

# **Determinants of Other-Regarding Behavior and the Private Provision of Public Goods**

A thesis submitted in fulfillment  
of the requirements for the degree of Dr. rer. pol.  
from the Faculty of Economic Sciences  
University of Göttingen

by

**Mattheus Florian Wilhelm Brenig**

born in Munich

Göttingen, 2021

First Supervisor: Prof. Dr. Kilian Bizer

Second Supervisor: Prof. Dr. Markus Spiwoks

Third Supervisor: Prof. Dr. Holger A. Rau

# Contents

<b>1</b>	<b>Introduction and Summary</b>	<b>1</b>
<b>2</b>	<b>Consumer Behavior under Benevolent Price Discrimination</b> <i>Alexander Erlei, Mattheus Brenig, and Nils Engelbrecht</i>	<b>23</b>
<b>3</b>	<b>Excusing selfish behavior with (willfully induced) uncertainty: An experimental study</b> <i>Mattheus Brenig, Nils Engelbrecht, and Daniel Hermann</i>	<b>75</b>
<b>4</b>	<b>Dishonest online: A distinction between observable and unobservable lying</b> <i>Daniel Hermann, and Mattheus Brenig</i>	<b>117</b>
<b>5</b>	<b>Can Gamification lead to Increased Paid Crowdworkers' Output?</b> <i>Sascha Lichtenberg, Tim-Benjamin Lembcke, Mattheus Brenig, Alfred Benedikt Brendel, and Simon Trang</i>	<b>127</b>
<b>6</b>	<b>Breathing life into consumer rights: Smartphone tools facilitating the “right to know” on substances of very high concern in REACH articles</b> <i>Julian Schenten, Mattheus Brenig, Martin Führ, and Kilian Bizer</i>	<b>145</b>
<b>7</b>	<b>Refining tools to bridge the gap between academia and chemical regulation: perspectives for WikiREACH</b> <i>Marlene Ågerstrand, Mattheus Brenig, Martin Führ, and Julian Schenten</i>	<b>175</b>



# CHAPTER I

## Introduction and Summary

**A**CCORDING to Robbins (1932), “[e]conomics is a science which studies human behaviour as a relationship between ends and scarce means which have alternative uses.” Since the beginning of the 20th century, with the formalization of rational choice theory (see Hicks and Allen 1934; Houthakker 1950; Samuelson 1947; Pareto 1906), economists have predominantly studied human behavior theoretically using an axiomatic approach. Moreover, they have often adopted a highly simplified notion of narrow self-interest, which is now referred to as the standard economic model. However, with the advent of behavioral economics in the late 20th century, the focus has shifted to an evidence-based understanding of behavior, using the standard economic model only as a benchmark (Thaler 2016). Along with limits to rationality and self-control, bounded self-interest is one of the key behavioral regularities found in this literature that depicts a more psychologically realistic picture of a representative agent (Mullainathan and Thaler 2000).

This thesis contributes to research that challenges the narrow self-interest assumption used in the standard economic model. The *first* part of the thesis comprises three experimental studies on other-regarding behavior. The *second* part of the paper includes three studies that fit into the literature on private provision of public goods. In the following, I will briefly introduce both parts of the thesis. The main goal is to highlight the unifying features of the studies and place them within the relevant literature. Within the sections describing the two parts, I will also briefly summarize each study, again with a focus on the overarching themes.

## Part I

Decades of experimental research with public goods, ultimatum, trust, and dictator games have shown that individuals forgo monetary payoff to the benefit of others, even in anonymous one-shot decisions, ruling out possible explanations based on a self-interested strategy (for reviews see e.g. Johnson and Mislin 2011; Chaudhuri 2011; Güth and Kocher 2014; Kagel and Roth 2016). The outcome-based preference models of Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) represent initial attempts to synthesize these findings and integrate them into formal economic theory. While these models have been able to explain most experimental findings using standard games up to this point, their simplicity means that other fairness considerations either cannot be accounted for or at least are not explicitly modeled. However, as noted by Kagel and Roth (2016), the two models have inspired new experiments designed to

capture that other-regarding behavior depends not only on the distribution of payoffs, but also on the process that leads to it. Process-related features include the intention of the involved decision-makers (e.g. Charness and Rabin 2002; Dufwenberg and Kirchsteiger 2004; Falk et al. 2008) and relatedly, the degree to which responsibility can be causally attribute to them (e.g. Coffman 2011; Falk et al. 2020; Hamman et al. 2010), the fairness of procedures (e.g. Bolton et al. 2005), or the adherence to specific social or moral norms such as telling the truth (e.g. Gneezy 2005; Gneezy et al. 2018; Abeler et al. 2019).

All three studies in the first part of this thesis examine behavior in situations where the decision-maker's choice is probably influenced by its consequences on the distribution of outcomes, but where other determinants are also likely to play a role. In the *first* study, we examine consumers' purchase decisions in response to price discounts for low income groups. Regarding the fairness of outcomes, price discrimination may be accepted as long as it redistributes outcomes in a manner that is consistent with what consumers perceive as fair. However, this ignores important procedural features that may also affect consumer acceptance of price discrimination. These include reciprocity concerns regarding the seller (e.g. Fehr and Gächter 2000), a dislike of unequal price treatment (e.g. Xia et al. 2004), and a lack of agency in the price-setting process (e.g. Alexopoulos et al. 2013; Choshen-Hillel and Yaniv 2011). The *second* study examines how different excuses emanating from uncertainty regarding the consequences of one's actions on others influence selfish behavior. Previous research has shown that individuals do not necessarily have preferences for a fair distribution of outcomes, but merely want to appear fair (e.g. Dana et al. 2007). Uncertainty blurs the causal responsibility for harm suffered by others and allows individuals to avoid the assignment of blame and responsibility by themselves or others (e.g. Bartling and Fischbacher 2012; Coffman 2011). The *third* study deals with dishonesty. While refraining from lying can also be explained by distributional preferences (Gneezy 2005), a recent meta-analysis concludes that a "preference for being seen as honest and a preference for being honest are the main motivations for truth-telling." (Abeler et al. 2019) That is, similar to study two, individuals are driven in part by how they are perceived by others. We test whether this preference to appear honest is sustained in an online setting where social identity concerns and reputational concerns are likely to be reduced.

*Consumer behavior under benevolent price discrimination*

In outlining their model of inequity aversion, Fehr and Schmidt (1999) note that they consider the egalitarian outcome to be the natural reference point of social comparison in their model because “subjects enter the laboratory as equals, they do not know anything about each other, and they are allocated to different roles in the experiment at random.” That is, in the specific context of standard laboratory experiments, the distribution of outcomes that subjects consider as “fair” is determined by equality. In a different context, however, this need not be the case. According to equity theory in social psychology (Festinger 1954; Adams 1963), a distribution is perceived as fair if outcomes are proportional to individuals’ input. That is, subjects follow a merit-based conception of fairness. When roles and income are randomly assigned, as in standard economic experiments, inputs do not differ and thus the egalitarian outcome constitutes the reference point. In contrast, when individuals must earn their income, the fair share is proportional to the individual’s effort or performance. Indeed, results from ultimatum (Hoffman and Spitzer 1985; Hoffman et al. 1996), double auction (Ball et al. 2001) and public good games (Clark 1998) show that behavior depends on whether initial endowments are assigned randomly or based on past performance. Merit-based notions of justice have also been used to explain cross-country differences in preferences for redistribution that result from different beliefs about equality of opportunity in a society (Alesina and Angeletos 2005; Krawczyk 2010; Durante et al. 2014).

In this study, we examine whether merit-based distributional preferences can also explain the acceptance of price discounts for certain low-income groups of the population, such as students and the elderly. When income inequality is not rooted in merit, such price discrimination leads to an equalization of outcomes that may be closer to what individuals perceive as fair. Although the application of merit-based equity preferences in the context of price discrimination is theoretically appealing, previous research using this theory to explain higher fairness perceptions of discounts for low-income groups (Wu et al. 2012) or incorporating this link into mathematical models via utility functions (Rotemberg 2011) relies on an assumption that has not yet been empirically tested. Alternative explanations for the prevalence of certain price discounts can therefore not be ruled out. For example, acceptance of lower prices for the elderly might instead be driven by habituation, learned societal norms or compensation for other types of inequality such as quality of life. Moreover, the results



of studies on redistributive preferences do not readily translate to the context of price discrimination in consumer markets. Most importantly, this literature abstracts from many process-related variables to focus on outcome distributions. This excludes many elements of the price-setting mechanism that might well interact with or mediate distributive preferences and their translation into behavior. For example, under price discrimination, redistribution is not actively chosen by consumers, but imposed by sellers. Reciprocity considerations thus become relevant (Fehr and Gächter 2000) and moral responsibility for bringing about a fair outcome is shifted away from consumers (Pizarro et al. 2003; Bandura 2002). At best, consumers can attribute to themselves a moral responsibility for influencing future prices and thus outcomes (Pigors and Rockenbach 2016).

We measure consumers' response to price discounts that benefit financially disadvantaged individuals, leading to a convergence of outcomes. Using consumers' purchase decisions and (costly) switching behavior to or away from a price discriminating store, we can quantify preferences both against and for price discrimination. To investigate whether consumer behavior is consistent with merit-based distributional preferences, we implement experimental variations of intentions, reciprocity, transparency, merited vs. unmerited income inequality, consumer agency through a market mechanism and transaction costs.

Our study provides insights into how sellers can differentiate prices in the market. These include a wider range of potential recipients of discounts, poaching prices, improved assessments of meaningful degrees of strategic price obfuscation and transparency, or the promotion of certain pricing strategies as part of corporate social responsibility communications.

#### *Excusing selfish behavior with (willfully induced) uncertainty: An experimental study*

While early models of fairness explain other-regarding behavior by a preference for equitable or fair outcomes, more recent economic studies propose that individuals do not want to *be* fair, but merely want to *appear* fair (e.g. Dana et al. 2007). In situations where the relationship between their actions and outcomes is clear, individuals are constrained by their desire to appear fair. In contrast, when this relationship is blurred by uncertainty, individuals are able to behave selfishly while maintaining the appearance of fairness and avoiding the assignment of blame and responsibility by themselves or others (e.g. Bartling and Fischbacher 2012; Coffman 2011). In

other words, moral norms act as constraints on selfish preferences rather than being preferences themselves (Konow 2000; Bolton et al. 1998). Uncertainty introduces “moral wiggle room” that allows individuals to relax these constraints (Dana et al. 2007). Uncertainty can arise from e.g. delegation of tasks to agents (e.g. Hamman et al. 2010), responsibility diffusion when outcomes depend on the interaction of multiple individuals (e.g. Falk et al. 2020) or willful ignorance (e.g. Bartling et al. 2014).

This study contributes to the above literature by examining the interplay of different excuses for selfish behavior associated with uncertainty. Our study is closely related to economic experiments in which unanimous group decisions are required to either cause or prevent an externality (e.g. Brütt et al. 2020; Falk et al. 2020; Irlenbusch and Saxler 2019; Behnk et al. 2017). In the first case, a single fair vote prevents the externality with certainty. Therefore, each individual can excuse a selfish vote on the grounds that there is a chance that the externality will be prevented by a fair vote of another group member (*bystander excuse*). In the second case (i.e. unanimity to prevent the externality) individuals cause an externality with certainty, but they can use the excuse “if I am not selfish, someone else probably will be” to justify their behavior (*replacement excuse*; Falk et al. 2020; Ziegler et al. 2021). In both cases, the perception of being pivotal, i.e. by how much the probability of the externality is affected by an individual’s decision, depends on an individual’s belief about the behavior of other group members.

In our study, we replace uncertainty due to the behavior of other decision-makers by moves of nature. This allows us to manipulate uncertainty precisely, while holding constant other dimensions of responsibility diffusion such as shared guilt (Rothenhäusler et al. 2018) or information on social norms. We further contribute to the literature by considering different combinations of excuses and different degrees of being pivotal. We also test excuse-driven behavior by restricting the scope of uncertainty so that subjects can always prevent part of the externality and therefore never justify that the ex-post outcome will be the same independent of their behavior. Finally, we examine whether decisions are sensitive to the possibility to ignore consequences for third parties.

Understanding exactly how excuse-driven behavior works in decisions that involve tradeoffs between self and others has far-reaching policy implications, particularly for markets. Although each individual market transaction may lead to comparatively small harmful effects, the ubiquity of markets, combined with extensive division of labor can lead to a wide diffusion of responsibility and increased opportunities for

willful ignorance, allowing for moral disengagement on both sides of the market and substantial aggregate effects. Our study can provide insight into whether measures that reduce (perceived) uncertainty and resulting excuses could be used as policy instruments to promote ethical purchasing behavior. Furthermore, our study has implications for the success of policies aimed at increasing transparency about the consequences of consumption decisions.

*Dishonest online: A distinction between observable and unobservable lying*

Standard economic models make clear predictions regarding dishonesty when information is private: Individuals will report what maximizes their profits and place no intrinsic value on telling the truth. People's honesty has far-reaching consequences because most economic interactions are characterized by asymmetric information. Examples include tax evasion (Allingham and Sandmo 1972), the used car market (Akerlof 1970), expert advice (Crawford and Sobel 1982), and dishonesty of entire economic sectors such as the Libor manipulation (Abrantes-Metz et al. 2012). However, recent economic experiments also show that individuals do not maximally lie about private information (see e.g. Gneezy 2005; Mazar et al. 2008; Abeler et al. 2019; Fischbacher and Föllmi-Heusi 2013). One possible explanation for this finding is that individuals care about the outcomes of others (either other participants or the experimenter) and therefore choose their extent of lying to achieve their desired distribution of payoffs. However, Gneezy (2005), in his early work on deception, suggests that distributive preferences are probably not the only explanation for the observed behavior. Indeed, Abeler et al. (2019) conclude in their recent meta-analysis that a "preference for being seen as honest and a preference for being honest are the main motivations for truth-telling."

Our study builds on the finding that individuals have a preference for *being seen* as honest, i.e., people are concerned about their social image or reputation. This motive has recently been studied theoretically and experimentally by Abeler et al. (2019), Gneezy et al. (2018), and Khalmetski and Sliwka (2019). All three studies show that lying behavior appears to be motivated in part by concern for social image. Building on this finding, we are interested in whether this result can be replicated in an online setting, where subjects are likely to have less concern about their social image or reputation.

Our experiment is conducted on the crowdsourcing platform “MTurk”. Due to the absence of a physically present experimenter, social identity concerns could be reduced. In addition, due to the flexibility and short-term nature of work arrangements in online labor markets (see Dube et al. 2020), reputational concerns might also be reduced. To study the role of social image on dishonesty, we focus on the effect of the observability of a lie. While subjects can be individually identified as liars for observable outcomes, this is not the case for unobservable outcomes. Instead, lies can only be identified at the group level, based on known prior probabilities of the possible outcomes. If individuals have a “pure” preference for truth-telling, the observability of a lie should not matter. If, however, individuals also have a preference for being seen as honest, we expect more lying on unobservable outcomes.

The study of dishonesty in online settings is important because economic interactions are increasingly characterized by (1) reduced physical interaction and (2) a greater number of potential trading partners due to technological advances. More recently, the ongoing digitization of economic transactions has been catalyzed by the COVID-19 pandemic.

## Part II

One of the first types of experiments used to study the self-interest assumption were public good games. These games create a social dilemma because the marginal utility of an individual contributing to the public good is negative, while the effect on social welfare is positive. Under the narrow self-interest assumption, no one contributes to the public good. The failure of markets to provide an efficient level of public goods is therefore seen as one of the main reasons for governmental intervention. However, decades of research in behavioral economics show that private contributions are generally higher than predicted by narrow self-interest and that a decay in contributions over repeated games can be prevented by means of either costly punishment of free-riders (e.g. Fehr and Gächter 2000; Fehr and Gächter 2002; Gürer et al. 2006) or non-monetary mechanisms such as moral suasion (e.g. Rege and Telle 2004; Masclet et al. 2003; Bochet and Putterman 2009) or assortative matching (e.g. Page et al. 2005; Gächter and Thöni 2005).

In addition to these ideal-type laboratory studies, scholars have also been interested in understanding the underlying preferences for private contributions to public goods in real-world settings. This research has scrutinized a range of motives to contribute

that go beyond narrow self-interest, but has also led to potential explanations that fall within the boundaries of self-interest, such as a short-term positive marginal utility (Lakhani and von Hippel 2003) or long-term signaling incentives. Lerner and Tirole (2002), e.g., posit several motives for open source software development, such as the career concern and ego gratification incentive. The literature since has largely supported their hypotheses (Bitzer and Schröder 2005; Xu et al. 2020; Mustonen 2005). Another prominent example of a digital public good that is privately provided is Wikipedia. Here, scholars have particularly focused on social benefits (Aaltonen and Seiler 2016; Zhang and Zhu 2011) and intrinsic motivation (Xu and Li 2015; Yang and Lai 2010) as an incentive to contribute. It is important to note that in the above examples, there is a positive marginal utility from private contributions, unlike standard public good games. However, the existence of positive externalities still leads to a sub-optimal provision of the public good. A better understanding of the underlying motives can help refine standard models of public goods and enable the development of policies that are more likely to lead to the desired behavioral changes and more cost-effective than alternative policy instruments. The second part of this thesis can be placed in said literature.

Study four is an experimental study of the influence of gamification on effort provision in a real-effort task that mimics important characteristics of micro-tasks typically found in crowd-sourcing projects. Crowdsourcing is increasingly used to generate public knowledge, and the use of gamification can increase intrinsic motivation and thus private contributions to the public good. Studies five and six both explore digital public goods in the context of the EU chemical regulation REACH<sup>1</sup>, evaluate how effectively the public good is provided and discuss possible adjustments to increase the incentives for individuals that freely contribute to the good.

#### *Can gamification lead to increased output from paid crowdworkers?*

Recently, advances in information and communications technology have enabled new forms of knowledge generation by larger, dynamically formed groups of individuals (Estellés-Arolas and González-Ladrón-de-Guevara 2012). This approach, referred to as crowdsourcing, is often used for tasks that are easy for humans but difficult for algorithms to solve, such as classifying text documents and images (Nakatsu et al.

---

<sup>1</sup>REACH concerns the registration, evaluation, authorisation and restriction of chemicals and established the European Chemicals Agency (ECHA). See Regulation (EC) No 1907/2006 of the European Parliament and of the Council, OJ 2006 L 396/1.

2014). Compared to the more complex tasks in software development or in the case of Wikipedia contributions, micro-tasks in crowdsourcing projects are often repetitive and therefore neither inherently enjoyable nor do they provide signaling incentives to contributors. Therefore, providers are increasingly using gamification, “the use of game design elements in non-game contexts”, to motivate contributions (Deterding et al. 2011; Morschheuser et al. 2017; Cooper et al. 2010).

In this study we experimentally study the effect of different gamification elements (progress bars, badges and leaderboards) on effort provision and task enjoyment in a real-effort task using a between-subject design. We contribute to an emerging literature on the influence of gamification in crowdsourcing, with a focus on micro-tasks. To mimic such tasks, we use a variation of the slider task by Gill and Prowse (2012). We pay subjects for a fixed number of completed tasks, but allow more tasks to be completed. We compare the number of completed tasks for the different gamified versions of the tasks with a control treatment without gamification. Moreover, we look at treatment differences in autonomous motivation, which we elicit in the post-experimental questionnaire using a modified version of the measure by Ryan and Connell (1989).

Crowdsourcing is increasingly used in the generation of public knowledge. Examples of citizen involvement in scientific research include protein folding (Cooper et al. 2010), classification of galaxies (Raddick et al. 2013) or animal sounds (Franzoni and Sauermann 2014). Outside of research, crowdsourcing is also used to provide evidence to influence policy decisions or to inform legal processes (Kullenberg and Kasperowski 2016). In this regard, crowdsourcing has the potential to lead to a democratization of regulatory capacity (Grabosky 2013). First examples of technology enabled quality assurance by citizen are e.g. the crowd-sensing of radiation levels near the Fukushima nuclear accident site to supplement governmental data, websites where patients report effects of medication to supplement to clinical trials (Weil et al. 2013) or crowd-sourced reviews of governmental documents for questionable claims (The Guardian 2009). Grabosky (2013) calls this increased citizen involvement in regulatory activity “wiki-regulation”. Gamifying public crowd-sourcing tasks potentially increases crowdworkers’ task motivation, engagement and enjoyment. This, in turn, can lead to a higher willingness to contribute, higher quality of work and long-term engagement.

*Breathing life into consumer rights: smartphone tools facilitating the “right to know” on substances of very high concern in REACH articles*

The generation of information about the toxicological properties of and exposure to chemicals is not an end in itself. Ultimately, its purpose is to inform the risk-management decisions of different societal actors. Public access to information is an important objective of REACH and Art. 33(2) gives consumers a “right to know” about toxic substances in articles. It provides the public with firm and product specific information and, therefore, has the potential to create market-based incentives for firms to reduce toxic substances. In the words of the European Commission, the “right to know” was included so “[c]onsumers can play an active role [...] by taking an interest in the safety of the products they buy”<sup>2</sup>. However, the provision’s operationalisation, most notably article suppliers’ response window of 45 days, does not support consumers to make immediate informed decisions—at the point of sale in particular (see reviews of empirical research on the effectiveness of information disclosure policies by Weil et al. 2006; Dranove and Jin 2010). Moreover, information provided by firms in response to requests by individual consumers is, a priori, not publicly available.

In an attempt to remove the limitations of Art. 33(2) REACH, the German NGO “Friends of the Earth Germany” (BUND) updated its smartphone application “ToxFox” in October 2016, to allow users to send “right to know” requests after having scanned a product’s barcode. Moreover, firms can voluntarily provide information to the database—a process that is called “unraveling” in the disclosure literature (Dranove and Jin 2010; Board 2009). Hence, information is not only crowdsourced through consumers. Information provided by suppliers (whether by email or through the database) is stored and thereafter immediately available. Thereby, the information is made public and waiting time is essentially reduced from the legal time limit of 45 days to less than a second.

We assess the impact of ToxFox in terms of the quantity of Art. 33(2) requests and publicly available information in the database. Moreover, we identify motivational, regulatory, legal and technical barriers to an increased use of ToxFox to send requests. Based on these barriers, we derive recommendations regarding (1) governance (including enforcement and implementation), (2) smartphone applications and (3) future research, all ultimately aimed at increasing user contributions to the public database.

---

<sup>2</sup>[http://ec.europa.eu/environment/chemicals/reach/right\\_en.htm](http://ec.europa.eu/environment/chemicals/reach/right_en.htm)

*Refining tools to bridge the gap between academia and chemical regulation: perspectives for WikiREACH*

To save resources, enforcement authorities often rely on firms to self-report information (Kaplow and Shavell 1994). This is also true for REACH. Indeed, REACH has been lauded for shifting the informational burden from regulators to firms (Führ and Bizer 2007; Silbergeld et al. 2015). However, if firms fear regulation or bad publicity, they have an incentive to withhold damaging information (Coria 2018) or to not report information truthfully (Jacob and Levitt 2003; Andreoni et al. 1998). Evaluation of registration dossiers (ECHA 2016; EC 2018) and information provided on SVHC in articles (Wursthorn 2018; KEMI 2016; SHRG 2013) revealed a low quality of self-reports by firms under REACH. According to the executive director of ECHA this “raises questions on the incentives (or lack of them) for complying with obligations [...]” (Buxton 2018). To counter-balance these incentives, a system of quality assurance has to be put in place. Traditionally, this is done by raising and properly enforcing sanctions (Kaplow and Shavell 1994; Shimshack and Ward 2005).

We propose a complementary quality assurance mechanism, whereby information on the effects of chemicals generated by academic research serves as a quality check for industry studies. However, screening the literature for appropriate studies and extracting data is burdensome. Therefore, the burden on regulators can be reduced by making data easily and readily accessible. We study WikiPharma<sup>3</sup>, a database that aggregates ecotoxicity data for pharmaceuticals extracted from peer-reviewed studies. The use of the WikiPharma database is explored to develop strategies on a similar tool in the context of REACH.

The main challenge of tools like WikiPharma is that they require constant updating as soon as new choice-relevant information becomes available. Therefore the wiki-function was added to crowdsource new data from academic researchers. However, researchers need to be rewarded within the academic system (in terms of career, income and internal recognition) to have incentives to provide information to the envisioned WikiREACH (Stephan 1996). Such incentives may arise if WikiREACH, similar to academic networking websites, increases the discoverability of research (Niyazov et al. 2016), emerges as a alternative measure of scholarly impact (Haustein et al. 2014) or funding agencies reward communication of results outside of academia (Franzoni et al. 2011). To ensure that such a tool is continuously maintained and compatible with the

---

<sup>3</sup><http://www.wikipharma.org/welcome.asp>



regulatory system, and thereby useful for hazard and risk assessments of chemicals, it would benefit from being developed in collaboration with the major stakeholders in the field, i.e. regulatory agencies, academia, industry, scientific journals, and providers of research network platforms.

## References

- Aaltonen, A. and S. Seiler (2016). "Cumulative Growth in User-Generated Content Production: Evidence from Wikipedia". In: *Management Science* 62.7, pp. 2054–2069.
- Abeler, J., D. Nosenzo, and C. Raymond (2019). "Preferences for Truth-Telling". In: *Econometrica* 87.4, pp. 1115–1153.
- Abrantes-Metz, R. M., M. Kraten, A. D. Metz, and G. S. Seow (2012). "Libor manipulation?" In: *Journal of Banking & Finance* 36.1, pp. 136–150.
- Adams, J. S. (1963). "Towards an understanding of inequity." In: *The Journal of Abnormal and Social Psychology* 67.5, pp. 422–436.
- Akerlof, G. A. (1970). "The Market for "Lemons": Quality Uncertainty and the Market Mechanism". In: *The Quarterly Journal of Economics* 84.3, pp. 488–500.
- Alesina, A. and G.-M. Angeletos (2005). "Fairness and Redistribution". In: *American Economic Review* 95.4, pp. 960–980.
- Alexopoulos, J., D. Pfabigan, F. Göschl, H. Bauer, and F. Fischmeister (2013). "Agency matters! Social preferences in the three-person ultimatum game". In: *Frontiers in Human Neuroscience* 7.312, pp. 1–10.
- Allingham, M. G. and A. Sandmo (1972). "Income tax evasion: a theoretical analysis". In: *Journal of Public Economics* 1.3-4, pp. 323–338.
- Andreoni, J., B. Erard, and J. Feinstein (1998). "Tax Compliance". In: *Journal of Economic Literature* 36.2, pp. 818–860.
- Ball, S., C. Eckel, P. J. Grossman, and W. Zame (2001). "Status in Markets". In: *The Quarterly Journal of Economics* 116.1, pp. 161–188.
- Bandura, A. (2002). "Selective Moral Disengagement in the Exercise of Moral Agency". In: *Journal of Moral Education* 31.2, pp. 101–119.
- Bartling, B., F. Engl, and R. A. Weber (2014). "Does willful ignorance deflect punishment? – An experimental study". In: *European Economic Review* 70, pp. 512–524.
- Bartling, B. and U. Fischbacher (2012). "Shifting the Blame: On Delegation and Responsibility". In: *The Review of Economic Studies* 79.1, pp. 67–87.
- Behnk, S., L. Hao, and E. Reuben (2017). *Partners in Crime: Diffusion of Responsibility in Antisocial Behaviors*. IZA Discussion Papers 11031. Institute of Labor Economics (IZA).
- Bitzer, J. and P. J. Schröder (2005). "Bug-fixing and code-writing: The private provision of open source software". In: *Information Economics and Policy* 17.3, pp. 389–406.

- Board, O. (2009). "Competition And Disclosure". In: *The Journal of Industrial Economics* 57, pp. 197–213.
- Bochet, O. and L. Putterman (2009). "Not just babble: Opening the black box of communication in a voluntary contribution experiment". In: *European Economic Review* 53.3, pp. 309–326.
- Bolton, G. E., J. Brandts, and A. Ockenfels (2005). "Fair Procedures: Evidence from Games Involving Lotteries". In: *The Economic Journal* 115.506, pp. 1054–1076.
- Bolton, G. E., E. Katok, and R. Zwick (1998). "Dictator game giving: Rules of fairness versus acts of kindness". In: *International Journal of Game Theory* 27.2, pp. 269–299.
- Bolton, G. E. and A. Ockenfels (2000). "ERC: A Theory of Equity, Reciprocity, and Competition". In: *American Economic Review* 90.1, pp. 166–193.
- Brütt, K., A. Schram, and J. Sonnemans (2020). "Endogenous group formation and responsibility diffusion: An experimental study". In: *Games and Economic Behavior* 121, pp. 1–31.
- Buxton, L. (2018). *Decade of REACH dossier evaluation reveals 70% non-compliance*. Shrewsbury, UK: Chemical Watch. Accessed 21 November 2018. URL: <https://goo.gl/AuXjQ8>.
- Charness, G. and M. Rabin (2002). "Understanding Social Preferences with Simple Tests". In: *The Quarterly Journal of Economics* 117.3, pp. 817–869.
- Chaudhuri, A. (2011). "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature". In: *Experimental Economics* 14.1, pp. 47–83.
- Choshen-Hillel, S. and I. Yaniv (2011). "Agency and the construction of social preference: Between inequality aversion and prosocial behavior." In: *Journal of personality and social psychology* 101.6, pp. 1253–1261.
- Clark, J. (1998). "Fairness in Public Good Provision: An Investigation of Preferences for Equality and Proportionality". In: *The Canadian Journal of Economics / Revue canadienne d'Économique* 31.3, pp. 708–729.
- Coffman, L. C. (2011). "Intermediation Reduces Punishment (and Reward)". In: *American Economic Journal: Microeconomics* 3.4, pp. 77–106.
- Cooper, S., F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, et al. (2010). "Predicting protein structures with a multiplayer online game". In: *Nature* 466, pp. 756–760.
- Coria, J. (2018). "Policy Monitor—The Economics of Toxic Substance Control and the REACH Directive". In: *Review of Environmental Economics and Policy* 12.2, pp. 342–358.
- Crawford, V. P. and J. Sobel (1982). "Strategic Information Transmission". In: *Econometrica* 50.6, pp. 1431–1451.

- Dana, J., R. A. Weber, and J. X. Kuang (2007). "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness". In: *Economic Theory* 33.1, pp. 67–80.
- Deterding, S., M. Sicart, L. Nacke, K. O'Hara, and D. Dixon (2011). "Gamification. Using Game-design Elements in Non-gaming Contexts". In: *CHI '11 Extended Abstracts on Human Factors in Computing Systems*. CHI EA '11. New York, NY, USA: ACM, pp. 2425–2428.
- Dranove, D. and G. Z. Jin (2010). "Quality Disclosure and Certification: Theory and Practice". In: *Journal of Economic Literature* 48, pp. 935–963.
- Dube, A., J. Jacobs, S. Naidu, and S. Suri (2020). "Monopsony in Online Labor Markets". In: *American Economic Review: Insights* 2.1, pp. 33–46.
- Dufwenberg, M. and G. Kirchsteiger (2004). "A theory of sequential reciprocity". In: *Games and Economic Behavior* 47.2, pp. 268–298.
- Durante, R., L. Putterman, and J. van der Weele (2014). "Preferences for Redistribution and Perception of Fairness: An Experimental Study". In: *Journal of the European Economic Association* 12.4, pp. 1059–1086.
- EC (2018). *Commission Staff Working Document Accompanying The Document Communication From The Commission To The European Parliament, The Council And The European Economic And Social Committee Commission General Report On The Operation Of Reach And Review Of Certain Elements Conclusions And Actions Swd/2018/058 Final*. Brussels, Belgium: European Commission. Retrieved 13 July 2018. URL: <https://goo.gl/xxyLAz>.
- ECHA (2016). *Report on the Operation of REACH and CLP 2016*. Helsinki, Finland: European Chemicals Agency. Retrieved 13 July 2018. URL: <https://goo.gl/5W9E5W>.
- Estellés-Arolas, E. and F. González-Ladrón-de-Guevara (2012). "Towards an integrated crowdsourcing definition". In: *Journal of Information Science* 38.2, pp. 189–200.
- Falk, A., T. Neuber, and N. Szech (2020). "Diffusion of Being Pivotal and Immoral Outcomes." In: *Review of Economic Studies* 87.5, pp. 2205–2229.
- Falk, A., E. Fehr, and U. Fischbacher (2008). "Testing theories of fairness — Intentions matter". In: *Games and Economic Behavior* 62.1, pp. 287–303.
- Fehr, E. and K. M. Schmidt (1999). "A Theory of Fairness, Competition, and Cooperation". In: *The Quarterly Journal of Economics* 114.3, pp. 817–868.
- Fehr, E. and S. Gächter (2000). "Fairness and Retaliation: The Economics of Reciprocity". In: *Journal of Economic Perspectives* 14.3, pp. 159–181.
- (2002). "Altruistic punishment in humans". In: *Nature* 415.6868, pp. 137–140.

- Festinger, L. (1954). "A Theory of Social Comparison Processes". In: *Human Relations* 7.2, pp. 117–140.
- Fischbacher, U. and F. Föllmi-Heusi (2013). "Lies in Disguise—An Experimental Study on Cheating". In: *Journal of the European Economic Association* 11.3, pp. 525–547.
- Franzoni, C. and H. Sauermann (2014). "Crowd science: The organization of scientific research in open collaborative projects". In: *Research Policy* 43, pp. 1–20.
- Franzoni, C., G. Scellato, and P. Stephan (2011). "Changing Incentives to Publish". In: *Science* 333.6043, pp. 702–703.
- Führ, M. and K. Bizer (2007). "REACH as a paradigm shift in chemical policy – responsive regulation and behavioural models". In: *Journal of Cleaner Production* 15.4, pp. 327–334.
- Gächter, S. and C. Thöni (2005). "Social Learning and Voluntary Cooperation among like-Minded People". In: *Journal of the European Economic Association* 3.2-3, pp. 303–314.
- Gill, D. and V. Prowse (2012). "A Structural Analysis of Disappointment Aversion in a Real Effort Competition". In: *American Economic Review* 102.1, pp. 469–503.
- Gneezy, U. (2005). "Deception: The Role of Consequences". In: *American Economic Review* 95.1, pp. 384–394.
- Gneezy, U., A. Kajackaite, and J. Sobel (2018). "Lying Aversion and the Size of the Lie". In: *American Economic Review* 108.2, pp. 419–53.
- Grabosky, P. (2013). "Beyond Responsive Regulation: The expanding role of non-state actors in the regulatory process". In: *Regulation & Governance* 7.1, pp. 114–123.
- Gürerk, Ö., B. Irlenbusch, and B. Rockenbach (2006). "The Competitive Advantage of Sanctioning Institutions". In: *Science* 312.5770, pp. 108–111.
- Güth, W. and M. G. Kocher (2014). "More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature". In: *Journal of Economic Behavior & Organization* 108, pp. 396–409.
- Hamman, J. R., G. Loewenstein, and R. A. Weber (2010). "Self-Interest through Delegation: An Additional Rationale for the Principal-Agent Relationship". In: *American Economic Review* 100.4, pp. 1826–46.
- Haustein, S., I. Peters, J. Bar-Ilan, J. Priem, H. Shema, and J. Terliesner (2014). "Coverage and adoption of altmetrics sources in the bibliometric community". In: *Scientometrics* 101.2, pp. 1145–1163.
- Hicks, J. R. and R. G. D. Allen (1934). "A Reconsideration of the Theory of Value. Part I". In: *Economica* 1.1, pp. 52–76.

- Hoffman, E., K. A. McCabe, and V. L. Smith (1996). "On expectations and the monetary stakes in ultimatum games". In: *International Journal of Game Theory* 25.3, pp. 289–301.
- Hoffman, E. and M. L. Spitzer (1985). "Entitlements, Rights, and Fairness: An Experimental Examination of Subjects' Concepts of Distributive Justice". In: *The Journal of Legal Studies* 14.2, pp. 259–297.
- Houthakker, H. S. (1950). "Revealed Preference and the Utility Function". In: *Economica* 17.66, pp. 159–174.
- Irlenbusch, B. and D. J. Saxler (2019). "The role of social information, market framing, and diffusion of responsibility as determinants of socially responsible behavior". In: *Journal of Behavioral and Experimental Economics* 80, pp. 141–161.
- Jacob, B. A. and S. D. Levitt (2003). "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating". In: *The Quarterly Journal of Economics* 118.3, pp. 843–877.
- Johnson, N. D. and A. A. Mislin (2011). "Trust games: A meta-analysis". In: *Journal of Economic Psychology* 32.5, pp. 865–889.
- Kagel, J. H. and A. E. Roth (2016). *The Handbook of Experimental Economics, Volume 2*. Princeton, US: Princeton University Press.
- Kaplow, L. and S. Shavell (1994). "Optimal Law Enforcement with Self-Reporting of Behavior". In: *Journal of Political Economy* 102, pp. 583–606.
- KEMI (2016). *Strategy for enforcement of chemicals in articles*. Bromma, Sweden: Swedish Chemicals Agency. Retrieved 13 July 2018. URL: <https://goo.gl/p3ujUb>.
- Khalmetski, K. and D. Sliwka (2019). "Disguising Lies—Image Concerns and Partial Lying in Cheating Games". In: *American Economic Journal: Microeconomics* 11.4, pp. 79–110.
- Konow, J. (2000). "Fair Shares: Accountability and Cognitive Dissonance in Allocation Decisions". In: *The American Economic Review* 90.4, pp. 1072–1091.
- Krawczyk, M. (2010). "A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution". In: *Journal of Public Economics* 94.1, pp. 131–141.
- Kullenberg, C. and D. Kasperowski (2016). "What Is Citizen Science? – A Scientometric Meta-Analysis". In: *PLOS ONE* 11.1, pp. 1–16.
- Lakhani, K. R. and E. von Hippel (2003). "How open source software works: "free" user-to-user assistance". In: *Research Policy* 32.6, pp. 923–943.
- Lerner, J. and J. Tirole (2002). "Some Simple Economics of Open Source". In: *The Journal of Industrial Economics* 50.2, pp. 197–234.

- Masclet, D., C. Noussair, S. Tucker, and M.-C. Villeval (2003). "Monetary and Non-monetary Punishment in the Voluntary Contributions Mechanism". In: *American Economic Review* 93.1, pp. 366–380.
- Mazar, N., O. Amir, and D. Ariely (2008). "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance". In: *Journal of Marketing Research* 45.6, pp. 633–644.
- Morschheuser, B., J. Hamari, J. Koivisto, and A. Maedche (2017). "Gamified crowdsourcing: Conceptualization, literature review, and future agenda". In: *International Journal of Human-Computer Studies* 106, pp. 26–43.
- Mullainathan, S. and R. H. Thaler (Oct. 2000). *Behavioral Economics*. Working Paper 7948. National Bureau of Economic Research.
- Mustonen, M. (2005). "When Does a Firm Support Substitute Open Source Programming?" In: *Journal of Economics & Management Strategy* 14.1, pp. 121–139.
- Nakatsu, R. T., E. B. Grossman, and C. L. Iacovou (2014). "A taxonomy of crowdsourcing based on task complexity". In: *Journal of Information Science* 40.6, pp. 823–834.
- Niyazov, Y., C. Vogel, R. Price, B. Lund, D. Judd, A. Akil, et al. (2016). "Open Access Meets Discoverability: Citations to Articles Posted to Academia.edu". In: *PLOS ONE* 11.2, pp. 1–23.
- Page, T., L. Putterman, and B. Unel (2005). "Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry and Efficiency". In: *The Economic Journal* 115.506, pp. 1032–1053.
- Pareto, V. (1906). *Manuale di economia politica, con una introduzione alla scienza sociale*. Vol. 16. Milano: Societa Editrice Libreria.
- Pigors, M. and B. Rockenbach (2016). "Consumer Social Responsibility". In: *Management Science* 62.11, pp. 3123–3137.
- Pizarro, D. A., E. Uhlmann, and P. Bloom (2003). "Causal deviance and the attribution of moral responsibility". In: *Journal of Experimental Social Psychology* 39.6, pp. 653–660.
- Raddick, M. J., G. Bracey, P. L. Gay, C. J. Lintott, C. Cardamone, P. Murray, et al. (2013). *Galaxy Zoo: Motivations of Citizen Scientists*.
- Rege, M. and K. Telle (2004). "The impact of social approval and framing on cooperation in public good situations". In: *Journal of Public Economics* 88.7, pp. 1625–1644.
- Robbins, L. (1932). *An Essay on the Nature and Significance of Economic Science*. London: Macmillan & Co., Limited.

- Rotemberg, J. J. (2011). "Fair Pricing". In: *Journal of the European Economic Association* 9.5, pp. 952–981.
- Rothenhäusler, D., N. Schweizer, and N. Szech (2018). "Guilt in voting and public good games". In: *European Economic Review* 101, pp. 664–681.
- Ryan, R. M. and J. P. Connell (1989). "Perceived locus of causality and internalization: Examining reasons for acting in two domains". In: *Journal of Personality and Social Psychology* 57.5, pp. 749–761.
- Samuelson, P. A. (1947). *Foundations of Economic Analysis*. Cambridge: Harvard University Press.
- Shimshack, J. P. and M. B. Ward (2005). "Regulator reputation, enforcement, and environmental compliance". In: *Journal of Environmental Economics and Management* 50.3, pp. 519–540.
- SHRG (2013). *Jaarrapportage Handhaving REACH en EU-GHS 2012*. Samenwerkingsverband Handhaving REACH en EU-GHS. Retrieved 13 July 2018. URL: <https://goo.gl/B4CCGu>.
- Silbergeld, E. K., D. Mandrioli, and C. F. Cranor (2015). "Regulating Chemicals: Law, Science, and the Unbearable Burdens of Regulation". In: *Annual Review of Public Health* 36.1, pp. 175–191.
- Stephan, P. E. (1996). "The Economics of Science". In: *Journal of Economic Literature* 34.3, pp. 1199–1235.
- Thaler, R. H. (2016). "Behavioral Economics: Past, Present, and Future". In: *American Economic Review* 106.7, pp. 1577–1600.
- The Guardian (2009). *How to crowdsource MPs' expenses*. London, England: The Guardian. Accessed 21 November 2018. URL: <https://goo.gl/MAKwHs>.
- Weil, D., A. Fung, M. Graham, and E. Fagotto (2006). "The effectiveness of regulatory disclosure policies". In: *Journal of Policy Analysis and Management* 25, pp. 155–181.
- Weil, D., M. Graham, and A. Fung (2013). "Targeting Transparency". In: *Science* 340, pp. 1410–1411.
- Wu, C.-C., Y.-F. Liu, Y.-J. Chen, and C.-J. Wang (2012). "Consumer responses to price discrimination: Discriminating bases, inequality status, and information disclosure timing influences". In: *Journal of Business Research* 65.1, pp. 106–116.
- Wursthorn, S. (2018). *Enforcement of SVHC*. Ministry of the Environment, Climate protection and Energy Sector Baden-Württemberg (Germany). Retrieved 13 July 2018. URL: <https://goo.gl/6U9LAz>.



- Xia, L., K. B. Monroe, and J. L. Cox (2004). "The Price is Unfair! A Conceptual Framework of Price Fairness Perceptions". In: *Journal of Marketing* 68.4, pp. 1–15.
- Xu, B. and D. Li (2015). "An empirical study of the motivations for content contribution and community participation in Wikipedia". In: *Information & Management* 52.3, pp. 275–286.
- Xu, L., T. Nian, and L. Cabral (2020). "What Makes Geeks Tick? A Study of Stack Overflow Careers". In: *Management Science* 66.2, pp. 587–604.
- Yang, H.-L. and C.-Y. Lai (2010). "Motivations of Wikipedia content contributors". In: *Computers in Human Behavior* 26.6, pp. 1377–1383.
- Zhang, X. ( and F. Zhu (2011). "Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia". In: *American Economic Review* 101.4, pp. 1601–1615.
- Ziegler, A., G. Romagnoli, and T. Offerman (2021). *Morals in multi-unit markets*. Tinbergen Institute Discussion Papers 20-072/I. Tinbergen Institute.



## CHAPTER II

### **Consumer Behavior under Benevolent Price Discrimination**

Alexander Erlei, Mattheus Brenig, and Nils Engelbrecht

Extensive research shows that consumers are generally averse to price discrimination. However, instruments of differential pricing can benefit consumer surplus and alleviate inequity through targeted price discounts. This paper examines whether and how these outcome considerations influence consumer reactions to price discrimination. The authors introduce the concept of benevolent price discrimination (BPD) as a downward-bound instrument that always benefits financially disadvantaged groups or individuals, thereby leading to a convergence of outcomes. Five experiments with 3415 participants show that a large share of consumers is willing to costly switch away from a store that offers a discount to low-income consumers. This happens irrespective of whether income differences are due to luck or merit. While the price-discriminating store does attract some new high-income consumers, it cannot compensate the loss of existing consumers. Simulating market transactions by endowing consumers with agency over future prices increases costly support for BPD. These results contrast previous findings on social preferences and inequity aversion, highlighting the importance of context and procedure for economic bargaining. Strong behavioral constraints persist even when price discrimination reduces unmerited income differences and no consumer experiences price increases.

# 1 Introduction

Technological progress, especially the advent of big data, has reduced the technological and informational constraints on sellers to differentiate prices among consumers. Despite these developments, sellers still hesitate to deviate from uniform pricing strategies (Fudenberg and Villas-Boas 2012). One commonly evoked explanation is that sellers anticipate the risk of antagonizing consumers (Fabiani et al. 2006), as exemplified by research showing that differential price strategies are often at odds with consumers' fairness perceptions (see e.g. Xia et al. 2004; Haws and Bearden 2006), which translates into shopping intentions (Bolton et al. 2010), purchase satisfaction (Shor and Oliver 2006), intentions to spread negative word-of-mouth (Ferguson et al. 2014) and consumption choices (Leibbrandt 2020). Thus, in addition to technological constraints, profit-seeking through demand-based price discrimination may also be inhibited by negative consumer reactions, i.e., "behavioral constraints" (Kahneman et al. 1986).

However, consumers do not appear to be universally averse to demand-based price discrimination. Most prominently, price discounts for specific groups of the population, like students and the elderly, are common in many countries. One explanation for this ostensible contradiction is that consumers accept price discrimination if it benefits lower income groups, thereby leading to an equalization of economic outcomes (Rotemberg 2011; Wu et al. 2012). Whilst theoretically appealing, this assertion has thus far not been empirically verified. Existing behavioral research primarily focuses on processes, without direct control or transparency over the realized distribution of economic outcomes between consumers. Alternative explanations for the prevalence of certain price discounts can therefore not be ruled out. Regarding e.g. the elderly, the acceptance of lower prices may instead be driven by habituation, learned societal norms, or compensation for other kinds of inequality like quality of living. If consumer reactions to price discrimination do, however, depend on the concurrent equalization of consumer outcomes, this insight could open up new possibilities for sellers to implement and market tools of differential price setting. This includes a broader range of potential price discount recipients, poaching prices, improved judgments regarding useful degrees of strategic price obfuscation and transparency, or promoting certain pricing strategies as part of corporate social responsibility communications.

This paper experimentally addresses whether and how consumer reactions depend on the effect of interpersonal price discrimination on the overall distribution of con-

sumer welfare. In particular, we introduce the concept of *benevolent price discrimination* (BPD). We define BPD as a policy of differential pricing that is downward-bound and always benefits financially disadvantaged groups or individuals, thereby leading to a convergence of outcomes. Whilst a seller can decrease prices for certain individuals or groups, they cannot increase prices for anyone. Thus, nobody is worse off as a consequence of price-discrimination. Consumers without access to the lower price experience *relatively disadvantageous*, rather than *disadvantageous* price discrimination. Most importantly, BPD always leads to a more equal outcome distribution. We thus quantify behavioral constraints towards price discrimination when implemented as a mechanism that increases joint overall welfare while creating a more equitable distribution of economic gains.

### 1.1 Procedural and distributive price fairness

The price fairness literature has long recognized the importance of outcome-related reasoning, identifying two distinct main criteria that determine consumer price judgments: procedural and distributive price fairness (Maxwell 2007). Procedural price fairness refers to the procedures or processes behind a pricing strategy, including intentions (Campbell 1999) and the specific variables determining differential pricing (Variable of Discrimination, VOD; Bayer 2010; Kuo et al. 2016). Distributive price fairness refers to comparisons with respect to the outcome of other actors, which has thus far often been interpreted as the offered price (see e.g. Ferguson et al. 2014; Xia et al. 2004). This conceptualization, while informative, does not address distributive fairness with regard to the distribution of consumer welfare. To integrate outcome equalization into the analysis of consumer behavior under price discrimination, research needs to relate price differences to differences in purchasing power and the subsequent distribution of e.g. goods or income, rather than equating offered prices with outcomes. For example, prices that are tailored towards different income groups might be seen as distributively unfair because people have to pay different amounts for the same good or service, but might simultaneously lead to a convergence of consumer welfare, which could increase distributive fairness. This article therefore refers to “outcomes” as the final distribution of consumer welfare, and interprets the unequal treatment of consumers through price differences as a procedural variable.

Despite these standing conceptual differences, prior research clearly establishes the importance of between-consumer comparisons for adverse reactions to price

discrimination. One central tenet is that a consumer explicitly or implicitly refers to an internal (e.g. Thaler 1985; Janiszewski and Lichtenstein 1999; Koszegi and Rabin 2006) or external (e.g. Mayhew and Winer 1992; Dholakia and Simonson 2005) reference point to assess whether an offered price is fair. Internal, memory-based reference points are the prices that the consumer themselves paid in similar previous transactions (Herz and Taubinsky 2018). External, interpersonal reference points are the prices paid by one's peers (Ho and Su 2009), i.e. other consumers (Haws and Bearden 2006; Jin et al. 2014), for an identical good. Although both factors are important, there is evidence to suggest that interpersonal comparison has a larger impact on consumers' fairness perception than self-comparison (see e.g. Xia et al. 2004), which is particularly relevant in modern online markets that exhibit high transparency and extensive information sharing (Anderson and Simester 2008).

However, moving the analysis to the level of realized economic outcome distributions requires a conceptual shift whereby potential preferences for equal outcomes are tested within a price setting framework that replicates the important prerequisites of consumer price comparisons.

## 1.2 Preferences for Equal Outcomes

There is ample evidence that people generally care about the equal distribution of monetary outcomes (Fehr et al. 1993; Harrington Jr et al. 2016), particularly if existing income inequalities are seen as arbitrary (Rutström and Williams 2000; Alesina and Angeletos 2005; Durante et al. 2014).

Two very influential economic models propose that a person's utility depends on the outcome of other people, either because people experience disutility when their outcome is different from other people's outcomes (Fehr and Schmidt 1999), or because their utility function depends on their share of total payoffs (Bolton and Ockenfels 2000). Thus, as long as differential pricing leads to a convergence of outcomes without any (or many) consumer losses, it may not only be accepted by consumers, but even be desirable.<sup>1</sup>

However, supporting results from the literature on inequity aversion and studies showing human preferences for redistribution under unequal economic outcomes

---

<sup>1</sup>See Krawczyk and Le Lec (2010) for a model of procedural and distributive fairness that predicts higher support of redistribution in such cases. See also Guo (2015) and Li and Jain (2016) for theoretical analyses on how consumer inequity aversion and social preferences affect optimal buyer strategies.

cannot be readily applied to the context of price discrimination in consumer markets. Foremost, this literature purposely abstracts from many process-related variables in order to focus exclusively on outcome distributions. It thus factors out many elements of the price setting mechanism, which could well interact with or mediate distributive preferences as well as their translation into behavior. For example, under price discrimination, redistribution is not actively chosen by consumers, but imposed by sellers. This evokes considerations regarding reciprocity (Fehr and Gächter 2000) and consumers act as second, rather than first movers, strongly inhibiting their agency in determining prices and thereby outcomes (Choshen-Hillel and Yaniv 2011; Alexopoulos et al. 2013). Moreover, moral responsibility for causing fairer market outcomes is shifted away from the consumer. At best, consumers may ascribe themselves moral responsibility through their influence on future prices and thereby future outcomes (Pigors and Rockenbach 2016).

### 1.3 Outcomes vs. Procedures

Overall, it remains unclear whether outcome equalization can alleviate behavioral constraints to price discrimination. This will crucially depend on the extent to which consumers weigh unequal price treatment (e.g., not being eligible for a price discount other consumers receive; Ho and Su 2009; Ho et al. 2014) against the equalization of consumer outcomes. This relationship is not obvious, since there is evidence from outside the price literature that both outcomes (Lerner and Whitehead 1980; Rutte and Messick 1995) and procedures (Folger and Konovsky 1989; Van den Bos, Vermunt, et al. 1997) could be more important in forming overall fairness judgments. In particular, a large body of work highlights the role of procedural elements e.g., in bargaining contexts (Trautmann 2009) or for contextual variables like norms (Garbarino and Maxwell 2010) and consumer loyalty (Anderson and Simester 2010). Other research finds equivalent effects on allocation acceptability (Bolton et al. 2005). According to Van den Bos, Lind, et al. (1997), the availability of outcome information could mediate the influence of procedural fairness. While procedures were significant in shaping fairness judgments without outcome information, subjects who knew about the outcomes of other participants did not appear to be significantly influenced by procedures. In Bolton et al. (2005), fair outcomes were accepted irrespective of procedural bias.

Thus, in so far as these results translate into purchasing environments, consumers who know that BPD only benefits those with lower incomes might accept or even

support it. However, for consumers who do not receive information on other consumers' outcomes and experience relatively disadvantageous price discrimination, the literature on price fairness and reference points suggests comparatively strong aversions.

## 1.4 Study Overview

To assess consumer reactions towards BPD, we conducted a series of incentivized, context-neutral, and controlled experimental studies on Amazon Mechanical Turk (MTurk). Sellers could offer a price discount for low-income consumers, leading to a convergence in consumer outcomes. Price discrimination was transparent and high-income consumers could either costly switch to a second seller who did (not) offer the discount, choose not to buy at all, or maximize their income by staying with the same seller across all rounds. By quantifying both, behavioral shifts away from and towards the price discriminating store, we were able to capture bidirectional consumer migration patterns which have so far been largely neglected by the experimental literature. This allowed us to make qualitative judgments about the net-effect of BPD, rather than focusing solely on the behavioral constraints elicited by existing consumers.

In the first study, high-income consumers received no explanation as to who benefited from the price discount or how it was implemented. Whilst this masked an important element of our framework, abstracting from differences in outcomes allowed us to establish a first benchmark of consumer reactions towards differential downward pricing. The second study addressed the role of reciprocity by keeping sellers out of the price-setting process. By shifting responsibility for the pricing decision from sellers to an algorithm, we were also able to examine the role of human intentionality in consumer reactions to price discrimination. Studies 3 to 5 then progressively introduced new process-related variables, endowing consumers with information and agency that mirror important contextual variables of real-world consumer markets and enable concrete between-consumer comparisons. In the third study, high-income consumers learned that price discounts were available for low-income consumers only. This change revealed the benevolent nature of our price discounts, and thus introduced the interaction between unequal price treatment and an equalization of consumer outcomes. Moreover, we used two mechanisms to allocate initial incomes. In one condition, income was randomly assigned to consumers, in the other, consumer



income was assigned based on their performance in a real-effort task. This allowed us to test the prediction of equity theory (i.e. distributive fairness), whereby an outcome distribution is perceived as fair if outcomes are proportional to an individual's input. In Study 4, we increased consumers' moral responsibility for their purchasing behavior by introducing a market mechanism which gave consumers agency over future prices and thereby outcomes. Finally, we conducted three robustness checks addressing transparency, the costs of switching and consumer motivations behind their purchasing decisions.

Our results suggest widespread consumer aversions towards BPD. Across all studies, between 30 %–40 % of consumers exhibited strictly BPD-averse purchasing patterns. Reactions were the strongest in Study 1, where consumers did not receive any contextual explanations for price discrimination. As shown in Study 2, this behavior can not be explained by consumers' assumptions about seller intentions. That is, holding seller intentions constant does not substantially affect consumer aversion towards BPD. Even in Study 3, with a clear understanding that price discrimination only benefited consumers who were arbitrarily assigned a lower income, every third consumer chose to costly switch away from the price discriminating store. Endowing consumers with agency through a market mechanism in Study 4 did slightly decrease costly switching, but only if income differences were generated randomly. Contrary to studies 1–3, a sizeable number of consumers was willing to *support* BPD by costly switching *towards* the price-discriminating seller when consumer purchases influenced the store's future pricing strategy. Finally, Study 5 subsumes three robustness checks that confirm the validity of our behavioral interpretations and show that results hold when consumers decide under full transparency as well as for varying switching costs.

Overall, we show that outcome-fairness related reasoning mediates behavioral consumer constraints only under very specific conditions. In particular, consumers need to participate in determining future prices. Highlighting the re-distributive nature of BPD has little effect on behavior, and contrary to much of the existing literature on social preferences, consumers do not differentiate between merited and unmerited income inequality. Even in the most benevolent case, a substantial share of consumers is willing to give up money in order to switch to a competitor without BPD despite no economic gains.

The complete data set, additional information and this project's code can be accessed under the following repository: <https://s.gwdg.de/mwmBdv>.

## 2 Study I

Study 1 first used a simple one-shot design to elicit general consumer preferences for price discrimination. Second, we examined whether these preferences generalize to a more comprehensive multiple round within-subject scenario with monetary “switching costs”. The latter reflects that in reality, punishing sellers for their pricing strategy by purchasing elsewhere usually entails transaction costs – even in the case of homogeneous goods (Nilssen 1992). Throughout both experiments, subjects did not receive any explanation about the introduction of price discrimination. The beneficiaries of BPD were kept private and consumers could not infer that price discounts led to more equal outcomes. In reality, it is likely that BPD would be used in situations where consumers cannot know how different prices are generated and who benefits. As we will show in later studies, the main effects of this study replicate to purchase environments where outcome equalization is transparent to consumers.

### 2.1 Experimental Design

Because the basic experimental setup was the same across all studies, we will describe it here in more detail and refer to this description in the remainder of the paper.

We ran purchasing experiments where consumers had the option to buy a homogeneous good at one of two stores, each run by another participant acting as store manager. We measure preferences for or against price discrimination by a consumer’s store choice. We included a post-experimental questionnaire measuring participants’ fairness perceptions, attitudes towards price discrimination as well as a questionnaire on social comparison (Gibbons and Buunk 1999). All experiments in this paper were conducted online using MTurk and oTree (Chen et al. 2016). Participants enrolled on their own accord and were randomly assigned to one treatment and one role. After the instructions, participants had two attempts to answer five comprehension questions correctly. Those who failed both attempts were excluded from participating. We used “Coins” as our experimental currency. Coins were later converted into dollars, where 10 coins equaled 3 cents. Additionally, subjects playing the role of “consumer” received a fixed payment of \$1.20 for completing the survey. There were two main experimental paradigms to analyze consumer behavior: *One-Shot* and *Repeated*.

### *One-Shot*

Upon arrival, participants were randomly assigned the role of either a *high-income* or a *low-income* consumer. High-income (low-income) consumers learned that they would receive an endowment of 100 (50) Coins. Coins could be used to buy a good at one of two stores (A, B). In both stores, the good returned a value of 150 Coins to a consumer's final payoff. Thus, consumers were incentivized to always buy the good. The two stores sold the same good and were identical except for one thing: one store offered all consumers the good for a price of 100 Coins. The other store offered low-income consumers the good for a discounted price of 50 Coins and charged high-income consumers the regular price of 100 Coins. Consumers learned that each store was run by a manager, who had decided on the pricing strategy of the store beforehand, and that managers earned money for each good sold in their store. This ensured that consumer purchasing choices were meaningful.

Participants decided at which of the two stores they wanted to purchase the good. High-income consumers could observe that one store offered the good at a price of 50 Coins and learned with a click that they were not eligible for that discount. However, in Studies 1 and 2, no explanation for the different prices was given.

We randomized the position (left, right) and name (A, B) of the store introducing price discrimination. Thus, if price discrimination had no impact, we would expect consumers to be equally distributed across both stores.

### *Repeated*

Compared to the one-shot design, consumers completed four purchasing rounds. We did not change the basic parameters of the experiment. High-income consumers received 100 Coins each round and could use them to purchase the good in either store A or store B. Low-income consumers were not able to purchase a good in the first two rounds and had to rely on an outside option that was a simple multiplier of one. Endowments could not be transferred from one round to another. Thus, we can rule out that consumers may respond to a switch in sellers pricing strategies by delaying their purchases in anticipation of future discounts (Coase 1972).

To make switching meaningful, we introduced monetary costs for switching between stores in two consecutive rounds. Consumers always started a round in the store they chose the previous round. For example, if a consumer purchased the good in store A in the first round, they started the second round in store A. If the consumer

then decided to switch and purchase the good in store B, they had to pay a fee of 10 Coins, i.e. 10% of their endowment per round. Thus, consumers were monetarily disincentivized to switch and maximized their payoff by always purchasing from the same store.

In the first two purchasing rounds, both stores offered the good for the same price. After the second round, consumers learned that one of the two store managers changed their pricing strategy and would offer low-income consumers the good for a price of 50 Coins in the remaining two rounds. Depending on treatment, this was either (i) the manager of the store consumers had purchased in during the second round (*Avoid*) or (ii) the manager of the store consumers had *not* purchased in during the second round (*Approach*). Price discrimination was transparent to all consumers.

Compared to the one-shot design, this setup imposes stricter conditions on observing consumer preferences against price discrimination. Subjects who switch away from (or towards) the price discriminating store are willing to substantially reduce their own payoff. Second, it allows us to analyze benevolent price discrimination, where consumers who already participate in the market do not experience a price increase and only low-income consumers who priorly could not purchase any goods benefit from lower prices. Whereas in the one-shot design, consumers might perceive the lower price to be at their own costs, the within-subject design establishes the reference price of 100 Coins over two rounds and no consumer loses money as a result of price discrimination. Importantly, low-income consumers miss out on substantial payoff over the first two rounds, and the price discount in rounds three and four serves to mitigate, but not fully eliminate unequal outcomes.

### *Manager*

In each treatment, two participants were assigned the role of a store manager. Before consumers made their choices, managers were free to decide whether they wanted to implement a price discrimination strategy (a discount for low-income consumers) in their store. In the *Repeated* treatments, price discounts could only be introduced after the second round. Managers had full information about the experimental setup. We gathered manager observations until we had one manager who decided for and one manager who decided against price discrimination for each experimental treatment. Hence, no deception was involved. Managers were rewarded with a base reward of \$0.50 and earned 1 Cent for each good sold in their store.

### Study 1

In Study 1, information on consumer endowment was private. High-income consumers did neither know about the existence of low-income consumers, nor the endowment of anybody else. They also did not receive any explanation about the price discount. Participants were randomly assigned to *One-Shot*, *Repeated-Avoid* or *Repeated-Approach*. We gathered data until we had 100 high-income consumers in each treatment and set a fixed probability of 10% for a participant to become a low-income consumer.

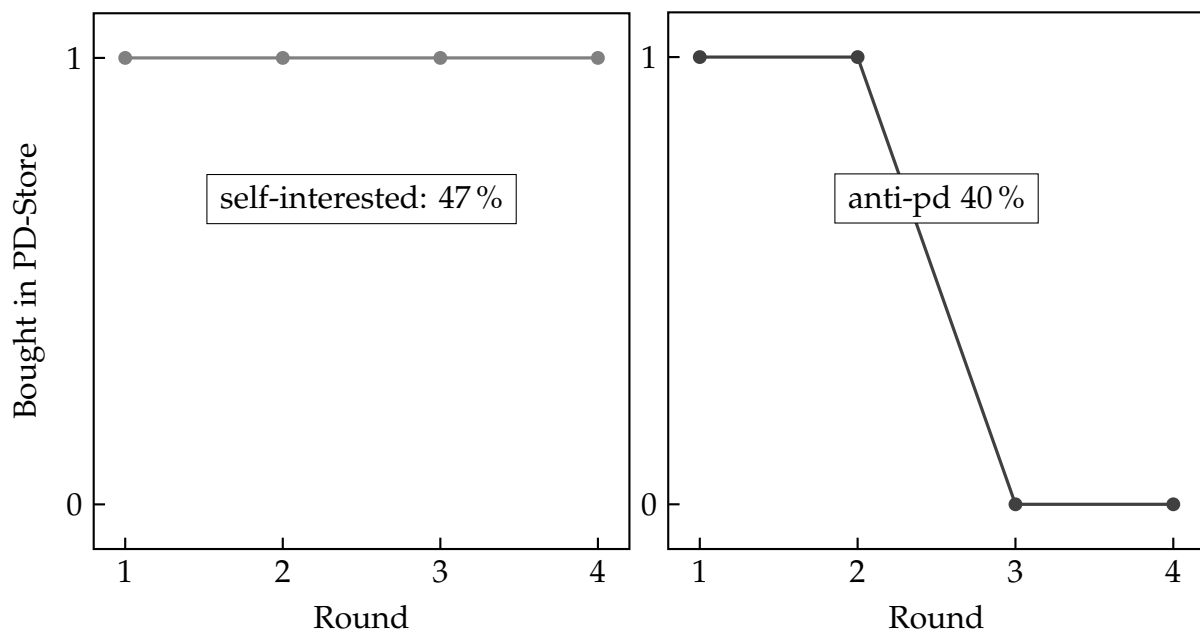
## 2.2 Results

We excluded all observations where a subject chose the outside option in the second round. This left us with 96 independent observations in *Avoid*, 97 independent observations in *Approach* and 100 in *One-Shot* (47% female). Our results confirm widespread consumer aversion to price discrimination. Even more so, we find that a large share of consumers is willing to incur costs in order to switch away from a benevolently price-discriminating store without any possibility of subsequent monetary gains.

### *One-Shot*

From now on, we will refer to high-income consumers as “consumers”. Low-income consumers were irrelevant for our analysis and only served to avoid deception. As hypothesized, a large majority of consumers preferred the non-price-discriminating store (*NoPD Store*). Without any explanation for the observed price discrimination, 86% of consumers purchased the good in the store that charged all consumers the same price, 9% purchased in store with price discounts and 5% chose the outside option. A one sample *t*-test also confirmed consumers to significantly prefer the non-price-discriminating store ( $t = 14.25, p = 0.000$ ). The name ( $\tilde{\chi}^2 = 1.33, p = 0.25$ ) and position ( $\tilde{\chi}^2 = 1.66, p = 0.20$ ) of the store introducing price discrimination did not have a significant effect on store choice.

**Result 1** *Under scarce information, consumers exhibit strong preferences against a price discriminating store.*



**Figure 1** — The two most common behavioral patterns in *Avoid*

*Note.* The two most common purchase patterns were “self-interested” and “anti-pd”. Consumers classified as self-interested never switched to the other store. Consumers classified as anti-pd stayed within the same store for the first two rounds and then switched after their store introduced price discrimination.

### *Repeated-Avoid*

In the first two rounds, both stores offered the good for the same price. After the second round, consumers learned that the store they chose in the second round offered some consumers lower prices. Figure 1 illustrates the two most common behavioral patterns in the *Avoid* treatment. We find that 47 % of consumers never switched and maximized their payoff. We classify those consumers as “self-interested”. Surprisingly, almost as many consumers decided to costly switch away from a benevolently price-discriminating store. We classify those consumers who stayed in the same store for the first two rounds, switched after the second round, and stayed in the other store for the third and fourth round as “anti-pd”.<sup>2</sup> Using this conservative measure, 40 % of consumers exhibited the corresponding pattern. To support this conjunction, we ran a paired *t*-test to test whether consumers were less likely to purchase in the price discriminating store after it had introduced price discrimination, i.e. in round three and four. On average, consumers purchased 1.91 out of 2 goods in round one and two in the price-discriminating store, and only 1.04 out of 2 in rounds three and four ( $t = -8.64, p = 0.000$ ).

Throughout this paper, we will use a random effects panel logistic regression model with clustered standard errors to confirm and expand on our findings. In accordance with the results above, we report a significant and large effect of the BPD-dummy on a subject’s probability to purchase in the price-discriminating store (see Tables 9 and 8 in the appendix).

**Result 2** *Under scarce information, roughly 40 % of consumers costly punish a benevolently price-discriminating store.*

### *Repeated-Approach*

In *Approach*, price discrimination was introduced by the store not chosen by a consumer in the second round. This setup was designed to capture positive reactions towards BPD. 80 % of consumers followed the self-interested pattern, whereas only 4 %

---

<sup>2</sup>We can rule out egalitarian motives towards the managers’ income distribution as a reason for consumer switching, because we did not tell participants how many rounds they would play. Thus, consumers who wanted both managers to earn the same would have already switched in the second round. A few consumers seemed to exhibit these egalitarian preferences. Text data from an open-ended question in the questionnaire of later studies also confirms the validity of our categorization.

switched to the price-discriminating store. A paired  $t$ -test showed no significant differences in the likelihood that a consumer purchased the good in the price-discriminating store between rounds 1 and 2 versus 3 and 4. On average, consumers purchased 0.07 out of 2 goods in round one and two in the price-discriminating store, and 0.12 out of 2 after the introduction of price discrimination in rounds three and four ( $t = 0.96, p = 0.339$ ).

**Result 3** *Consumers are generally not willing to costly support a benevolently price-discriminating store under scarce information.*

### *Consumer Attitudes*

Table 1 shows subjects' evaluations for a selection of post-experimental questions from *Avoid*.<sup>3</sup> There are only few differences between self-interested and anti-pd consumers. Descriptively, those who costly switched away from the price-discriminating store found price discrimination on average fairer (but still very unfair) than those who decided to stay in the same store throughout all four purchasing rounds. Since both groups also did not differ on the Social Comparison Orientation scale, the recorded attitudes suggest that in the no-information context, peer-induced fairness concerns cannot explain adverse consumer reactions towards BPD. While price-discrimination was seen as unfair by almost everybody, perceptions of unfairness do not appear to be a sufficient condition for costly switching.

Instead, consumer behavior might be partially driven by perceptions of intention and exploitation. Subjects categorized as *anti-pd* were less likely to ascribe good intentions to the introduction of price discrimination and thought more often that their manager wanted to take advantage of them ( $t = 1.80, p = 0.075$ ). They also found costly punishment to reduce the manager's income fairer. This interpretation would be in line with attribution theory (Fiske and Taylor 2013), whereby unfair acts by a causal agent can cause the attribution of blame and motivational beliefs.

## 2.3 Discussion

We show that a large share of consumers costly punishes store managers even when they are not negatively affected by the introduction of price discrimination. Forty per-

---

<sup>3</sup>Because participants in *Approach* did not significantly react towards price discrimination, we refrain from reporting the results here.



**Table 1** — Consumer attitudes in *Avoid* for full sample and by behavioral pattern

Short description of questionnaire item	By behavioral pattern				
	Full sample		<i>anti-pd</i>		<i>self-interested</i>
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Diff. <i>t</i> -stat.
Fairness of PD <sup>1</sup>	2.21 (1.55)	2.47 (1.67)	2.11 (1.54)	2.11 (1.54)	1.03
Fairness of punishment by switching <sup>2</sup>	5.81 (1.60)	6.32 (0.93)	5.58 (1.86)	5.58 (1.86)	2.22**
Good intentions of PD manager <sup>3</sup>	2.68 (1.52)	2.42 (1.43)	2.96 (1.62)	2.96 (1.62)	-1.58
Exploitation by PD manager <sup>4</sup>	4.79 (1.84)	5.18 (1.67)	4.44 (2.01)	4.44 (2.01)	1.80*
SCO <sup>5</sup>	4.50 (1.24)	4.42 (1.17)	4.71 (1.25)	4.71 (1.25)	-1.07

*Note.* All questions were answered on a 7-point scale. Superscripts 3, 4: scale ranged from 1 (unfair) to 7 (fair). Superscripts 1, 2, 5: scale ranged or from 1 (strongly disagree) to 7 (strongly agree); 1<sup>1</sup>Please rate how fair you think it is that manager A/B decided to offer some consumers lower prices for the same good.; 2<sup>2</sup>The more goods a store sells, the higher the income of the participant acting as its manager. Do you consider it fair to switch from Store A/B to Store B/A in order to reduce manager A/B's profit because of their pricing strategy?"; 3<sup>3</sup>Manager A/B's intention for introducing interpersonal price differences was good."; 4<sup>4</sup>Manager A/B intended to take advantage of me (the consumer)."; 5<sup>5</sup>Mean of Social comparison orientation scale by Gibbons and Buunk (1999); The reported *t*-statistic results from a two-sample mean-comparison *t*-tests by behavioral pattern; \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

cent of consumers exhibit strictly averse behavioral patterns by engaging in inefficient switching and punishing the differential price decrease. In total, the number of goods sold by the price-discriminating store decreases by roughly 45 %.

These results cannot be explained by fairness concerns or social comparison. Rather, subjects appear to assign bad intentions to the introduction of price discrimination. One reason might be that subjects in this study received no information about either the goal, the reasons or the beneficiaries of price discrimination. Without any salient point of comparison or justification mechanism, subjects were free in their interpretation of the manager's actions. These were largely judged to be unfair, regardless of the subject's behavior. This interpretation is coherent with a number of studies on attribution theory, showing that people have a strong tendency to infer causes and assign responsibility, especially when outcomes are perceived to be unfair (Heider 1982; Blount 1995; Falk et al. 2008).

In order to better understand the motives behind consumer aversions towards BPD, we conducted a follow-up study designed to control for manager intentions. If rather than fairness, the ascription of intentions and bad motives drive consumer punishment, we would expect less inefficient switching as a response to price discrimination. Study 2 also controls for negative reciprocity towards the store manager as a reason for switching between stores.

### 3 Study II

We eliminated any intentionality behind price discrimination by keeping managers out of the price-setting process. Consumers learned that both managers decided to let a pricing algorithm determine store prices, instead of setting prices themselves. They were told that managers had no information about how the algorithm determined prices and could therefore not predict changes caused by the algorithm. This setup ensures that we isolate the effect of differing intentions: subjects can still punish the manager of the price-discriminating store for bearing responsibility for choosing the algorithm in the first place.<sup>4</sup> To the extent that subjects are willing to attribute motives,

---

<sup>4</sup>Importantly, this setup ensures meaningful punishment, since switching decreases the income of a related actor, rather than an innocent bystander. The popular choice of substituting the human decision-maker with e.g. a computer or an algorithm would not only render punishment arbitrary, but eliminate a whole host of other variables we cannot account for.

e.g. carelessness, these motives should not differ between both managers since they made the same decision under the same information.

### 3.1 Experimental Design

To test for the effect of intentionality, we ran two treatments (*Intention* and *No Intention*) of a slightly adjusted version of our (repeated) *Avoid* setup from Study 1.<sup>5</sup> Whereas *Intention* simply replicated the first study, *No Intention* introduced the pricing algorithms mentioned in the previous paragraph.<sup>6</sup> One algorithm was randomly selected to price discriminate, the other one did not change prices throughout the task. We gathered 200 consumer observations per treatment and extended the post-experimental questionnaire to better evaluate the effectiveness of our manipulation (see the online appendix for an overview).

### 3.2 Results

We excluded all observations where a subject chose to rely on the outside option in the second round. This left us with 198 independent observations in *Intention* and 193 independent observations in *No Intention* (52% female). Our results indicate that intentions are not the main driver of consumer aversions against BPD. Across both treatments, a significant share of consumers is willing to costly punish the price discriminating store manager.

#### *Manipulation Check*

Our manipulation of intentionality should have two effects: First, when asked about the intentions behind the manager action triggering price discrimination<sup>7</sup>, we expected subjects in *No Intention* to ascribe more “neutral” intentions. Second, since both managers made the exact same decision, we expected no differences in the ascribed intentions between both managers in *No Intention*. Table 2 shows average consumer

<sup>5</sup>Because consumers appeared to have no preferences for price discrimination in Study 1, we omit a follow-up analysis for *Approach*.

<sup>6</sup>Before launching the main experiment, we gathered two manager observations per treatment. In *No Intention*, both managers decided to delegate price-setting to their respective algorithm.

<sup>7</sup>i.e. the intentions behind offering some consumers lower prices in *Intention*, and the intentions behind choosing the pricing algorithm in *No Intention*

sentiments towards the price-discriminating manager over the whole sample by treatment.

Subjects in *Intention* evaluated the intention of the price-discriminating store manager to be both significantly less good as well as significantly more bad. For *No Intention*, the averages are closer to the “neutral” point of 4 on the response scale. Furthermore, subjects in *Intention* were more likely to think that the price-discriminating manager had the intention of exploiting them. Thus, our manipulation seems to have worked in the desired direction.

The differences in ascribed intentions between the two managers in *No Intention* are more ambivalent. For all three questions capturing a manager’s intentions, subjects ascribed more positive intentions to the manager whose algorithm did not introduce the lower price for some consumers. That was true when asked whether (i) the manager’s intention to choose the pricing algorithm was bad (PD: 3.39, NoPD: 2.76;  $t = 4.96$ ,  $p = 0.000$ ), (ii) the manager’s intention to choose the pricing algorithm was good (PD: 4.47, NoPD: 5.11;  $t = -5.58$ ,  $p = 0.000$ ) or (iii) the manager tried to take advantage of the consumer (PD: 3.57, NoPD: 2.34;  $t = 8.49$ ,  $p = 0.000$ ). However, when compared to *Intention*, the differences between the evaluation of both managers are much smaller. For differences of 0.63, 0.65 and 1.23 in *No Intention*, we record differences of 2.64, 2.47 and 2.64 in *Intention*. Hence we still affirm that our manipulation worked in the desired direction.

### *Main Results*

First, we replicate our results from Study 1. In *Intention* (*No Intention*), 37% (30%) of consumers were willing to costly switch away from a benevolently price-discriminating store, while 45% (49%) acted monetarily self-interested. We further identify a third group of subjects who switched within the first two rounds and settled for the non-price discriminating store afterwards (10%). Since subjects did not know about the number of rounds beforehand, this might be interpreted as egalitarian preferences, i.e. a preference for both managers to earn the same, until price discrimination swayed them towards one store. The average number of goods sold by the price-discriminating store decreased significantly with the introduction of price discrimination in round three and fourth in both *Intention* ( $t = -13.23$ ,  $p = 0.000$ ) and *No Intention* ( $t = -11.02$ ,  $p = 0.000$ ).

**Table 2** — Consumer attitudes by experimental condition

Short description of questionnaire item	By experimental condition		
	<i>Intention</i>	<i>No intention</i>	Diff.
	Mean (SD)	Mean (SD)	<i>t</i> -stat.
Fairness of PD <sup>1</sup>	2.27 (1.44)	2.41 (1.55)	-0.90
Good intentions of PD manager <sup>2</sup>	2.56 (1.51)	4.47 (1.66)	-11.91***
Bad intentions of PD manager <sup>3</sup>	5.31 (1.60)	3.39 (1.82)	11.11***
Exploitation by PD manager <sup>4</sup>	5.02 (1.75)	3.57 (2.01)	7.63***
Accountability of PD manager <sup>5</sup>	5.69 (1.70)	3.81 (2.08)	9.78***

*Note.* All questions were answered on a 7-point scale. Superscript 1: 1 (unfair) to 7 (fair). Superscripts 2–4: 1 (strongly disagree) to 7 (strongly agree). Superscript 5: 1 (not at all) to 7 (very much); <sup>1</sup>“Please rate how fair you think it is that [*Intention*: manager A/B; *No intention*: Store A/B’s algorithm] decided to offer some consumers lower prices for the same good”, <sup>2</sup>“Manager A/B’s intention for [*Intention*: introducing interpersonal price differences; *No intention*: choosing the pricing algorithm] was good.”, <sup>3</sup>Wording is equal to row 2, except last word is “bad”, <sup>4</sup>“Manager A/B intended to take advantage of me (the consumer).”, <sup>5</sup>“How much do you hold Manager A/B accountable for not being offered the good at the lower price?”; The reported *t*-statistic results from a two-sample mean-comparison *t*-tests by behavioral pattern; \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Second, we find moderate evidence for small treatment differences. While the share of subjects classified as *anti-pd* drops by seven percentage points in *No Intention*, a *t*-test comparing the average number of goods sold by the price-discriminating store during rounds three and four in *Intention* (0.98) and *No Intention* (1.09) shows no significant difference ( $t = 1.1, p = 0.270$ ). The regression model reveals a significant interaction between the introduction of BPD and *Intention* (see Table 9), suggesting that subjects were less likely to purchase in the price-discriminating store after the introduction of BPD when the decision was made by a store manager instead of a pricing algorithm. Overall, the main results hold irrespective of whether price-discrimination is introduced by the manager themselves or an autonomous pricing algorithm. Our manipulation succeeded in partially neutralizing consumer perceptions of manager intentions and significantly reduced the perceived differences in intentions between the two managers. We take these results as evidence that consumer aversion towards benevolent price-discrimination is largely independent of manager intentions. Despite controlling for intentions, all three behavioral patterns are largely consistent across the two treatments.

**Result 4** *Consumer aversion to benevolent price discrimination under information scarcity is not driven by perceived manager intentions.*

### *Consumer Attitudes*

Table 3 shows subjects' attitudes for a selection of post-experimental questionnaire items. Notably, subjects categorized as *anti-pd* were significantly more neutral in their assessment of the price-discriminating manager's intentions when they learned that the manager had delegated price-setting to a "black-box" algorithm. This again suggests that manager intentions were not the primary driver of costly switching. Subjects in *No Intention* also tended to hold the manager less accountable. However, there were no differences in the perception of fairness regarding the introduction of price discrimination as well as perceived levels of exploitation. Irrespective of treatment, subjects judged price discrimination to be very unfair and felt exploited. However, *anti-pd* consumers found price discrimination significantly more unfair than *self-interested* consumers (*Intention*:  $t = -3.66, p = 0.000$ ).<sup>8</sup> The fact that both groups found price discrimination unfair does suggest, however, that fairness perceptions are neither the sole, nor a sufficient driver of costly punishment. For all other items, the differences between treatments and categories are as expected and in line with Study 1.

### 3.3 Discussion

We show that consumer aversions to benevolent price discrimination under scarce information persist even when we control for manager intentions. The share of consumers willing to costly switch away from a store that introduces downward price discrimination remains high, and the store sells between 48% (*Intention*) and 40% (*No Intention*) less as a result. In sum, the two studies suggest consumer aversion to benevolent price discrimination to be a widespread phenomenon that does not depend on the ascription of bad motives or intentions. Rather, a large share of consumers seems to object to the general concept, at least in a situation of information scarcity. While the prevalence of costly switching may appear high, the finding itself is in line with the existing literature on consumer attitudes towards and perceptions of

---

<sup>8</sup>For *No Intention*, the results are not significant ( $t = -1.77, p = 0.079$ ).

**Table 3** — Consumer attitudes by behavioral patterns and experimental condition

<i>Behavioral pattern</i>	By experimental condition		
	<i>Intention</i>	<i>No intention</i>	Diff.
Short description of questionnaire item	Mean (SD)	Mean (SD)	<i>t</i> -stat.
<i>self-interested</i>			
Fairness of PD <sup>1</sup>	2.62 (1.57)	2.51 (1.58)	0.48
Fairness of punishment by switching <sup>2</sup>	5.30 (1.64)	4.29 (1.85)	3.92***
Good intentions of PD manager <sup>3</sup>	2.89 (1.64)	4.72 (1.57)	-7.76***
Bad intentions of PD manager <sup>4</sup>	4.99 (1.79)	3.18 (1.69)	7.05***
Exploitation by PD manager <sup>5</sup>	4.69 (1.91)	3.23 (1.95)	5.12***
Feel exploited by PD manager <sup>6</sup>	5.08 (1.64)	4.90 (1.78)	0.69
Accountability of PD manager <sup>7</sup>	5.12 (1.90)	3.00 (1.93)	7.52***
SCO <sup>8</sup>	4.26 (1.23)	4.27 (1.08)	-0.04
<i>anti-pd</i>			
Fairness of PD <sup>1</sup>	1.82 (1.13)	2.05 (1.47)	-1.01
Fairness of punishment by switching <sup>2</sup>	6.57 (0.97)	5.89 (1.63)	2.94***
Good intentions of PD manager <sup>3</sup>	2.19 (1.28)	4.11 (1.69)	-7.40***
Bad intentions of PD manager <sup>4</sup>	5.62 (1.32)	3.88 (1.91)	6.17***
Exploitation by PD manager <sup>5</sup>	5.54 (1.42)	4.30 (2.04)	4.12***
Feel exploited by PD manager <sup>6</sup>	6.00 (1.27)	6.02 (1.32)	-0.08
Accountability of PD manager <sup>7</sup>	6.49 (0.97)	4.95 (1.85)	6.16***
SCO <sup>8</sup>	4.57 (1.12)	4.31 (1.06)	1.30

*Note.* All questions were answered on a 7-point scale. Superscript 1: 1 (unfair) to 7 (fair). Superscripts 2-4: 1 (strongly disagree) to 7 (strongly agree). Superscript 5: 1 (not at all) to 7 (very much); <sup>1</sup>"Please rate how fair you think it is that [*Intention*: manager A/B; *No intention*: Store A/B's algorithm] decided to offer some consumers lower prices for the same good", <sup>2</sup>"The more goods a store sells, the higher the income of the participant acting as its manager. Do you consider it fair to switch from Store A/B to Store B/A in order to reduce manager A/B's profit because of their pricing strategy?", <sup>3</sup>"Manager A/B's intention for [*Intention*: introducing interpersonal price differences; *No intention*: choosing the pricing algorithm] was good.", <sup>4</sup>Wording is equal to row 2, except last word is "bad", <sup>5</sup>"Manager A/B intended to take advantage of me (the consumer).", <sup>6</sup>"The fact that [*Intention*: Manager A/B, *No intention*: Store A/B's algorithm] decided to offer some consumers lower prices makes me feel taken advantage of.", <sup>7</sup>"How much do you hold Manager A/B accountable for not being offered the good at the lower price?", <sup>8</sup>Mean of Social comparison orientation scale by Gibbons and Buunk (1999); The reported *t*-statistic results from a two-sample mean-comparison *t*-tests by behavioral pattern; \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

price discrimination. What happens, however, when we allow for some amount of information on the consumer side? It is likely that consumers often have some information about the rules by which interpersonal price differences are determined. Intuitively, one probable reason for the wide acceptance of student discounts is that students tend to have less money. Consumers might thus be more forgiving to the practice of downward price adjustments, because it decreases inequality. More generally, one might hypothesize that benevolent price discrimination becomes more accepted once consumers are aware who benefits, i.e. consumers with a lower income who otherwise would not be able to participate in the market. If true, transparency could be an efficient way of dealing with negative consumer reactions in the context of price discrimination. To test these assertions, the following studies extend our analysis of consumer behavior under BPD to situations where consumers know how the seller discriminates and who the beneficiaries are.

## 4 Study III

We extended the experimental design by two crucial features. First, we made the variable of discrimination, i.e. income, transparent to all consumers. Second, we introduced two different causes of income inequality. Endowments were either allocated randomly (*Random*) or based on the performance (*Effort*) in a real-effort slider task (Gill and Prowse 2012). Consumers only knew by which mechanism endowments were allocated, but not the exact endowment levels. That is, consumers in *Effort* knew that their performance would influence their endowment, but were not informed about the exact functional relationship. Consumers received either the high or low endowment based on a fixed performance threshold.<sup>9</sup>

High-income consumers in Study 3 knew that (i) low-income consumers existed, (ii) low-income consumers could not afford to buy a good for the initial uniform price, (iii) only low-income consumers benefited from price discrimination and (iv) high- and low-income consumers could afford the same amount of goods under price discrimination. Because price discounts do not negatively affect high-income consumer welfare, increase the number of consumers able to participate in the market and increase overall welfare, we expected relatively little switching away from the price

---

<sup>9</sup>The threshold was 31 sliders over three rounds of 60 seconds each. We followed a pretest with 100 participants and selected the cut-off for the 10th percentile.



discriminating store (*Avoid*). Likewise, we expected some high-income consumers to support the price discriminating store by switching towards it (*Approach*). We also expected high-income consumers in *Effort* to exhibit stronger aversions towards BPD than those in *Random*. That is because high-income consumers might feel entitled to higher payoffs when they are earned by greater effort in the slider task. BPD, in that sense, undermines a merit-based advantage.

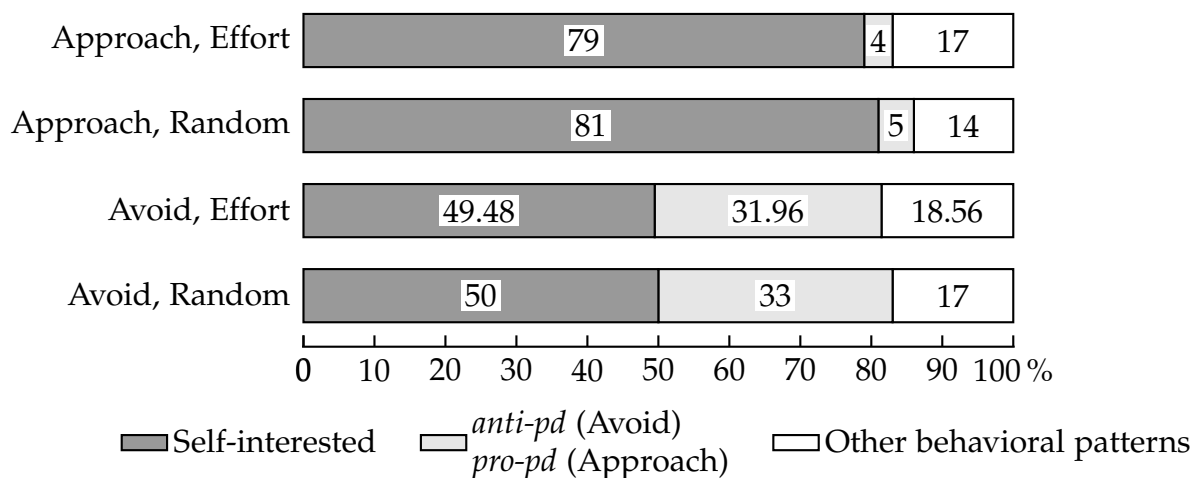
## 4.1 Experimental Design

The experiment consisted of six treatments. We gathered observations until we had 100 high-income consumers in each treatment with a fixed probability of 10% to become a low-income consumer in *Random*.

The basic procedure of the experiment mirrored Study 1. In *Random*, we added one page where participants learned about the random endowment mechanism. In *Effort*, participants first completed three rounds of an effort task based on sliders. After being informed about their endowment, consumers indicated their agreement to two statements on a seven point Likert-scale to validate our manipulation. In both treatments, we asked whether it would be fair if every consumer received the same endowment, and whether it was fair that (1) endowments were allocated randomly or (2) better performances in the slider task were rewarded with a higher endowment. Consumers then proceeded with the original instructions. Participants who completed the effort task but failed to answer the comprehension questions correctly were paid the base reward of \$1.20.

## 4.2 Results

After excluding all observations where a subject chose the outside option in the second round, we were left with 597 participants (47% female). Our results show that consumer aversions towards BPD persist under full transparency and are not affected by the cause of income inequality. Even with the information that price discrimination only benefits consumers who otherwise cannot participate in the market, a large share of consumers is willing to engage in costly, inefficient switching to punish the price-discriminating store. Similar to Study 1, almost no consumer costly supports BPD.



**Figure 2** — Behavioral patterns across the four conditions from Study 3.

### *One-Shot*

In *Effort*, consumers found a hypothetically equal distribution of endowments significantly more unfair than in *Random* (*Effort*: 2.52, *Random*: 6.42;  $t = 18.31$ ,  $p = 0.000$ ). They also found the mechanism by which differences in endowments were achieved significantly fairer (*Effort*: 6.37, *Random*: 4.22;  $t = -9.34$ ,  $p = 0.000$ ). This affirms that subjects in *Effort* perceived differences in income and thereby different purchasing abilities as more fair, presumably because they were rooted in merit.

Regarding store choices, consumers in the one-shot experiment still exhibited strong aversions against price discrimination, albeit weaker than in Study 1. In *Random*, almost 67% of consumers chose to purchase in the non-price-discriminating store. In *Effort*, the share was around 72%. The difference was not significant ( $t = 1.46$ ,  $p = 0.146$ ).

### *Repeated-Avoid*

Results for the manipulation check are in line with the ones from *One-Shot* and confirm the success of our intervention. In *Avoid*, we largely replicate the results from Study 1 for both *Effort* and *Random* (see Figure 2).

Surprisingly, we find no differences between price discrimination over merit-based compared to arbitrary income inequality. In both treatments, one third of consumers was willing to costly switch to a non-price-discriminating store. Note that this was with the information that only low-income consumers who otherwise were not able to participate in the market benefited from price discrimination, while nobody in-

curred any income loss. Consumers purchased significantly less goods in the price discriminating store after the introduction of price discrimination (*Effort*: 1.11 vs. 1.86,  $t = -7.61, p = 0.000$ ; *Random*: 1.11 vs. 1.88,  $t = -8.22, p = 0.000$ ). There was no significant treatment difference ( $t = -0.10, p = 0.923$ ) and a random effects logistic regression revealed no significant interaction between *Effort* and the introduction of BPD over the whole sample (see Table 9).

**Result 5** *Consumer aversions to BPD persist under full information about the VOD and the beneficiaries.*

**Result 6** *Consumers do not differentiate between BPD over merit-based inequality and BPD over arbitrary inequality.*

#### *Repeated-Approach*

In accordance with Study 1, almost no consumer was willing to costly support a benevolently price discriminating store. Instead, the majority of consumers exhibited self-interested behavior patterns (*Effort*: 79%, *Random*: 81%). Only four and five percent respectively followed the pro-pd pattern. Paired *t*-tests suggest that consumers did not significantly change their behavior once the other store introduced price discrimination (*Effort*: 0.17 vs. 0.11,  $t = 1.10, p = 0.275$ ; *Random*: 0.15 vs. 0.11,  $t = 0.68, p = 0.495$ ). There are no significant treatment differences.

#### *Consumer Attitudes*

Table 4 shows consumer attitudes of the whole sample as well as segmented by treatment. Differences in *Avoid* point in the expected direction, but are not significant. For *Approach*, despite equal switching behavior, subjects in *Random* perceived price discrimination as significantly fairer and ascribed more positive intentions to the manager of the price discriminating store. These results are in accordance with Studies 1 and 2, whereby feelings of unfairness and the ascription of intentions are not sufficient to explain consumer behavior under BPD.

### 4.3 Discussion

Consumer aversions towards BPD are not restricted to decision-making under scarce information. Our results indicate that a large share of consumers is willing to costly

**Table 4** — Consumer attitudes by behavioral patterns and experimental condition

<i>Experimental condition</i>	By experimental condition		
	<i>Effort</i>	<i>Random</i>	Diff.
Short description of questionnaire item	Mean (SD)	Mean (SD)	<i>t</i> -stat.
<i>Avoid</i>			
Fairness of PD <sup>1</sup>	3.71 (2.09)	4.07 (2.12)	-1.20
Fairness of punishment by switching <sup>2</sup>	5.05 (1.88)	4.79 (1.97)	0.95
Good intentions of PD manager <sup>3</sup>	4.33 (1.84)	4.38 (1.93)	-0.19
Exploitation by PD manager <sup>4</sup>	3.25 (1.99)	3.55 (2.10)	-1.04
SCO <sup>5</sup>	4.38 (1.29)	4.33 (1.42)	0.26
<i>Approach</i>			
Fairness of PD <sup>1</sup>	3.68 (1.93)	4.80 (2.08)	-3.95***
Fairness of switching to increase profit <sup>2</sup>	4.47 (1.80)	4.80 (1.85)	-1.28
Good intentions of PD manager <sup>3</sup>	4.17 (1.77)	5.18 (1.78)	-4.02***
Exploitation by PD manager <sup>4</sup>	3.27 (1.76)	2.79 (1.89)	1.86*
SCO <sup>5</sup>	4.54 (1.21)	4.31 (1.26)	1.33

*Note.* All questions were answered on a 7-point scale. Superscripts 3, 4: scale ranged from 1 (un-fair) to 7 (fair). Superscripts 1, 2, 5: scale ranged or from 1 (strongly disagree) to 7 (strongly agree); <sup>1</sup>"Please rate how fair you think it is that manager A/B decided to offer consumers who [*Effort*: earned a lower endowment in the slider task, *Random*: randomly received a lower endowment] lower prices for the same good.", <sup>2</sup>"The more goods a store sells, the higher the income of the participant acting as its manager. Do you consider it fair to switch from Store A/B to Store B/A in order to [*Avoid*: reduce, *Approach*: Increase] manager [*Avoid*: A/B's, *Approach*: B/A's] profit because of their pricing strategy?", <sup>3</sup>"Manager A/B's intention for introducing interpersonal price differences was good.", <sup>4</sup>"Manager A/B intended to take advantage of me (the consumer).", <sup>5</sup>Mean of Social comparison orientation scale by Gibbons and Buunk (1999); The reported *t*-statistic results from a two-sample mean-comparison *t*-tests by behavioral pattern; \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

switch away from a benevolently price-discriminating store, even when price discrimination only benefits low-income consumers. Given that it was, by design, impossible for low-income consumers to earn a higher total payment than high-income consumers, this behavior is largely inconsistent with inequality aversion and the concept of distributive fairness. Although we document a qualitative drop in *anti-pd* subjects compared to the first two studies, equalizing outcomes does not seem to substantially alleviate behavioral constraints towards BPD. As before, net consumer switching is clearly negative for the store that introduces price discrimination. What is more, consumers appear to disregard the cause of income inequality. In contrast to the existing literature on economic inequality and re-distributive preferences, consumer choice under BPD does not depend on whether differences in initial endowments are earned or arbitrary. One possible explanation might simply be that procedural fairness judgments, i.e. charging different prices for the same good, are much more important in determining consumer behavior than outcome-related reasoning. If consumers do not think about the distribution of economic outcomes but are instead purely focused on the process, it follows that the source of income inequality is a negligible factor. This would also explain why so many consumers are willing to give up a substantial portion of their endowment to decrease the income of a seller who did nothing to reduce their outcome.

An alternative explanation is that the top-down approach by which sellers impose differential pricing reduces both the moral responsibility and agency of consumers in causing more equal outcomes. Because in our setup the “future market” is unaffected by consumer decisions, and consumers do not decide whether low-income consumers receive discounts, they might exhibit patterns that are different from e.g. voting for re-distributive policies. To test this explanation, Study 4 extends our design by a second period to simulate agency through a market environment.

## 5 Study IV

We examine whether consumer aversions towards BPD as well as consumers’ disregard of arbitrary vs. merit-based inequality generalize to a market-analogous situation where consumers’ purchasing decisions affect a store’s future pricing strategy. By endowing current consumers with agency over the price-setting process for a later period, we induce externalities on their purchasing decisions and thus agency regard-

ing future consumer outcomes. Consumers are responsible for future overall welfare as well as the potential equalization of outcomes induced by BPD. Those who still switch away from the price discriminating store essentially pay to decrease seller and low-income consumer profits as well as to prevent a more equal outcome distribution.

## 5.1 Experimental Design

We implemented two (*Avoid* vs. *Approach*) 2x2 mixed factorial designs manipulating the source of income inequality (*Random* vs. *Effort*) as well as the agency of consumers regarding the price-discriminating store's pricing strategy in a future task (*Agency* vs. *No agency*). The *No agency* treatments were equivalent to the corresponding *Avoid* conditions in Study 3. For the agency treatments, we added one screen after the store's introduction of price discrimination. On this screen, subjects were informed that the discriminating store's sales in the remaining two rounds would determine its price-setting in a future HIT on MTurk. If sales exceeded a certain threshold, a rule would automatically implement the lower prices for low-income consumers in the last two rounds of a future setup-equivalent task. Otherwise, future low-income consumers would not be able to purchase the good at the reduced price. We thereby foreclosed any considerations regarding differing manager income levels or reverse outcome inequalities.<sup>10</sup>

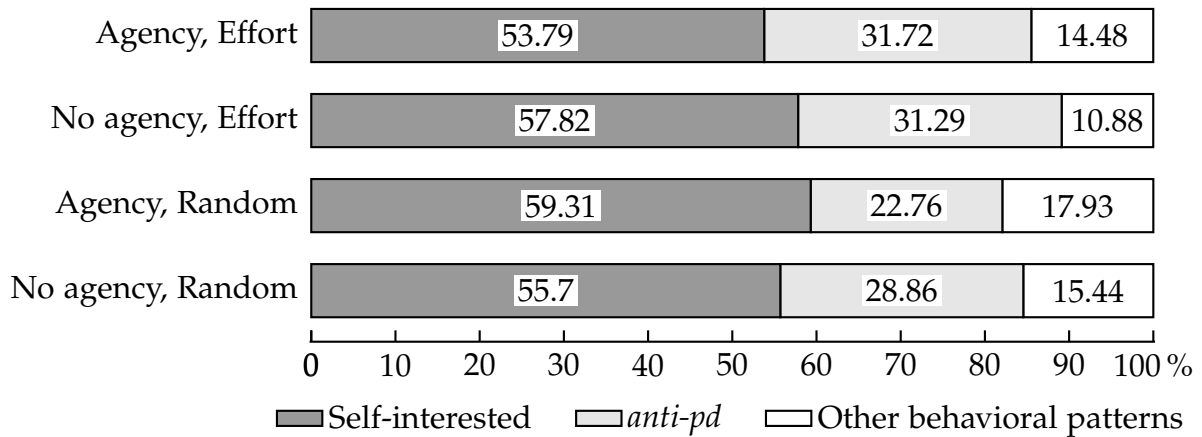
For both *Avoid* and *Approach*, we gathered 150 high-income consumer observations per treatment and set a fixed probability of two percent to become a low-income consumer in *Random*.

## 5.2 Results Avoid

After dropping subjects who chose the outside option in round two, the final sample consisted of 586 observations (49% female). Our results largely replicate Study 3, once again establishing the robustness of consumer aversions towards BPD, but also hint that under agency, consumers are more likely to reject BPD when income differences are rooted in merit.

---

<sup>10</sup>Specifying that low-income consumers would receive the discount only for two rounds ensures that it is impossible for them to earn more money in total than regular consumers.



**Figure 3** — Behavioral patterns across the four conditions from Study 4, *Avoid*

#### *Manipulation Check Avoid*

To check that subjects understood our manipulation, we extended the post-experimental questionnaire by a control question about the future impact of their purchase decisions, asking subjects to choose the correct statement out of four options. In total, 76 % of subjects answered the questions correctly. Since there could be multiple reasons why participants might answer the question incorrectly, we do not drop observations based on the manipulation check. Instead, we will use the results from the sub-sample of subjects who correctly answered the question as a robustness check.

#### *Main Results Avoid*

In line with Study 3, 58 % (56 %) of subjects in the *No agency* treatments of *Effort* (*Random*) decided to stay in the same store for the whole experiment. Similarly, 31 % and 29 % respectively exhibited strictly anti price-discrimination behavioral patterns by switching after the introduction of discounted prices. Endowing subjects with agency regarding future consumer welfare does not appear to fundamentally shift behavior (see Figure 3).

In *Agency, Effort*, 32 % are categorized as *anti-pd*. In *Random*, that share drops to 23 %. Contrary to the *No agency* treatments, we document a significant difference in the drop of sales conditional on the source of endowment differences. In *Random*, the store introducing BPD experienced a significantly smaller decline in demand (*Effort*: -0.72, *Random*: -0.54;  $t = -1.56$ ,  $p = 0.06$ ), which is accentuated when restricting the sample to subjects who correctly answered the control question ( $t = -1.89$ ,  $p = 0.03$ ). The

regression model replicates this effect across the whole sample through a significant interaction between the *Effort*-Dummy and the BPD-Dummy. Further and in line with the descriptive results, agency and benevolent price discrimination are positively – but not significantly – correlated in predicting purchases at the price-discriminating store (see Table 9).

Overall, there is moderate evidence that under agency, the source of endowment differences becomes more meaningful and consumers tend to switch less. However, differences are small and only marginally significant, depending on the sample.

**Result 7** *Consumer aversions to BPD persist when consumer choices determine future prices.*

As expected given the procedural nature of the intervention, subject answers in the post-experimental questionnaire did not differ substantially (see Table 6 in the appendix).

### 5.3 Results Approach

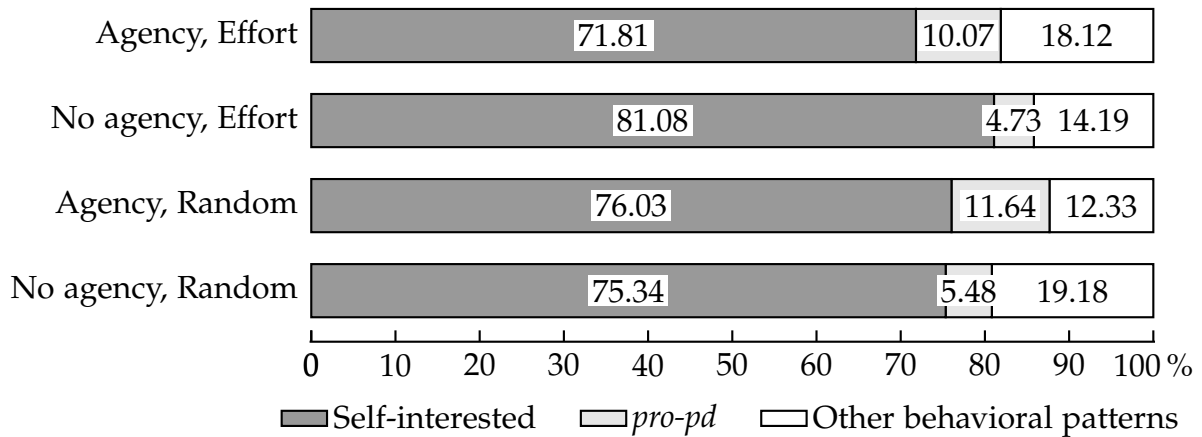
Our final sample consisted of 589 (50 % female) observations. 76 % of subjects answered the control question correctly. In *No agency*, consumers generally do not support BPD by switching towards the store. Once consumers are endowed with some agency, the price-discriminating store sells significantly more goods in rounds three and four. There are no differences between *Effort* and *Random*.

#### *Main Results Approach*

The *No agency* treatments closely replicate our results from Study 3 (see Figure 4).

The majority of consumers maximized their income by staying in the same store for all four rounds and the price-discriminating store sold roughly the same amount before and after the introduction of discounts for low-income consumers (*Effort*:  $t = 0.29, p = 0.769$ ; *Random*:  $t = 0.27, p = 0.787$ ). Once consumers were endowed with some agency regarding future store prices, their purchasing behavior changed significantly. The share of subjects classified as *pro-pd* roughly doubled with 10 % in *Effort-Agency* and 12 % in *Random-Agency*. In *Effort-Agency*, the price discriminating store sold 0.09 goods on average over the first two rounds, and 0.30 goods over rounds three and four – a significant increase ( $t = 3.89, p = 0.000$ ). The results for *Random-Agency* are similar (0.08 vs. 0.32;  $t = 4.04, p = 0.000$ ) and there are no significant





**Figure 4** — Behavioral patterns across the four conditions from Study 4, *Approach*

treatment differences between the two ( $t = 0.25, p = 0.802$ ). The seller introducing BPD after the second round sold significantly more in *Effort-Agency* than in *Effort-No agency* ( $t = -2.94, p = 0.013$ ) and more in *Random-Agency* than in *Random-No agency* ( $t = -2.21, p = 0.028$ ). Concurrently, the regression model reveals a significant interaction between BPD and the *Agency* treatment dummy on store choice. The source of endowment does not predict switching.

**Result 8** *Endowing consumers with agency over future prices increases costly support for BPD.*

Differences in self-reported consumer attitudes between agency treatments are again minimal (see Table 7 in the appendix).

## 5.4 Discussion

The results show that a salient market mechanism, which endows consumers with agency over future prices for low-income consumers significantly increases costly support of BPD, while aversions persist. A large share of consumers continues to switch away from the price discriminating store, and there is only little evidence that costly switching might decrease as a function of consumer influence. However, we do find a relatively large effect of agency on consumer switching *towards* the price discriminating store. This effect holds irrespective of the cause for income differences.

We think that there are two main reasons for the increased acceptance and support of BPD. First, some consumers might feel morally obligated to accept or support lower

prices for poor consumers once their actions determine future outcomes. Second, introducing a second period allows (or forces) consumers to ascribe themselves some responsibility for changing the outcome of low income consumers. Thus, introducing agency increases the significance of outcome-related reasoning, as consumers are forced to take the future impact of their purchase decisions into account. We deduce that greater consumer agency increases the importance of outcome equalization and distributive fairness for judging BPD, consequently leading to more support and potentially less aversion. The more differences in income are based on merit, the higher the behavioral constraints on BPD.

## 6 Study V

Study 5 combines three robustness checks to our main results. First, we increased the salience of low-income consumers and disclosed their endowment of 50 Coins per round to all consumers. Before, consumers only knew that low-income consumers received a lower endowment, they did not know their exact endowment. This information does not change any behavioral predictions regarding distributive fairness and outcome equalization. However, it made it harder for high-income consumers to e.g. strategically ignore that low-income consumers could not purchase anything before the introduction of BPD. It also highlighted that price discrimination merely led to a convergence in payoff – low-income consumers remained considerably poorer than high-income consumers. Moreover, while consumers were treated unequally with respect to absolute prices, they were treated equally in that everybody was priced relative to their endowment.

Second, we varied switching costs to assess how sensitive consumer choices were towards the ease of switching. In reality, this could for instance be determined by the purchase environment (online vs. analog), the level of market competition or the availability of substitutes.

Third, we summarize and analyse text data from an open-ended question on consumer behavior to demonstrate data quality, confirm the validity of our behavioral labels and highlight some alternative consumer reasoning. We show that consumer animosity towards BPD exceeds quantifiable switching rates and that economic self-interest disciplines inefficient consumer reactions.

## 6.1 Robustness Check: Salience

We implemented a 2 (*Approach* vs. *Avoid*) × 2 (*Effort* vs. *Random*) mixed-factorial design that replicated the agency-treatments from Study 4 with additional information on the income of low-income consumers as well as their ability to purchase exactly one good per round after the introduction of price discounts. In such a framework, we would expect maximal impact of outcome-related fairness concerns. We collected 150 independent observations per treatment. Dropping all subjects who chose the outside option in the second round left us with 592 (51 % female) total subjects.

### *Results*

Results closely follow the findings from Study 4 (see Table 5). Consumer still exhibited substantial aversions towards BPD while the majority followed a self-interested behavioral pattern. There was both significant consumer emigration from (*Avoid*), as well as migration to (*Approach*) the price discriminating store (see also Tables 9 and 8 in the appendix). However, differences for average number of goods sold in *Avoid* (*Effort*:  $-0.67$ ; *Random*:  $-0.53$ ) are much larger than in *Approach* (*Effort*:  $0.31$ ; *Random*:  $0.25$ ). Hence, the price discriminating store still experienced a net loss of consumers. Looking only at the 78% who correctly answered the control question does not change the results. We thus affirm that our main results are not dependent on subjects' failure to understand the implications for low-income consumers or a wrong attribution of differential procedural treatment between consumers regarding the variable of discrimination.

## 6.2 Robustness Check: Switching Costs

To assess the influence of switching costs on purchasing behavior, we opted for the *Avoid* paradigm where subjects earned their endowment through the effort task and had agency over the store's future prices. We implemented three treatments with switching costs of either 5, 15 or 30 Coins. From 125 independent observations per treatment, we dropped 8, leaving us with 367 independent observations in total (54 % female).

### Results

Results reveal moderate evidence for downward elasticity of consumer BPD aversions. Increasing switching costs does not affect behavior (see Tables 5 and 9). For switching costs of 5 Coins (5% of a participant's endowment), the drop in average goods sold after the introduction of price discounts was significantly higher than for switching costs of 15 ( $-0.91$  vs.  $-0.70$ ;  $t = -1.71$ ,  $p = 0.045$ ) and weakly significantly higher than for switching costs of 30 Coins ( $-0.91$  vs.  $-0.74$ ;  $t = -1.40$ ,  $p = 0.082$ ). There was no difference between 15 and 30 Coins ( $t = 0.30$ ,  $p = 0.619$ ).<sup>11</sup> Finally, the regression analysis reveals no significant treatment dummy (Table 9), confirming that consumer aversions towards BPD are relatively inelastic to changes in switching costs.

**Table 5** — Behavioral patterns from the robustness checks

<i>Robustness check</i> Treatment	Behavioral pattern (%)	
	<i>anti-pd in Avoid</i> <i>pro-pd in Approach</i>	<i>self-interested</i>
<i>Salience</i>		
Avoid, Random	19.59	62.16
Approach, Random	10.00	65.33
Avoid, Effort	28.57	58.50
Approach, Effort	14.97	76.19
<i>Switching costs</i>		
5 Coins	38.21	43.09
15 Coins	31.40	53.72
30 Coins	33.33	51.22

*Note.* Results are consistent with the main studies. Aversions towards BPD are substantially stronger than respective support, and behavior is relatively inelastic to changes in switching costs.

### 6.3 Robustness Check: Consumer Reasoning

The post-experimental questionnaire of Studies 2, 4, and 5 included an open-ended question that allowed participants to explain why they “did (not) decide to switch

<sup>11</sup>Restricting the sample to subjects who correctly answered the control question about consumer purchases affecting store prices in a future HIT (75%) amplifies these results for switching costs of 15 (5:  $-0.82$  vs. 15:  $-0.51$ ;  $t = -2.30$ ,  $p = 0.011$ ), but not for switching costs of 30 (5:  $-0.82$  vs. 30:  $-0.65$ ;  $t = -1.23$ ,  $p = 0.110$ ).

stores after [PD-Store] offered some consumers lower prices for the same good” in their own words.<sup>12</sup>

First, the data largely confirmed the authenticity of subject feedback and thus validates the quality of our data. A large majority of participants offered straightforward reasoning for their choices which demonstrated their understanding of the task, the payoff scheme and the actors involved. Each participant wrote an average of 114 characters, and across all considered studies, only 136 out of 3229 stated reasons (4%) were independently judged to be nonsensical.<sup>13</sup>

Second, subject answers confirmed that our interpretation of the two predominant behavioral patterns as *self-interested* and *anti-pd/pro-pd* was overwhelmingly in line with subject reasoning. From 590 identified reasons that were stated by subjects categorized as *anti-pd* in *Avoid*, the majority (334) explicitly stated that they switched because they were snubbed by the price-discriminating store. The second most cited reason was that consumers wanted to punish the manager who introduced price discounts (73). For the 1160 reasons from subjects labeled as *self-interested* in *Avoid*, the two main recurring explanations were switching costs (444) and a lack of economic benefits to switching (389). Thus, it appears that economic considerations disciplined many consumers to abstain from inefficient switching. The analysis further highlighted that a significant portion of consumers “silently” supported BPD (165), which could not be captured by the *Avoid* paradigm. All these 165 subjects participated in studies where consumers were able to influence the price setting in a future market and knew about low-income consumers as well as income as the variable of discrimination (Study 4 and 5).

For *Approach*, the most prominent reason by subjects classified as *pro-pd* (104 identified reasons) was support for the price-discriminating store (51). Beyond that, 14 answers mentioned their agency over future prices as well as the option to equally split manager profits over the four rounds. Consumers who stayed in the same store over all rounds (880 reasons) predominantly mentioned that there were either no benefits to switching (459), that switching was costly (225) or that they felt loyalty towards the other store (66).

---

<sup>12</sup>To draw conclusions from these data, we opted for an inductive and manual coding scheme to identify different response labels (coding framework). Further, we used the independent coder method to heighten the reliability of our coding. Results from three coders were compared for consistency and amended by the authors if needed.

<sup>13</sup>We extracted up to two reason from each subject’s answer. In case of two reasons, each received equal weight in the analysis. For the full data set, please refer to the web appendix.

Overall, the data supports our prior interpretations while underlining the importance of some crucial design choices like switching costs and the consideration of both switching directions.

## 6.4 Demographics

Table 10 summarizes the results of a pooled ordered logistic regression containing observations from all studies segregated by *Avoid* and *Approach*. Because differences in the first two rounds were negligible, we concentrate on behavior after the introduction of BPD in rounds three and four. Results show that BPD attitudes are mostly independent from specific demographic variables. The main outliers are sex and individuals with very high education. Subjects who identified as female were significantly more likely to switch away from the price-discrimination store, whereas there is no difference for *Approach*. This is coherent with recent evidence suggesting that compared to men, women are more likely to be inequity averters that prefer egalitarian pay, but less likely to be surplus maximizers (Kamas and Preston 2015). The fact that there is no such effect for *Approach* underscores that the decision to (not) costly reject something is distinctly different from the decision to (not) costly support something. Apart from that, reportedly out-of-work subjects as well as those with particularly high education were more likely to stay within the same store and thereby maximize their payoff.

## 7 General Discussion

In this paper, we introduced the concept of *benevolent price discrimination* as a downward-bound implementation of differential pricing that always benefits financially disadvantaged groups or individuals and thus leads to more equal economic outcomes while no consumer incurs any losses. Across five studies, we show that a large share of consumers rejects BPD and pays to switch to a competitor that treats all consumers equally. Effects are the strongest when the benevolent nature of price discounts is obfuscated. However, even when consumers know that everybody is priced according to the same rule, and that differences in endowment are purely due to random chance, the main results hold. This behavior is not driven by reciprocity, intentions, a lack of transparency or a perception of merited outcome inequality. Costly support for BPD is rare, and largely contingent on high levels of transparency as well as salient agency over a store's future prices. While consumers do move bidirectionally, net

migration for the BPD-store is consistently negative. Thus, distributive fairness of realized economic outcomes appears to play a relatively small role in determining consumer reactions towards differential pricing.

Our results partially contrast well-researched phenomena from the literature on social preferences and inequity aversion. First, there is good evidence to suggest that consumers should be expected to support BPD as a less invasive analog to other re-distributive policies. Second, many consumers do not differentiate whether pricing equalizes outcome differences based on luck or based on merit, despite being similarly salient as in other studies on social preferences relating to e.g. economic redistribution (Alesina and Giuliano 2011), inequity aversion (Engelmann and Strobel 2004), the fair process effect (Van den Bos, Vermunt, et al. 1997) or ultimatum bargaining (Camerer 2011). Under price discrimination, effects on the overall distribution of economic welfare appear to be uniquely minuscule. Our results support the interpretation that this is partially due to consumers being second movers and thus unable to ascribe themselves agency. The introduction of a second period increases costly support for BPD, while decreasing consumer aversion if income is allocated by a random mechanism as opposed to an effort task. Hence, caring about other consumer's payoff appears to be partially dependent on one's own direct influence. Put differently, second-mover social preferences differ from first-mover social preferences.

One reason why price discrimination might elicit unique counter reaction lies in people's general aversion towards demand-driven pricing changes. We think that many people neither judge prices, nor price-setting as something endogenously determined by supply and demand, but rather something that should reflect the costs of a good. This has also been proposed by Kahneman et al. (1986) in their seminal work on price fairness perceptions. Setting different prices for the same good might be viewed as uniquely undesirable by relatively adversely affected consumers, since (i) there is no objective reason why prices should differ in any way and (ii) price discounts signal that the "true" value of a product lies below what has been charged before.

### *Implications*

While some studies have found that obfuscating interpersonal price differences reduces consumer fairness concerns and thereby increases sellers' pricing power (see e.g. Allender et al. 2021), other researchers have suggested overt transparency to signal benevolence if price discrimination is based on consumer income (Rotemberg 2011).

Our results indicate that the latter is, by and large, not necessarily a good strategy for sellers. At the very least, firms should expect a drop in demand by current consumers, which could become quite large and almost certainly exceeds the attraction of new consumers who would not benefit from price discounts. Given that price discrimination in our study is downward-bound, further segmentation by increasing prices would presumably culminate in even larger counter reactions. Instead, our findings are more in line with results from Li and Jain (2016). Firms that are expecting consumer fairness concerns reduce differential pricing, e.g. low poaching prices, which reduces inefficient switching and consumer disutility from perceived price unfairness, but hurts overall consumer surplus. Thus, consumer reactions induce constraints for price interventions that increase market participation as well as joint welfare. We conclude that seller hesitation to deviate from uniform pricing appears to be a well-adapted strategic reaction towards substantial behavioral constraints that cannot be alleviated by equalizing unequal outcome distributions and transparently targeting low-income consumers.

#### *Limitations and Future Research*

Our results support the conjunction that contextual and procedural elements inherent to the price setting framework play a considerable role for consumer behavior and, in our setting, clearly outweigh outcome-related reasoning. Nevertheless, contextual and procedural elements may naturally differ between markets or even be strategically influenced by sellers, possibly tipping the scales in favor of outcome-related reasoning. For instance, our results cannot explain the widely observed acceptance of price discounts for specific low-income groups. Since the experiments were explicitly designed to be context-neutral, effects elicited by e.g. social norms, cultural norms, inertia or habituation are beyond the scope of this paper. They might, however, moderate the influence of outcome reasoning on consumer behavior. It is also possible that people are more accepting of different prices for population groups they (i) have been part of in the past (e.g. people below 25) or (ii) will be part of in the future (e.g. people above 65).

One explanation for our results may be that consumers expect prices to correlate with costs, which is violated by demand-based price discrimination. If that is the case, the introduction of price *increases* for high-income consumers might counter-intuitively *reduce* behavioral constraints to price discrimination, because they signal a



potential necessity to compensate price discounts, rather than relatively large profit margins at the original price. One hidden implication is that consumers could react differently to price discrimination over physical products than price discrimination over services. Presumably, it is much harder to define the “true” value of a service than that of a product, as the former’s cost function more saliently captures intangibles like opportunity costs, rather than costs of production. This question, however, requires a less abstract approach to the quantification of consumer behavior.

Finally, we were able to identify consumers’ agency over prices and thereby outcomes, as one relevant procedural element that interacts with outcome-driven behavioral changes. Although consumers’ agency was already quite salient in our study, there may be other interventions that are more effective in increasing perceptions of agency. First, a logical extension of our setup is to implement a multi-round dynamic market environment in which consumers can “experience” their agency over market outcomes. Other interventions may increase agency by involving consumers more directly in the price setting process, e.g. by some form of voluntary extra payment or “Pay-as-You-Wish” pricing (Chen et al. 2017). Here, we hope that our results help future researchers in exploring innovative pricing mechanisms that better fit the social and economic preferences of consumers while avoiding the pitfalls of antagonizing consumers against one another. Consumer social and fairness preferences are a demonstrably complicated phenomenon, influenced by a multitude of procedural and distributive variables, that often translate into non-obvious behavioral patterns. Without careful considerations, instruments of differential pricing might provoke consumer reactions with at least ambiguous, if not outright negative welfare effects. An increasing public recognition of price discrimination could be detrimental to consumer welfare, which should caution policy calls for more transparency.

## **Acknowledgments**

The authors gratefully acknowledge support from the “Ministry of Science and Arts” of Lower Saxony (Germany), the Volkswagen Foundation (grant number ZN3492, “Vorab”), and the Center for Digital Innovations (ZDIN).

## References

- Alesina, A. and G.-M. Angeletos (2005). "Fairness and Redistribution". In: *American Economic Review* 95.4, pp. 960–980.
- Alesina, A. and P. Giuliano (2011). "Preferences for redistribution". In: *Handbook of social economics*. Vol. 1. Elsevier, pp. 93–131.
- Alexopoulos, J., D. Pfabigan, F. Göschl, H. Bauer, and F. Fischmeister (2013). "Agency matters! Social preferences in the three-person ultimatum game". In: *Frontiers in Human Neuroscience* 7.312, pp. 1–10.
- Allender, W. J., J. Liaukonyte, S. Nasser, and T. J. Richards (2021). "Price Fairness and Strategic Obfuscation". In: *Marketing Science* 40.1, pp. 122–146.
- Anderson, E. T. and D. I. Simester (2008). "Research note—does demand fall when customers perceive that prices are unfair? The case of premium pricing for large sizes". In: *Marketing Science* 27.3, pp. 492–500.
- (2010). "Price stickiness and customer antagonism". In: *The quarterly journal of economics* 125.2, pp. 729–765.
- Bayer, R.-C. (2010). "Intertemporal price discrimination and competition". In: *Journal of Economic Behavior & Organization* 73.2, pp. 273–293.
- Blount, S. (1995). "When social outcomes aren't fair: The effect of causal attributions on preferences". In: *Organizational behavior and human decision processes* 63.2, pp. 131–144.
- Bolton, G. E., J. Brandts, and A. Ockenfels (2005). "Fair Procedures: Evidence from Games Involving Lotteries". In: *The Economic Journal* 115.506, pp. 1054–1076.
- Bolton, G. E. and A. Ockenfels (2000). "ERC: A Theory of Equity, Reciprocity, and Competition". In: *American Economic Review* 90.1, pp. 166–193.
- Bolton, L. E., H. T. Keh, and J. W. Alba (2010). "How Do Price Fairness Perceptions Differ across Culture?" In: *Journal of Marketing Research* 47.3, pp. 564–576.
- Camerer, C. F. (2011). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.
- Campbell, M. C. (1999). "Perceptions of Price Unfairness: Antecedents and Consequences". In: *Journal of Marketing Research* 36.2, pp. 187–199.
- Chen, D. L., M. Schonger, and C. Wickens (2016). "oTree—An open-source platform for laboratory, online, and field experiments". In: *Journal of Behavioral and Experimental Finance* 9, pp. 88–97.

- Chen, Y., O. Koenigsberg, and Z. J. Zhang (2017). "Pay-as-You-Wish Pricing". In: *Marketing Science* 36.5, pp. 780–791.
- Choshen-Hillel, S. and I. Yaniv (2011). "Agency and the construction of social preference: Between inequality aversion and prosocial behavior." In: *Journal of personality and social psychology* 101.6, pp. 1253–1261.
- Coase, R. H. (1972). "Durability and monopoly". In: *The Journal of Law and Economics* 15.1, pp. 143–149.
- Dholakia, U. M. and I. Simonson (2005). "The effect of explicit reference points on consumer choice and online bidding behavior". In: *Marketing Science* 24.2, pp. 206–217.
- Durante, R., L. Putterman, and J. van der Weele (2014). "Preferences for Redistribution and Perception of Fairness: An Experimental Study". In: *Journal of the European Economic Association* 12.4, pp. 1059–1086.
- Engelmann, D. and M. Strobel (2004). "Inequality aversion, efficiency, and maximin preferences in simple distribution experiments". In: *American economic review* 94.4, pp. 857–869.
- Fabiani, S., M. Druant, I. Hernando, C. Kwapil, B. Landau, C. Loupias, et al. (2006). "What Firms' Surveys Tell Us about Price-Setting Behavior in the Euro Area". In: *International Journal of Central Banking*.
- Falk, A., E. Fehr, and U. Fischbacher (2008). "Testing theories of fairness — Intentions matter". In: *Games and Economic Behavior* 62.1, pp. 287–303.
- Fehr, E. and K. M. Schmidt (1999). "A Theory of Fairness, Competition, and Cooperation". In: *The Quarterly Journal of Economics* 114.3, pp. 817–868.
- Fehr, E. and S. Gächter (2000). "Fairness and Retaliation: The Economics of Reciprocity". In: *Journal of Economic Perspectives* 14.3, pp. 159–181.
- Fehr, E., G. Kirchsteiger, and A. Riedl (1993). "Does fairness prevent market clearing? An experimental investigation". In: *The quarterly journal of economics* 108.2, pp. 437–459.
- Ferguson, J. L., P. S. Ellen, and W. O. Bearden (2014). "Procedural and Distributive Fairness: Determinants of Overall Price Fairness". In: *Journal of Business Ethics* 121.2, pp. 217–231.
- Fiske, S. T. and S. E. Taylor (2013). *Social cognition: From brains to culture*. Sage.
- Folger, R. and M. A. Konovsky (1989). "Effects of procedural and distributive justice on reactions to pay raise decisions". In: *Academy of Management journal* 32.1, pp. 115–130.

- Fudenberg, D. and J. M. Villas-Boas (2012). "Price discrimination in the digital economy." In: *The Oxford handbook of the digital economy*, pp. 254–272.
- Garbarino, E. and S. Maxwell (2010). "Consumer response to norm-breaking pricing events in e-commerce". In: *Journal of Business Research* 63.9-10, pp. 1066–1072.
- Gibbons, F. X. and B. P. Buunk (1999). "Individual differences in social comparison: Development of a scale of social comparison orientation." In: *Journal of Personality and Social Psychology* 76.1, pp. 129–142.
- Gill, D. and V. Prowse (2012). "A Structural Analysis of Disappointment Aversion in a Real Effort Competition". In: *American Economic Review* 102.1, pp. 469–503.
- Guo, L. (2015). "Inequity aversion and fair selling". In: *Journal of Marketing Research* 52.1, pp. 77–89.
- Harrington Jr, J. E., R. H. Gonzalez, and P. Kujal (2016). "The relative efficacy of price announcements and express communication for collusion: Experimental findings". In: *Journal of Economic Behavior & Organization* 128, pp. 251–264.
- Haws, K. L. and W. O. Bearden (2006). "Dynamic Pricing and Consumer Fairness Perceptions". In: *Journal of Consumer Research* 33.3, pp. 304–311.
- Heider, F. (1982). *The psychology of interpersonal relations*. Psychology Press.
- Herz, H. and D. Taubinsky (2018). "What makes a price fair? An experimental study of transaction experience and endogenous fairness views". In: *Journal of the European Economic Association* 16.2, pp. 316–352.
- Ho, T.-H. and X. Su (2009). "Peer-Induced Fairness in Games". In: *American Economic Review* 99.5, pp. 2022–2049.
- Ho, T.-H., X. Su, and Y. Wu (2014). "Distributional and Peer-Induced Fairness in Supply Chain Contract Design". In: *Production and Operations Management* 23.2, pp. 161–175.
- Janiszewski, C. and D. R. Lichtenstein (1999). "A range theory account of price perception". In: *Journal of Consumer Research* 25.4, pp. 353–368.
- Jin, L., Y. He, and Y. Zhang (2014). "How Power States Influence Consumers' Perceptions of Price Unfairness". In: *Journal of Consumer Research* 40.5, pp. 818–833.
- Kahneman, D., J. L. Knetsch, and R. Thaler (1986). "Fairness as a constraint on profit seeking: Entitlements in the market". In: *The American economic review*, pp. 728–741.
- Kamas, L. and A. Preston (2015). "Can social preferences explain gender differences in economic behavior?" In: *Journal of Economic Behavior & Organization* 116, pp. 525–539.
- Koszegi, B. and M. Rabin (2006). "A Model of Reference-Dependent Preferences". In: *The Quarterly Journal of Economics* 121.4, pp. 1133–1165.

- Krawczyk, M. and F. Le Lec (2010). “‘Give me a chance!’ An experiment in social decision under risk”. In: *Experimental economics* 13.4, pp. 500–511.
- Kuo, A., D. H. Rice, and P. Fennell (2016). “How fitting! The influence of fence-context fit on price discrimination fairness”. In: *Journal of Business Research* 69.8, pp. 2634–2640.
- Leibbrandt, A. (2020). “Behavioral constraints on price discrimination: Experimental evidence on pricing and customer antagonism”. In: *European Economic Review* 121, p. 103303.
- Lerner, M. J. and L. A. Whitehead (1980). “Procedural justice viewed in the context of justice motive theory”. In: *Justice and social interaction: Experimental and theoretical contributions from psychological research*, pp. 219–256.
- Li, K. J. and S. Jain (2016). “Behavior-based pricing: An analysis of the impact of peer-induced fairness”. In: *Management Science* 62.9, pp. 2705–2721.
- Maxwell, S. (2007). *The price is wrong: Understanding what makes a price seem fair and the true cost of unfair pricing*. John Wiley & Sons.
- Mayhew, G. E. and R. S. Winer (1992). “An empirical analysis of internal and external reference prices using scanner data”. In: *Journal of consumer Research* 19.1, pp. 62–70.
- Nilssen, T. (1992). “Two kinds of consumer switching costs”. In: *The RAND Journal of Economics*, pp. 579–589.
- Pigors, M. and B. Rockenbach (2016). “Consumer Social Responsibility”. In: *Management Science* 62.11, pp. 3123–3137.
- Rotemberg, J. J. (2011). “Fair Pricing”. In: *Journal of the European Economic Association* 9.5, pp. 952–981.
- Rutström, E. and M. B. Williams (2000). “Entitlements and fairness: an experimental study of distributive preferences”. In: *Journal of Economic Behavior & Organization* 43.1, pp. 75–89.
- Rutte, C. G. and D. M. Messick (1995). “An integrated model of perceived unfairness in organizations”. In: *Social Justice Research* 8.3, pp. 239–261.
- Shor, M. and R. L. Oliver (2006). “Price discrimination through online couponing: Impact on likelihood of purchase and profitability”. In: *Journal of Economic Psychology* 27.3, pp. 423–440.
- Thaler, R. (1985). “Mental accounting and consumer choice”. In: *Marketing science* 4.3, pp. 199–214.
- Trautmann, S. T. (2009). “A tractable model of process fairness under risk”. In: *Journal of Economic Psychology* 30.5, pp. 803–813.

- Van den Bos, K., E. A. Lind, R. Vermunt, and H. A. Wilke (1997). "How do I judge my outcome when I do not know the outcome of others? The psychology of the fair process effect." In: *Journal of personality and social psychology* 72.5, pp. 1034–1046.
- Van den Bos, K., R. Vermunt, and H. A. Wilke (1997). "Procedural and distributive justice: What is fair depends more on what comes first than on what comes next." In: *Journal of Personality and Social Psychology* 72.1, pp. 95–104.
- Wu, C.-C., Y.-F. Liu, Y.-J. Chen, and C.-J. Wang (2012). "Consumer responses to price discrimination: Discriminating bases, inequality status, and information disclosure timing influences". In: *Journal of Business Research* 65.1, pp. 106–116.
- Xia, L., K. B. Monroe, and J. L. Cox (2004). "The Price is Unfair! A Conceptual Framework of Price Fairness Perceptions". In: *Journal of Marketing* 68.4, pp. 1–15.

## Appendix

**Table 6** — Consumer attitudes by experimental condition (Study 4, *Avoid*)

<i>Experimental condition</i>	By experimental condition		
	<i>Agency</i>	<i>No agency</i>	<i>Diff.</i>
Short description of questionnaire item	Mean (SD)	Mean (SD)	<i>t</i> -stat.
<i>Effort</i>			
Fairness of PD <sup>1</sup>	3.79 (1.96)	3.41 (1.90)	1.68*
Fairness of punishment by switching <sup>2</sup>	5.21 (1.68)	5.03 (1.87)	0.89
Good intentions of PD manager <sup>3</sup>	4.24 (1.78)	4.19 (1.97)	0.23
Bad intentions of PD manager <sup>4</sup>	3.65 (1.84)	3.75 (2.04)	-0.44
Exploitation by PD manager <sup>5</sup>	3.66 (2.00)	3.47 (2.08)	0.81
Feel exploited by PD manager <sup>6</sup>	3.82 (2.07)	3.90 (2.07)	-0.35
SCO <sup>7</sup>	4.52 (1.24)	4.47 (1.11)	0.34
<i>Random</i>			
Fairness of PD <sup>1</sup>	4.26 (1.88)	4.26 (2.07)	0.00
Fairness of punishment by switching <sup>2</sup>	4.78 (1.81)	4.91 (1.97)	-0.57
Good intentions of PD manager <sup>3</sup>	4.62 (1.75)	4.75 (1.80)	-0.63
Bad intentions of PD manager <sup>4</sup>	3.30 (1.83)	3.09 (1.91)	0.99
Exploitation by PD manager <sup>5</sup>	3.55 (1.95)	3.26 (2.09)	1.26
Feel exploited by PD manager <sup>6</sup>	3.72 (1.99)	3.68 (2.18)	0.16
SCO <sup>7</sup>	4.70 (1.05)	4.33 (1.19)	2.81***

*Note.* All questions were answered on a 7-point scale. Superscript 1: 1 (unfair) to 7 (fair). Superscripts 2–4: 1 (strongly disagree) to 7 (strongly agree). Superscript 5: 1 (not at all) to 7 (very much); <sup>1</sup>“Please rate how fair you think it is that manager A/B decided to offer consumers who [*Effort*: earned a lower endowment in the slider task, *Random*: randomly received a lower endowment] lower prices for the same good.”, <sup>2</sup>“The more goods a store sells, the higher the income of the participant acting as its manager. Do you consider it fair to switch from Store A/B to Store B/A in order to reduce manager A/B’s profit because of their pricing strategy?”, <sup>3</sup>“Manager A/B’s intention for [*Intention*: introducing interpersonal price differences; *No intention*: choosing the pricing algorithm] was good.”, <sup>4</sup>Wording is equal to row 2, except last word is “bad”, <sup>5</sup>“Manager A/B intended to take advantage of me (the consumer).”, <sup>6</sup>“The fact that [*Intention*: Manager A/B, *No intention*: Store A/B’s algorithm] decided to offer some consumers lower prices makes me feel taken advantage of.”, <sup>7</sup>Mean of Social comparison orientation scale by Gibbons and Buunk (1999); The reported *t*-statistic results from a two-sample mean-comparison *t*-tests by behavioral pattern; \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

**Table 7** — Consumer attitudes by experimental condition (Study 4, *Approach*)

<i>Experimental condition</i> Short description of questionnaire item	By experimental condition		
	<i>Agency</i> Mean (SD)	<i>No agency</i> Mean (SD)	Diff. <i>t</i> -stat.
<i>Effort</i>			
Fairness of PD <sup>1</sup>	3.94 (2.02)	4.03 (1.88)	-0.42
Fairness of switching to increase profit <sup>2</sup>	4.71 (1.65)	4.55 (1.79)	0.82
Good intentions of PD manager <sup>3</sup>	4.50 (1.66)	4.57 (1.66)	-0.40
Bad intentions of PD manager <sup>4</sup>	3.52 (1.84)	3.28 (1.71)	1.16
Exploitation by PD manager <sup>5</sup>	3.32 (1.82)	2.93 (1.79)	1.86*
Feel exploited by PD manager <sup>6</sup>	3.91 (2.02)	3.25 (1.91)	2.90***
SCO <sup>7</sup>	4.67 (1.08)	4.63 (1.11)	0.29
<i>Random</i>			
Fairness of PD <sup>1</sup>	4.42 (1.84)	4.66 (1.85)	-1.11
Fairness of switching to increase profit <sup>2</sup>	4.66 (1.77)	4.50 (1.69)	0.78
Good intentions of PD manager <sup>3</sup>	4.79 (1.55)	4.93 (1.60)	-0.78
Bad intentions of PD manager <sup>4</sup>	3.10 (1.66)	2.97 (1.70)	0.70
Exploitation by PD manager <sup>5</sup>	3.25 (1.81)	2.81 (1.80)	2.11**
Feel exploited by PD manager <sup>6</sup>	3.66 (1.87)	3.30 (1.90)	1.64
SCO <sup>7</sup>	4.66 (1.09)	4.66 (1.17)	-0.02

*Note.* All questions were answered on a 7-point scale. Superscript 1: 1 (unfair) to 7 (fair). Superscripts 2–4: 1 (strongly disagree) to 7 (strongly agree). Superscript 5: 1 (not at all) to 7 (very much); <sup>1</sup>“Please rate how fair you think it is that manager A/B decided to offer consumers who [*Effort*: earned a lower endowment in the slider task, *Random*: randomly received a lower endowment] lower prices for the same good.”, <sup>2</sup>“The more goods a store sells, the higher the income of the participant acting as its manager. Do you consider it fair to switch from Store A/B to Store B/A in order to reduce manager A/B’s profit because of their pricing strategy?”, <sup>3</sup>“Manager A/B’s intention for [*Intention*: introducing interpersonal price differences; *No intention*: choosing the pricing algorithm] was good.”, <sup>4</sup>Wording is equal to row 2, except last word is “bad”, <sup>5</sup>“Manager A/B intended to take advantage of me (the consumer).”, <sup>6</sup>“The fact that [*Intention*: Manager A/B, *No intention*: Store A/B’s algorithm] decided to offer some consumers lower prices makes me feel taken advantage of.”, <sup>7</sup>Mean of Social comparison orientation scale by Gibbons and Buunk (1999); The reported *t*-statistic results from a two-sample mean-comparison *t*-tests by behavioral pattern; \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .



**Table 8** — Panel logistic regression using random effects for *Approach*

	Study 1	Study 3	Study 4b	Study 5a
	Coef. (se)	Coef. (se)	Coef. (se)	Coef. (se)
PD	0.636 (0.666)	0.409 (0.586)	0.120 (0.437)	1.689*** (0.411)
PD × <i>Effort</i>		0.184 (0.774)	0.026 (0.654)	1.619** (0.770)
PD × <i>Agency</i>			1.923*** (0.650)	
PD × <i>Effort</i> × <i>Agency</i>			-0.258 (0.913)	
<i>Effort</i>		-0.012 (0.568)	-0.384 (0.448)	-2.031*** (0.705)
<i>Agency</i>			-0.992** (0.503)	
<i>Effort</i> × <i>Agency</i>			0.527 (0.709)	
SCO	-0.308 (0.239)	-0.076 (0.193)	-0.064 (0.121)	0.410** (0.193)
Constant	-2.731** (1.229)	-3.709*** (1.037)	-3.420*** (0.675)	-6.070*** (1.140)
N	388	800	2356	1188
AIC	153.150	381.708	1246.513	728.275
BIC	168.994	409.816	1304.160	758.755

*Note.* Table reports results of panel logistic regressions using random effects and a cluster-robust VCE estimator. Dependent variable is a dummy variable that equals 1 if participant bought a good in the price discriminating store, 0 otherwise — Independent variables: “PD” equals 1 in Round 2 and 3, 0 otherwise; *Effort* and *Agency* are dummy variables for the experimental conditions; Reference groups are omitted from the table; “SCO” is mean of Social comparison orientation scale by Gibbons and Buunk (1999) — \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 9 — Panel logistic regression using random effects for *Avoid*

	Study 1	Study 2	Study 3	Study 4a	Study 5a	Study 5b
	Coef. (se)	Coef. (se)	Coef. (se)	Coef. (se)	Coef. (se)	Coef. (se)
PD	-5.353*** (1.206)	-3.935*** (0.459)	-4.700*** (0.806)	-3.994*** (0.587)	-4.136*** (0.674)	-5.273*** (0.726)
PD × <i>Intention</i>		-1.543** (0.617)				
PD × <i>Effort</i>			0.507 (0.852)	-1.300* (0.754)	-1.074 (0.852)	
PD × <i>Agency</i>				0.380 (0.739)		
PD × <i>Effort</i> × <i>Agency</i>				0.351 (1.075)		
PD × 15 Coins						0.598 (0.849)
PD × 30 Coins						0.732 (0.812)
<i>Intention</i>						
<i>Effort</i>		0.969* (0.520)				
<i>Agency</i>			-0.412 (0.731)	1.119 (0.683)	0.506 (0.772)	
<i>Effort</i> × <i>Agency</i>				0.270 (0.646)		
15 Coins				-1.050 (0.951)		0.504 (0.731)

*continued*

Table 9 — continued

	Study 1		Study 2		Study 3		Study 4a		Study 5a		Study 5b	
	Coef. (se)	Coef. (se)	Coef. (se)	Coef. (se)	Coef. (se)	Coef. (se)	Coef. (se)	Coef. (se)	Coef. (se)	Coef. (se)	Coef. (se)	
<i>30 Coins</i>												
SCO	0.513* (0.304)	-0.237 (0.172)	-0.086 (0.199)	-0.375*** (0.144)	0.316 (0.216)							0.096 (0.683)
Constant	3.383** (1.584)	5.582*** (0.965)	5.647*** (1.263)	6.907*** (0.958)	4.754*** (1.279)							5.758*** (0.992)
N	384	1564	788	2344	1180							1468
AIC	292.194	1267.133	634.181	1734.051	819.050							1129.807
BIC	307.996	1299.263	662.198	1791.647	849.489							1172.140

Note. Table reports results of panel logistic regressions using random effects and a cluster-robust VCE estimator. Dependent variable is a dummy variable that equals 1 if participant bought a good in the price discriminating store, 0 otherwise — Independent variables: “PD” equals 1 in Round 2 and 3, 0 otherwise; *Intention*, *Effort*, *Agency*, *15 ECU* and *30 ECU* are dummy variables for the experimental conditions; Reference groups are omitted from the table; “SCO” is mean of Social comparison orientation scale by Gibbons and Buunk (1999) — \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

**Table 10** — Ordered logistic regression over all studies by *Approach*

	<i>Avoid</i>		<i>Approach</i>	
	Coef.	se	Coef.	se
Year of birth	0.008*	(0.005)	0.034***	(0.010)
<i>Gender</i>				
Male		Reference group		
Female	-0.306***	(0.097)	0.141	(0.174)
Prefer not to answer	-0.259	(0.533)	1.270	(1.005)
<i>Employment status</i>				
Working (paid employee)		Reference group		
Working (self-employed)	0.090	(0.137)	-0.031	(0.276)
Not working (temporary layoff)	-0.099	(0.517)	-0.216	(0.436)
Not working (looking for work)	-0.005	(0.177)	0.372	(0.266)
Not working (retired)	0.649**	(0.313)	0.973	(0.593)
Not working (disabled)	0.299	(0.334)	-0.527	(1.148)
Not working (other)	0.285	(0.180)	-0.294	(0.312)
Prefer not to answer	0.006	(0.371)	-13.237***	(0.437)
<i>Education</i>				
Less than high school degree	0.063	(0.957)	0.065	(1.246)
High school graduate	0.071	(0.169)	-0.147	(0.344)
Some college but no degree	-0.186	(0.120)	-0.267	(0.234)
Associate degree in college (2-year)	-0.242	(0.172)	-0.600*	(0.358)
Bachelor's degree in college (4-year)		Reference group		
Master's degree	0.516***	(0.144)	0.086	(0.251)
Doctoral degree	1.311***	(0.413)	-0.382	(0.700)
Professional degree (JD, MD)	-0.005	(0.324)	0.342	(0.542)
<i>Social Comparison Orientation Scale</i>				
Item 1	0.041	(0.035)	-0.045	(0.064)
Item 2	-0.039	(0.043)	0.057	(0.078)
Item 3	0.027	(0.045)	-0.008	(0.085)
Item 4	0.008	(0.042)	-0.125*	(0.073)
Item 5	-0.101***	(0.037)	0.088	(0.077)
Item 6	0.047	(0.043)	0.072	(0.091)
Item 7	0.040	(0.044)	0.112	(0.080)
Item 8	0.049	(0.054)	-0.004	(0.108)
Item 9	-0.044	(0.052)	-0.180	(0.113)
Item 10	-0.075*	(0.044)	0.016	(0.089)
Item 11	0.009	(0.033)	0.020	(0.062)
<i>Ethnicity</i>				
African American	0.185	(0.211)	0.343	(0.411)

*continued*

Table 10 — *continued*

	<i>Avoid</i>		<i>Approach</i>	
	Coef.	se	Coef.	se
American Indian	0.265	(0.466)	1.063*	(0.614)
Asian	0.480**	(0.221)	-0.949*	(0.497)
Hispanic/Latino	0.032	(0.221)	-0.021	(0.387)
White/Caucasian	0.379*	(0.199)	0.067	(0.395)
Other	0.158	(0.512)	-0.398	(1.169)
N	1932		1183	
AIC	3328.196		1260.092	
BIC	3528.583		1442.821	

*Note.* Table reports results of ordered logistic regressions using a robust VCE estimator. Dependent variable is the number of goods sold in the PD store in rounds 3 and 4 — Reference groups of independent variables are omitted; Items of social comparison orientation scale by Gibbons and Buunk (1999): Item 1: “I often compare how my loved ones (boy or girlfriend, family members, etc.) are doing with how others are doing”, Item 2: “I always pay a lot of attention to how I do things compared with how others do things”, Item 3: “If I want to find out how well I have done something, I compare what I have done with how others have done”, Item 4: “I often compare how I am doing socially (e.g., social skills, popularity) with other people”, Item 5: “I am not the type of person who compares often with others” (reverse coded), Item 6: “I often compare myself with others with respect to what I have accomplished in life”, Item 7: “I often like to talk with others about mutual opinions and experiences”, Item 8: “I often try to find out what others think who face similar problems as I face”, Item 9: “I always like to know what others in a similar situation would do”, Item 10: “If I want to learn more about something, I try to find out what others think about it”, Item 11: “I never consider my situation in life relative to that of other people.” (reverse coded) — \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .



## CHAPTER III

### **Excusing selfish behavior with (willfully induced) uncertainty: An experimental study**

Mattheus Brenig, Nils Engelbrecht, and Daniel Hermann

In most decisions, the causal relationship between one's actions and the consequences for others is subject to uncertainty. We examine how uncertainty due to moves of nature affects selfish behavior in a decision that involves a trade-off between one's own and others' payoff. We consider three specific excuses for selfish behavior associated with uncertainty: A diffusion of being pivotal, the bystander excuse, and the replacement excuse. The replacement excuse can be used when there is uncertainty about whether an unselfish decision prevents the externality. The bystander excuse can be used when there is uncertainty about whether a selfish decision results in the externality. We extend our analysis of excuse-driven behavior by varying the scope of uncertainty and introducing the opportunity for willful ignorance. We find that introducing uncertainty increases selfish behavior by approximately one-fourth. This result is robust to our experimental variations of pivotality, the scope of uncertainty, and possibilities to invoke the bystander and/or replacement excuses. Once pivotality is reduced to zero, selfish behavior increases further. Two-thirds of subjects disclose information about their influence on the externality, resulting in no significant effect when willful ignorance is allowed for.

## 1 Introduction

In his seminal work on moral disengagement, Bandura (1999) identifies several psychosocial mechanisms that decision-makers use to justify harmful effects of their actions on others. These mechanisms include denial of personal agency through diffusion of responsibility, as well as consciously ignoring effects on others. While Bandura mainly applies his theory to war crimes and military atrocities, it also appears to apply to decisions involving more mundane trade-offs between self and others, such as everyday market transactions. Although each individual market transaction may lead to comparatively small harmful effects, the ubiquity of markets, combined with extensive division of labor can lead to a wide diffusion of responsibility and increased opportunities for willful ignorance, allowing for moral disengagement on both sides of the market and substantial aggregate effects.

Consider climate change. Because climate is a global public good, responsibility is maximally diffused and each consumer or producer may justify selfish behavior with a negligible impact on the outcome. Moreover, sociological studies show that diffused responsibility is accompanied by a reluctance to learn about climate change (Stoll-Kleemann et al. 2001; Norgaard 2006). Diffusion of responsibility also seems to apply in markets with fewer participants, like the arms trade market. Tony Blair, then prime minister of the United Kingdom (UK), in a 2002 press conference used the prospect of replacement by market competitors and the resulting uncertainty about the UK's ability to prevent military atrocities as an excuse not to restrict UK's arms industry and risk jobs.<sup>1</sup>

More rigorous evidence on moral disengagement in markets comes from an emerging strand of economic research that indicates that decision-makers are more selfish once the relationship between their actions and others' consequences is blurred through e.g. (i) delegating tasks to agents (Hamman et al. 2010; Bartling and Fischbacher 2012), (ii) responsibility diffusion when outcomes depend on the interaction of multiple persons (Brütt et al. 2020; Falk et al. 2020; Ziegler et al. 2021; Falk and Szech 2013), and (iii) willful ignorance (Bartling et al. 2014; Grossman and Weele 2016; Reczek et al. 2018). All these examples introduce ex-ante uncertainty about the consequences of one's actions on others. This allows decision-makers to justify selfish behavior

---

<sup>1</sup>"If we want to stop the defence industry operating in this country we can do so, and the result incidentally will be that someone else supplies the arms that we supply, [...]" ; see <https://webarchive.nationalarchives.gov.uk/20100105173918/http://www.number10.gov.uk/Page3000>.



and avoid attribution of blame and responsibility not only by themselves, but also by others (e.g. Bartling and Fischbacher 2012; Coffman 2011). Dana et al. (2007) have fittingly described such behavior as exploiting “moral wiggle room”.

This study contributes to the aforementioned literature, by investigating how different excuses emanating from uncertainty about ex-ante causal responsibility for a negative externality suffered by third parties influence selfish behavior. In addition, we are interested in whether, once uncertainty is introduced, the ability of decision-makers to remain ignorant of the consequences for third parties leads to further increase in selfishness.

We consider the interplay of three specific excuses for selfish behavior associated with uncertainty: A diffusion of being pivotal, the bystander excuse and the replacement excuse. The extent of (ex-ante) diffusion of being pivotal can be described as the weight of an individual’s decision on the externality, i.e., how much the probability of the externality is affected by an individual’s decision (Engl 2018). At the extremes, subjects are either fully pivotal or not pivotal at all. In the first case, there is no moral wiggle room, since a selfish decision leads to the externality with certainty. In the second case, subjects have no influence at all and only individuals following a deontological moral reasoning would still refrain from choosing the selfish option (Falk et al. 2020). In the intermediate case, with ex-ante uncertainty, subjects can use the bystander and/or replacement excuse. The replacement excuse can be applied when there is uncertainty as to whether an unselfish decision prevents the externality. For example, market participants may cause an externality with certainty if they participate in a transaction, but they can use the excuse “if I don’t make the transactions, someone else will” to justify their behavior (Falk et al. 2020; Ziegler et al. 2021). The bystander excuse can be applied when there is uncertainty as to whether a selfish decision results in the externality. For example, a bystander can help a victim and prevent harm with certainty, but may rely on other bystanders to respond to the emergency (Darley and Latane 1968; Fischer et al. 2011). In the above examples, the uncertainty is associated with only one choice option and therefore the excuses do not occur together. However, both excuses can occur simultaneously. To stick with the market analogy: In terms of a consumer’s influence on future market outcomes, uncertainty is probably prevalent in both the selfish and unselfish choice option.

Subjects in our experiment make a decision in a non-strategic environment where they can either maximize their own payoff or forgo some of it. Depending on subjects ex-ante causal responsibility, forgoing payoff may or may not prevent a negative

externality for third parties. While this describes our experimental design in abstract terms, the details replicate basic aspects of market transactions, as outlined in Section 2. Our design is motivated by how consumers perceive their causal responsibility for externalities triggered by market transactions. Moreover, in introducing strategic ignorance, we closely mirror how consumers in markets retrieve information, e.g., via product labels, just before making a purchase. Our analysis is conducted in three steps. *First*, we vary the ex-ante uncertainty regarding the consequences of a subject's choice on the externality along two properties. The first property is the pivotality of the subject, i.e. by how much the probability of the externality increases with a selfish choice. The second property is the "inherent" probability of the externality, which is independent of the subject's choice. Based on these two properties, we can vary a subject's pivotality and whether the replacement and/or bystander excuse can be applied.<sup>2</sup> *Second*, we extend the analysis by restricting the scope of uncertainty to only half of the externality, while subjects can prevent the other half with certainty. As subjects are now fully pivotal for part of the externality, this may limit their ability to use excuses and thus reduce selfish behavior. That is, subjects can never justify that the ex-post outcome will be the same, independent of their behavior. *Third*, we add the possibility of strategic ignorance in both aforementioned decision regimes. We expect that the possibility of willful ignorance in and of itself increases selfish behavior. However, depending on how excuse-driven behavior works, adding further justifications may only lead to a small increase in selfishness.

Our study is closely related to economic experiments in which unanimous group decisions are required to either cause or prevent an externality. In Falk et al. (2020) and Brütt et al. (2020), groups need a unanimous vote to prevent a negative externality. In Hauser et al. (2014), a single group member can exploit a common-pool resource to the point that it is not replenished for future generations. Since a single selfish vote or choice is sufficient to cause the negative externality, the ex-ante probability of each individual being pivotal is reduced. Similarly, in the multi-unit market treatment in Ziegler et al. (2021), one buyer-seller pair can trade all available units in a market, leading to the largest possible negative externality. In all the above-mentioned studies, subjects can use the replacement excuse to justify selfish behavior. Other studies have investigated unanimity voting in favor of a negative externality (Dana et al. 2007;

---

<sup>2</sup>In our paper, we use the term uncertainty as a synonym for risk, i.e. a known probability distribution over a range of known outcomes. If either probabilities or outcomes are unknown, we use the term ambiguity.

Irlenbusch and Saxler 2019; Behnk et al. 2017). Here, subjects can apply the bystander excuse. Irrespective of whether individuals can veto in favor or against the selfish outcome, studies find increased selfishness compared to individual decision making, i.e. with maximum pivotality. Brütt et al. (2020) compare both types of veto rights and find no significant difference in selfishness.

We contribute to this literature in various ways. First, we replace uncertainty through the behavior of other decision-makers by moves of nature. This allows us to manipulate uncertainty in a more fine-grained and controlled manner, while holding other dimensions of responsibility diffusion constant like shared guilt (Rothenhäusler et al. 2018) or information on social norms. In particular, for group decisions, perception of uncertainty depends on individuals' endogenous beliefs about the behavior of other group members and is thus the result of "fuzzy" probabilistic reasoning. The formation of belief is difficult to control and does not lend itself to precise manipulation through experimental treatments. We further contribute to the literature by considering different combinations of excuses and different degrees of being pivotal. Most situations, unlike unanimous votes, have uncertain consequences regardless of an individual's decision. In such situations, decision-makers can combine both the replacement and bystander excuse. We further test excuse-driven behavior by restricting the scope of uncertainty such that subjects can always prevent part of the externality and therefore never justify that the ex-post outcome will be the same independent of their behavior. Finally, we examine whether decisions are sensitive to the possibility to ignore consequences for third parties.

## 2 Experimental design

### 2.1 Setting

Subjects participate in groups of five in an ultimatum bargaining situation in which there are three roles: proposers ( $A$ ), responders ( $B$ ), and inactive third parties ( $C$ ). In each group, there are two proposers  $A_1$  and  $A_2$ , two inactive third parties  $C_1$  and  $C_2$ , and one responder  $B$ . Each proposer is assigned one inactive third party. The game is played for 2 rounds. Both rounds are relevant for payoffs. The group composition is fixed over both rounds. In each round, all subjects first receive an endowment of

$a = b = c = 100$  coins.<sup>3</sup> As explained in more detail below, proposers and responders have varying degrees of influence over the final distribution of coins in each round.

In round one, each proposer is first given a binary choice between offering the responder a distribution of either  $(a_i = 110, b = 120, c_i = 70)$  or  $(a_i = 110, b = 110, c_i = 100)$  coins. For both possible offers, the sum of the payments that proposer and responder receive is 30 or 20 coins greater than the sum of their initial endowments of 100 coins each. However, one distribution does not affect the inactive third party's endowment, while the other entails a loss of 30 coins. Throughout this paper, we will refer to the first and the second offer as selfish and unselfish, respectively. In the instructions, we used the neutral terms A and B and the colors blue and red to differentiate the offers.<sup>4</sup> Matching of proposers and responders only takes place after the first stage of round one is implemented, and is not random. Instead, groups are matched in a way that each responder can choose between both, the unselfish and selfish offer.<sup>5</sup> The responder has to accept one of the two offers. Thereby, the final allocation of coins to all group members is determined. The responder and the "winning" proposer receive the coins associated with the accepted offer. The "losing" proposer keeps the initial endowment of 100 coins. The determination of the payoffs for the inactive third parties depends on the experimental condition, as explained below.

The process of round two is identical to round one, except that offers are no longer chosen by proposers. Instead, offers depend on a move of nature in which the probabilities of future offers depend on the choice of the responder. The parameters of the move of nature depend on the experimental condition, as explained below. Independent of the specific parameter values, the move of nature leads to either two selfish or unselfish offers.

---

<sup>3</sup>Throughout the study, we used the term "coins" for our experimental currency, with a conversion rate of 100 coins to \$1.

<sup>4</sup>We randomized both colors, to control for possible color preferences.

<sup>5</sup>Responders are not informed about this detail of the matching process. However, it is important to emphasize that no deception is involved. To control for possible positional preferences on the computer screen (i.e. left and right), we randomized the position of the two offers.

## 2.2 Experimental conditions

### *Move of nature*

While offers in round one are chosen by proposers, the offers in round two are determined by a move of nature. The move leads to one of two possible “market structures” in round two: a market where both proposers make the selfish or unselfish offer. We call the former the selfish and the latter the unselfish market. The move of nature is determined by the two parameters  $0 \leq \alpha \leq 1$  and  $0 \leq \beta \leq 1$ . The parameter  $\alpha$  determines the pivotality or weight of the responder’s choice in terms of how much the probability of a selfish market is increased if the responder accepts the selfish offer. Weighted with  $(1 - \alpha)$ , the parameter  $\beta$  determines the inherent probability of the externality. For example if  $\alpha = 0.1$  and  $\beta = 0.5$ , the inherent probability of the externality is  $45\% = (1 - 0.1) \times 0.5 \times 100$ , and a selfish choice increases the probability by 10 percentage points to  $55\%$ .

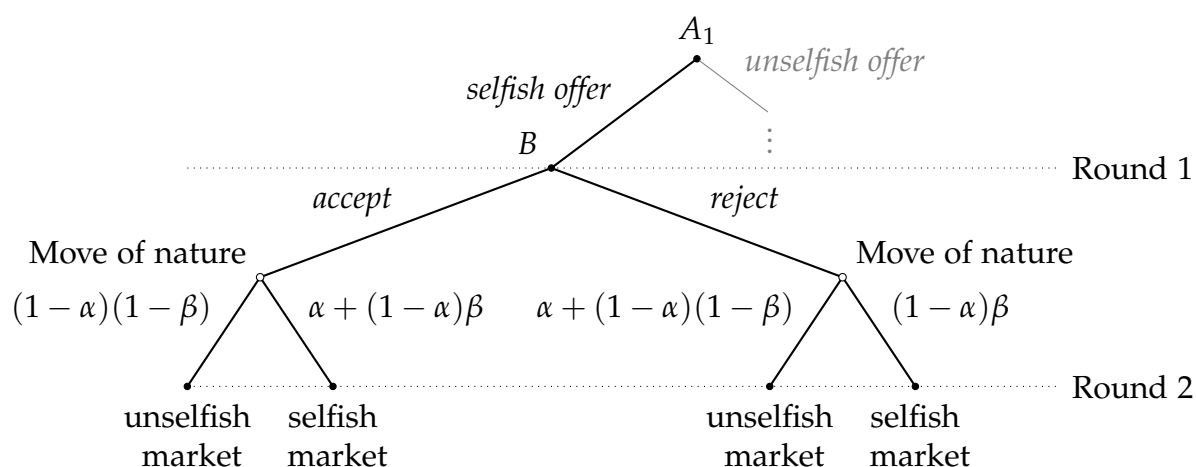
Round one and two in extensive form are represented by Figure 1.<sup>6</sup> In reading the game tree, remember that groups are matched so that both the unselfish and selfish offer are available to the responder in round one. Hence, if a responder accepts a selfish offer by e.g.  $A_1$ , an unselfish offer by  $A_2$  is rejected and vice versa. Due to this symmetry, only half of the game tree is shown in detail. If  $\alpha$  equals 1, the value of  $\beta$  is irrelevant and the responder is solely responsible for the market structure. If the responder chooses the selfish offer in round one, the market will be selfish in round two. The opposite is true for  $\alpha = 0$ . In this case, the responder has no influence on the market structure in round two. For any  $0 < \alpha < 1$ , varying  $\beta$  effects whether the responder can apply the replacement and/or bystander excuse. If  $\beta$  equals 1, there is uncertainty as to whether an unselfish decision prevents the externality (replacement excuse). If  $\beta$  equals 0, there is uncertainty as to whether a selfish decision results in the externality (bystander excuse). In the intermediate case of  $0 < \beta < 1$ , both choice options are subject to uncertainty, which is why both excuses can be used.

### *Scope of uncertainty*

Proposers’ and responder’s payoffs are always determined once the responder accepts an offer. The timing of the payoff consequences for third parties, on the other hand,

---

<sup>6</sup>Payoffs are not shown and can be deduced from Figure 2. Instead, the game tree is intended to show which path of the decision tree leads to which market structure.



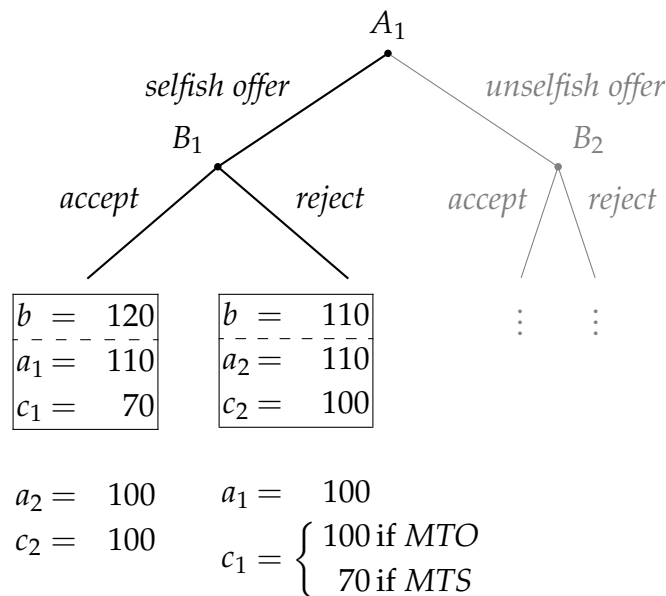
**Figure 1** — Game tree over both rounds

depends on the experimental condition. This applies to both rounds, regardless of whether the proposer has actively selected the offer (round one) or whether the offer for the proposer was determined by the market rule (round two). Analogous to the Made-to-Stock (*MTS*) and Made-to-Order (*MTO*) manufacturing processes, pivotality for externalities of offers that have already been made is implemented as a binary variable, i.e., responders are either pivotal or not.<sup>7</sup> In *MTO*, a negative externality arises by consumption. In *MTS* all the injustices associated with the production process have already occurred once a product is offered to consumers. Experimental studies on socially responsible behavior in markets have applied either *MTO* or *MTS* (e.g., *MTO* in Bartling et al. 2015 and *MTS* in Pigors and Rockenbach 2016). In *MTS*, a third party payoff is determined as soon as the assigned proposer chooses the offer. In *MTO*, all payoffs are determined once the responder accepts an offer. Hence, while responders cannot prevent the externality by rejecting the selfish offer in *MTS*, they can do so in *MTO*. Round one in extensive form is represented by Figure 2 for *MTO* and *MTS*. The payoffs associated with the “winning” offer are highlighted by a box.

Importantly, *MTO* splits the avoidable externality over both rounds. If a responder accepts the selfish offer in round one, 30 coins of externality are created. When the selfish market occurs in round two, two selfish offers are made, but only 30 coins

<sup>7</sup>Traditionally, goods are produced ahead, i.e. made-to-stock (*MTS*). However, due to new production technologies and more differentiated consumer demand, goods are increasingly made-to-order (*MTO*). While *MTO* was predominantly used for complex and highly customized products such as industrial machines or cars, it is now applied to mass products such as clothes or even groceries (Gilmore and Pine 1997; Holweg and Pil 2001).

of externality are created by consumption. In contrast, all avoidable externalities in *MTS* are realized in round two. When round two is the selfish market, two selfish offers are made, resulting in 60 coins of externality regardless of responder choice in round two. Because the move of nature only applies to offers in round two, the scope of uncertainty is reduced by half in *MTO*. Because this experimental condition depends on subjects understanding the consequences of their choices on payoffs, we included three comprehension questions that asked about all relevant payoffs for a selfish choice in round one and the payoffs for both round two scenarios (see Online Appendix A.5 for all comprehension questions).



**Figure 2** — Game tree for round 1

*Strategic ignorance*

In the *Full information* condition, subjects are informed about the payoff consequences for all roles in the instructions and their understanding of these consequences is tested with several comprehension questions. In the *Strategic ignorance* condition, payoff consequences for third parties are hidden from the subjects in the instructions. That is, subjects do not know the direction or magnitude of the payoff consequences. However, subjects are informed that they have the opportunity to reveal the payoff consequences before making a decision. To reveal the information, subjects only have to click a button on the screen, i.e. there are no transaction costs.

### 2.3 Treatments and behavioral hypotheses

Table 1 gives an overview of the treatments. In the first step of our analysis, we investigate behavior in *MTS*, i.e. when the externality of 60 coins is realized in round two and therefore in its entirety subject to the move of nature.

**Table 1** — Treatments

Name	Experimental setup			
	Pivotality $\alpha$	Inherent probability $(1 - \alpha) \times \beta$	MTO/MTS	Strategic ignorance
<i>Baseline</i>	100	0	MTS	No
<i>Bystander</i>	50	0	MTS	No
<i>Replacement</i>	50	50	MTS	No
<i>Compound</i>	10	45	MTS	No
<i>Deontological</i>	0	50	MTS	No
<i>CompoundMTO</i>	10	45	MTO	No
<i>BaselineMTO</i>	100	0	MTO	No
<i>Compound*</i>	10	45	MTS	Yes
<i>CompoundMTO*</i>	10	45	MTO	Yes

We hypothesize that selfishness increases in the order of treatments as shown in the first section of Table 1, i.e. rows one to five. In our *Baseline* condition, subjects are maximally pivotal, i.e. a selfish choice leads to the externality with certainty and vice versa. In this case, there is no moral wiggle room. In our *Replacement* condition, we set the inherent probability of the externality and pivotality to fifty percentage points. A selfish choice leads to the externality with certainty, while there is a fifty percent probability of the externality even if the subject is not selfish. Although pivotality is still relatively high, subjects may use the uncertainty regarding their ability to prevent the externality as an excuse for selfishness. In our *Bystander* condition, the intrinsic probability corresponds to zero and pivotality to fifty percentage points. In this case, subjects can prevent the externality. However, they can excuse selfish behavior on the grounds that there is a chance that the externality will be prevented either way. Pivotality is equally strong in *Bystander* and *Replacement*, the only difference being the inherent probability of the externality. This is analog to unanimity voting in favor (*Bystander*) or against (*Replacement*) the externality. To our knowledge, only Brütt et al. (2020) compare both types of voting. They find slightly lower selfishness in a



bystander-like situation, but no significant differences. In our *Compound* condition, we set the inherent probability to forty-five and pivotality to ten percentage points. In *Compound*, pivotality is reduced substantially and subjects can apply both the bystander and replacement excuse. Accordingly, we expect increased selfishness compared to *Bystander* and *Replacement*. Finally, in *Deontological*, the inherent probability is fifty percent and subjects have no influence on the market at all. In this scenario, only individuals following a deontological moral reasoning would still refrain from choosing the selfish option (Falk et al. 2020).

**Hypothesis 1** *Selfishness is lowest in Baseline and largest in Deontological. The treatments Bystander, Replacement, and Compound lie in-between, the latter exhibiting relatively strong selfish behavior.*

In step two, we extend our main analysis by investigating behavior in *MTO*. In this choice setting, subjects always generate half of the externality, i.e. 30 coins, by choosing the selfish offer in round one. If round two is a selfish market, another 30 coins of externality arise by consumption. Essentially, the move of nature now only applies to half of the externality, while subjects can prevent the other half with certainty. That is, subjects can never justify that the ex-post outcome will be the same, independent of their behavior. We chose the *Compound* condition for the *MTO* setting, because it generates the largest difference between *MTO* and *MTS* in terms of expected externality, while allowing for the largest number of excuses due to low pivotality as well as the bystander and replacement excuse. We call this treatment *CompoundMTO*.

**Hypothesis 2** *Selfish behavior is lower in CompoundMTO compared to Compound.*

Furthermore, as a control, we also run a *BaselineMTO* treatment with full pivotality. For the purpose of our study, we are only interested in the differences between *MTO* and *MTS* in terms of the scope of uncertainty. However, there may be a concern that the two conditions introduce other differences that are relevant to behavior. For example, simply by framing the timing of the externality differently, subjects may feel like second movers in *MTS* and, as a result, less responsible for the externality. At full pivotality, there is no uncertainty and thus no difference between *MTO* and *MTS* on the dimension of interest for our study. In both treatments, subjects have full pivotality for 60 coins of externality. Comparing *Baseline* and *BaselineMTO* thus serves as a control for potential confounds due to unintended changes.

In a third and last step, we allow for strategic ignorance in both *CompoundMTO* and *Compound*. *Ceteris paribus*, giving subjects the opportunity to remain strategically ignorant about the consequences of their decisions on third parties should increase selfishness.

**Hypothesis 3** *Selfishness is higher in CompoundMTO\* and Compound\*, compared to CompoundMTO and Compound, respectively.*

## 2.4 Experimental procedures

The experiment was programmed with oTree (Chen et al. 2016) and administered to subjects via MTurk between December 2020 and February 2021. We prevented retakes and restricted participation to US residents.<sup>8</sup> Since we were only interested in the responder decisions and needed to ensure that both offers were made in each group, we collected the proposer and third party observations beforehand. We applied a many-to-one matching, collecting at least one selfish and unselfish proposer and two third party observations per treatment. These participants received a fixed payment immediately, and (if matched) the rest of their payment after all responder observations were collected.

In total, 1,130 subjects in the role of responder participated. After reading a plain language statement and giving consent to participate, subjects read the instructions (see Online Appendix A.4) and had a maximum of two attempts to answer a few comprehension questions correctly. Subjects who failed twice were excluded from participation. We performed two additional tests to check the validity of our experimental design and subjects' understanding of the experiment and instructions (see Online Appendix A.2). After completion of the main task, subjects completed a post-experimental questionnaire. Following our preregistration and based on arrival time, we excluded observations in excess of 125 observations per treatment. Our final sample thus comprises 1125 observations. About 49% of subjects were female. Participants are on average 36 years old. For more information on participants' socio-demographics, see Online Appendix A.1. Participants took an average of 21 minutes, with a mean payment of \$3.05, which equals \$13.30 per hour.<sup>9</sup>

---

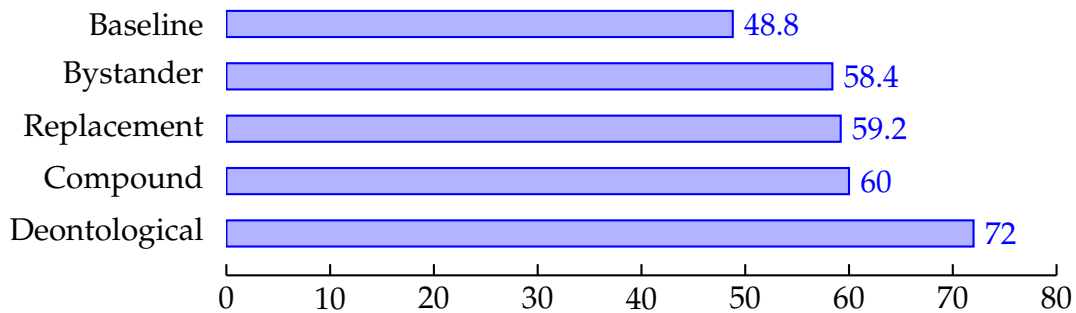
<sup>8</sup>In addition, due to readability issues, we excluded participants who reported accessing the study via smartphone or tablet.

<sup>9</sup>We preregistered the experiment at <https://aspredicted.org/blind.php?x=be7c8c>. The complete data, do-files that contain all the commands we execute for the analysis, oTree app as well as

## 3 Results

### 3.1 Main results

Figure 3 shows the share of selfish choices per treatment in *MTS* and *Full information*. As expected, subjects are least selfish in *Baseline* (48.8%), with deterministic influence on second-round offers. Once uncertainty is introduced, selfish behavior increases about ten percentage points to 58.4%, 59.2%, and 60%. This increase is on the borderline to significance in a pairwise comparison (chi<sup>2</sup>-test,  $p = 0.13, 0.10,$  and  $0.08$ ). There is no significant difference (chi<sup>2</sup>-test,  $p = 0.967$ ) among the three treatments with uncertainty. That is, it does not seem to matter for selfish behavior whether uncertainty allows subjects to use either the replacement or bystander excuse. Even more surprisingly, reducing pivotality from 50% to 10% does not significantly increase selfishness either. However, subjects do seem to differentiate between uncertain causal responsibility and no causality at all. Once responders have no influence on the market in *Deontological*, selfishness increases to 72%. This increase is significant with respect to all other treatments in a pairwise comparison (chi<sup>2</sup>-test,  $p = 0.05, 0.03, 0.02,$  and  $0.00$ ). The treatment *Deontological* also reveals that 28% of responders still reject the selfish offer, which may be due to deontological reasons or to punish the proposer.



**Figure 3** — Share of selfish choices in treatments

**Result 1** *Introducing uncertainty about a responder’s influence on future market outcomes increases selfish behavior.*

**Result 2** *Subjects behavior seems to be inelastic to changes in pivotality and different combinations of the replacement and bystander excuse.*

additional information about the data structure can be accessed via the following repository: [https://osf.io/md7wu/?view\\_only=02fe724834be413d8ef84370adb20f97](https://osf.io/md7wu/?view_only=02fe724834be413d8ef84370adb20f97).

**Result 3** *Selfish behavior is highest in Deontological. Still, 28 % of responders reject the selfish offer.*

Similar to the different excuses relating to uncertainty, reducing its scope from *Compound* to *CompoundMTO* does not impact behavior significantly. Despite the possibility to prevent half of the externality with certainty and therefore considerable differences in aggregate pivotality, selfish behavior only decreases slightly to 58.4 %, i.e. by 1.6 percentage points. Concordantly, a Pearson's  $\chi^2$  test is not significant ( $p = 0.8$ ). A comparison of *Baseline* (48.8 %) and *BaselineMTO* (47.2 %) shows that there are no other confounding differences between *MTS* and *MTO*.

**Result 4** *Reducing the scope of uncertainty by half does not significantly affect selfish behavior.*

Allowing responders to remain ignorant about the consequences of their decisions on third parties does not significantly affect selfish behavior in our experiment. While selfishness increases in *MTS* as expected (from 60 % to 63.2 %), there is a comparable decrease in *MTO* (from 58.4 % to 54.4 %). About two-third of subjects chose to reveal the payoff consequences for third parties, in both *MTO* (68.80 %) and *MTS* (61.60 %). This difference is not statistically significant ( $\chi^2 = 1.43, p = 0.23$ ). Subjects who actively disclose the payoff information are less likely to choose the selfish offer than subjects who get (passively) informed in the instructions. In *MTO* the difference is 16.54 percentage points (41.86 % vs. 58.4 %;  $\chi^2 = 5.58, p = 0.02$ ); in *MTS* the difference is 10.65 percentage points (49.35 % vs. 60 %;  $\chi^2 = 2.19, p = 0.14$ ).

**Result 5** *Allowing responders to remain ignorant about the consequences of their decision on third parties does not significantly affect selfishness.*

**Result 6** *About two-thirds of subjects reveal information on the consequences of their actions on third parties.*

### 3.2 Individual heterogeneity

Table 2 shows the results of a logistic regression including all observations. Model 1 includes the treatments as explanatory variables and a dummy variable that equals one for subjects revealing the payoff information in the *Strategic ignorance* condition. Controlling for those who actively revealed the information, *Strategic ignorance* leads to a significant increase in selfish behavior. Unsurprisingly, subjects who remain ignorant predominantly choose the offer that maximizes their payoff.

Model 2 adds several socio-demographic variables from the post-experimental questionnaire. Overall the coefficients and standard errors of the treatments remain fairly stable. Regarding employment status, those self-employed and those not working, but looking for a job are significantly less selfish than those employed (reference group). Regarding education, those holding an associate degree are significantly less selfish than those holding a bachelor's degree (reference group). Ethnicity, age, and gender all have no significant effect on selfishness. We also asked subjects about their position on the political spectrum on a nine-item scale ranging from 1 (left-wing) to 9 (right-wing). Each unit increase on the scale significantly decreases the likelihood of a selfish choice by about 6%. That is, going from one to nine on the scale, increases the likelihood of a selfish choice by almost 56%. Furthermore, we asked subjects a few questions regarding their income situation (see verbatim questions and response scale below Table 2). Subjects who reported that the COVID-19 pandemic improved their financial situation are significantly less likely to choose selfishly.

As mentioned in the description of our experimental design, we randomized both the color and screen position of the two offers. While the latter had no effect, subjects were about 23% less likely to choose the selfish offer when it was visually highlighted by the color blue.

**Table 2** — Logistic regression

	Model 1		Model 2	
	Odds ratio	se	Odds ratio	se
<i>Experimental conditions</i>				
Baseline	Reference group		Reference group	
Bystander	1.473	0.375	1.529	0.403
Replacement	1.522*	0.389	1.534	0.404
Compound	1.574*	0.402	1.580*	0.416
Deontological	2.698***	0.722	2.429***	0.669
CompoundMTO	1.473	0.375	1.412	0.371
BaselineMTO	0.938	0.237	0.866	0.227
CompoundMTO*	4.695***	1.737	4.800***	1.834
Compound*	6.275***	2.318	6.370***	2.433
<i>Decisions</i>				
Revealed Payments	0.162***	0.054	0.149***	0.051

*continued*

**Table 2** — *continued*

	Model 1		Model 2	
	Odds ratio	se	Odds ratio	se
<i>Employment status</i>				
Working (paid employee)			Reference group	
Working (self-employed)			0.577***	0.111
Not working (temporary layoff)			1.174	0.530
Not working (looking for work)			0.628**	0.147
Not working (retired)			0.684	0.330
Not working (disabled)			2.079	1.363
Not working (other)			0.925	0.223
Prefer not to answer			0.933	0.522
<i>Education</i>				
Less than high school degree			0.406	0.509
High school graduate			0.805	0.225
Some college but no degree			0.896	0.162
Associate degree in college			0.669*	0.156
Bachelor's degree in college			Reference group	
Master's degree			0.793	0.142
Doctoral degree			0.711	0.388
Professional degree (JD, MD)			0.884	0.387
<i>Ethnicity</i>				
African american			1.216	0.408
American indian			0.741	0.397
Asian			1.741	0.604
Hispanic			0.778	0.266
White caucasian			0.937	0.290
<i>Gender</i>				
Male			Reference group	
Female			0.835	0.110
Prefer not to answer			0.271*	0.203
<i>Other socio-demographics</i>				
Age			0.990	0.007
Political spectrum			1.057*	0.032
<i>Income</i>				
Relative Income			1.021	0.066
Effect COVID-19 (current)			0.847*	0.083
Effect COVID-19 (until end 2021)			1.104	0.117

*continued*

**Table 2** — *continued*

	Model 1		Model 2	
	Odds ratio	se	Odds ratio	se
<i>Other controls</i>				
Deal B is blue			0.774**	0.099
Deal B is left			0.917	0.119
<i>N</i>	1125		1125	
LR $\chi^2(8)$	59.61		105.66	
<i>p</i> -value	0.00		0.00	
Pseudo- $R^2$	0.04		0.07	

*Note.* Table reports results of a logistic regression. Dependent variable is a binary variable that equals 1 if a responder chose the fair offer, 0 otherwise. Independent variables: Education (“What is the highest level of school you have completed or the highest degree you have received?”); Employment status (“Which statement best describes your current employment status?”); Ethnicity (“Choose one or more ethnicities that you consider yourself to be:”); Political spectrum (“Where would you classify yourself on the left/right political spectrum?”), response scale ranges from 1 (left-wing) to 9 (right-wing); Relative Income (“How do you think your income and financial situation currently compare to those of others in the U.S. who are of similar age?”), response options are Don’t know/No answer, Much below average, Somewhat below average, About the average, Somewhat above average, Much above average; Effect COVID-19 (current, “How did the COVID-19 pandemic affect your financial situation?”, response options are Worsened a lot, Worsened, Remained unchanged, Improved, Improved a lot; Effect COVID-19 (until end 2021, “How do you expect the COVID-19 pandemic will affect your financial situation until the end of 2021?”), response options are Worsen a lot, Worsen, Remain unchanged, Improve, Improve a lot — \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

### 3.3 Behavioral intentions

In the post-experiment questionnaire, we asked subjects to indicate on a scale from 1 (“not at all important”) to 7 (“very important”), how important a number of given reasons were for their decision.

Table 4 shows the means of the seven reasons for the full sample and for the different combinations of  $\alpha$  and  $\beta$ . It also reports the results of a Kruskal-Wallis equality-of-populations rank test for each item by experimental condition. Looking at the full sample, punishment of the unfair proposer was the least important reason (1.96). Most important was the maximization of one’s own payoff (4.95) and the impact on the second round (4.41). For all items but punishment of the unfair proposer and deontological reasoning, a Kruskal-Wallis equality-of-populations rank test by experimental condition (i.e. combinations of  $\alpha$  and  $\beta$ ) is significant. Overall, the order of means for the experimental conditions is consistent with expectations. For example, subjects in  $\alpha = 0$  report the lowest importance of the reasons “Reduce

inequality”, “Reduce externality” and “Impact on second round”. In contrast, subjects in  $\alpha = 1$  report the highest importance of the reasons “Reduce inequality” and “Reduce externality”.

Besides differences between experimental conditions, the items can also be used to compare motives between subjects that chose the selfish versus the unselfish offer. One way to make this comparison is to identify for each subject the reason that was attributed the greatest (least) importance and to determine the relative frequencies depending on choice.<sup>10</sup> Reducing payout inequality (30.23 %) and the externality (28.75 %) were most frequently valued as the most important reason by unselfish subjects. Punishing (33.5 %) or rewarding (28.96 %) the unfair proposer were regarded as least important most frequently. The most and the least important reason most frequently chosen by selfish subjects were the maximization of one’s own payoff (61.04 %) and the impact on the second round (17.33 %) compared to punishment of the proposer of the selfish offer (33.28 %) and deontological reasoning (24.69 %). See Tables 10–12 in Appendix A.3 for more information.

Table 3 shows the means for the items depending on the choice in round one. A Mann-Whitney test is significant for all reasons, but “Impact on second round”. This reason, of course, should be equally important regardless of whether someone is driven by material self-interest or concern for the third party. The difference for all other items is in the expected direction. Subjects choosing the unselfish offer are driven relatively strongly by the intention to (i) reduce payoff inequality (5.79 vs. 2.64), (ii) reduce the externality (5.74 vs. 2.43), (iii) “not get their hands dirty” (4.43 vs. 2.09), and (iv) punish the unfair proposer (2.39 vs. 1.65). In contrast, they are relatively weakly driven by the intention to maximize their own payoff (3.18 vs. 6.23) and reward the unfair proposer (2.47 vs. 4.11).

## 4 Discussion and conclusion

This paper examines how selfish behavior is affected by uncertainty about ex-ante causal responsibility for a negative externality suffered by third parties. Using moves of nature, we investigate the role of pivotality, the bystander and replacement excuse, and the scope of uncertainty.

---

<sup>10</sup>If subjects gave equal importance to more than one reason, ties were randomly broken.



**Table 3** — Behavioral intentions by choice

Behavioral intention	Choice		Mann-Whitney test	
	Selfish Mean	Unselfish Mean	<i>z</i>	<i>p</i> -value
Reduce inequality	5.79	2.64	22.18	0.00
Punish unfair proposer	2.39	1.65	6.96	0.00
Reduce externality	5.74	2.43	22.69	0.00
Impact on second round	4.45	4.37	0.75	0.45
Deontological reasoning	4.43	2.09	17.99	0.00
Maximize own payoff	3.18	6.23	-22.83	0.00
Reward unfair proposer	2.47	4.11	-12.37	0.00

*Note.* Wording of question: For each of the following reasons, please indicate how important it was for your decision to accept Deal [A/B] by Proposer [1/2] in Round 1; Item 1: "Reduce the inequality of payments between participants."; Item 2: "Punish Proposer [A/B] for choosing a deal associated with a loss for Third Party [1/2]"; Item 3: "Reduce or eliminate the loss for the third parties."; Item 4: "Increase the probability of an 'Only [A/B]' scenario in Round 2."; Item 5: "Not 'get my hands dirty' by accepting an offer that reduced the payment of a third party."; Item 6: "Maximize my own payoff."; Item 7: "Reward Proposer [A/B] for an offer that was financially advantageous for me."; All questions were answered on a 7-point Likert-type scale, where 1 is "not important at all" and 7 is "very important."

**Table 4** — Behavioral intentions by combinations of  $\alpha$  and  $\beta$

Behavioral intention	Experimental condition ( $\alpha, \beta$ )								Kruskal-Wallis test <sup>a</sup>
	Full sample	0, 0.5	0.1, 0.5	0.5, 1	0.5, 0	1, —	$\chi^2(4)$	<i>p</i> -value	
Reduce inequality	3.97	3.54	3.94	3.70	4.01	4.35	13.46	0.01	
Punish unfair proposer	1.96	1.89	1.88	2.02	2.22	2.00	3.14	0.53	
Reduce externality	3.82	3.25	3.79	3.53	3.94	4.27	19.94	0.00	
Impact on second round	4.41	2.52	4.32	5.20	4.98	4.83	120.61	0.00	
Deontological reasoning	3.07	2.78	3.09	2.82	3.26	3.22	7.48	0.11	
Maximize own payoff	4.95	5.17	5.02	5.18	4.94	4.59	9.30	0.05	
Reward unfair proposer	3.42	3.82	3.46	3.34	3.58	3.12	10.16	0.04	

*Note.* <sup>a</sup> $\chi^2$  adjusted for ties; Wording of question: For each of the following reasons, please indicate how important it was for your decision to accept Deal [A/B] by Proposer [1/2] in Round 1; Item 1: "Reduce the inequality of payments between participants."; Item 2: "Punish Proposer [A/B] for choosing a deal associated with a loss for Third Party [1/2]"; Item 3: "Reduce or eliminate the loss for the third parties."; Item 4: "Increase the probability of an 'Only [A/B]' scenario in Round 2."; Item 5: "Not 'get my hands dirty' by accepting an offer that reduced the payment of a third party."; Item 6: "Maximize my own payoff."; Item 7: "Reward Proposer [A/B] for an offer that was financially advantageous for me."; All questions were answered on a 7-point Likert-type scale, where 1 is "not important at all" and 7 is "very important."

We find that, once uncertainty is introduced, selfish behavior increased by about one-fourth. Interestingly, neither the degree of being pivotal, nor which specific excuse can be applied, nor the scope of uncertainty seem to affect this result. This inelasticity to changes in risk can be seen as an indication that people employ self-serving interpretations of risk and can thus be interpreted as excuse-driven behavior (Haisley and Weber 2010; Exley 2015). The only other study that directly compares the bystander and replacement excuse also finds no significant difference between the two (Brütt et al. 2020). Regarding pivotality, our results stand in contrast to Falk et al. (2020), who find that perceived diffusion of being pivotal in group decisions is significantly correlated with selfish behavior. Differences in the decision context between their and our their study may explain these results. Ex-ante pivotality in groups is ambiguous and depends on subjective beliefs about the behavior of other group members. A belief of high pivotality thus corresponds to the belief that a relatively large share of group members votes against the externality. That is, perceptions of ex-ante pivotality reflect the perceived social norm and level of morality in the group. Beliefs about pivotality thus “carry” social information. In contrast, in our experiment, uncertainty, and hence pivotality, is the result of moves of nature. Thus, our results indicate that depending on how strongly perceptions of pivotality are imbued with social information, behavior may vary.

We find that behavior is inelastic to changes in uncertainty. However, once subjects have no impact on the externality, selfishness increases significantly. Nevertheless, 28 % of subjects forgo payoff and thus seem to follow deontological moral reasoning. The results of other studies also suggest deontological reasoning (Casal et al. 2019; Falk et al. 2020), but the precise numbers are probably context-specific and therefore difficult to compare. Falk et al. (2020) e.g. find that about 18 % of subjects who believe that they are not pivotal at all in a simultaneous group decision forgo payoff by voting against the externality. One implication of our finding of a significant increase in selfishness at a pivotality of zero is that policies or factors that make consumers feel that they have at least some influence on market outcomes may lead to a reduction in selfish behavior. In markets, perceptions of consumer pivotality may arise simply due to observation of changes in supply. In addition, advances in information and communication technologies allow consumers to coordinate more easily and may increase perceptions of pivotality. In a similar vein, Brütt et al. (2020) find that subjects have a strong tendency to self-select into group decision-making if they can guarantee

a pro-social outcome. This indicates that individuals seek decision environments that reduce uncertainty and in which they feel that their decision has an impact.

In our *Strategic ignorance* condition, two-thirds of subjects do not use the opportunity to remain ignorant about the consequences of their actions on third parties. Overall, there is no significant effect of *Strategic ignorance* on selfish behavior. This is in contrast to findings in previous literature. One explanation could be that subjects in our experiment had no information about the possible direction of influence on third parties. The high percentage of information disclosure may thus be explained by ambiguity. In studies finding widespread strategic ignorance, like Dana et al. (2007), Bartling et al. (2014), and Grossman and Weele (2016), subjects know the potential consequences for third parties; only which action leads to which outcome is hidden information. Thus, for decisions where individuals have no prior information or expectations about the potential impact of their actions on others, strategic ignorance may not be that prevalent. From a policy perspective, our result implies that a large fraction of consumers are interested in the consequences of their actions for others and that information disclosure policies that allow easy access to information at the time of purchase (Weil et al. 2006), e.g. standards and labels, can have a considerable impact on consumer behavior.

We see a couple of interesting avenues for future research. First, in this study, we examined comparatively large changes in pivotality from one hundred percent down to fifty, ten, and zero percent. Consistent with excuse-driven behavior, our results indicate an inverse S-shaped relationship between pivotality and selfish behavior. A more fine-grained examination of this relationship may further support this finding and could ultimately reveal differences between the replacement and bystander excuse, as well as a combination of the two. Second, changing the scope of uncertainty had no influence in our experimental setup. However, we only considered one variation in scope. This finding could also be explored in a more fine-grained study. Third, the role of social information for the perception of ex-ante causal responsibility represents an interesting area of research. How decision-makers perceive social cues may depend on the decision environment, such as the (shared) identity with other individuals or the importance of moves by nature compared to human decisions. Moreover, belief formation and susceptibility to social information may be correlated with personality traits. Fourth, the high proportion of information disclosure in our experiment underscores the dependence of strategic ignorance on the experimental setup and specifically on the information that decision-makers have about the possible outcomes

for others. Future research could examine how strategic ignorance changes with different prior information about others' outcomes, including uncertainty or vague cues about the direction of influence. Rarely do people know the exact outcomes for others.

## **Acknowledgements**

Daniel Hermann acknowledges financial support from Deutsche Forschungsgemeinschaft (grant number: HE 8478/1-1)

## References

- Bandura, A. (1999). "Moral Disengagement in the Perpetration of Inhumanities". In: *Personality and Social Psychology Review* 3.3, pp. 193–209.
- Bartling, B., F. Engl, and R. A. Weber (2014). "Does willful ignorance deflect punishment? – An experimental study". In: *European Economic Review* 70, pp. 512–524.
- Bartling, B. and U. Fischbacher (2012). "Shifting the Blame: On Delegation and Responsibility". In: *The Review of Economic Studies* 79.1, pp. 67–87.
- Bartling, B., R. A. Weber, and L. Yao (2015). "Do Markets Erode Social Responsibility?" In: *The Quarterly Journal of Economics* 130.1, pp. 219–266.
- Behnk, S., L. Hao, and E. Reuben (2017). *Partners in Crime: Diffusion of Responsibility in Antisocial Behaviors*. IZA Discussion Papers 11031. Institute of Labor Economics (IZA).
- Brütt, K., A. Schram, and J. Sonnemans (2020). "Endogenous group formation and responsibility diffusion: An experimental study". In: *Games and Economic Behavior* 121, pp. 1–31.
- Casal, S., F. Fallucchi, and S. Quercia (2019). "The role of morals in three-player ultimatum games". In: *Journal of Economic Psychology* 70, pp. 67–79.
- Chen, D. L., M. Schonger, and C. Wickens (2016). "oTree—An open-source platform for laboratory, online, and field experiments". In: *Journal of Behavioral and Experimental Finance* 9, pp. 88–97.
- Coffman, L. C. (2011). "Intermediation Reduces Punishment (and Reward)". In: *American Economic Journal: Microeconomics* 3.4, pp. 77–106.
- Dana, J., R. A. Weber, and J. X. Kuang (2007). "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness". In: *Economic Theory* 33.1, pp. 67–80.
- Darley, J. M. and B. Latane (1968). "Bystander intervention in emergencies: Diffusion of responsibility". In: *Journal of Personality and Social Psychology* 8.4, Pt.1, pp. 377–383.
- Engl, F. (2018). "A theory of causal responsibility attribution". In: *Available at SSRN* 2932769.
- Exley, C. L. (2015). "Excusing Selfishness in Charitable Giving: The Role of Risk". In: *The Review of Economic Studies* 83.2, pp. 587–628.
- Falk, A., T. Neuber, and N. Szech (2020). "Diffusion of Being Pivotal and Immoral Outcomes." In: *Review of Economic Studies* 87.5, pp. 2205–2229.
- Falk, A. and N. Szech (2013). "Morals and Markets". In: *Science* 340.6133, pp. 707–711.

- Fischer, P., J. I. Krueger, T. Greitemeyer, C. Vogrincic, A. Kastenmüller, D. Frey, et al. (2011). "The bystander-effect: A meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies." In: *Psychological Bulletin* 137.4, pp. 517–537.
- Gilmore James H., I. and B. J. Pine (1997). "The four faces of mass customization". In: *Harvard Business Review* 75, pp. 91–102.
- Grossman, Z. and J. J. van der Weele (2016). "Self-Image and Willful Ignorance in Social Decisions". In: *Journal of the European Economic Association* 15.1, pp. 173–217.
- Haisley, E. C. and R. A. Weber (2010). "Self-serving interpretations of ambiguity in other-regarding behavior". In: *Games and Economic Behavior* 68.2, pp. 614–625.
- Hamman, J. R., G. Loewenstein, and R. A. Weber (2010). "Self-Interest through Delegation: An Additional Rationale for the Principal-Agent Relationship". In: *American Economic Review* 100.4, pp. 1826–46.
- Hauser, O. P., D. G. Rand, A. Peysakhovich, and M. A. Nowak (2014). "Cooperating with the future". In: *Nature* 511.7508, pp. 220–223.
- Holweg, M. and F. K. Pil (2001). "Successful build-to-order strategies start with the customer". In: *MIT Sloan Management Review* 43.1, pp. 74–83.
- Irlenbusch, B. and D. J. Saxler (2019). "The role of social information, market framing, and diffusion of responsibility as determinants of socially responsible behavior". In: *Journal of Behavioral and Experimental Economics* 80, pp. 141–161.
- Norgaard, K. M. (2006). "'People Want to Protect Themselves a Little Bit': Emotions, Denial, and Social Movement Nonparticipation". In: *Sociological Inquiry* 76.3, pp. 372–396.
- Pigors, M. and B. Rockenbach (2016). "Consumer Social Responsibility". In: *Management Science* 62.11, pp. 3123–3137.
- Reczek, R. W., J. R. Irwin, D. M. Zane, and K. R. Ehrich (2018). "That's Not How I Remember It: Willfully Ignorant Memory for Ethical Product Attribute Information". In: *Journal of Consumer Research* 45.1, pp. 185–207.
- Rothenhäusler, D., N. Schweizer, and N. Szech (2018). "Guilt in voting and public good games". In: *European Economic Review* 101, pp. 664–681.
- Stoll-Kleemann, S., T. O'Riordan, and C. C. Jaeger (2001). "The psychology of denial concerning climate mitigation measures: evidence from Swiss focus groups". In: *Global Environmental Change* 11.2, pp. 107–117.
- Weil, D., A. Fung, M. Graham, and E. Fagotto (2006). "The effectiveness of regulatory disclosure policies". In: *Journal of Policy Analysis and Management* 25, pp. 155–181.

Ziegler, A., G. Romagnoli, and T. Offerman (2021). *Morals in multi-unit markets*. Tinbergen Institute Discussion Papers 20-072/I. Tinbergen Institute.



## Appendix

### A.1 Socio-demographics

**Table 5** — Education

What is the highest level of school you have completed or the highest degree you have received?	N	%
Less than high school degree	3	0.27
High school graduate (high school diploma or equivalent including GED)	70	6.22
Some college but no degree	221	19.64
Associate degree in college (2-year)	104	9.24
Bachelor's degree in college (4-year)	474	42.13
Master's degree	212	18.84
Doctoral degree	16	1.42
Professional degree (JD, MD)	25	2.22

**Table 6** — Employment status

Which statement best describes your current employment status?	N	%
Working (paid employee)	668	59.38
Working (self-employed)	177	15.73
Not working (temporary layoff from a job)	26	2.31
Not working (looking for work)	104	9.24
Not working (retired)	23	2.04
Not working (disabled)	12	1.07
Not working (other)	97	8.62
Prefer not to answer	18	1.60

**Table 7** — Ethnicity

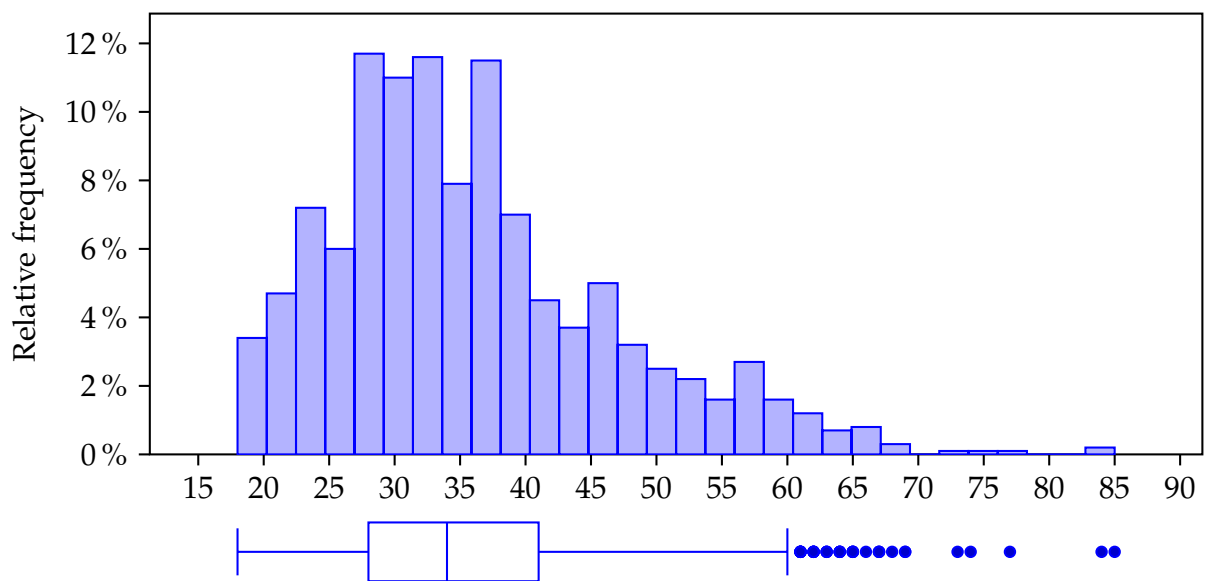
Choose one or more ethnicities that you consider yourself to be:	N	%
African American	86	7.64
American Indian	16	1.42
Asian	109	9.69
Hispanic	58	5.16
White Caucasian	918	81.60
Other	13	1.16

**Table 8** — Political spectrum

Where would you classify yourself on the left/right political spectrum?	N	%
1 - left-wing	152	13.51
2	199	17.69
3	172	15.29
4	110	9.78
5	220	19.56
6	104	9.24
7	83	7.38
8	58	5.16
9 - right-wing	27	2.40

**Table 9** — Relative income

How do you think your income and financial situation currently compare to those of others in the U.S. who are of similar age?	N	%
Don't know/No answer	25	2.22
Much below average	181	16.09
Somewhat below average	266	23.64
About the average	355	31.56
Somewhat above average	267	23.73
Much above average	31	2.76



**Figure 4** — Age of participants

*Note.* Subjects were asked about their year of birth. To approximate age, we assumed that subjects who participated in December 2020 already had a birthday. For subjects who participated between January and February 2021, we assumed the opposite.

## A.2 Data quality checks

### *Quality of answers to an open-ended question*

As a test of the quality of our data, we asked subjects to explain their decision using an open-ended question in the post-experimental questionnaire. Overall, the data quality for this question is quite good, both linguistically and in terms of scope. The vast majority of participants gave comprehensible reasons for their decisions, which demonstrated their understanding of the experiment and payoff scheme. There were no copied texts or meaningless combinations of strings. Less than 5 percent gave answers that were inconsistent with their choices in the experiment or that showed a misunderstanding of the experiment. On average, subjects wrote 18 words (with a maximum of 151) or 94 characters (with a maximum of 820). Subject that chose the unselfish offer wrote significantly more (104.81 vs. 86.13 characters;  $t$ -stat = 3.71,  $p$ -value < 0.00).

### *Sensitivity to changes in incentives*

To test the validity of our experimental set-up and subjects' understanding of the instructions and experiment, we tested subjects' sensitivity to changes in incentives under full pivotality in MTS. We ran an additional study with three variations of payoffs for the third party associated with the selfish offer: 70 Coins (as in the main study), 40 Coins, and 10 Coins. We collected 50 observations per treatment. We preregistered this study at <https://aspredicted.org/blind.php?x=xr7e79>. The proportion of selfish decisions is 56, 38, 28 percent for the third-party payoffs of 70, 40, 10. Pearson  $\chi^2 = 8.34$  with  $p = 0.015$ .

### A.3 Behavioral intentions

**Table 10** — Percentage of *selfish* subjects reporting reason as most important

Behavioral intention	%
Maximize own payoff <sup>6</sup>	61.81
Impact on second round <sup>4</sup>	15.18
Reward unfair proposer <sup>7</sup>	13.04
Reduce inequality <sup>1</sup>	5.21
Reduce externality <sup>3</sup>	2.91
Deontological reasoning <sup>5</sup>	1.23
Punish unfair proposer <sup>2</sup>	0.61

*Note.* Ties were randomly broken; Wording of question: For each of the following reasons, please indicate how important it was for your decision to accept Deal [A/B] by Proposer [1/2] in Round 1; <sup>1</sup>: “Reduce the inequality of payments between participants.”; <sup>2</sup>: “Punish Proposer [A/B] for choosing a deal associated with a loss for Third Party [1/2]”; <sup>3</sup>: “Reduce or eliminate the loss for the third parties.”; <sup>4</sup>: “Increase the probability of an ‘Only [A/B]’ scenario in Round 2.”; <sup>5</sup>: “Not ‘get my hands dirty’ by accepting an offer that reduced the payment of a third party.”; <sup>6</sup>: “Maximize my own payoff.”; <sup>7</sup>: “Reward Proposer [A/B] for an offer that was financially advantageous for me.”; All questions were answered on a 7-point Likert-type scale, where 1 is “not important at all” and 7 is “very important.”

**Table 11** — Percentage of *selfish* subjects reporting reason as least important

Behavioral intention	%
Punish unfair proposer <sup>2</sup>	33.74
Deontological reasoning <sup>5</sup>	24.08
Reduce inequality <sup>1</sup>	13.65
Reduce externality <sup>3</sup>	13.65
Impact on second round <sup>4</sup>	7.36
Reward unfair proposer <sup>7</sup>	6.60
Maximize own payoff <sup>6</sup>	0.92

*Note.* Ties were randomly broken; Wording of question: For each of the following reasons, please indicate how important it was for your decision to accept Deal [A/B] by Proposer [1/2] in Round 1; <sup>1</sup>: "Reduce the inequality of payments between participants."; <sup>2</sup>: "Punish Proposer [A/B] for choosing a deal associated with a loss for Third Party [1/2]"; <sup>3</sup>: "Reduce or eliminate the loss for the third parties."; <sup>4</sup>: "Increase the probability of an 'Only [A/B]' scenario in Round 2."; <sup>5</sup>: "Not 'get my hands dirty' by accepting an offer that reduced the payment of a third party."; <sup>6</sup>: "Maximize my own payoff."; <sup>7</sup>: "Reward Proposer [A/B] for an offer that was financially advantageous for me."; All questions were answered on a 7-point Likert-type scale, where 1 is "not important at all" and 7 is "very important."

**Table 12** — Percentage of *unselfish* subjects reporting reason as most important

Behavioral intention	%
Reduce externality <sup>3</sup>	31.71
Reduce inequality <sup>1</sup>	30.02
Impact on second round <sup>4</sup>	13.53
Deontological reasoning <sup>5</sup>	11.63
Maximize own payoff <sup>6</sup>	8.46
Punish unfair proposer <sup>2</sup>	2.96
Reward unfair proposer <sup>7</sup>	1.69

*Note.* Ties were randomly broken; Wording of question: For each of the following reasons, please indicate how important it was for your decision to accept Deal [A/B] by Proposer [1/2] in Round 1; <sup>1</sup>: "Reduce the inequality of payments between participants."; <sup>2</sup>: "Punish Proposer [A/B] for choosing a deal associated with a loss for Third Party [1/2]"; <sup>3</sup>: "Reduce or eliminate the loss for the third parties."; <sup>4</sup>: "Increase the probability of an 'Only [A/B]' scenario in Round 2."; <sup>5</sup>: "Not 'get my hands dirty' by accepting an offer that reduced the payment of a third party."; <sup>6</sup>: "Maximize my own payoff."; <sup>7</sup>: "Reward Proposer [A/B] for an offer that was financially advantageous for me."; All questions were answered on a 7-point Likert-type scale, where 1 is "not important at all" and 7 is "very important."

**Table 13** — Percentage of *unselfish* subjects reporting reason as least important

Behavioral intention	%
Punish unfair proposer <sup>2</sup>	34.46
Reward unfair proposer <sup>7</sup>	27.27
Maximize own payoff <sup>6</sup>	14.59
Impact on second round <sup>4</sup>	10.99
Deontological reasoning <sup>5</sup>	7.19
Reduce externality <sup>3</sup>	3.81
Reduce inequality <sup>1</sup>	1.69

*Note.* Ties were randomly broken; Wording of question: For each of the following reasons, please indicate how important it was for your decision to accept Deal [A/B] by Proposer [1/2] in Round 1; <sup>1</sup>: "Reduce the inequality of payments between participants."; <sup>2</sup>: "Punish Proposer [A/B] for choosing a deal associated with a loss for Third Party [1/2]"; <sup>3</sup>: "Reduce or eliminate the loss for the third parties."; <sup>4</sup>: "Increase the probability of an 'Only [A/B]' scenario in Round 2."; <sup>5</sup>: "Not 'get my hands dirty' by accepting an offer that reduced the payment of a third party."; <sup>6</sup>: "Maximize my own payoff."; <sup>7</sup>: "Reward Proposer [A/B] for an offer that was financially advantageous for me."; All questions were answered on a 7-point Likert-type scale, where 1 is "not important at all" and 7 is "very important."

## A.4 Experimental instructions

*Text*

*Group composition and role assignment*

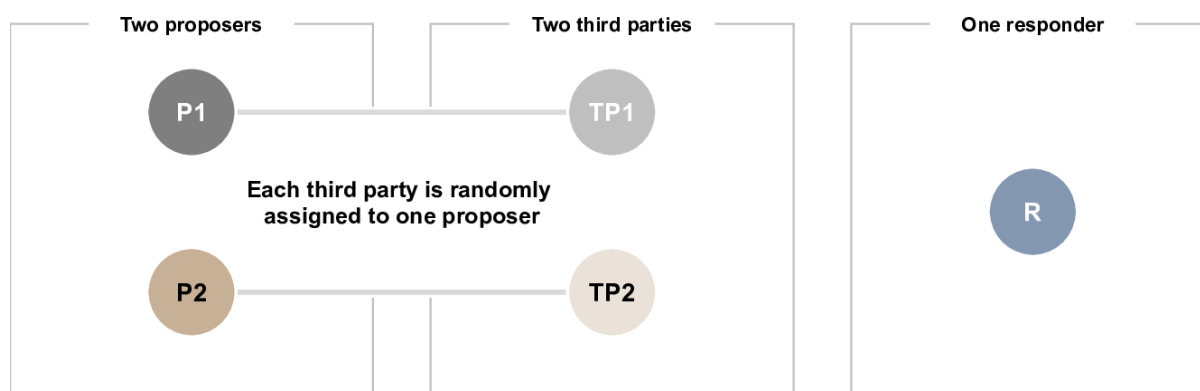
After you have read the instructions and answered the comprehension questions correctly, you will be matched into groups of five participants.

All participants in your group will be assigned to one of three different roles. In each group, there are

- two proposers (P1 and P2),
- two third parties (TP1 and TP2) and
- one responder (R).

Each third party is randomly assigned to one proposer. To indicate this assignment, we will denote the two pairs as “Third Party 1, Proposer 1” and “Third Party 2, Proposer 2”. The figure below illustrates the group composition.

The group composition and role assignment remain unchanged throughout the experiment.



*The task*

This task lasts for two rounds. In each round, each participant initially receives 100 coins.



The final earnings of all participants per round may differ from this initial endowment and depend on the decisions of the two proposers and the responder. Third parties cannot make decisions.

Each round comprises two consecutive steps:

- **Step 1:** Both proposers offer a deal to the responder.
- **Step 2:** The responder decides which of the two offered deals he/she wants to accept.

Deals are always beneficial for proposers and the responder, but they can also affect the earnings of third parties. The earnings of a third party are [MTO: only affected by a deal if it is accepted by the responder; MTS: affected by a deal as soon as it is offered by a proposer. Therefore, irrespective of the responder's choice, the earnings of third parties are determined by the deals offered by the proposers.]

The following illustration shows the two consecutive steps of the deal-making process and at which step payment consequences occur for each participant. As an example, the illustration assumes that the responder accepts the deal offered by Proposer 1.

[MTO: Figure 5; MTS: Figure 6]

#### *The two deals*

Proposers and responder can make two types of deals: [Deal A](#) and [Deal B](#). The table below shows the payment consequences of the deals and the resulting earnings per round, taking into account the initial endowment of 100 coins. [Strategic ignorance: As you can see from the table, the third parties' values are not visible.]

After the comprehension questions but before any decisions are made, each participant has the opportunity to reveal the payment consequences of the two deals for the third party. A participant's decision whether to reveal or not is not shared with other participants.]

[Full information: Table 14; Strategic ignorance: Table 15]

#### *Earnings per round*

**Proposers:** The proposer of the accepted deal receives 110 coins (+10 coins). The proposer of the rejected deal, keeps the initial endowment of 100 coins ( $\pm 0$  coins).

**Third parties:** Third party earnings are [MTO: only affected by a deal if it is accepted by the responder. If a responder accepts Deal A, the assigned third party receives [Full information: 70 coins (-30 coins); Strategic ignorance: ??? coins (??? coins)]. If a responder accepts Deal B, the assigned third party receives [Full information: 100 coins ( $\pm 0$  coins); Strategic ignorance: ??? coins (??? coins)]. The third party assigned to the rejected deal remains unaffected and keeps the initial endowment of 100 coins ( $\pm 0$  coins); MTS: affected by a deal as soon as it is offered by a proposer and therefore they do not depend on the responder's choice. If a proposer offers Deal A, the assigned third party receives [Full information: 70 coins (-30 coins); Strategic ignorance: ??? coins (??? coins)]. If a proposer offers Deal B, the assigned third party receives [Full information: 100 coins ( $\pm 0$  coins); Strategic ignorance: ??? coins (??? coins)].]

**Responder:** If the responder accepts Deal A, he/she receives 120 coins (+20 coins). If the responder accepts Deal B, he/she receives 110 coins (+10 coins).

### *Offers in Round 2*

While deals offered in Round 1 are deliberately chosen by proposers, the deals offered in Round 2 are determined for the proposers by a computerized urn draw. The computerized urn draw can lead to one of two possible scenarios:

- **Only A:** Both proposers offer Deal A
- **Only B:** Both proposers offer Deal B

### *Composition of the urn and the responder's influence on the offers in Round 2*

The urn is filled with 20 balls. Each ball is labeled with either A or B and has the same chance of being drawn. If an A ball is randomly drawn, Round 2 will be the Only A scenario. If a B ball is randomly drawn, Round 2 will be the Only B scenario.

The following illustration shows the composition of the urn depending on the responder's choice in Round 1. In addition, the probabilities for both scenarios are shown.

[Figure 7 for  $\alpha = 0.1$  and  $\beta = 0.5$  as an example]

If the responder chooses Deal A in Round 1, Round 2 will be either an Only A scenario with a probability of 55 percent or an Only B scenario with a probability of 45 percent.

If the responder chooses **Deal B** in Round 1, Round 2 will be either an **Only A** scenario with a probability of 45 percent or an **Only B** scenario with a probability of 55 percent.

*Tables and figures*

**Table 14** — Tabular representation of the payoffs associated with the deals in *Full information*

	Deal A		Deal B	
	Payment consequences	Earnings	Payment consequences	Earnings
Proposer	+10 coins	110 coins	+10 coins	110 coins
Third Party	-30 coins	70 coins	±0 coins	100 coins
Responder	+20 coins	120 coins	+10 coins	110 coins

**Table 15** — Tabular representation of the payoffs associated with the deals in *Strategic ignorance*

	Deal A		Deal B	
	Payment consequences	Earnings	Payment consequences	Earnings
Proposer	+10 coins	110 coins	+10 coins	110 coins
Third Party	??? coins	??? coins	??? coins	??? coins
Responder	+20 coins	120 coins	+10 coins	110 coins

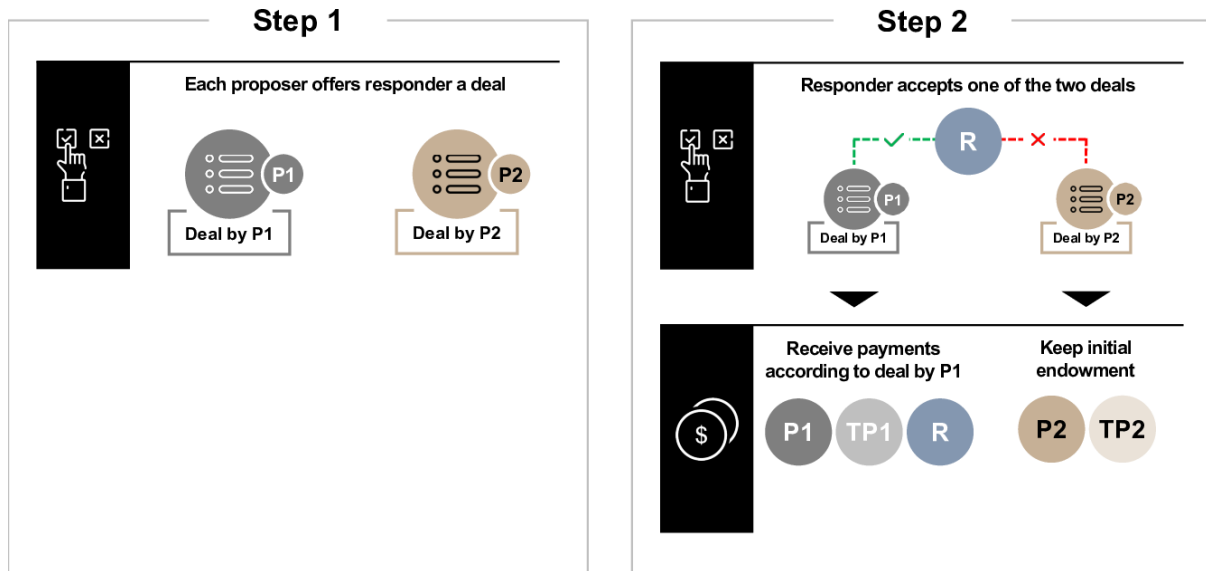


Figure 5 — Graphical representation of the steps in *MTO*

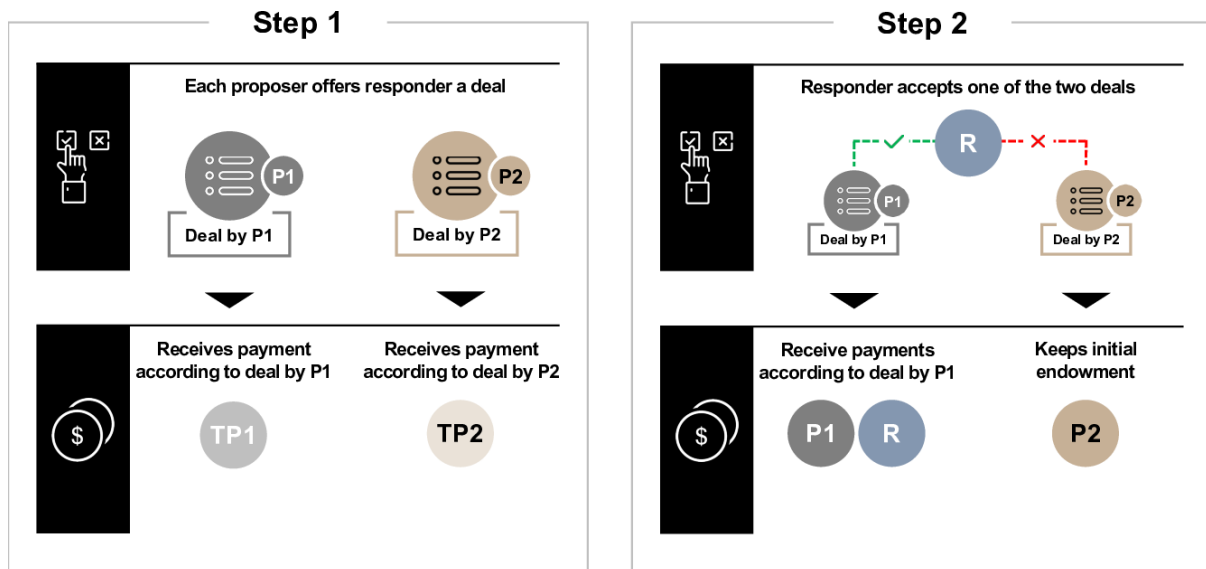


Figure 6 — Graphical representation of the steps in *MTS*

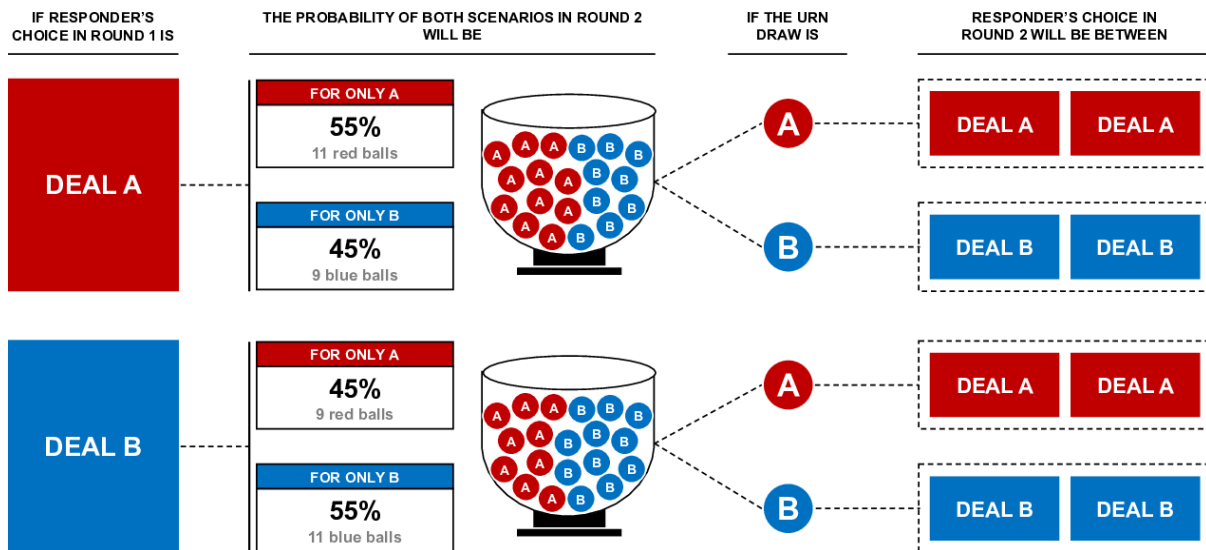


Figure 7 — Graphical representation of the urn drawing for  $\alpha = 0.1$  and  $\beta = 0.1$

## A.5 Comprehension questions

**Table 16** — Comprehension questions

Question	Type
<p>At which step of the deal-making process do third parties bear the payment consequences of a deal?</p> <ul style="list-style-type: none"> <li>• Third parties bear no payment consequences at all.</li> <li>• At step 1, as soon as a proposer offers a deal.</li> <li>• At step 2, only if the responder accepts a deal.</li> </ul>	Single-answer multiple choice question
<p>The deals offered in Round 1 are ...</p> <ul style="list-style-type: none"> <li>• deliberately chosen by proposers</li> <li>• determined by a computerized urn drawing</li> <li>• negotiated between proposers and third parties</li> </ul>	Single-answer multiple choice question
<p>Imagine that in Round 1 Proposer 1 offers Deal A and Proposer 2 offers Deal B. The responder accepts Deal B by Proposer 2. How many Coins does each participant earn in Round 1?<sup>a</sup></p> <ul style="list-style-type: none"> <li>• Responder</li> <li>• Proposer 1</li> <li>• Third Party 1</li> <li>• Proposer 2</li> <li>• Third Party 2</li> </ul>	Dropdown question; List of answers ranges from 0 to 150 Coins in 10 Coin increments and “???”

*continued*

**Table 16** — *continued*

Question	Type
<p>Imagine an Only A scenario, where a responder accepts the deal offered by Proposer 1. How many Coins do the third parties earn in that round?<sup>a</sup></p> <ul style="list-style-type: none"> <li>• Third Party 1</li> <li>• Third Party 2</li> </ul>	<p>Dropdown question; List of answers ranges from 0 to 150 Coins in 10 Coin increments and “???”</p>
<p>Imagine an Only B scenario, where a responder accepts the deal offered by Proposer 1. How many Coins do the third parties earn in that round?<sup>a</sup></p> <ul style="list-style-type: none"> <li>• Third Party 1</li> <li>• Third Party 2</li> </ul>	<p>Dropdown question; List of answers ranges from 0 to 150 Coins in 10 Coin increments and “???”</p>
<p>Which scenario, i.e. which combination of deals is possible in Round 2?<sup>b</sup></p> <ul style="list-style-type: none"> <li>• Deal A and Deal B: One proposer offers Deal A and the other proposer offers Deal B</li> <li>• Only A: Both proposers offer Deal A</li> <li>• Only B: Both proposers offer Deal B</li> </ul>	<p>Multiple choice question</p>

*continued*

**Table 16** — *continued*

Question	Type
If the responder chooses Deal A in Round 1, what is the probability of an Only A scenario in Round 2? <sup>c</sup>	Open-ended question, answers were limited to integers between 0 and 100
If the responder chooses Deal B in Round 1, what is the probability of an Only B scenario in Round 2? <sup>c</sup>	Open-ended question, answers were limited to integers between 0 and 100

*Note.* For some questions, we provided an additional hint, which was shown directly below the question in italics: <sup>a</sup>“Please take into account the initial endowment of 100 Coins, i.e. if a participant gains 10 Coins as a result of a deal, choose ‘110’. Remember that Third Party 1 is assigned to Proposer 1 and Third Party 2 is assigned to Proposer 2. [*Strategic ignorance*: If a payment is unknown, please choose the option ‘???’.]”, <sup>b</sup>“Remember that the deals offered by the proposers in Round 2 are determined by a computerized urn drawing.”, <sup>c</sup>“Please enter the probability in percent. Remember that the deals offered by the proposers in Round 2 are determined by a computerized urn drawing.”



## CHAPTER IV

### **Dishonest online: A distinction between observable and unobservable lying**

Daniel Hermann, and Mattheus Brenig

This paper investigates dishonesty in an online experiment. A distinction is made between lies that are observable to the experimenter and those that are not. Participants are asked whether (1) a randomly generated number between 0 and 9 and (2) their mother's birth year is even or odd; only the answer "even" is associated with a payoff. Responses of "even" are significantly more frequent when lying is unobservable. The order of these two questions significantly affects only the frequency of observable lying.

## 1 Introduction

In many decisions, lying can result in an economic advantage for the dishonest person. Examples range from dishonesty by entire industries such as the emission control defeat device in the automotive sector, to individual lies such as tax evasion. Due to technological advances, economic interactions are increasingly characterized by (1) reduced physical interaction and (2) a greater number of potential trading partners; as opposed to repeated face-to-face interactions. This could reduce reputational and social identity concerns that otherwise constrain lying. More recently, the ongoing digitization of economic transactions has been catalyzed by the COVID-19 pandemic.

While there is interest in dishonesty and lying in the economic literature, few studies investigate such behavior online (Gerlach et al. 2019; Gomes et al. 2021). This study presents the results of an experiment conducted on the crowdsourcing platform “MTurk” that tests for differences between observable and unobservable lying in an online setting. Due to the absence of a physically present experimenter, social identity concerns of observable lying could be reduced. In addition, due to the flexibility and short-term nature of work arrangements in online labor markets (see Dube et al. 2020), reputational concerns may also be reduced. Thus, in contrast to findings from laboratory experiments (see Gneezy et al. 2018; Abeler et al. 2019; Crede and Bieberstein 2020), the observability of lying may not matter in online settings.

This study adds four contributions to the literature: First, lying that is observable and unobservable to the experimenter in an online experiment is investigated. Second, an order effect for these two treatments is tested. In the literature, there is no comparison of order effects for observed and unobserved lying, while there are some indications that lying increases (slightly) in repeated one-shot decisions (Kocher et al. 2018; Ben-Ner and Hu 2021). Third, compared to previous studies (Gneezy et al. 2018; Abeler et al. 2019; Crede and Bieberstein 2020), the experimental procedure of measuring observed and unobserved lying is held constant, allowing for a true *ceteris paribus* comparison.<sup>1</sup> Finally, this study contributes to a small body of literature on lying in an online setting (Schild et al. 2020) and moves beyond the usual student sample.

---

<sup>1</sup>In Gneezy et al. (2018), participants either had to draw an object from an envelope (unobserved) or tick a box on a computer (observed). In Abeler et al. (2019), payments were either handed out by the experimenter (observed) or taken by participants (unobserved). In Crede and Bieberstein (2020), a die was rolled either physically (observed) or digitally (unobserved).

## 2 Experimental design and procedure

A simple experimental design is implemented in which participants are asked whether a number is odd or even (Heinicke et al. 2019). Thus, there are only two possible answers, of which only one is true. To provide an economic incentive for dishonesty, each answer of “even” is incentivized with \$0.5.<sup>2</sup> An answer of “odd” leads to no payment. Participants are asked for two different numbers specifying the treatments:

**Observable:** The last digit of a randomly generated completion code. Participants have to enter the completion code at the end of the experiment and are therefore aware that it is observable to the experimenters.

**Unobservable:** The year of birth of the participant’s mother. This is participants’ private information. There is no question regarding the birth year of the mother or any other relatives within the experiment.

Each participant is asked both questions on separate pages in random order.<sup>3</sup> To avoid strategic answers, we did not mention to participants ex-ante that they had to answer two incentivized questions.

The experiment was conducted online from November 2020 to February 2021 using the software oTree (Chen et al. 2016) via Amazon Mechanical Turk.<sup>4</sup> The design and hypotheses of this study are preregistered just like another part of the experiment, not dealing with research questions on honesty. This study was only available to US residents. A total of 1,130 participants completed the experiment and are considered for this study.<sup>5</sup> To describe our sample, some characteristics are illustrated in Table 1, Appendix A.1.

## 3 Results

To compare observable and unobservable lying, we use only the first randomly assigned question per participant. Therefore, we restrict our analysis to a between-

---

<sup>2</sup>Throughout the study, we used the term “Coins” for our experimental currency, with a conversion rate of 100Coins/\$1.

<sup>3</sup>For the verbatim questions see Appendix A.3.

<sup>4</sup>For a discussion of running experiments on Amazon Mechanical Turk please refer to Paolacci et al. (2010).

<sup>5</sup>For more detailed information on the experimental procedure see Appendix A.2.

subjects comparison and neglect within-subjects data that might be affected by order effects (which will be analyzed later).

In the *unobservable* treatment, the share of even answers has to be compared to an underlying uniform distribution.<sup>6</sup> In contrast, the percentage of participants lying about the *observable* completion code can be exactly quantified. Participants with an even completion code, which was the case for 52.73 %, had no incentive to lie.<sup>7</sup> A total of 268 participants were assigned an odd completion code, of which 107 or 39.93 % untruthfully reported “even”, while 60.07 % of answers were honest.

**Result 1** *If lying is observable, 39.93 % of participants lie.*

When asked as the first question ( $n = 563$ ), 82.24 % of participants report an even maternal birth year. A binomial-probability-test of the observed proportion against the expectation (50 %) reveals a significant ( $p < 0.001$ ) deviation.

**Result 2** *If lying is unobservable, 82.24 % of responses are “even”, which is significantly different from the expected relative frequency.*

Since even and odd numbers are equally likely in both the unobservable and observable treatment, a simple test for treatment differences is to directly compare the proportion of even responses. A binomial-probability-test reveals a significant difference ( $p < 0.001$ ) between these two relative frequencies (82.24 % vs. 71.60 %). A more accurate test takes into account the actual frequency of even numbers (52.73 %) in *observable*. More specifically, the assumed frequency of true “even” responses (50 %) has to be added to the observed frequency of dishonest “even” responses ( $39.93 \% \times 0.5$ ). Since an even completion code was randomly generated in a little more than half of the cases, the “naive” test slightly underestimates the actual difference between treatments (82.24 % vs. 69.97 %).

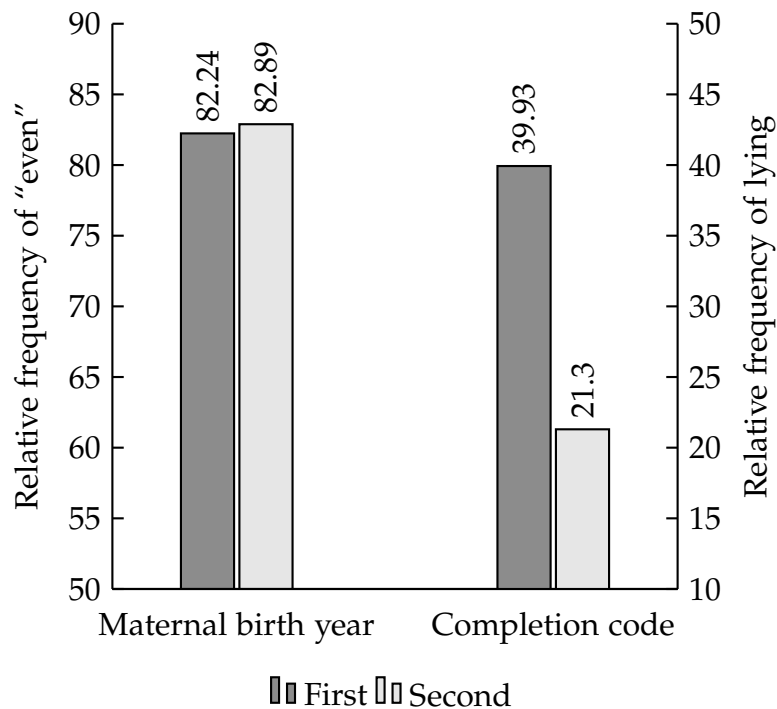
**Result 3** *Lying is more frequent in unobserved than in observed.*

Thus far, we have only compared answers to the first question. We now include the responses to the second question to investigate order effects. There may be order effects

<sup>6</sup>We do not correct for the marginal change of the uniform distribution through leap years. However, we tested for it and results do not differ.

<sup>7</sup>In the whole sample, there are four participants who state “odd”, while the last number was even. All four participants had 0 as last number. We assume that these participants are not sure whether 0 is odd or even, and therefore treat these observations as if the number was odd and therefore answered honestly.

that are specific to the degree of observability of a lie or simply because lying changes over repeated tasks. Figure 1 shows the relative frequency of “even” in *unobserved* and of lying in *observed* by order. There is no significant difference between the proportion



**Figure 1** — Proportion of “even” (lying) in *unobserved* (*observed*) by order

of “even” responses for maternal birth year by order (chi<sup>2</sup>-test;  $p = 0.772$ ). In contrast, there is an order effect for the question regarding the completion code. A statistical comparison of the share of lies—39.93 % if *observed* is the first question and 21.30 % if *observed* is the second question—reveals a significant difference (chi<sup>2</sup>-test;  $p < 0.001$ ).

**Result 4** *No order effect for the proportion of “even” responses in unobserved.*

**Result 5** *Lying in observed depends on the order, with more lying if the question is asked first.*

Another finding is that almost every participant who lies in *observed* reports an even number in *unobserved*, regardless of the order: 97.20 % and 98.31 % when *observed* is asked first and second, respectively. The proportion of “even” responses in *unobserved* is significantly (Fisher’s-Exact-Test;  $p < 0.001$ ) higher for liars in *observed* (97.59 %) compared to those honest in *observed* (79.98 %). Similarly, almost all participants

who report an odd maternal birth year report their completion code truthfully. The proportion of liars is significantly (Fisher's-Exact-Test;  $p < 0.001$ ) lower for participants who report odd in *unobserved* (4.49 %) compared to those who report even (35.53 %).

Although there is a significant difference between females and males for both observable and unobservable lying, the main results also hold for both genders.<sup>8</sup> Overall, males lie significantly more often than females in *observable* (35.71 % vs. 24.71 %; chi<sup>2</sup>-test,  $p = 0.006$ ). The gender difference is less pronounced in *unobservable* (84.78 % vs. 80.36 %; chi<sup>2</sup>-test,  $p = 0.051$ ), and a binomial-probability-test affirms that both genders lie significantly more in *unobservable* ( $p < 0.001$  for both genders). The main results for the order effect also replicate for both genders: there is no order effect in *unobservable*, and *observable* lying is significantly reduced when asked second. Interestingly, males lie significantly more than females when *observable* is asked first (50.37 % vs. 30.23 %; chi<sup>2</sup>-test,  $p = 0.001$ ), but not second (22.07 % vs. 19.23 %; chi<sup>2</sup>-test,  $p = 0.562$ ). Thus, the order effect is more pronounced for males.

## 4 Conclusion and Discussion

Online economic interaction leads to reduced physical and personal interaction and therefore might reduce reputational and social identity concerns for observable lies. The results of this study, however, contradict the expectation that the observability of a lie has no influence on the tendency to lie in an online setting. A lower proportion of dishonesty can be found in observable compared to unobservable lying. Thus, reputational concerns or the fear of negative consequences in any way matter in online experiments.

Furthermore, based on the observable (dis)honesty of the participants and assuming an even distribution of odd and even maternal birth years, three groups can be defined: (1) 31.20 %<sup>9</sup> of participants are always honest; (2) 28.99 %<sup>10</sup> are always dishonest, while

<sup>8</sup>Ten participants who "Prefer not to answer" the question about their gender are excluded from the following analysis.

<sup>9</sup>69.54 % of participants with an odd completion code report truthfully. Of those, 22.43 % report an odd maternal birth year. Assuming a uniform distribution of odd and even birth years, 22.43 of 50, or 44.86 % tell the truth. Consequently,  $69.54 \% \times 44.86 \% = 31.20 \%$  are honest under both degrees of observability.

<sup>10</sup>30.46 % of participants with an odd completion code report untruthfully. Of those, 2.41 % report an odd maternal birth year. Assuming a uniform distribution of odd and even birth years, only 2.41 of 50, i.e. 4.82 % tell the truth. Hence,  $30.46 \% \times (100 - 4.82) \% = 29.2 \%$  always lie.

for; (3) 39.81 % behavior depends on whether a lie is observable or not.<sup>11</sup> Thus, for the first two groups, our manipulation of observability did not change behavior, but for the third group it did. From previous studies, there is evidence that participants lie even though their behavior can be observed (see e.g. Kocher et al. 2018), suggesting that there may be individuals who lie under both conditions. However, comparisons of observable and unobservable dishonesty have previously been conducted between subjects. Thus, it has not been possible to actually identify groups of participants who lie in only one or both conditions in previous studies.

Another finding is that an order effect can only be found for observable dishonesty. Economically, this could be explained by higher (psychological) costs of observable compared to unobservable dishonesty and a negative marginal utility of earnings. If someone has already earned money through the first question (due to honest or dishonest response), the psychological cost of an observable lie to the second question may outweigh the lower marginal benefit of additional earnings. In contrast, the psychological costs of an unobservable lie may be comparatively small and always outweighed by its benefits. Another explanation could be that the order of the two questions produces a salience/contrast effect leading to varying perceived costs: the observability of a lie is more salient due to the unobservability of truthfulness of responses to the first question. Since our results indicate a pronounced order effect, future studies may want to investigate the causes of this result. Finally, our results do not support a relevant extent of (moral) restitution cleansing (see West and Zhong 2015) due to a second truthful answer. Such behavior would imply an order effect for the unobservable treatment as well.

## Acknowledgements

Daniel Hermann acknowledges financial support from Deutsche Forschungsgemeinschaft (grant number: HE 8478/1-1)

---

<sup>11</sup>These are almost exclusively participants who lie only when it is unobservable.

## References

- Abeler, J., D. Nosenzo, and C. Raymond (2019). "Preferences for Truth-Telling". In: *Econometrica* 87.4, pp. 1115–1153.
- Ben-Ner, A. and F. Hu (2021). "Lying in a finitely repeated game". In: *Economics Letters* 201, p. 109741.
- Chen, D. L., M. Schonger, and C. Wickens (2016). "oTree—An open-source platform for laboratory, online, and field experiments". In: *Journal of Behavioral and Experimental Finance* 9, pp. 88–97.
- Crede, A. K. and F. von Bieberstein (2020). "Reputation and lying aversion in the die roll paradigm: Reducing ambiguity fosters honest behavior". In: *Managerial and Decision Economics* 41.4, pp. 651–657.
- Dube, A., J. Jacobs, S. Naidu, and S. Suri (2020). "Monopsony in Online Labor Markets". In: *American Economic Review: Insights* 2.1, pp. 33–46.
- Gerlach, P., K. Teodorescu, and R. Hertwig (2019). "The truth about lies: A meta-analysis on dishonest behavior". In: *Psychological Bulletin* 145.1, pp. 1–44.
- Gneezy, U., A. Kajackaite, and J. Sobel (2018). "Lying Aversion and the Size of the Lie". In: *American Economic Review* 108.2, pp. 419–53.
- Gomes, H. S., D. P. Farrington, I. N. Defoe, and Â. Maia (2021). "Field experiments on dishonesty and stealing: what have we learned in the last 40 years?" In: *Journal of Experimental Criminology*, pp. 1–31.
- Heinicke, F., S. Rosenkranz, and U. Weitzel (2019). "The effect of pledges on the distribution of lying behavior: An online experiment". In: *Journal of Economic Psychology* 73, pp. 136–151.
- Kocher, M. G., S. Schudy, and L. Spantig (2018). "I lie? We lie! Why? Experimental evidence on a dishonesty shift in groups". In: *Management Science* 64.9, pp. 3995–4008.
- Paolacci, G., J. Chandler, and P. G. Ipeirotis (2010). "Running experiments on Amazon mechanical Turk". In: *Judgment and Decision Making* 5.5, pp. 411–419.
- Schild, C., M. Moshagen, K. A. Ścigała, and I. Zettler (2020). "May the odds — or your personality — be in your favor: Probability of observing a favorable outcome, honesty-humility, and dishonest behavior". In: *Judgment and Decision Making* 15.4, pp. 600–610.
- West, C. and C. B. Zhong (2015). "Moral cleansing". In: *Current Opinion in Psychology* 6, pp. 221–225.



## Appendix

### A.1 Socio-demographics

**Table 1** — Socio-demographics

Variable	Mean	SD	N
Female in %	49.55 <sup>a</sup>	—	1120 <sup>b</sup>
Age in years	36.80 <sup>c</sup>	10.90	1130
Income compared to others in the US <sup>d</sup>	2.73	1.09	1104 <sup>e</sup>

*Note.* <sup>a</sup> For comparison the average age in the U.S. is 38.4 years (United States Census Bureau, 2019); <sup>b</sup> Ten participants choose the option “Prefer not to answer”; <sup>c</sup> Subjects were asked about their year of birth. To approximate age, we subtracted the response from the year of data collection; For comparison the percentage of females in the U.S. is 50.8% (United States Census Bureau, 2019); <sup>d</sup> Response to the question: “How do you think your income and financial situation currently compare to those of others in the U.S. who are of similar age?” on a 5-point-Likert-scale: 1 = Much below average; 2 = Somewhat below average; 3 = About the average; 4 = Somewhat above average; 5 = Much above average; <sup>e</sup> Twenty six participants choose the option “Don’t know/no answer”.

### A.2 Experimental procedure

This experiment was part of the post-experimental questionnaire of an otherwise unrelated experiment. We prevented retakes and restricted participation to US residents. In addition, due to readability issues, we excluded participants who reported accessing the study via a smartphone or tablet. After reading a plain language statement and giving consent to participate, subjects read the instructions for the (unrelated) main task. Subjects had a maximum of two attempts to answer a few comprehension questions. Subjects who failed twice were excluded from participation. After completion of the main task and as part of the questionnaire, two questions relevant to this study were presented on separate web-pages in a randomized order.

Participants that finished the questionnaire received a fixed payment of \$1 in addition to the earnings based on their decisions in the main task and the two honesty questions. Participants earned an average of \$3.05 or \$/h13.30.

### A.3 Verbatim questions

The following question is about a specific number. Your task is to report whether the answer to the question is odd or even. Your answer determines your payoff for this task. You receive a payment of

- 50 Coins, if you report “even” and
- 0 Coins, if you report “odd”.

[*observable*: This is your completion code for this HIT: [Completion code]. Don’t worry, you don’t have to remember your completion code. It will be displayed again on the last page of this HIT. Is the last number of your completion code odd or even?]

[*unobservable*: Please think of the birth year of your mother. Is the last number of this year odd or even?]

## CHAPTER V

### **Can Gamification lead to Increased Paid Crowdworkers' Output?**

Sascha Lichtenberg, Tim-Benjamin Lembcke, Mattheus Brenig,  
Alfred Benedikt Brendel, and Simon Trang

Gamifying serious work environments, such as paid crowdsourcing platforms, potentially increases crowdworkers' task motivation, engagement and enjoyment. This, in turn, can lead to a higher willingness to contribute, higher quality of work and long-term engagement. However, it remains unclear how crowdworkers behave when gamification is applied to motivate them to do more tasks than being paid for. In this study, we conducted an experiment on Amazon Mechanical Turk to investigate this relationship in a controlled setting that allows us to isolate the effects of gamification. Using 320 crowdworkers, we examine the effect of different gamification affordances (progressbar, badges and leaderboard) on autonomous motivation and task performed. We find that some gamification affordances (namely badges and leaderboard) can lead crowdworkers to do more work than they are paid for. However, this is not necessarily associated with autonomous motivation, as we did not consistently observe an increase in autonomous motivation along with more tasks performed.

S. Lichtenberg et al. (2020). "Can Gamification lead to Increase Paid Crowdworkers Output?" In: *WI2020 Zentrale Tracks*. GITO Verlag, pp. 1188–1202.

## CHAPTER VI

### **Breathing life into consumer rights: Smartphone tools facilitating the “right to know” on substances of very high concern in REACH articles**

Julian Schenten, Mattheus Brenig, Martin Führ, and Kilian Bizer

#### *Background*

The EU chemicals regulation “Registration, Evaluation, Authorisation and Restriction of Chemicals” (REACH) aims to reduce usage of substances of very high concern (SVHCs) by firms. Therefore, a consumer right-to-know about SVHCs in articles is intended to create market-based incentives. However, awareness of the right-to-know among EU citizens is low. Moreover, the response window of 45 days afforded to suppliers impedes immediate, informed decisions by consumers. Consequently, despite being in effect for more than 10 years, only few consumers send requests. Civil society actors have developed smartphone applications reducing information search costs, allowing users to send right-to-know requests upon scanning an article’s barcode. Answers are stored in a database and made available to the public immediately. This paper assesses to which extent smartphone tools contribute to increased use of the right-to-know by undertaking a case study of the application “ToxFox” by the German non-profit organisation Bund für Umwelt und Naturschutz Deutschland (BUND).

#### *Results*

An analysis of the data from the BUND database for the period 2016 to 2018 reveals that about 20 thousand users have sent almost 49 thousand requests. This has led to more than 9 thousand database entries, including 189 articles which contain SVHCs above the legal threshold. The data also indicate that receiving information on requested articles encourages further use of the application. Many suppliers accept the application and pro-actively provide information on articles without SVHCs. However, most consumers use the application only for a short time, and suppliers are struggling to reply to right-to-know requests.

## *Conclusion*

Evaluating the results, the study identifies options to enhance the application's design in terms of user motivation and legal certainty, and to enhance the framework governing barcode assignments to articles with a view to better contributing to transparency. As for policy implications, a lack of consumer requests can in part be traced back to design flaws of the right-to-know and a lack of implementation and enforcement of REACH. In addition, suppliers have to increase their supply chain communication efforts to make sure they are in the position to properly answer consumer requests. We recommend several policy options addressing these and additional aspects, thus contributing to the legislative review of Art. 33 REACH.

J. Schenten et al. (2020). "Breathing life into consumer rights: smartphone tools facilitating the "right to know" on substances of very high concern in REACH articles". In: *Environmental Sciences Europe* 32.1, p. 114.

## CHAPTER VII

### **Refining tools to bridge the gap between academia and chemical regulation: perspectives for WikiREACH**

Marlene Ågerstrand, Mattheus Brenig, Martin Führ, and Julian Schenten

Regulatory hazard and risk assessments of chemical substances have to include all reliable and relevant data to be credible and complete. However, screening the literature for appropriate studies and extracting data is burdensome. Therefore, reducing impediments by making data easily and readily accessible to risk assessors could result in more comprehensive hazard and risk assessments. In this paper, we study WikiPharma, a database that aggregates ecotoxicity data for pharmaceuticals, extracted from peer-reviewed studies. The use of the WikiPharma database is explored to develop strategies on how similar tools can bridge between science and policy by providing risk assessors with easily accessible summary data. Specifically, adapting the concept of WikiPharma to industrial chemicals regulated under the REACH regulation is discussed. Experiences with WikiPharma show that there is interest in using peer-reviewed studies in regulatory decision-making. However, tools like WikiPharma require constant updates. Hence, as for “WikiREACH”, effective incentives are needed to motivate researchers to feed in relevant data for regulatory assessments. Besides, support by automated processes can aid in the labour-intensive activity of gathering data. To ensure that such a tool is continuously maintained and compatible with the regulatory system, and thereby useful for hazard and risk assessments of chemicals, it would benefit from being developed in collaboration with the major stakeholders in the field, i.e. regulatory agencies, academia, industry, scientific journals, and providers of research network platforms.

M. Ågerstrand et al. (2017). “Refining tools to bridge the gap between academia and chemical regulation: perspectives for WikiREACH”. in: *Environ. Sci.: Processes & Impacts* 19, pp. 1466–1473.

## **Declaration of contribution to each essays of the cumulative dissertation (based on the CRediT taxonomy by Brand et al. 2015<sup>1</sup>)**

To the six essays of the cumulative dissertation I contributed as follows:

1. Consumer Behavior under Benevolent Price Discrimination

*co-authored with Alexander Erlei, and Nils Engelbrecht:*

Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data curation, Writing (Original Draft), Visualization.

2. Excusing selfish behavior with (willfully induced) uncertainty:

An experimental study

*co-authored with Nils Engelbrecht, and Daniel Hermann:*

Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data curation, Writing (Original Draft), Visualization, Supervision, Project Administration.

3. Dishonest online: A distinction between observable and unobservable lying

*co-authored with Daniel Hermann:*

Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data curation, Writing (Original Draft), Visualization.

4. Can Gamification lead to Increased Paid Crowdworkers' Output?

*co-authored with Sascha Lichtenberg, Tim-Benjamin Lembcke, Alfred Benedikt Brendel, and Simon Trang:*

Conceptualization, Methodology, Software, Validation, Investigation, Resources, Data curation, Writing (Original Draft).

---

<sup>1</sup>A. Brand et al. (2015). "Beyond authorship: attribution, contribution, collaboration, and credit". In: *Learned Publishing* 28.2, pp. 151–155.

5. Breathing life into consumer rights: Smartphone tools facilitating the “right to know” on substances of very high concern in REACH articles

*co-authored with Julian Schenten, Martin Führ, and Kilian Bizer:*

Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data curation, Writing (Original Draft, Review & Editing), Visualization, Project Administration.

6. Refining tools to bridge the gap between academia and chemical regulation: perspectives for WikiREACH

*co-authored with Marlene Ågerstrand, Martin Führ, and Julian Schenten:*

Conceptualization, Writing (Original Draft, Review & Editing).

---

Date, Signature



**Ph.D. program in Economics**  
**Declaration for admission to the doctoral examination**

I confirm

1. that the dissertation that I submitted “Determinants of Other-Regarding Behavior and the Private Provision of Public Goods” was produced independently without assistance from external parties, and not contrary to high scientific standards and integrity,
2. that I have adhered to the examination regulations, including upholding a high degree of scientific integrity, which includes the strict and proper use of citations so that the inclusion of other ideas in the dissertation are clearly distinguished,
3. that in the process of completing this doctoral thesis, no intermediaries were compensated to assist me neither with the admissions or preparation processes, and in this process,
  - No remuneration or equivalent compensation were provided
  - No services were engaged that may contradict the purpose of producing a doctoral thesis
4. that I have not submitted this dissertation or parts of this dissertation elsewhere.

I am aware that false claims (and the discovery of those false claims now, and in the future) with regards to the declaration for admission to the doctoral examination can lead to the invalidation or revoking of the doctoral degree.

---

Date, Signature