2012. "Detecting Price and Search Discrimination on the Internet." In *Proceedings of the 11th ACM Workshop on Hot Topics in Networks*, 79–84. Redmond, WA: ACM. doi:10.1145/2390231.2390245.

Nagel, Rosemarie. 1995. "Unraveling in Guessing Games: An Experimental Study." *The American Economic Review* 85 (5): 1313–26.

Neubersch, Delf, Hermann Held, and Alexander Otto. 2014. Operationalizing climate targets under learning: an application of cost-risk analysis. *Climatic Change* 126 (3-4), 315–328. doi:10.1007/s10584-014-1223-z

New York Times. 2012. "Acxiom, the Quiet Giant of Consumer Database Marketing." *The New York Times*, June 16, sec. Technology. http://www.nytimes.com/2012/06/17/technology/acxiom-the-quiet-giant-of-consumer-database-marketing.html.

Nielsen. 2012. "Which Smartphone Apps Do Savvy Shoppers Use Most." *Newswire - Mobile*. The Nielsen Company. http://www.nielsen.com/us/en/insights/news/2012/top-mobile-shopping-apps.html.

Odlyzko, Andrew. 2003. "Privacy, Economics, and Price Discrimination on the Internet." In *Proceedings of the 5th International Conference on Electronic Commerce - ICEC 2003*, edited by Norman M. Sadeh, Mary Jo Dively, Robert J. Kauffman, Yannis Labrou, Onn Shehory, Rahul Telang, and Lorrie Faith Cranor, 355–66. ICEC 2003. Pittsburgh, Pennsylvania, USA: ACM Press. doi:10.1145/948005.948051.

Peters, Ellen, Daniel Västfjäll, Paul Slovic, C. K. Mertz, Ketti Mazzocco, and Stephan Dickert. 2006. Numeracy and Decision Making. *Psychological Science* 17 (5): 407–413. doi: 10.1111/j.1467-9280.2006.01720.x

Pigou, Arthur C. 1920. *The Economics of Welfare*. London: Macmillan and Co., limited.

Posner, Richard A. 1978. "An Economic Theory of Privacy." *Regulation* 2: 19.

Posner, Richard A. 1981. "The Economics of Privacy." *The American Economic Review* 71 (2): 405–9. doi:10.2307/1815754.

Prüfer, Jens, and Christoph Schottmüller. 2017. "Competing with Big Data." CentER Discussion Paper No. 2017-007.

Schmidt, Matthias G.W., Alexander Lorenz, Hermann Held, and Elmar Kriegler. 2011. Climate targets under uncertainty: challenges and remedies. *Climatic Change* 104 (3-4): 783–791. doi:10.1007/s10584-010-9985-4

Selten, Reinhard, Klaus Abbink, and Ricarda Cox. 2005. "Learning Direction Theory and the Winner's Curse." *Experimental Economics* 8 (1): 5–20. doi:10.1007/s10683-005-1407-5

Shiller, Benjamin Reed. 2013. "First Degree Price Discrimination Using Big Data." Working Paper 58. Working Paper Series. Brandeis University, Department of Economics and International Businesss School. http://ideas.repec.org/p/brd/wpaper/58.html.

Shopkick. 2014. "Shopkick Surpasses $1 Billion Generated for Partners." *PR Newswire*. October 15. http://www.prnewswire.com/news-releases/shopkick-surpasses-1-billion-generated-for-partners-279269882.html.

Sinn, Hans-Werner. 2008. Public policies against global warming. *International Tax and Public Finance* 15 (4): 360–394.

Smith, Vernon L.. 1976. Experimental Economics: Induced Value Theory. *The American Economic Review (Papers & Proceedings)* 66 (2): 274–279.

Stahl, Dale O., and Paul W. Wilson. 1994. "Experimental Evidence on Players' Models of Other Players." *Journal of Economic Behavior & Organization* 25 (3): 309–27. doi:10.1016/0167-2681(94)90103-1.

Stahl, Dale O., and Paul W. Wilson. 1995. "On Players' Models of Other Players: Theory and Experimental Evidence." *Games and Economic Behavior* 10 (1): 218–54. doi:10.1006/game.1995.1031.

Stehr, Nico. 2015. Democracy is not an inconvenience. *Nature* 525, 449–450. doi:10.1038/525449a

Stigler, George J. 1980. "An Introduction to Privacy in Economics and Politics." *Journal of Legal Studies* 9 (4): Article 2.

Sutter, Matthias, Haigner, Stefan, and Martin G. Kocher. 2010. Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations. *Review of Economic Studies* 77 (4): 1540–1566. https://doi.org/10.1111/j.1467-937X.2010.00608.x

Tavoni, Alessandro. 2014. Evolutionary escape from the climate dilemma. *Physics of Life Reviews* 11 (4): 587–588. doi:10.1016/j.plrev.2014.06.021

Taylor, Curtis R. 2004. "Consumer Privacy and the Market for Customer Information." *The RAND Journal of Economics* 35 (4): 631–50. doi:10.2307/1593765.

Thaler, Richard H., and Eric J. Johnson. 1990. Gambling with the house money and trying to break even: the effect of prior outcomes on risky choice. *Management Science* 36 (6): 643–660. doi:10.1287/mnsc.36.6.643

Tirole, Jean. 1988. *The Theory of Industrial Organization*. 1st ed. MIT Press.

Tsai, Janice Y., Serge Egelman, Lorrie Cranor, and Alessandro Acquisti. 2011. "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study." *Information Systems Research* 22 (2): 254–68. doi:10.1287/isre.1090.0260.

Tucker, Catherine E. 2014. "Social Networks, Personalized Advertising, and Privacy Controls." *Journal of Marketing Research* 51 (5): 546–62. doi:10.1509/jmr.10.0355.

Van der Ploeg, Frederick. 2016. Second-best carbon taxation in the global economy: the Green Paradox and carbon leakage revisited. *Journal of Environmental Economics and Management*, 78: 85–105.

Varian, Hal R. 1997. "Economic Aspects of Personal Privacy." Privacy and Self-Regulation in the Information Age. U.S. Department of Commerce - National Telecommunications and Information Administration. http://www.ntia.doc.gov/page/chapter-1-theory-markets-and-privacy.

Villas-Boas, J. Miguel. 2004. "Price Cycles in Markets with Customer Recognition" *RAND Journal of Economics* 35 (3): 486–501. http://www.jstor.org/stable/1593704.

Wall Street Journal Blog. 2014. "Amazon Wants to Ship Your Package Before You Buy It." *Digits – Tech News & Analysis From the WSJ.* January 17. http://blogs.wsj.com/digits/2014/01/17/amazon-wants-to-ship-your-package-before-you-buy-it.

Warren, Samuel D., and Louis D. Brandeis. 1890."The Right to Privacy." *Harvard Law Review* 4 (5): 193–203.