The strategic role of information in markets and games

Essays in behavioral economics

Ziegler, A.G.B.

Publication date
2023

Document Version
Final published version

tinbergen institute

GRADUATE PROGRAM

The Strategic Role of Information in Markets and Games:

Essays in Behavioral Economics

Andreas G. B. Ziegler

Universiteit van Amsterdam

The Strategic Role of Information in Markets and Games

Andreas G. B. Ziegler

818

This thesis consists of three essays that study how the public revelation of information affects strategic interaction across three different environments. Using laboratory experiments to test theoretical conjectures, the essays analyze how humans process information, reason about how others process information, and what their choices reveal about the information they privately hold.

The first chapter compares auction formats that differ in the degree of information revealed to bidders during the auction. Bidders in the laboratory are not able to learn about the value of the item for sale from revealed information, and popular open auction formats trigger behavioral biases, which increases the auctioneer's revenues.

The second chapter studies how interacting in markets affects the participants' concern about causing externalities. Markets overall decrease our concern for damages to third parties. However, markets that allow each participant to trade repeatedly while not being pivotal for aggregate outcomes most strongly erode such concerns.

The third chapter studies how to best persuade a strategically interacting audience by using public or private communication. Empirically, public communication is particularly effective. In line with theoretical predictions, senders also benefit from adjusting their persuasion to the presence of coordination motives in the strategic interaction of the audience.

Andreas Ziegler holds a BSc in Management, Philosophy & Economics from the Frankfurt School of Finance & Management (2014), a Master in Economic Analysis from the Universidad Carlos III de Madrid (2016), and an MPhil in Economics from the Tinbergen Institute (2018). In 2018, he joined CREED at the Amsterdam School of Economics at the University of Amsterdam as a PhD student, supervised by Theo Offerman and Giorgia Romagnoli. Andreas currently works as a lecturer (assistant professor) at the University of Essex.

The Strategic Role of Information in Markets and Games: Essays in Behavioral Economics

ISBN: **978 90 361 0711**

Cover illustration: Andreas Ziegler using Midjourney

Cover design: Crasborn Graphic Designers bno, Valkenburg a.d. Geul

This book is no. **818** of the Tinbergen Institute Research Series, established through cooperation between Rozenberg Publishers and the Tinbergen Institute. A list of books which already appeared in the series can be found here.

The Strategic Role of Information in Markets and Games:
Essays in Behavioral Economics

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op vrijdag 8 september 2023, te 16.00 uur

door Andreas Günter Bernd Ziegler
geboren te Bayreuth

Promotiecommissie

Promotor:	prof. dr. T.J.S. Offerman	Universiteit van Amsterdam
Copromotor:	dr. G. Romagnoli	Universiteit van Amsterdam
Overige leden:	prof. dr. M. Bigoni	University of Bologna
	prof. dr. A.M. Onderstal	Universiteit van Amsterdam
	prof. dr. R. Sloof	Universiteit van Amsterdam
	prof. dr. J. Tuinstra	Universiteit van Amsterdam
	prof. dr. J. van de Ven	Universiteit van Amsterdam

Faculteit Economie en Bedrijfskunde

Acknowledgement of financial support

The research for this doctoral thesis received financial assistance from the Research Priority Area Behavioral Economics of the University of Amsterdam, the A Sustainable Future platform of University of Amsterdam, and the Amsterdam Center for Behavioural Change of the University of Amsterdam.

Acknowledgements

I started my PhD with high hopes, yet very much underestimated how much I would enjoy this period of my life. For this, I am deeply grateful to my two supervisors, Theo and Giorgia. I've learned a lot, often in unexpected ways and areas. I've learned a lot about being ambitious, about thinking deeply yet communicating clearly. Maybe even more importantly, I had a lot of fun with my ever-youthful advisors. I will miss stumbling into your always-open doors and soon or not-so-soon after leaving with a smile on my face. I feel very lucky to have had the opportunity to work with and learn from you.

Next to that, I am very grateful to everyone at CREED, UvA and beyond whom I met along my journey. While many faces changed during my years here, it was amazing to be part of a group so dedicated to running great experiments yet so welcoming that I will always fondly remember the conferences which felt like joint school trips, our outings and foosball matches. Thank you to Johan, Jan, Ollie, Ivo, Kostas, Julian, Nils, Matthieu, Jantsje, Sander, Charlotte, Shaul, Julian, Aljaz, Ailko, Boris, Frans, Matthijs, Linh and Andrej for making this time unforgettable. For all your advice, feedback and help on the job market, thank you to Natalie, Arthur, Randolph, Paul, Joep, Joël and especially Jeroen. With many people I met during my PhD, I also had a lot of fun away from work, when sharing drinks, dinners, bike trips, Bayern games, or unicorns and boardgames, with Dianna, Adam, Chris, Ayse, Jan, Thao, Margarita, Maria, Davide, Wally, Julia, Ivan and Oda, when traveling to Istanbul for Asli's and Rik's wedding, or when off-roading with Stephan and playing tennis with Huaiping. Finally, the office with the best view obviously had the best people, so thank you to Cristina, Alejandro and Junze for always solving all PhD problems that came up on the spot.

Working with the group at the NKI on actual policies for the market for pharmaceuticals, and running experiments with four labs synchronously, was extremely interesting and exciting. Thank you to you, Nora, for making this possible.

I also want to thank Uri for hosting me in San Diego. I, and especially my job market paper, benefitted a lot from the discussions thanks to that visit. It also

provided the opportunity to enjoy one-too-many tacos and many road trips to all the incredible nature around. Thank you also to Nathalie, Alex, Daniel and René for our time together.

Before starting the PhD, I by now fondly remember the days and nights at the Tinbergen Institute. That time, and the following years in Amsterdam, was a lot more fun thanks to the many friends I met there: Luisa, Paul, Diana, Eva, Alina, Leonard and Stefan.

My parents and my sister Katha were always there for me and remained curious and supportive about what I love to spend my time on, in spite of the increasing abstractness and number of equations. Thank you for that.

Most important during my time in Amsterdam, and long before that, were you, Kathi. I could always count on you to support me when I doubted myself the most, but also on shooting down my stupidest of ideas in seconds. I wouldn't be where I am without you. I can't wait for the next decades of life (and sometimes work) together.

Contents

CHAPTER 1

Introduction

We often possess only imprecise information about the world around us. Our ability to learn from the private information we acquire is crucial to navigating this uncertainty. While the processing of private information in individual decision-making has been extensively explored in behavioral economics (see Benjamin, 2019, for a review), less is known about how people reason about others' reasoning and the use of private information in strategic settings, as well as the impact of individual biases in such interactions. In this thesis, I employ laboratory experiments to investigate this aspect in three contexts: auctions, markets, and persuasion.

The common theme is the experimental analysis of mechanisms that publicly disclose private information to multiple agents. This information will be the value of items for sale in an auction, others' valuation of externalities in markets, or the strategic revelation of information to an audience. By using control treatments that keep information private, I can examine how the availability of public information influences behavior. Chapters 2 and 4 will also compare empirical behavior and outcomes to benchmarks from economic theory. In these benchmarks, agents are expected to make rational inference about others' private information, or to account for the fact that others will process their private and public information. The papers in this thesis use the laboratory to study the empirical counterparts of these settings.

As a first strategic environment, we focus on auctions in Chapter 2. We investigate explanations for the prevalence of open auction formats in practice. In open ascending auctions, also called English auctions, public information is disclosed through the visibility of others' bids to all bidders. Therefore, in the English auction, information is made public endogenously. In a control treatment, we investigate how bidders bid in sealed-bid auctions, in which they place their bids without observing others' choices. We focus on auctions where the item for sale has a value that is common to all bidders, with each bidder possessing uncertain information about that value. In this common-value setting, others' private information about the item's value is useful, as it enables bidders to estimate the common value more precisely. This private information can be revealed and inferred from the earlier, publicly observable bids in open auctions. The disclosure of information is the primary theoretical explanation for the prevalence of open ascending auctions in practice. According to the *linkage principle* (Milgrom and Weber, 1982), providing more information to bidders increases revenue, as greater precision reduces the winners' curse and permits the winning bidders to bid more assertively.

A competing explanation for the popularity of open auctions is their capacity to trigger behavioral biases. For instance, they may induce a quasi-endowment effect. Bidders who were the highest bidder for a period may attribute a higher value to the item for sale, akin to an endowment effect that arises for owned items (Thaler, 1980). As a result, these bidders might bid more than the item's underlying value warrants, thus raising revenue. Moreover, open auctions in practice permit bidders to submit jump bids, where a bidder significantly outbids the current highest bid. While jump bidding might deter other bidders from competing, it could also stimulate naïve overbidding due to an auction fever.

In our laboratory auctions, the highest auction revenue is observed in an open auction format resembling real-world auction designs. Our laboratory experiment enables us to rule out the theoretical literature's proposed mechanism for the increase in revenue, and the aggregation of information in open auctions. Contrarily, we find that less information is aggregated in open auctions compared to sealed-bid auctions, where bidders rely solely on their private information and cannot incorporate public information. This information aggregation failure occurs despite others' bids revealing valuable information, albeit less than what the theory predicts. Instead, the open auction's success is driven by behavioral factors: early bids are elevated by spiteful bidders, bidders with robust endowment effects lose the most money, and naïve jump bids placed by impatient bidders inflate prices.

In Chapter 3, we focus on markets that publicly disclose information through the endogenous formation of prices of goods. We are particularly interested in markets where trading generates negative externalities, causing harm to third parties. Examples of such markets include those for airplane tickets or meat, where the activities contribute to increased CO_2 emissions and global warming, or illegal weapons trading. Activity in these markets provides insights into the morality of their participants: are they willing to harm third parties in exchange for their own financial gains?

Importantly, the strategic interaction among market participants may influence the value we assign to the damages to these third parties. The mere act of engaging in market contexts could erode our morality (Falk and Szech, 2013). Furthermore, the strategic interaction in markets might introduce additional excuses for selfish behavior. For instance, the replacement logic suggests that selfish behavior in markets can be justified with the belief that others would step in and trade if one abstained from trading. As a result, one's actions would not impact aggregate outcomes or total negative externalities. The replacement logic becomes relevant in the unrestricted multi-unit markets we study, where each trader is permitted to trade additional units, regardless of their previous trading, up to the total tradable quantity in the market. In contrast, the previous literature only examined the erosion of morals in single-unit markets, where each trader is limited to trading a maximum of one unit. This feature prevents traders who have already traded from trading again, which means that these traders' activity cannot provide an excuse in the sense of the replacement logic. Therefore, multi-unit markets may both be particularly erosive to their participants' morality as well as realistic in many markets in practice, such as the markets for flights or illicit weapons trading. In a control treatment, we examine how people value the same negative externality in an individual decision-making task, without strategic interaction and without a market frame.

In a laboratory experiment, we compare behavior in multi-unit markets to single-unit markets and an individual control task. We observe a slight erosion of morals in single-unit markets compared to the control task, an erosive effect of markets that has recently been contested (Sutter, Huber, Kirchler, Stefan, and Walzl, 2020; Bartling, Fehr, and Özdemir, 2023). However, we find substantially stronger erosion in multi-unit than in single-unit markets. In fact, trading behavior in unrestricted multi-unit markets is consistent with a complete erosion of morals. Trading is indistinguishable from the selfish competitive equilibrium predictions, where traders are assumed to disregard the negative externalities they cause entirely. This profound erosion suggests that the current literature may have underestimated the extent to which markets can erode morals, by neglecting the fact that in most markets traders can trade multiple units.

We find that the replacement logic can account for this complete erosion of morals. A vast majority of participants continue to trade units even when the moral costs associated with causing negative externalities, measured in individual decision-making, would far outweigh the minimal gains from trade that can be made. This widespread trading is consistent with the idea that everyone's selfish trading is acceptable since everyone else is trading. Each trader's actions do not affect aggregate outcomes or total negative externalities. Crucially, this behavior occurs only in unrestricted multi-unit markets, which fully enable the use of replacement logic, but not in single-unit markets, which limit the power of this excuse. Moreover, the market composition does not impact trading behavior. Participants interacting with homogeneous groups are as selfish as those interacting in heterogeneous groups. This behavior is consistent with the erosion driven by the replacement logic but contradicts competing explanations, where the most selfish traders may disproportionately affect trading behavior only in heterogeneous groups. Additionally, we demonstrate that the market participants' beliefs align with replacement logic reasoning. First, traders in unrestricted multi-unit markets accurately anticipate more active others. Second, those traders who believe that more others are active are also more likely to trade themselves.

In Chapter 4, I explore how to best persuade audiences comprised of multiple receivers. Think, for example, about a government that may want to convince citizens to adopt measures against the spread of a disease, such as vaccinations, or managers who may want to motivate employees to increase their effort. When multiple receivers can be persuaded simultaneously, the sender can opt to use public signals, which convey the same information to all receivers, or private signals. These private signals can provide different information to individual receivers while keeping them uninformed about the signals sent to others. Therefore, in this chapter, a sender can strategically choose between public and private information to effectively encourage receivers to implement the actions the sender recommends.

The theoretical literature on information design (Bergemann and Morris, 2019) suggests that the success of public or private signals may hinge on the receivers' strategic environment. When there are coordination motives between receivers with strategic complementarities, public signals facilitate that all receivers take the same action, which in turn increases each receiver's incentive to adopt that same action, as desired by the sender. In contrast, miscoordination motives between receivers, with strategic substitutes, are best targeted with private signals. These signals promote that different receivers take different actions, which maximizes the receivers' incentives to follow the sender's recommended action.

In a first laboratory experiment, I exogenously vary the receivers' strategic environment and whether computerized senders use public or private signals. I demonstrate that the relative success of public and private signals indeed relies on the receivers' strategic environments, as predicted by theory. Importantly, however, public signals are especially persuasive, and more so than anticipated, in comparison to private signals. Empirically, senders never perform worse when using public signals as opposed to private signals, a finding that current theory does not fully capture. In addition, and in agreement with theory, senders are more persuasive with public signals than with private signals when the receivers' strategic environment involves coordination motives. The empirical advantage of public signals can be attributed to their simplicity and to their equal treatment of the receivers. In a second experiment, senders in the laboratory can select the type of signals they want to employ. Senders exhibit considerable sophistication, as they both adapt their persuasion strategy to the receivers' strategic environment and predominantly use public signals, aligning with their empirical success.

In conclusion, this thesis contributes to the understanding of how information is utilized in strategic settings through a series of laboratory experiments. By focusing on the empirical effects of public and private information disclosure across different strategic settings, this thesis offers insights into the role of both individual biases and strategic reasoning in shaping economic outcomes. It highlights the importance of considering public and private information when designing auctions, markets, and persuasion strategies.

CHAPTER 2

Why are open ascending auctions popular? The role of information aggregation and behavioral biases

This chapter is based on Offerman, Romagnoli, and Ziegler (2022).

2.1 Introduction

Open ascending auctions are routinely preferred to sealed-bid formats by both private platforms (e.g., Amazon, eBay, Catawiki) and policy makers, for example in the allocation of spectrum rights (Milgrom, 1989; McMillan, 1994; Milgrom, 2004). One compelling theoretical reason for their popularity is that open ascending auctions allow bidders to endogenously aggregate dispersed information due to the observability of the bids. Standard theory predicts information aggregation to have two advantages: it allows for a more precise estimate of the value and it leads to higher revenues in expectations. In single-unit auctions with affiliated values, buyers who are better informed bid more aggressively (Milgrom and Weber, 1982). This is implied by the linkage principle, according to which average revenues are increased by providing bidders with more information about the value of the item for sale. To this date, the linkage principle remains highly influential and is often cited as the reason why open auctions are and should be preferred over sealed-bid formats.[1]

Empirically, however, it remains an open question whether open ascending auctions are indeed capable of aggregating information. One challenge is that the single-unit setup with affiliated values hosts multiple equilibria (Bikhchandani, Haile, and Riley, 2002). This multiplicity may impede information aggregation (Milgrom, 2004, p. 197). Another challenge is that some open ascending auctions allow for jump bidding, which may obfuscate information (Avery, 1998; Ettinger and Michelucci, 2016). Also, in every-day auctions, particularly those involving non-professional bidders, the reasoning required to infer information from the bidding of others may be too demanding.

Aside from their potential for information aggregation, open ascending auctions may also differ from closed formats in the extent to which they activate or mitigate behavioral biases. Some of these biases provide alternative mechanisms for raising revenues. For instance, it is common for open ascending auctions to provisionally award the item during the auction to the bidder who submits the highest standing bid. As a result, auction fever may be activated, which encourages overbidding and leads to a quasi-endowment effect (Heyman, Orhun, and

[1] In a policy report on the question whether the spectrum auctions ran in the UK in 2018 should use an open or sealed-bid design, PowerAuctions (2015, p. 6) writes: "..., an auction should be structured in an open fashion that maximizes the information made available to each participant at the time she places her bids (Paul R. Milgrom and Robert J. Weber, 1982a). When there is a common value component to valuation and when bidders' signals are affiliated, an open ascending-bid format may induce participants to bid more aggressively (on average) than in a sealed-bid format, since participants can infer greater information about their opponents' signals at the time they place their final bids." In a footnote they explain that the text is quoted from Ausubel (2004), and add that "Its assessment is typical of the consensus of the auction literature today." The NERA (2017) report also favors an open ascending auction and echoes the same view on page 11: "Auction theory tells us that price discovery can ease common-value uncertainty, and encourage bidders to bid a higher proportion of value ...".

Ariely, 2004; Ehrhart, Ott, and Abele, 2015). Another possibility is that open ascending auctions encourage naïve jump bidding, for instance when bidders are impatient and want to terminate the auction quickly. In contrast to when jump bidding is motivated by strategic reasons, naïve jump bidding may easily enhance revenues.[2] Open ascending auctions may also encourage spiteful bidding because bidders can condition their overbidding on the presence of other remaining active bidders (Andreoni, Che, and Kim, 2007; Bartling, Gesche, and Netzer, 2016). There is, however, also a possibility that open ascending auctions mitigate behavioral biases. For example, the higher transparency of open formats may lead to buyers becoming aware of the winner's curse and tame the overbidding (Levin, Kagel, and Richard, 1996). When the winner's curse is mitigated, lower revenue may be the result in an open ascending auction.

In this paper we explore whether open auctions do raise higher revenues than sealed-bids formats. Moreover, we disentangle whether this is due to information being successfully aggregated or other behavioral mechanisms.

eBay provides a natural setting to explore information aggregation and revenues in open auctions. eBay uses an open ascending format which allows for jump bidding and provisionally awards the good to the highest standing bidder. Thus, both information aggregation and revenue-enhancing biases are possible in this format. We collected eBay data for one of the most frequently auctioned cellphones at the time of the study. The field-data analysis that we report in the Appendix, Section 2.A.1, offers suggestive evidence that information endogenously generated during the auction (proxied by the price reached halfway through the auction) and jump bidding (proxied by the average increment per bidder) correlate positively with final prices. On the basis of a median split, we find that above median bidding in the first half of the auction corresponds to an increase of 67% in the final price. Likewise, with a median split on the average increment per bidder, we find above median increments between consecutive bids correspond to an increase of 14% of the final price. The findings are consistent with information aggregation and also with the presence of revenue-enhancing naïve jump bidding. However, such data has severe limitations. First, the direction of causality is unclear. Second, such data is lacking crucial insights about bidders' information and the value of the item for sale, which makes it impossible to separate behavioral mechanisms from information aggregation. Third, we miss data from an appropriate control condition, i.e., a counterfactual auction which does not allow for information aggregation.

[2] Probably the most preposterous auction ever was decided by a naïve jump bid. After murdering the Roman emperor Pertinax (A.D. 193), the praetorian guard offered the Roman empire for sale in an ascending auction. Julianus topped Sulpicianus' highest bid of 20,000 sesterces per soldier by a winning bid of 25,000 sesterces. The winning bid corresponded to 5 years of wage of each of the 10,000 praetorians. After Julianus defaulted on his bid, he was murdered after a reign of only 66 days (Klemperer and Temin, 2001).

To overcome these limitations, we employ a laboratory experiment where we randomly assign subjects to three different auction formats. These differ in the information revealed during the bidding process, and, possibly, also in the extent to which different behavioral biases can be triggered. To ensure comparability, all formats use a second-price rule.

The first auction format is the Japanese-English auction, an open ascending auction with irrevocable exits. In this format, a clock tracks the ascending price and bidders withdraw from the auction until a single bidder remains, who wins the auction and pays the last exit price. The exit prices of other buyers are publicly observed. These bids then allow to infer other bidders' private signals, which are informative about the common value.

The second auction format is the ascending Vickrey auction, a sealed-bid ascending auction. It is implemented identically to the Japanese-English auction with an ascending clock and irrevocable exits. However, exits are not observable by others, thereby eliminating the possibility of information aggregation.

The third format we run is the Oral Outcry auction, modeled to fit popular auction designs. It falls between the other two in terms of its potential for information aggregation. In this auction, bidders can control how much information is revealed. They can engage in the informative, incremental bidding that characterizes the Japanese-English auction. They can also engage in jump bidding, i.e. out-bid the standing bid by a non-negligible amount. Jump bidding can be used rationally, for instance to obfuscate information (Ettinger and Michelucci, 2016) or to signal to other bidders that it is better to back off (Avery, 1998). Jump bidding could also be used naïvely by impatient bidders. The Oral Outcry auction, while still allowing for information aggregation, may also be the most conducive to revenue-enhancing biases. This is the only format that allows bidders to submit naïve jump bids, and it is also the only format that can activate auction fever by provisionally awarding the good during the auction.

The comparison between the ascending Vickrey auction and the Japanese-English auction provides a clean comparison of the role of information aggregation, since these formats differ only in the public revelation of exits. Theoretically, rational bidders use the information revealed in the auction to form a more precise estimate of the common value, which makes them less fearful of the winner's curse (Milgrom and Weber, 1982). As a result, the Japanese-English auction is expected to raise higher revenue than the ascending Vickrey auction. Remarkably, this prediction is reversed if bidders are naïve and tend to fall prey to the winner's curse. By gradually revealing the exit prices of bidders with low signals, the Japanese-English auction could make the risk of suffering from the winner's curse more transparent, thus taming the overbidding and reducing revenues compared to the ascending Vickrey auction. This intuition is captured by signal averaging models, which we describe more precisely in Section 2.3.

When information is successfully aggregated, remaining bidders' uncertainty about the common value is reduced and prices approximate the underlying common value more closely (Wilson, 1977; Kremer, 2002). We evaluate information aggregation by comparing the squared distance between the price and the common value across formats.

We further decompose information aggregation into two components: (i) the extent to which bids are objectively informative of the common value (*information revelation*); and (ii) the extent to which bidders actually use this information effectively in their own bidding (*information processing*).

We find that in the Japanese-English auction, less information than expected is generated. One factor that contributes to this finding is that some bidders with a low signal display spiteful behavior and stay in the auction longer than they would in the ascending Vickrey auction. Such heterogeneity is not observable by the remaining bidders and degrades the quality of the revealed information. In addition, bidders are processing the available information sub-optimally. Even though bidders are responding appropriately to the fact that early bids are revealing little information by largely disregarding them, the potential to aggregate the information actually available is mostly not realized. Instead, the processing of information is qualitatively in agreement with signal averaging heuristics. This combination of noisy early bids and sub-optimal information processing leads to a failure of information aggregation. Although subjects have only access to their private information in the ascending Vickrey auction, more information is aggregated: the squared distance between prices and common value is lower in the ascending Vickrey than in the Japanese-English auction, in which additional information is available.

Surprisingly, bids in the Oral Outcry and Japanese-English auction reveal a similar amount of information about the common value. That is, bidders do not make extensive use of the potential to strategically hide their information via jump bidding. However, in the Oral Outcry auction, the available information is processed to an even smaller extent than in the Japanese-English auction. Here, final bids are substantially distorted by the quasi-endowment effect and rash jump bidding. Subjects who are prone to endowment effects on a questionnaire measure tend to stay too long in the auction and earn substantially lower payoffs. Additionally, this auction encourages many bidders to submit unfounded jump bids. These forces result in systematic overbidding and a price which is the poorest predictor of the common value across our auction formats.

The interplay of all aforementioned factors leads to similar revenues in the Japanese-English auction and the ascending Vickrey auction. Highest revenues are observed in the Oral Outcry auction. The rationale for why the Oral Outcry auction is most often observed in the field may be quite different from the understanding in the theoretical and policy-oriented literature. Instead of leading to information aggregation, it triggers behavioral biases such as the quasi-endowment effect and reckless jump bidding.

In many ways, the laboratory provides the ideal environment to study how information is generated and processed. An important question is whether experimental results generalize to the field. Our experiments use non-professional bidders (students) that bid for objects with moderate values (of approximately € 25). We think that this situation is representative for most online auctions in the field. Beyond everyday auctions involving consumers, some of our results may also extrapolate to some situations involving professional bidders. For instance, Dyer, Kagel, and Levin (1989) find that professional bidders in the construction industry fall prey to the winner's curse in the same way as students do. We do not claim that our results generalize to spectrum auctions where bidders seek the advice of game theorists.[3]

The remainder of the paper is organized in the following way. Section 2.2 reviews the literature, Section 2.3 presents the game and some theoretical benchmarks, Section 2.4 describes how information aggregation is evaluated. Section 2.5 presents the experimental design and procedures. Section 2.6 discusses the experimental results and Section 2.7 concludes.

2.2 Related literature

Previous laboratory studies have documented how people succumb to the winner's curse in common value auctions. For an overview, see Kagel and Levin (2014). Eyster and Rabin (2005) and Crawford and Iriberri (2007) present behavioral models to explain the winner's curse. Recent studies have studied pathways behind the winner's curse, highlighting that problems with contingent reasoning (Charness and Levin, 2009) and disentangling the importance of belief formation and non-optimal best responses (Charness and Levin, 2009; Ivanov, Levin, and Niederle, 2010; Camerer, Nunnari, and Palfrey, 2016; Koch and Penczynski, 2018). We compare whether open auctions mitigate or worsen the importance of behavioral biases such as the winner's curse. Levin, Peck, and Ivanov (2016) find that a Dutch auction lessens a winner's curse compared to sealed bid formats.

An important strand of literature investigates whether markets are capable of aggregating dispersed information. A series of experiments have investigated information aggregation in asset markets. Results have been mixed. Plott and Sunder (1988) find that information aggregation only occurs when preferences are ho-

[3]Nevertheless, it is interesting to note that also in those auctions bidders sometimes engage in bidding that is merely motivated to drive up the price for a competitor. Such bidding may be driven by a spiteful motivation, or by a predatory desire to weaken the competitor in a future market (Levin and Skrzypacz, 2016). When bidding behavior may be driven by such considerations, it becomes very hard to infer valuable information from competitors' bids.

mogeneous or when a complete set of contingent claims can be traded. Forsythe and Lundholm (1990) find that information aggregation only succeeds with trading experience and common knowledge of dividends. Hence, information aggregation seems to fail when the inference task is complicated by the presence of several dimensions of uncertainty, or when the information conveyed by prices in equilibrium is less naturally interpretable.

How information is processed is also studied in the context of auctions, a particularly important form of a market. Several papers study the effect of an auctioneer exogenously revealing information in auctions. Kagel and Levin (1986) and Kagel, Levin, and Harstad (1995) show that there are ambiguous effects of revealing information in first-price and second-price sealed-bid auctions. In a setting with both private and common value elements, Goeree and Offerman (2002) find that high-quality reports of the auctioneer can positively affect efficiency and revenue, but to a lower extent than predicted by theory.[4] In contrast to this work, our paper explores *endogenous* information aggregation. Aside from shedding light on revenue effects, we uncover the process of how bidding generates information in auctions, and how bidders process the available information.

Close to our work, Levin et al. (1996) compare the performance of the Japanese-English auction and the first-price auction in a common value setting. They find that the revenue comparison of the Japanese-English auction and the first-price auction depends on the experience of the bidders: with inexperienced bidders the first-price auction raises more revenue. However, with experience this effect disappears and is sometimes reversed. Changing the price-rule and the auction format across treatments simultaneously complicates identifying the effect that information aggregation has on the outcomes. As a result, their paper remains silent about the extent to which the endogenous information revealed in the Japanese-English auction allows bidders to actually aggregate information. On an individual bidder level, they cannot use the sealed-bid auction as a benchmark to measure the degree of information processing in the Japanese-English auction. Their focus is more on a revenue comparison of their two auction formats, instead of evaluating the extent to which information is aggregated empirically. Shedding light on this phenomenon is a key contribution of our paper. We also contribute by showing that bidders process revealed information, as our design allows to compare Japanese-English and second-price sealed-bid auctions that only differ in the observability of information. Levin et al. (1996) only provide evidence that bids correlate with previous dropouts in their Japanese-English auction, which may be driven by mechanical correlation introduced by arranging bids into order statistics (as we explain in Section 2.6.2). They do not, and due to the differences in pricing rules

[4] Dufwenberg and Gneezy (2002) study another form of exogenous information disclosure. They find that the disclosure of losing bids after first-price sealed-bid common value auctions reduces revenue.

cannot, provide evidence that bids do respond to revealed dropouts. Another important difference is that their analysis does not include the Oral Outcry auction, which triggers the revenue enhancing biases that may explain their actual popularity. A less important difference is that Levin et al. (1996) adopt uniformly distributed values and signals, a knife-edge case where in equilibrium rational bidders will only process the lowest dropout price and disregard all other exit decisions in the Japanese-English auction.

A related literature compares different auction formats when bidders have interdependent valuations. In such environments, the linkage principle does not hold; with symmetric bidders, expected revenue and efficiency are predicted to be the same across auction formats (Goeree and Offerman, 2003a). Some experimental papers introduce specific asymmetries that break the revenue and efficiency equivalence results. For instance, Kirchkamp and Moldovanu (2004) compare efficiency between the Japanese-English and second-price sealed-bid auctions in a particular setup with interdependent values, where a bidder's value is the sum of the own private signal and one specific signal of the other bidders. In that setup, they find that the Japanese-English auction generates higher efficiency.
Boone, Chen, Goeree, and Polydoro (2009) and Choi, Guerra, and Kim (2019) compare open and sealed-bid auctions with interdependent values in the presence of insiders, to whom the value of the item for sale is revealed. In line with their theoretical predictions, revenue and efficiency increases in the Japanese-English auctions.[5]

In contrast to this work, our paper sheds light on how bidders process information in the more common case where signals are affiliated. We investigate the case in which the linkage principle applies and information revelation occurs with symmetric bidders. As Perry and Reny (1999) note, "The linkage principle has come to be considered one of the fundamental lessons provided by auction theory." Another distinction between our approach and this literature is that we study how information is aggregated directly, instead of by relying on comparative statics effects which are predicted by information aggregation. We do so by employing measures of information aggregation frequently used to theoretically evaluate information aggregation in auctions, see, for example, Wilson (1977), Pesendorfer and Swinkels (2000) and Kremer (2002). Our results show that although revenue is increased in some of our formats, this occurs while information aggregation *de-*

[5] A different kind of interdependence is studied in the multi-unit auction experiments of Betz, Greiner, Schweitzer, and Seifert (2017). They consider the sale of multi-unit private values emission certificates of this year (good A) and of next year (good B). Interdependence is created because units of type A can be used as type B unit, but not vice-versa. Their treatment variables are the type of auction and whether goods are auctioned sequentially or simultaneously. When items are auctioned simultaneously, they find that open ascending auctions are more efficient than sealed-bid auctions. Auctioning the items sequentially enhances the performance of sealed-bid auctions but leaves the efficiency of ascending auctions unaffected. In each auction format, total revenues are higher when items are sold sequentially.

creases, opposite to the theoretical prediction.

We also contribute to the literature on the Oral Outcry auction. Roth and Ockenfels (2002) study the impact of different rules for ending internet auctions at eBay and Amazon on bidders' propensity for late bidding. Amazon's rule to extend bidding deadlines if new bids are submitted resembles our procedure. In the lab, Ariely, Ockenfels, and Roth (2005) find that Amazon's rule to extend bidding deadlines generates higher revenue than eBay's in a private value setting. Cho, Paarsch, and Rust (2014) provide field evidence and show that in the comparison of two open auction formats, an open outcry English auction format raises more revenue, which they attribute to endogenous information revelation. It can however not be excluded that the higher revenue in the open outcry auction is actually due to behavioral factors. Close to our experiment, Gonçalves and Hey (2011) compare a Japanese-English and an Oral Outcry auction and find that they result in approximately equal revenue. However, they focus on auctions with only two bidders, which means that the potential of the Japanese-English auction to generate endogenous information is excluded by design.

It is also instructive to contrast what can be learned from our work compared to a structural approach that uses field data. For instance, Haile and Tamer (2003) use data from Oral Outcry auctions of timber-harvesting contracts held by the U.S. Forest Service to infer information about bidders' valuations. In a private values model, they show what can be learned from two simple assumptions (i) bidders do not bid above value, and (ii) bidders do not drop out unless the price is higher than their value. Their approach allows the researchers to find bounds on the valuations of bidders. Such information is useful, for instance to investigate whether reserve prices are set optimally. In contrast, in our laboratory experiment, we observe the common value and the signals. This allows us to investigate how information is revealed, processed and aggregated in strategically more complicated common value auctions, and how this depends on the auction format. More importantly, where the structural approach takes rationality as a given, our approach makes it possible to identify potential behavioral biases. In fact, we find that behavioral biases are key to explain the popularity of Oral Outcry auctions vis-a-vis other second-price formats.

Finally, we relate to the literature on endogenous information processing in stylized games. Anderson and Holt (1997) initiated a literature on informational cascades. Eyster, Rabin, and Weizsacker (2018) find that subjects' social learning depends on the complexity of the underlying problem. Magnani and Oprea (2017) investigate why subjects violate no-trade theorems and find that over-weighting of one's private information contributes to such violations. Hossain and Okui (2018) study how subject's correlation neglect (Enke and Zimmermann, 2019) explains information processing. Other studies show that biased inference can arise in in-transparent problems where subjects display a lack of contingent reasoning (Esponda and Vespa, 2014; Ngangoué and Weizsäcker, 2021; Martínez-Marquina, Niederle, and Vespa, 2019). Our take-away from this literature is that subjects do pay attention to the behavior of others, but that their sophistication depends on specifics of the problem, such as the transparency of its presentation and its complexity. There is no single result that generalizes across all contexts. In our view, this implies that social learning should be studied in the particular setup of interest. How information is processed and aggregated in the canonical affiliated values setup of Milgrom and Weber (1982) is therefore still an open question. While this setup not only inspired a vast body of theoretical work, it also was and continues to be very influential in advice on actual auction design (McMillan (1994, p. 151-152), Cramton (1998)).

2.3 Auction formats and theoretical benchmarks

In the following, we describe the auctions implemented in the laboratory, present Nash equilibria as well as behavioral heuristics and explain revenue predictions.

2.3.1 General setup: Bidders and payoffs

All our formats are common value auctions with five bidders and a second-price rule. The common value of the object for sale is unknown to bidders, who only receive a private signal about the value. More precisely, the good has value V, where $V \sim N(\mu, \sigma_V) = N(100, 25)$. Each bidder $i \in \{1, 2, \ldots, 5\}$ receives a signal X_i of the common value V. This signal is the sum of the underlying value and an individual error ϵ_i:

$$X_i = V + \epsilon_i$$

This error is *i.i.d.* across bidders and normally distributed with mean 0 and standard deviation σ_ϵ: $\epsilon_i \sim N(0, \sigma_\epsilon) = N(0, 35)$.

In all formats, the winner of the auction is the bidder who submits the highest bid. This bidder receives a payoff equal to V minus the second highest bid. All the other bidders receive a payoff of 0. For notational purposes, define a signal realization x_i for bidder i. Let $Y_{i,(k)}$ represent the k-th highest of the signals received by any other bidder $j \neq i$, so e.g. $Y_{i,(1)}$ is the highest signal received by any bidder other than bidder i.

2.3.2 Auction formats

We now provide details for each of the three auction formats we study.

The ascending Vickrey auction (AV)

We implement the ascending Vickrey auction (AV) with a clock procedure. After bidders have been privately informed of their signals, the price rises simultaneously from 0 for all participants. At any integer price $0, 1, 2, 3, \dots$, bidders can decide to leave the auction by pressing the "EXIT"-button. In the AV, no bidder observes whether any other bidder has left. The auction stops as soon as four bidders have exited the auction. The last remaining bidder wins the auction and pays the price at which the fourth bidder leaves. In case multiple bidders leave last at the same price, one of them is randomly selected to be the winner and pays the price at which she left. In this format, a bid is the price at which the bidder decided to leave the auction.

The Japanese-English auction (JEA)

The Japanese-English auctions (JEA) makes use of the same clock procedure. Differently from the AV, all remaining bidders are notified in real time of other bidders' exit prices. Like in the AV, the winning bidder is the last remaining bidder after four bidders exit. This bidder pays the price at which the fourth bidder left the auction.

The Oral Outcry auction (OO)

In the Oral Outcry auction (OO) bidders can outbid each other repeatedly and by arbitrary amounts until no more out-bidding takes place and the good is awarded to the highest standing bidder. In our implementation, bidding proceeds in bidding rounds. In each bidding round, all bidders have 15 seconds to submit a maximum bid. As soon as one bid is submitted, the bidding round is interrupted. At this point, the bidder who submitted the highest bid becomes the standing bidder, the provisional winner in case the auction would stop afterwards. The current price is set to the second highest bid at this moment. A new bidding round starts, the clock is reset to 15 seconds and the standing bidder is excluded from submitting a new bid.[6] During the auction, bidders are notified of the highest maximum bid of each of the other bidders, with the exception of the current standing bidder, about whom it is only revealed that her highest bid is at least as high as the current price. The auction ends as soon as the countdown elapses without further bidding. At this point, the last standing bidder wins the auction. She pays the last current price, which is the second highest bid at the end of the auction.

[6] This leads to an auction ending time being determined endogenously. Such a rule is a feature of online auctions at amazon.com, yahoo.com and catawiki.com.

2.3.3 Nash equilibrium predictions and behavioral forces

In this Section, we use game theoretic results, behavioral theories and recent experimental findings to contextualize our research questions. We start with presenting the Nash equilibrium predictions, according to which the JEA should aggregate information and consequently lead to higher revenues than the AV.

AV and JEA: Nash Equilibria and the *Linkage Principle*

Symmetric Nash equilibria in single-unit auctions with affiliated values have been derived in Milgrom and Weber (1982). In the AV, a bidder's strategy can be described by a reservation price, which makes this format strategically equivalent to the standard second-price sealed-bid auction (see Milgrom, 2004, p. 187-188). A symmetric equilibrium of the AV is given by bids $b(x_i)$:

$$b(x_i) = \mathbb{E}\left[V|X_i = x_i, Y_{i,(1)} = x_i\right]$$

That is, each bidder exits the auction as soon as the clock reaches the expected value of the good for sale conditional on her signal and assuming that the highest signal obtained by other bidders is also x_i.[7]

In the symmetric Nash equilibrium of JEA, bidders include endogenously revealed information into their bidding strategies. The first bid is given by (see Milgrom and Weber, 1982):

$$b_1(x_i) = \mathbb{E}\left[V|X_i = x_i, Y_{i,(1)} = x_i, \dots, Y_{i,(4)} = x_i\right]$$

Just like in the AV, the first exit bid is obtained via a conditional expectation, assuming that all other bidders hold an equally high signal. However, as soon as the first bidder drops out at p_1, the remaining bidders perfectly infer the signal of the exiting bidder, from $p_1 = b_1\left(Y_{i,(4)}\right)$. All bidders dropping out subsequently base their j-th bid (for $j > 1$) on their private information and the signals inferred from the $j-1$ observed dropouts. The remaining bidders bid $b_j(x_i)$:

$$b_j(x_i) = \mathbb{E}\Big[V|X_i = x_i, Y_{i,(1)} = x_i, \dots, Y_{i,(5-j)} = x_i, p_1 = b_1\left(Y_{i,(4)}\right), \dots$$
$$\dots, p_{j-1} = b_{j-1}\left(Y_{i,(5-j+1)}\right)\Big]$$

This equilibrium allows to iteratively back out all information except the one con-

[7]In our experimental setup with 5 bidders and normally distributed values and signals, Goeree and Offerman (2003b) show that the above conditional expectation is equal to: $b(x_i) = \mathbb{E}\left[V|X_i = x_i, Y_{i,(1)} = x_i\right] = x_i - \frac{\int_{-\infty}^{\infty} \epsilon\phi_V(x_i-\epsilon)\phi_\epsilon^2(\epsilon)\Phi_\epsilon^3(\epsilon)d\epsilon}{\int_{-\infty}^{\infty} \phi_V(x_i-\epsilon)\phi_\epsilon^2(\epsilon)\Phi_\epsilon^3(\epsilon)d\epsilon}$, where $\phi_V(\cdot)$ denotes the pdf of the common value distribution, $\phi_\epsilon(\cdot)$ the pdf of the error distribution, with its cdf $\Phi_\epsilon(\cdot)$.

tained in the highest signal.[8] According to the *linkage principle*, the information revealed in the JEA leads to more aggressive bidding, the fourth bid in the JEA is on average higher than the fourth bid in the AV (Milgrom and Weber, 1982). Bikhchandani et al. (2002) have identified other symmetric Nash equilibria that implement the same outcome. In such equilibria, the first three bidders drop out at a fraction $\alpha \in (0, 1)$ of the bids at which they dropped out in the just described equilibrium, and the last two bidders bid as before.[9]

AV and JEA: A behavioral perspective

Overbidding is often observed in experimental common value auctions, suggesting that in practice bids may not align well with Nash equilibrium predictions. Even in the AV, bidding in agreement with a symmetric equilibrium is quite sophisticated and requires bidders to (i) use their prior about the distribution of the value; (ii) account for the fact that the bidder with the highest signal is predicted to win the auction. Thus, to avoid the winner's curse, bids need to be shaded.

Simpler behavioral rules have been proposed in alternative to Nash equilibrium bidding. For example, bidders in the AV who ignore both (i) and (ii), and only rely on their private signal, may adopt the "bid signal"-heuristic (Goeree and Offerman, 2003b): $b(x_i) = x_i$, which leads to expected overbidding.

The JEA, on the other hand, allows bidders to observe early exits of other bidders with low signals. This could make (ii), i.e., the fact that winning bidders receive higher signals than their peers, transparent to bidders in a natural way. The "bid signal"-heuristic remains available in the JEA. However, by raising awareness about the winner's curse, the JEA can lead to less overbidding. The "signal averaging rule" proposed by Levin et al. (1996) captures this intuition. According to this rule, bidders bid an equally weighted average of their own signal and the signals of their fellow bidders, revealed from the previous dropouts. After $j - 1$ bidders dropped out, with the vector of revealed signals being $\mathbf{Y}_i = \{Y_{i,(4)}, \dots, Y_{i,(5-j+1)}\}$, this implies the following bid: $b_j(x_i, \mathbf{Y}_i) = \frac{1}{j}x_i + \frac{1}{j}\sum_{k=1}^{j-1} Y_{i,(5-k)}$.[10] In expectation, the "signal aver-

[8] We determine Nash equilibrium bids in our setup, using a result by DeGroot (2005, p. 167). For inferred or assumed signal realizations by bidder i, define $\bar{x}_i = \frac{1}{5}\left(\sum_{j=1}^{4} Y_{i,(j)} + x_i\right)$. Then in equilibrium each bidder i bids: $\mathbb{E}[V|x_i, Y_{i,(1)} \dots, Y_{i,(4)}] = \frac{\frac{\mu}{\sigma_V^2} + \frac{5\bar{x}_i}{\sigma_\epsilon^2}}{\frac{1}{\sigma_V^2} + \frac{5}{\sigma_\epsilon^2}} = \frac{5\bar{x}_i\sigma_V^2 + \mu\sigma_\epsilon^2}{5\sigma_V^2 + \sigma_\epsilon^2}$. On request, we provide derivations showing that equilibrium bids can be inverted such that they depend linearly on the signal and observed bids. This also applies to all other models considered in this paper. We therefore restrict ourselves to linear information use in all estimations.

[9] Bikhchandani and Riley (1991) study asymmetric Nash equilibria and show that they can lead to different revenue rankings than those established by Milgrom and Weber (1982). In our experiment, all bidders are treated symmetrically and there is nothing that facilitates coordination on an asymmetric equilibrium. In this sense, a symmetric equilibrium is more plausible.

[10] Note that this rule can be plugged in iteratively, such that bidding depends only on the most recent dropout, which is an average of all previously revealed signals. This yields $b_j(x_i, b_{j-1}) =$

aging rule" corrects for the overbidding observed in the "bid signal"-heuristic. If bidders follow these two behavioral rules in the JEA and the AV respectively, then the former format is predicted to raise lower revenues.

Somewhat more sophisticated bidders could process information about the prior distribution of the value, and thereby accommodate (i), incorporating information on the prior. This would lead to a slightly modified versions of the two rules above, the "Bayesian bid signal"-heuristic, and the "Bayesian signal averaging rule". By anchoring bidding to the prior, these rules lead to less extreme under- and overbidding. However they continue to predict that the JEA raises lower revenues than the AV.

In the "Bayesian bid signal"-heuristic bidders bid the expected value of the good for sale, conditional on one's signal: $b(x_i) = \mathbb{E}[V|x_i] = x_i - \mathbb{E}[\epsilon_i|x_i]$. Goeree and Offerman (2003b) show that $b(x_i) = \frac{\sigma_V^2 x_i + \sigma_\epsilon^2 \mu}{\sigma_V^2 + \sigma_\epsilon^2}$. According to the "Bayesian signal averaging rule", bidders combine Bayes rule with the symmetric signal averaging rule.[11] After $j-1 > 0$ observed dropouts, bidder i calculates the average of available signals $\bar{x}_i = \frac{1}{j} x_i + \frac{1}{j} \sum_{k=1}^{j-1} Y_{i,(5-k)}$ and bids $b(\bar{x}_i) = \frac{\sigma_V^2 \bar{x}_i + \sigma_\epsilon^2 \mu}{\sigma_V^2 + \sigma_\epsilon^2}$.

Nash equilibrium predictions and predictions based on behavioral rules now lead to conflicting effects of information revelation on revenues. While private signals can be inferred in both types of benchmarks, revenue ranking predictions with the behavioral rules are driven by the degree to which bidders' are made aware of the winners' curse in the JEA relative to the AV.

Using our parameterization and draws, Table 2.1 summarizes the revenue predictions for the Nash Equilibrium and the behavioral models that we discussed.[12]

Table 2.1: Revenue predictions

	AV	JEA
Nash equilibrium	95.8	97.4
Bid signal	117.4	117.4
Signal averaging rule	117.4	91.1
Bayesian bid signal	105.9	105.9
Bayesian signal averaging	105.9	94.0

Nash equilibrium revenues are only slightly higher in the JEA than in the AV. This is not an artifact of our parameter choices. As we show in Appendix Section 2.A.2,

$\frac{1}{j} x_i + \frac{j-1}{j} b_{j-1}$.

[11] A peculiar feature of the setup of Levin et al. (1996) with uniformly distributed values and signals is that a Bayesian will form the same belief as a naïve bidder who ignores the prior. This is not the case in our setup with normally distributed values and errors.

[12] Note that the revenue prediction of a model only depend on the revenue-determining bidder using the particular model. Theoretically, in the JEA, bidders are able to infer all other bidders' signals irrespective of the model these other bidders are using, as long as all bidders hold correct beliefs on which model others are using.

similar minor revenue differences result for various combinations of variances of the values and errors. In both formats, the winners capture some information rents and make positive profits, as the price-determining bidder in equilibrium slightly underestimates the value by design of the equilibrium bidding strategies.

The differences in predictions for the behavioral models are much larger. Moreover, the behavioral rules yield losses for the winners in the AV. In the JEA, bidders make substantial profits if they use (Bayesian) signal averaging rules.[13]

The Oral Outcry: information aggregation and behavioral biases

The Oral Outcry auction format is very rich and there are no clear Nash equilibria for this format. Still, we can make some observations about the potential of the Oral Outcry for information aggregation and revenues. In this format, bidding may proceed incrementally as in the Japanese-English auction. That is, bidders may constantly be active until their reservation price is reached, which would allow for similar inference as in the JEA.

This format can also encourage jump bidding. From a strategic point of view, jump bidding can be used to signal a high estimated value of the item and deter other bidders from continuing to bid. Avery (1998) shows how strategic jump bidding can be supported in an equilibrium of a game that is much simpler than ours. Similarly, jump bidding may obfuscate information, as shown in a stylized auction game in Ettinger and Michelucci (2016). In either case, severe jump bidding suppresses information aggregation and its revenue-enhancing effects.

On the other hand, recent experimental findings suggest that some features in Oral Outcry may be particularly prone to revenue-enhancing behavioral biases, such as auction fever (Heyman et al., 2004; Ehrhart et al., 2015). Similarly, jump bidding might not be used in the sophisticated way studied theoretically, e.g. it might rather be driven by bidders' impatience.

2.4 Information aggregation: Measure and benchmarks

When information is successfully aggregated, bidding and prices move closer to the underlying common value (Wilson, 1977; Kremer, 2002). We measure the degree of information aggregation with the squared distance between the price and the common value and compare it across formats (Hanson, Oprea, and Porter, 2006). A distance of 0 would imply perfect information aggregation in the sense that bidders inferred the exact true value.

[13] Note that our experimental setup leads to low expected revenue with signal averaging-rules. This allows us to test the rules beyond what was possible in Levin et al. (1996). In their setup, signal averaging-rules lead to predictions more similar to Nash equilibrium revenues.

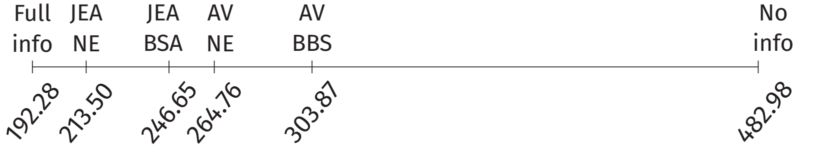
The possibility of perfect inference is curtailed by the noisiness of the signals. We account for the maximal information potentially available, the one contained in the five signals, by computing the Full Information benchmark. In it, all five signals are revealed and bidders bid the conditional expected value of the item given these signals. Additionally, we model the lowest degree of aggregation with the No Information benchmark, where bidders bid the prior average common value, thus ignoring also their own private signal.

We illustrate the Full and No Information benchmark as the lower and upper bounds of a segment measuring information aggregation. On this segment, lower values indicate a better approximation of the common value by the price, hence improved information aggregation.

In the segment, we also show how much information aggregation is predicted in Nash equilibrium and by some exemplary behavioral models. In the Nash equilibrium of the JEA, we see that the Full Information benchmark is almost attained.[14] In the Nash equilibrium of the AV, the squared distance to the common value is higher, as less information aggregation is possible. By comparing the Nash equilibrium predictions of the two formats, we see the theoretical impact of information aggregation: If dropouts are observable, bidders obtain a more precise estimate of the value and the price follows the common value more accurately.

The prediction that the JEA leads to higher information aggregation compared to the AV generalizes to the behavioral models of bidding behavior. The Bayesian bid signal heuristic (BBS) in the AV auctions predicts a larger dispersion around the common value compared to Bayesian signal averaging (BSA) in the JEA.[15] Therefore, even when processing information in a sub-optimal manner, bidders are predicted to improve their estimate of the value when they observe others' bids.

Figure 2.1: Squared distance to common value - JEA and AV



[14]It is not fully attained for two reasons: (i) the bid determining the price is based on 4, rather than 5, signals; (ii) bidders maximize expected profit, with information rents for the winner.

[15]This also holds for the comparison of signal averaging- and bid signal-heuristics, which are omitted for brevity.

2.5 Experimental design and procedures

The computerized laboratory experiment was conducted in July and October 2018 at the CREED laboratory of the University of Amsterdam. In total, we ran 30 sessions with 10 subjects each. We preregistered this experiment (Offerman, Romagnoli, and Ziegler, 2019a). Most subjects were students of business, economics or other social sciences, with 50.7% being male and an average age of 23. Each subject participated in only one session.

The experiment was conducted in a laboratory with soundproof cubicles. As a consequence, information revelation was entirely controlled as intended in the experimental design. In Appendix B, we present the instructions together with screenshots of the auction interface for all formats. Subjects read the computerized instructions at their own pace, and they had to correctly complete a set of test questions before they could proceed to the experiment. Before the experiment started, subjects received a handout with a summary of the instructions. At the end of the experiment, subjects filled out a brief questionnaire.

In the experiment, 30 auction rounds were played. Payment was based on five rounds randomly selected at the end of the experiment. Subjects earned points that were exchanged according to a rate of € 0.25 for each point. Subjects earned on average € 24.28 (standard deviation: 6.02, minimum earnings were set to € 7) in approximately two hours.[16]

We run three between-subject treatments, each corresponding to one auction format. In each ten-subject session, subjects were randomly rematched into groups of five every round, therefore a matching group of 10 subjects coincides with the session size. Common values and corresponding signals were drawn before sessions started. Draws are *i.i.d.* across rounds for common values, and error draws are also *i.i.d.* across subjects. For the experiment, we use identical draws in the identical order across treatments. Thus, treatment differences are not driven by differences in random draws. In the experiment, we truncate common value and signal draws between 0 and 200 and also only allow for bids between 0 and 200.[17]

We communicated the distributions of values and signals with the help of density plots and we allowed subjects to generate example draws for the common value and corresponding signals. At the start of each round in each auction, subjects were privately informed about their signals and the auction started as soon as all bidders in a session indicated that they were ready.

The rules of the auction formats were described in Section 2.3. The auction

[16]In the experiment, only one subject had a negative payment balance if calculating total earnings across *all* rounds. In the pre-registration, we announced that we also analyze our data without bankrupted subjects. However, excluding this one subject does not affect results.

[17]We discarded a set of draws whenever a common value or signal exceeded our bounds. This occurred for 0 out 600 common value draws, and 121 out of 6000 drawn signals. Due to the small scale of this phenomenon, we ignore truncation in our analysis.

procedure was visualized with a thermometer. In the AV and the JEA, the price increased from 0 by one point every 650 milliseconds. Approximately three times per second, the program checked whether any bidder dropped out. In the JEA, bidders were shown the prices at which the first, second and third dropout occurred. After a dropout in this auction, there was a pause of four seconds where the price did not rise to allow the remaining bidders to process the information.

In all three treatments, at the end of each round all subjects were shown the price which the winner paid and the common value that was drawn. In each round, each bidder was endowed with 20 points, and the winning bidder was additionally paid the difference between the common value and the second highest bid. When negative, the difference was deducted.

In the 13 sessions ran in October 2018, we included two additional incentivized tasks at the end to investigate some conjectures developed after the first sessions. First, we used a measure adapted from Goeree and Yariv (2015) to elicit a subject's tendency to conform to others' choices in an environment where these choices contain no information. Subjects had an incentive to guess an unknown binary state. Their choice was to either receive a noisy but informative signal of the state, or to sample the uninformative decisions from three previous subjects. Crucially, these previous subjects had no access to any information about the true state, and subjects were made aware of this fact. Second, we obtained a measure of subjects' social preferences by using the circle test to measure their value orientation (Sonnemans, van Dijk, and van Winden, 2006). We included these measures to test some conjectures about the exit decisions of subjects with low signals in the Japanese-English auction. In addition, in the oral outcry auction we included two unincentivized questionnaire measures of subjects' tendency to succumb to endowment effects to further investigate the role of the quasi-endowment effect in this auction.[18]

Many features of our experimental design are motivated by the theoretical model with affiliated signals (Milgrom and Weber, 1982). The situation that we study is stylized, and our setup may offer more opportunities for learning than bidders would have outside of the laboratory when they bid on real commodities. In auctions outside of the laboratory, it may be much less clear to the winner that he suffered a loss, which may impede learning. In addition, our conjecture is that bidders may suffer more from endowment effects when they are bidding on a real commodity than when they are bidding on a fictitious good with induced value. From this perspective, we expect that biases may be larger outside of the labora-

[18]Question 1 was: "Suppose you paid € 30 for 5 cello lessons. After the first lesson you realize that you really don't like it. How many of the remaining lessons do you attend? You cannot get the money back." Question 2 was: "Suppose that tickets are on sale for the National Lottery to be played out in one week, with a prize of € 100.000 and you just bought one ticket for € 2.50. A colleague offers you money to buy the ticket from you. What is the minimum price at which you are willing to sell the ticket to him?"

tory.

2.6 Experimental results

In this Section we present the experimental results. We first present an overview of the revenues generated in the three auctions. Next, we discuss information use in the Japanese-English auction (JEA). Then, we compare the level of information aggregation in all three formats. Finally, we present evidence on jump bidding and the quasi-endowment effect in the Oral Outcry auction (OO).

In our analysis, we use data from all 30 rounds. We present results on experienced bidders in the Appendix Section 2.A.7. Results are mostly in line with the main analysis, otherwise we address this within the main text.

2.6.1 Revenue

Figure 2.2 and Table 2.2 present mean revenues by treatment.[19] Average revenues are quite similar in the AV and the JEA, but are substantially larger in the OO. Differences are most pronounced in the first 15 rounds, but differences continue to be significant also for experienced bidders in the last 15 rounds. Table 2.2 also reports test results of comparisons of revenue across treatments together with test results of the comparisons of revenues with the Nash benchmark.[20]

We find strongly significant revenue differences between the OO and both other auction formats. While the theory predicts higher revenue in the JEA than in the AV, we cannot reject equality of revenues between the two formats. In both the AV and the JEA, actual revenues deviate systematically from the Nash benchmark.

One explanation for the failure of rejecting equality of revenues between the AV and the JEA is that bidders simply ignore the information that is revealed in the JEA. Another possibility is that the more transparent JEA activates different behavioral forces that offset each other. In the next Section we explore these possible explanations.[21]

[19] In one auction in the AV, the auction unintentionally ended after only three, not four, bidders dropped out. We remove the data from this particular auction.

[20] Treatment results are robust to using parametric tests and the non-significance of a treatment difference is not arising from comparing matching group averages. When regressing revenues on treatment dummies, clustering standard errors on a matching group-level (600 observations per treatment), we find that compared to a baseline of the AV, the dummy on the JEA is not significant with a p-value=.778, whereas the dummy on the OO is significant at a p-value=.005.

[21] In the preregistration plan, we announced that we would compare how well rational and behavioral models organize actual bidding. It turns out that none of the models comes even close to explaining the early dropouts in the auction. As a result, we have chosen to relegate this analysis to the Appendix Section 2.A.5.

Figure 2.2: Mean revenue, Nash equilibrium predictions and common values

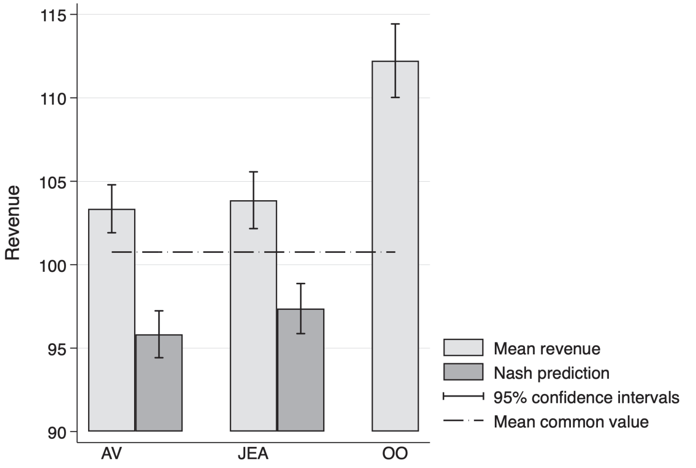


Table 2.2: Revenue statistics by treatment

		Revenue		
		Mean (Standard deviation)		
Round		1-30	1-15	16-30
AV		103.4 (17.9)	106.1 (19.5)	100.6 (15.7)
JEA		103.9 (21.2)	106.5 (20.9)	101.3 (21.3)
OO		112.2 (27.5)	118.0 (31.2)	106.5 (21.7)
		Treatment effects: p-values		
Round		1-30	1-15	16-30
AV vs.	JEA	.597	.940	.734
	OO	.003	.011	.049
JEA vs.	OO	.009	.003	.059
		Revenue difference to Nash eq'm: p-values		
Round		1-30	1-15	16-30
AV vs.	Nash eq'm	.001	.002	.010
JEA vs.	Nash eq'm	.001	.000	.049

Notes: Mean and standard deviation of revenues by treatment, over time. Test results (p-values) of revenue comparisons across treatments and to the Nash equilibrium prediction. For each test, we use the averages per matching group as independent observations for the Mann-Whitney U-tests (MWU). This gives 10 observations per treatment.

2.6.2 Information processing in JEA

We find that bidders overbid both in the JEA and in the AV compared to the rational benchmark. Our data also do not agree with the revenue prediction of (Bayesian) signal averaging, according to which revenue in the JEA must be lower compared to the AV. These findings raise the question whether subjects make use in any way of the information released in the auction. One possibility is that bidders in the JEA disregard the bidding of others and only use their private information. In this Section we show that this is not the case. We start by comparing how bids correlate in the JEA with previous dropouts, and contrast this to information use in the theoretical benchmarks. Then we proceed by showing that bidders' dropouts correlate more with previous dropouts in the JEA than in the AV, in which endogenous information of others' bids is not available.

Table 2.3 presents the results of a fixed-effects regression analysis that models how bids correlate with available information. Define as $b_{j;i,t}$ the dropout price of bidder i in round t, where, for ease of exposition, j denotes the dropout order corresponding to that observation. Further denote with $\mathbf{b}_{j-1,t}$ the vector collecting the $j-1$ dropout prices preceding the $j-th$ bid in round t. For each $j \in \{1, \dots, 4\}$ we pool data for each dropout order j and separately estimate the models:

$$b_{j;i,t} = \alpha + \beta x_{i,t} + \gamma \mathbf{b}_{j-1;t} + \delta t + \eta_i + \epsilon_{i,t}$$

where $x_{i,t}$ is the private signal of bidder i and t is the auction round. η_i is a bidder-specific fixed effect and $\epsilon_{i,t}$ is a bidder-round error. We use the within-estimator, where we are demeaning the variables with their time-averaged counterparts. This allows us to interpret the constant as the average intercept across bidders, and each bidder's fixed effect as the deviation in this bidder's bidding level from the average.

Models (1) to (4) provide fixed effects estimates of dropout prices regressed on available information, similar to the analysis by Levin et al. (1996). There is a recurring pattern in how subjects' bids correlate with available information: Bidders' dropouts depend significantly only on their own signal and the just preceding dropout.[22] The most recent dropout receives much more relative weight than bidders' signals. Thus, bids appear to react quite strongly to the auction proceedings.[23]

All theoretical models considered in this paper process information linearly

[22]Conditional on using information summarized in the previous dropouts, earlier bids do not add additional explanatory power. There is indeed a correlation to earlier bids, which is fully captured in the reaction to the current dropout. Repeating (3) and (4) without $b_{j-1,t}$ yields significant coefficients on $b_{j-2,t}$.

[23]This analysis does not shed light on the possibility that the strong weight on the most recent dropout is due to correlation neglect (Enke and Zimmermann, 2019). With correlation neglect, information in early dropouts is double-counted in later dropouts. In the Appendix Section 2.A.9, we present regressions similar to the above, while excluding bidders' private information. We then predict residuals in this estimations, which capture bidders' private information (their signals and noise). We then regress later bids on all residuals. We find little evidence for strong correlation neglect, as especially residuals from late dropout orders most strongly explain variation in bids. This suggests that subjects understand that the most recent dropout contains information of the signals conveyed in the earlier dropouts.

Table 2.3: Bidders' use of information in the JEA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	b_1	b_2	b_3		b_4			V	$\hat{BR}$
	Observed	Observed	Observed	Observed	Nash	SA	BSA		
x	0.294	0.267	0.172	0.118	0.287	0.250	0.168	0.250	0.288
	(0.057)	(0.034)	(0.027)	(0.016)	(.)	(.)	(.)	(0.020)	(0.001)
b_1		0.372	0.023	0.025	0.100	0	0	-0.009	0.032
		(0.035)	(0.018)	(0.015)	(.)	(.)	(.)	(0.025)	(0.003)
b_2			0.552	-0.038	0.167	0	0	-0.003	0.060
			(0.044)	(0.037)	(.)	(.)	(.)	(0.052)	(0.003)
b_3				0.709	0.333	0.750	0.832	0.291	0.151
				(0.072)	(.)	(.)	(.)	(0.070)	(0.003)
t	-0.316	-0.122	-0.083	-0.075				0.295	0.087
	(0.281)	(0.114)	(0.074)	(0.031)				(0.073)	(0.002)
Constant	35.185	41.823	32.049	26.290	11.265	0	0	41.882	44.804
	(8.628)	(2.723)	(2.933)	(3.619)	(.)	(.)	(.)	(3.799)	(0.361)
Observations	600	600	600	600				600	600
Adj. R^2	0.119	0.491	0.756	0.817				0.362	0.996
Adj. R^2 absorb. i	0.425	0.592	0.768	0.821					
Rounds	1-30	1-30	1-30	1-30				1-30	1-30
Estimation	FE	FE	FE	FE				OLS	OLS

Notes: b_j: dropout price at order j; V: common value; x: own signal. (1) to (4) are fixed effects estimates (within estimation) of information use. Dependent variables (in columns) are dropout prices at each order, e.g. (1) are all bidders dropping out first in an auction. Regressors (in rows) are the available information at each dropout order, i.e., the signal x and the preceding dropout prices $\mathbf{b}_{j-1}$. (5) to (7) show how information is used in three canonical models, only for the fourth dropout. SA refers to the signal averaging-rule, BSA to the Bayesian signal averaging-rule. Note that these show how theoretical bids respond to earlier bids, where these bids are also calculated to follow the theoretical models. (8) shows how the price-setting bidder would have to use information to predict the common value after observing three dropouts. (9) shows how the bidder dropping out fourth would weigh information in an empirical best response. We provide adjusted R^2 of the original within-estimated model, as well as from estimating standard OLS where we include subject-specific absorbing indicators. The latter also includes fit obtained from subject fixed effects. Standard errors in parentheses, clustered at the matching group level.

(derivations available on request).[24] In models (5) to (7), we provide theoretical benchmarks for the fourth dropouts, representing informational weights implied by these models. These models show how bids would react to (theoretical) earlier dropouts, and are purely theoretical, not estimated.[25] By comparing estimated information use to the use implied by these models we can evaluate whether bidding strategies are consistent with any of the models, which can be helpful to predict outcomes in other auction environments.

In model (5), Nash equilibrium, bidders do not ignore information from the first and second dropouts when they choose the fourth dropout conditional on the third dropout, contrary to information use in our data. Instead, the observed pattern is more in agreement with the signal averaging rules (models (6) due to Levin et al. (1996) and (7)). Both signal averaging rules correctly predict that the last dropout is a sufficient statistic for all previously revealed information, as this bid summarizes all previously revealed information. Qualitatively, the Bayesian signal averaging rule (model (7)) performs particularly well, as it approximates the relative weight on last dropout compared to the own signal more closely than in (6). A further pattern in favor of Bayesian signal averaging is that bidders do not ignore the prior. In the AV, which offers the cleanest view on whether subjects use the prior, bids are anchored towards the mean common value. Bidders who receive a signal above 100 bid on average 72.4% of their signal, while bidders with a signal of at most 100 bid on average 117.4% of their signal.

Still, the bids predicted by the Bayesian signal averaging rule do differ significantly from observed behavior. The intercepts across all dropout orders are quite large and lead to the observed overbidding.[26] As later bids are incorporating revealed information, constant overbidding early on carries over to later bids, which then determine revenue.

One remaining question is whether observed early dropouts are informative for subsequent bidders, and in how far bidders could use these bids to improve their estimates of the common value. In Nash equilibrium, all available information should be used when best responding, see model (5). However, early bids differ systematically from Nash equilibrium bids, and are potentially less informative of the common value than they are in Nash equilibrium. The informativeness of early bids should determine how later bids should respond to early bids. We proceed by using two types of analyses: studying (i) how informative bids are of the value and (ii) how information is used in an empirical best reply.

In estimation (8) we provide an analysis of the informational content of observed bids. We regress the common value on the information available to the

[24] We verified that our findings are not driven by the linear impact of information, by repeating (4) and (8) with the additional regressors x^2 and $(b_3)^2$. Both are not significant in either model.

[25] Applying OLS to simulated bids also recovers the coefficients presented in Table 2.3.

[26] In fact, we can reject the coefficient restrictions implied by (5) to (7) in F-tests based on the estimated equation (4), with p-values=.000.

bidder dropping out fourth. This analysis studies how the information available to the bidder determining the price is predictive of the common value, which at the end of each round is revealed to the subjects. Thus, model (8) provides a benchmark of what information is useful to bidders when attempting to predict the value using a linear rule.[27] In model (8), we observe that it is sufficient for bidders to attach positive weights only to the third dropout and own signal to predict the common value. This implies that early bids are not useful to predict the common value, which in fact our subjects appear to incorporate by disregarding this information. However, the relative weights attached to the third dropout relative to the own signal differ strongly from the rule predicting the value, as bidders appear to react too much to the third dropout given the informational content of these bids.

In (9), we study how information would be weighted in an empirical best response. In this, we assume that the two bidders that remain in the auction longest bid the expected value of the item for sale, conditional on the other remaining bidder holding an equally high signal as the own signal, and incorporating information revealed in the previous dropouts. To infer signals from early dropouts, we use linear regressions in which we regress signals on observed bids, round, session fixed effects and signals predicted from earlier bids if available.[28] The empirical best response then equals the conditional expected value calculated on the basis of the inferred signals, under the assumptions that the other remaining bidder has a signal that equals the own signal, using the result by DeGroot (2005).[29] By assuming that the other remaining bidder has a signal that equals the own signal, the bidder beats types that are below the own type, and by doing so wins in cases where the expected profit is positive, and loses against types that are above the own type, and thereby avoids winning in cases where the expected profit is negative. Notice that the procedure is quite similar to how bidders bid in the symmetric Nash equilibrium. The difference lies in how signal are inferred from earlier bids. In the Nash equilibrium, bidders infer the signals of bidders that previously dropped out from their actual (Nash equilibrium) bidding strategies. In our empirical best response, signals are estimated from previous dropouts. We then regress the obtained empirical best response on the same set of observables

[27] Note that the positive coefficient on t is a mechanical effect of all bids decreasing in t (see (1) to (4)), as V is in expectation constant over time. From experience, bidders learn that the amount of overbidding by others decreases over time (at the end of each round the common value of a round is communicated). To accommodate for this downward trend in the bidding, given the same previous dropouts, a bidder who estimates the common value will form a higher prediction of the common value in later rounds compared to early rounds. Such a compensating factor would have been absent if there had not been a trend in subjects' bidding. Allowing for a more flexible time trend in (8) with squared round or round fixed effects does not affect estimates on information use (b_1, b_2, b_3, x).

[28] We reproduce these estimations in the Appendix, Table 15.

[29] In calculating the conditional expected value, we invoke the assumption that signals inferred from previous dropouts are distributed as the true signals are (that is, conditional on the value they are *i.i.d.*, $N(0, 35)$).

for the second-highest bidder.

Consistent with the findings of model (8), (9) shows that early bids optimally receive little weight in an empirical best response. Due to early bidding being less informative than in Nash equilibrium, the optimal weights are below the weights on observed bids in model (5). However, even if the estimated coefficients are small, they are significant and positive. Again similar to (8), (9) shows that bidders do not rely sufficiently strongly on their own signal when bidding, and disregard valuable information in bidding.[30]

Importantly, this analysis in itself does not provide evidence that bidders actively incorporate information. This is the case as the regressions in Table 2.3 organize bids into order statistics and this mechanically produces some degree of correlation, even if bidders were to ignore entirely the bidding behavior of others. Given that a bidder's bid is noisy and not completely determined by the own signal, information will be conveyed in the previous dropout(s). As an illustration, consider the case in which the previous dropout is very high, in fact higher than the expected current dropout conditional on own signal. Then, by definition, the expected current dropout conditional on previous dropout and own signal will be higher than the expected dropout level conditional on own signal only, thus leading to positive residual correlation between dropout orders.

This produces a mechanical correlation between dropouts and previous dropouts even if bidders do not pay any attention to the previous dropouts.

In order to use correlations among dropout prices as evidence for information processing, we need to move from an absolute to a comparative approach. In Table 2.4, we show excerpts from regressions where we pool data from the AV and the JEA and regress bids on the previous dropouts, signals, and interactions for the JEA. We refer to Table 16 in the Appendix for the full results. In the AV, where by design no information can be extracted from the unobservable bidding of others, we observe the mechanical correlation in dropout order statistics, as all coefficients on the just preceding dropouts $b_{j-1,t}$ are significant at conventional levels. Using the bidding in the AV as a benchmark, we measure the amount of information processing in the JEA by computing the additional correlation observed in the JEA compared to the AV. Table 2.4 shows that the slope parameters on every just-preceding bid are statistically larger in the JEA compared to the AV at each dropout order. As bids in the JEA are more strongly correlated than in the AV, we can conclude that bidders do react to the information contained in the bids of others.

To sum up, we conclude that subjects' bidding is consistent with them paying attention and responding to the bids of others in the JEA. Compared to the empirical best response, subjects pay too much attention to the most recent dropout and underweigh their own signal. How subjects' bidding weighs information in the own

[30] Note that R^2 is mechanically high in this regression because the best response is calculated as a linear function of the bids.

Table 2.4: Comparing information use in the AV and the JEA

	b_2	b_3	b_4
b_{j-1}	0.285	0.357	0.465
	(0.0309)	(0.0319)	(0.0440)
JEA × b_{j-1}	0.0871	0.195	0.244
	(0.0463)	(0.0533)	(0.0827)
Observations	1199	1199	1199
Adjusted R^2	.502	0.732	0.777

Notes: b_{j-1} denotes the just preceding dropout, e.g. it is b_1 for b_2. JEA is a dummy equal one for JEA auctions. Additional variables omitted from the table: all regressions include signal x, round t, all preceding dropouts (b_{j-k} for all $k \in \{1, \ldots, j-1\}$) as well as all these variables interacted with the JEA-dummy and a constant. For the full regression results, see Table 16 in the Appendix. Standard errors in parentheses and clustered at the matching group level.

signal relative to the observed dropout is qualitatively in line with Bayesian signal averaging. Still, our data does not accord with the prediction of the Bayesian signal averaging model that lower revenue will result in the JEA than in the AV. In the next Section we address how heterogeneity in early bidding contributes to understanding this puzzle.

2.6.3 Exploring heterogeneity in bidding

In this Section, we investigate whether individual-specific characteristics correlate with bidders' behavior in early dropouts. Bidding behavior in the JEA is quite heterogeneous, and especially so at early dropouts - in Table 2.3, we see that the R^2 increases in dropout orders. Additionally, especially at early dropout orders, subject-level fixed effects bring in significant additional explanatory power. Our finding that individual-specific characteristics matter more at early stages of bidding in the JEA agrees with the observation that deviations from the theoretical benchmark are less costly at these early stages in this auction format. For instance, a bidder who considers dropping out first may choose to overbid almost without costs: even when overbidding, the bidder can avoid winning by immediately dropping out when others do so. Likewise, if this bidder decides to drop somewhat earlier than the theoretical benchmark, this also happens almost without costs because the chances that all the others would drop before the theoretical benchmark is negligible if no other bidder has dropped out yet.

To shed light on whether there are systematic patterns in this heterogeneity in bidding behavior, we elicited subjects' social value orientation and their tendency for imitation at the end of the experiment for the last 13 sessions.[31] For the

[31]Another candidate to explain deviations from risk neutral Nash bidding is risk aversion. Because all auctions use the second-price rule and there is uncertainty about the value, risk aversion

imitation measurement, subject could choose to sample non-informative social information of prior participants instead of obtaining an informative signal. This behavior is consistent with a desire to imitate others. Participants that chose to reveal uninformative choices are classified as imitators, which applies to 26.9% of our participants.[32] Social value orientation is measured as an angle, where 0° correspond to a dictator keeping all to herself, 45° giving an equal amount to recipient and herself and 90° giving everything to the recipient. We find an average SVO of 21.13°, with a standard deviation of 19.93°.

To investigate whether these measures correlate with heterogeneity in bidding behavior we exploit that the estimations in Table 2.3 provide us with estimates of bidder fixed effects. In this context, the bidder fixed effect captures bidder-specific level shifts of bids, holding the use of information constant across bidders. Crucially, identical bidders may behave differently between different auction formats, especially as behavioral motives may be differentially triggered. Note that our within-estimations impose that the average bidder fixed effects have a mean of zero. This means that any bidder's fixed effect can be interpreted as a deviation from the average bidding behavior within our sample for each treatment.

Per participant, we average the fixed effects of the first and second dropouts as well as the fixed effects from the third and fourth dropout. For the AV and the JEA separately, we then regress the averaged fixed effects on subjects' social value orientation and imitation proneness. Table 2.5 presents OLS estimates.[33]

In both treatments, estimates imply that imitators are willing to bid higher than non-imitators. The effects ares similar in size but only significant in the AV, which may be due to a lack of power. In any case, the fact that the bids of imitators are not higher in the JEA than in the AV hints at the possibility that this measure may not only capture a tendency to imitate but also general overbidding caused by confusion.[34] From this perspective, imitation is not a good candidate to explain

will have a downward pressure on Nash equilibrium bids (see also Levin et al. (1996)). Given that observed bids tend to be higher than risk neutral Nash equilibrium bids, we think that risk aversion is a less important force in our experiment. Similarly, the heterogeneous behavior of early dropouts is not only incompatible with the symmetric equilibrium in Milgrom and Weber (1982), but also with the asymmetric equilibria in open auctions identified by Bikhchandani and Riley (1991). In addition, the asymmetric equilibria predict lower revenues in the JEA, while we observe revenue in excess of the symmetric Nash equilibrium.

[32] In a similar setting, Goeree and Yariv (2015) find that 34% of subjects chose such information.

[33] Note that the fixed effects are estimated, and thus may contain noise from the first stage in this estimation procedure. In the Appendix, Section 2.A.10, we show that point estimates are similar using WLS, which addresses concerns that some fixed effects might be estimated more noisily than others. These observations receive less weight in variance-weighted WLS. The estimates on SVO and Imitator in (2) are significant in this specification, which suggests that the noise in estimating fixed effects may be important. Point estimates with experienced bidders are mostly similar, see Table 13 in the Appendix. The coefficient on Imitator is insignificant across specifications (1) to (3), and the coefficient on SVO is significant and positive in (3) and (4).

[34] There are also situational factors that affect the extent of overbidding. For instance, Levin et al. (1996) and Goeree and Offerman (2002) find that subjects' overbidding enhances with the variance of the noise term in the signals.

Table 2.5: Bidder fixed effects and their characteristics

	(1)	(2)	(3)	(4)
	Average bidder fixed effect			
	b_1 & b_2		b_3 & b_4	
	AV	JEA	AV	JEA
SVO	0.125	-0.202	0.005	0.027
	(0.045)	(0.146)	(0.120)	(0.078)
Imitator	5.699	5.213	6.528	1.575
	(1.479)	(3.823)	(3.121)	(0.471)
Constant	-1.876	6.225	-4.674	-1.080
	(1.919)	(2.363)	(1.678)	(2.681)
Observations	50	40	50	40
Adjusted R^2	0.031	0.014	0.048	-0.031

Notes: Average fixed effects from regressing bids on available information for first and second vs. third and fourth dropout. SVO is a subject's social value orientation, in degrees. Imitator is a dummy variable equal one if a subject chose to retrieve social information when this contains no valuable information on the true state. Standard errors in parentheses, clustered at the matching group level.

differential bidding in the early dropouts between the two auction formats.

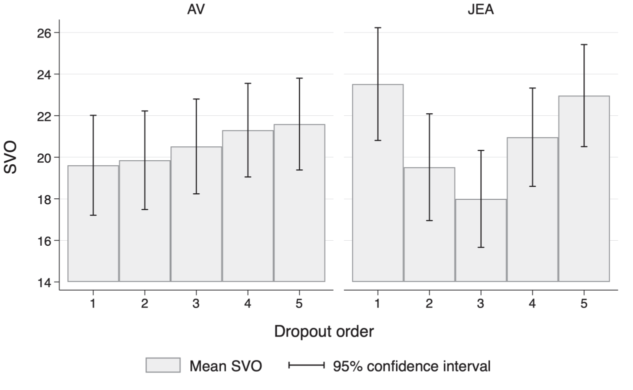
Our conjecture was that SVO would explain differences in the early bidding between the two auctions. The coefficient for SVO in column (2) of Table 2.5 is in line with the conjecture that spiteful bidders bid higher early in the JEA to drive up the price for others: only in JEA the coefficient is negative. However, the standard error is large, and we cannot conclude whether there is a negative effect or no effect of SVO on early bidding in JEA.[35]

Given that the evidence in Table 2.5 is not conclusive about the effect of SVO on early bidding in JEA, we looked further into how SVO affects bidding in the two auction formats. As we expected when we decided to measure SVO, competitors, those with below-median SVO, bid on average 71.3 in the first two dropouts in JEA, significantly more than in the AV where they bid on average 56.4 in the first two dropouts (Mann-Whitney U-test, 9 observations, p-value=0.086). This finding reflects that driving up the price for others is relatively cheap in the JEA, because this format allows bidders to enhance the price for others without much risk of actually winning the good. To put things into perspective, it is not clear that cooperators bid significantly more in the early dropouts of the JEA than in the AV (average bid of 59.6 in the AV versus 70.8 in the JEA; Mann-Whitney U-test, 9 observations, p-value=0.327).

[35]Somewhat surprisingly, more pro-social bidders bid slightly higher on average in the early bidding in the AV. Note that for the SVO, inequality averse participants are classified as pro-social. Therefore, bidding higher initially in the AV can be consistent with bidders trying to minimize payoff inequality, which might arise if an opponent wins at a low price. Pro-social bidders' behavior is not significantly different in the early bidding across auctions.

Figure 2.3 displays for each of the two auction formats the SVO per dropout order. Whereas there is a slight increase of SVO over dropout orders in the AV, there is a surprising but intuitive pattern in the JEA: Bidders who drop out first or last have on average a higher SVO than bidders who drop out in the middle. This suggests that cooperators decide at the start to either be nice and drop out first or to go all-in in a serious attempt to win the auction. By doing so they would refrain from driving up the price for others when they do not win. In the cases where they decide to win the auction, cooperators have to outbid spiteful bidders, who are bidding more aggressively than they would have in the AV. We find that cooperators (with an SVO above the median) end up significantly more often in an extreme position (either first or last) than competitive bidders (those with an SVO below the median): Mann-Whitney U-test, 8 observations, p-value=0.043. This pattern only materializes in the JEA: the same test for the AV is insignificant (p-value=0.917, 10 observations).[36]

Figure 2.3: SVO by dropout order



Overall, our suggestive evidence is consistent with the following picture of how SVO may affect bidding in the two auction formats. In the JEA, spiteful bidders tend to bid higher at the start than they would have in the AV, because the information about how many other bidders are still active makes it cheap for them to overbid. Without too much risk they can stay longer in the auction and drive up the price without actually winning the object. Cooperators on the other hand decide at the start of the auction whether or not they want to compete and win the object for sale. If their signal makes them decide it is better not to win, they drop out early

[36] To verify that the difference between treatments is significant, we run a logistic regression. We regress the binary dependent variable (0 if dropping out first or last, 1 otherwise) on SVO, Imitator and signal, a treatment dummy as well as interactions of all independent variables and treatment. While the coefficient on SVO is not significant (p-value=0.817), the coefficient on the interaction of JEA and SVO is negative and (weakly) significant (p-value=0.071).

and by doing so refrain from enhancing the price for others. If they decide to compete, they relatively often end up winning the auction. In this case, they have to outbid spiteful bidders who tend to bid higher than they would have in the AV.

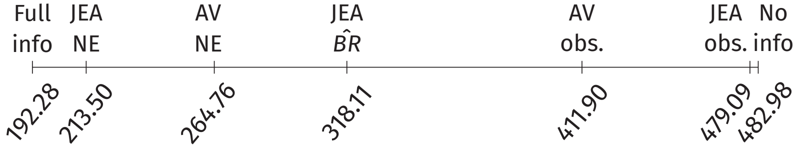
2.6.4 Information aggregation

Previously we showed that bidders engage in overbidding (Figure 2.2). Even bidders who depart from rationality can convey information in their bids, or infer information from others' bids. For instance, if bidders follow a behavioral model, then their bids will still convey information about their signals. If this is anticipated by other bidders, bidders can still process this information in their own bids. In this Section, we investigate the extent to which bidders aggregate information in the different auction formats. The measure of information aggregation is the squared distance between the price and the common value, as discussed in Section 2.4.

We first present a comparison between the JEA and the AV, the two auctions that differ only in the information on previous dropouts. Both rational and behavioral benchmarks predict that additional information improves bidders' precision in estimating the value. This prediction, however, is not borne out in our data. Figure 2.4 plots the distance between price and value that is actually observed in the data. For a comparison, it also includes Nash equilibrium predictions.

As it turns out, the theoretically predicted ranking is reversed in our data. The observed squared distance in the AV is 411.9, and *increases* to 479.1 when *more* information is available in the JEA. This difference is statistically significant (p-value=0.028, MWU, 10 observations per treatment). Actually, the JEA aggregates almost no information. The observed squared distance of the JEA is not statistically different from the No Information benchmark, where the price is set equal to the prior mean of the common value, ignoring all information contained in signals.[37]

Figure 2.4: Squared distance to common value



There can be two reasons why information aggregation fails in open ascending auctions: i) there is not sufficient informational content in observable bids (*infor-*

[37]We verified that the same ordering in our results on information aggregation is observed when using the squared distance to the Full information benchmark as a measure, instead of the squared distance to the common value. The latter does not directly control for variance in signals conditional on the common value. In our analysis, this is captured by the distance to the common value measured in the Full information benchmark.

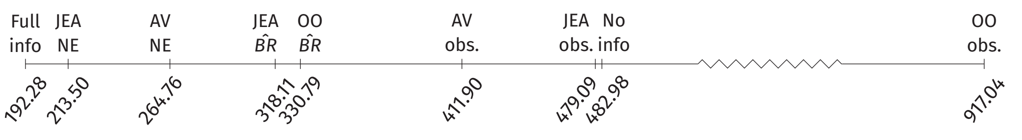
mation revelation) (ii) bidders do not process the available information as a rational bidder would (*information processing*). To isolate the two forces, we use an empirical best response $\widehat{BR}$ as described in Section 2.6.2, given observed bidding behavior of early dropouts.

Note that $\widehat{BR}$ is a statistic that separates between information processing and revelation. It represents the level at which the two remaining bidders best respond to each other, when they incorporate information available in the experiment. The gap between the observed level of information aggregation (JEA obs.) and the maximal level of aggregation achievable given the available information ($\widehat{BR}$) serves as our measure of the failure of information processing. Failure in information revelation is measured by the distance between $\widehat{BR}$ and JEA NE, as in Nash equilibrium signals from earlier dropouts can be inferred perfectly. From inspecting the segment, it is apparent that both forces play a role: Information in the JEA is dissipated by noisy early dropouts and further processed in a sub-optimal way.

Using the empirical best response, we can also provide a lower bound for the importance of heterogeneity in early dropouts on the failure of information aggregation. Using bidder fixed effects, instead of only session fixed effects, when estimating signals from observed bids, the squared distance of the empirical best response to the common value reduces to 303.0. The difference of this new benchmark to the empirical best response is significant (Wilcoxon signed-rank test, 10 observations, p-value=0.047). Note that this is a lower bound due to the role played by individual heterogeneity, as it ignores the additional gains brought about by bidders iteratively making the intermediate dropouts more precise, something they cannot do as the identity of other bidders is not observed.

Lastly, when it comes to our third auction format, the OO, the higher revenue that we observe is not caused by a higher degree of information aggregation in this format. To the contrary, in the OO overbidding is so severe that the price is a highly inaccurate predictor of the common value, resulting in a very imprecise measure of information aggregation, with a squared distance of 917.0. If bidders had simply ignored their private signal and the bidding of others, and bid the prior mean value according to the no-information benchmark, this distance would shrink to 483.0. Figure 2.5 presents the information aggregation benchmarks of the OO in comparison to the other auction formats.

Figure 2.5: Squared distance to common value, including the OO



This lack of information aggregation cannot be attributed to information in bids being obfuscated. The same decomposition as performed for the JEA shows that the second-highest bidder in the OO would be able to predict the common value relatively well if they attempted to bid the conditional expected value as in the JEA, by incorporating the own signal and the maximal bids of the three non-winners. This is a conservative measure of how much information is potentially available in the OO, because it ignores other, possibly informative, observables such as the time elapsed between bids, the size of the jump bids, or the number of returning bidders.

2.6.5 Bidding in Oral Outcry auctions

We have previously seen that revenue is higher in the OO than in the other two formats. Also, information aggregation in this format fails.

The OO differs from the two clock-formats in how bids can be submitted. In both the AV and the JEA, the price rises at an exogenously set pace and bidders can only decide whether to leave or remain at every price. In the OO, bidders can submit their own bids. In the following, we discuss two ways in which this change matters: it may trigger a quasi-endowment effect in bidders, as well as allow for non-incremental jump bidding.

During an Oral Outcry auction, a standing bidder is identified, who is the highest bidder at that moment. The previous literature has established that this can induce a so-called auction fever (Heyman et al., 2004; Ehrhart et al., 2015). A standing bidder may get used to the feeling of winning the good and become prepared to bid higher than she originally intended. If that happens, auction fever triggers a quasi-endowment effect.

Auction fever is in agreement with the fact that, beyond the average revenue already being significantly higher, we also observe relatively many extreme auction revenues in the OO compared to the other two formats. For example, only 1.3% of all common values are in the right hand tail of the common value distribution, at values above 150. In both the AV and the JEA, less than 1% of auctions end up at revenues above 150. In the OO in turn, 7.3% of auctions conclude at prices above 150, suggesting that especially this format triggers strong mispricing.

To evaluate the impact of auction fever, we use bidder's exogenously measured inclination to succumb to the endowment effect, and perform a median split based on this measure.[38] There are two main effects: (i) bidders do not systematically differ in how often they win auctions (MWU-test, p-value=.773), thus bidding behavior appears similar at first; (ii) whenever they win an auction, bidders with stronger endowment effects generate higher losses than their peers, as their total profits

[38]We normalize both measures to mean 0, variance 1, then take the average response as a measure of the endowment effect. We compare matching group averages of those bidders with above and below median endowment effects, yielding 8 observations (4 matching groups, one observation above and below the median each).

are significantly different (MWU-test, p-value=.083)[39], thus when becoming active and winning an auction, bidders with strong endowment effects lose more money. This evidence provides support for the conjecture that the OO activates auction fever among people who suffer from the endowment effect.

A second important feature of the OO is that bidders can submit non-incremental jump bids. Theoretical analyses of jump bidding suggest that this may be a profitable strategy for a jump-bidder. Avery (1998) derives equilibria in which jump bidding is used for signaling high value estimates, which predicts increased profits for the winner. Ettinger and Michelucci (2016) show that jump bidding can be used to obfuscate information. Naturally, behavioral factors may also affect jump bidding. For example, impatient bidders who are determined to win an auction quickly might frequently submit jump bids which lead them to win auctions in cases in which they have initially overestimated the value, an error which could have been corrected in the price discovery of an incremental bidding process. These behavioral factors suggest that jump bidding may also be costly and reduce winners' profits. In the following, we evaluate the effect of jump bidding in the OO auctions, focusing on whether jump bidding increases profits.

Note that within our auctions and due to the second-price rule in setting the current price, jump bids are only revealed if at least one other bidder continues to bid. While submitting additional bids, other bidders learn that the jump bidder has entered an aggressive jump bid, as the jump bidder continues to be the standing bidder. The level of the jump bid is revealed at the moment that some other bidder enters a bid higher than the jump bid. This feature captures how jump bidding in popular auction formats occurs. As such, we expect weaker effects of jump bidding than in first-price formats, where the level of a jump bid is revealed immediately. In our analysis, we will show that even this subtle effect of jump bidding matters for outcomes.

As a measure of jump bidding we construct the total jump bid of each bidder in each round. To do so, we first calculate the increment of a new bid above the current price, the second highest bid submitted in previous bidding rounds, at the moment the new bid was submitted. By the rules of the auction, this increment varies between 1 point, which is the minimum increment, and 200 points, if the maximum possible bid was submitted straight at the start of the auction. Often, the same bidder submits multiple bids. We denote the sum of all increments for one bidder across one auction as the total jump bid of this bidder.

We observe extensive jump bidding: 21.6% of bids exceed the current price by at least 20 points, and 11.2% by at least 50 points. Jump bidding is most prevalent at the start of an auction, where 81.7% of entered bids are at least 20 points, and

[39] This analysis is robust to performing a median split based on the first principal component obtained from the two measures of the endowment effect, with p-values of .564 and .083, respectively (MWU-tests, 8 observations).

60.4% are at least 50 points high. Jump bidding also gains in popularity over time: in the first 15 rounds, the average jump bid at the start of an auction is 53.8, this increases to an average of 61.6 in the last 15 rounds.[40]

In Table 2.6, we show regression results on the use and effect of jump bids. The main regressor of interest is the total jump bid, the sum of all bid increments by each bidder in an auction. However, in regressions studying the effect of jump bids, these bids are likely endogenous as strategies adjust to observed jump bids submitted earlier. To account for this, we rely on instruments generated from other rounds, which capture an individual bidder's proneness for jump bidding. As instruments, we use the average total jump bid of each bidder across all other rounds, as well as the maximum bid increment in any of the other rounds. Using 2SLS, we then predict in a first stage the total jump bid in the current round using the two instruments and other variables, such as the signal x. In the second stage, we regress our dependent variables of interest on the predicted total jump bid and some other variables. This provides a clean identification of the effect of jump bids. For relevance, we here assume that a bidder's proneness to jump bid in other rounds correlates with this bidder's jump bidding in the particular round. For the exclusion restriction, we assume that other rounds' jump bids only affect outcomes through the bidding in that particular round. We think that this is plausible for two reasons. First, the only way of affecting a particular round's outcomes is only through bidding in that round, while other rounds' bids (our instruments) cannot directly affect outcomes by the auction rules. Second, as for potential indirect effects, this exclusion is reinforced by our experimental design, as every round bidders draw new random signals and are allocated to new random groups within the matching group, which limits the effects other rounds' behavior may have on this round's competitors. In the Appendix Section 2.A.11, we present first-stage regression results in combination with a robustness check based on the use of only the average total jump bids across all other rounds as instrument. We show that the instruments are relevant, as all first stage regressions are significant at conventional levels, with Kleibergen-Paap F-statistics of 96.4 or greater. In addition, we show that we cannot reject the null hypothesis that the instruments are valid, with p-values of the Hansen J-statistic of .582 or higher.

Column (1) presents results of regressing these jump bids on bidders' information. As predicted by theoretical models, bidders with higher signals submit higher jump bids. The size of the jump bid is not significantly increasing over time. Inter-

[40] In the first six sessions, the bidding rounds at which a bid was submitted was not saved correctly due to a programming mistake. We reconstructed this data by the time stamp at which bids were submitted. In 10.7% of the bids in these sessions, this classification is potentially ambiguous, we assumed that bids were submitted in a later bidding round in these cases, which leads to potentially fewer bids being considered for our type of analysis. The results we present are robust to instead assuming that these bids were submitted simultaneously, or randomizing this classification. Also, only using data from the last four sessions, where this error was corrected, yields similar results.

Table 2.6: Effect of jump bids in the OO

	(1)	(2)	(3)	(4)
	Jump bid	Pr(win)	Profits	Winners' profits
Total jump bid (IV)		0.350	-0.261	-0.316
		(0.083)	(0.115)	(0.133)
x	0.276	0.144	-0.067	-0.029
	(0.031)	(0.038)	(0.037)	(0.042)
t	-0.138		0.877	0.784
	(0.124)		(0.169)	(0.154)
V			0.624	0.633
			(0.046)	(0.064)
Constant	30.433	-12.306	-66.653	-58.996
	(5.897)	(2.656)	(7.000)	(9.917)
Observations	2687	2687	2687	600
Adjusted R^2	0.070	0.102	0.291	0.287
Estimation	OLS	2SLS	2SLS	2SLS

Notes: Jump bid is the increment of a bid beyond the current price at the moment the bid was submitted. In (1), we regress total jump bid on bidders' signals and round t. In (2) to (4), we use 2SLS, where we instrument using the average total jump bid and the maximum bid increment in other rounds. (2) is the ex-post probability of winning, which is a dummy equal to 100 if a bidder wins the auction, 0 otherwise. Mean earnings are a participants' average earning across all auctions, winners' profits are the earnings for the auctions which a participant won. x is the submitting bidder's signal in round t. V represents the common value. Standard errors in parentheses, clustered at the matching group level.

estingly, this suggests that bidders with more experience shift their jump bids to the start of the auction, as we do observe a significant increase in jump bidding at the start over time while overall jump bidding remains constant.[41]

Table 2.6 also presents an analysis of the effects of jump bids. In (2), the dependent variable is a dummy equal to one when a bidder wins the auction, 0 otherwise. Here we show that, controlling for own signal, a larger jump bid increases the likelihood to win the auction. This is consistent with the signaling motive in the theoretical literature.

Models (3) and (4) then study how profits are affected by the size of the jump bid. Contrary to theoretical predictions, profits are significantly decreasing in the size of the jump.

Winners on average lose money in the OO and, by submitting a jump bid, participants select into this group of winners making a loss. Model (4) studies whether this selection effect is the full reason beyond the negative relation between jump bidding and profits. We do so by restricting the analysis to bidders who end up winning the auction. We find that even within this group of bidders, the size of the

[41] In the last 4 sessions, we elicited how much participants agreed with several motives for jump bidding in the questionnaire, see Appendix Section 2.A.12 for details. If we include those in (1) as controls, the only statement that correlates significantly with the size of the jump bid is "I tried to deter other bidders from bidding by entering a bid much higher than the current price."

jump bid decreases profits further.

Results for experienced bidders are similar, see Table 14 in the Appendix. In later rounds, jump bidding has a slightly less pronounced effect on earnings and profits. Still, jump bidding continues to be a disadvantageous strategy also with more experience, while jump bidding is in fact used more extensively later on.

2.7 Conclusion

In this paper, we study some salient factors that can contribute to the popularity of open ascending auctions. In particular, we assess the roles that endogenous information aggregation and behavioral biases play in explaining their prevalence. In a common value setting, we compare two clock auctions, the ascending Vickrey auction (AV) and the Japanese-English auction (JEA), which differ in irrevocable exits of bidders being observable only in the latter. We also study the Oral Outcry auction (OO), an auction format modeled to approximate popular designs, in which bidders choose how much information they want to reveal through bids.

In agreement with their popularity, we find that the OO is most successful in raising revenue. The JEA and the AV both raise higher revenue than expected in the symmetric Nash equilibrium. In contradiction to some behavioral models that predict higher revenue for the AV, we do not reject equality of revenue between the JEA and the AV. We find that information aggregation fails in the JEA. Bidding in the JEA reflects a worse estimate of the common value than in the AV.

It is not the case that bidders do not pay attention to early exits in the JEA. To the contrary, bids correlate more strongly with the most recent dropout than in the AV benchmark.[42] The bidding pattern, however, deviates from what would be observed when bidders bid according to the Nash equilibrium benchmark, and also from what would be observed when they choose empirical best responses. The relative weight of how bidders incorporate information is best captured by a Bayesian signal averaging heuristic. However, all models incorporating public information underestimate bid levels and bidders in the JEA do not use public information sufficiently to tamper the winner's curse, as predicted by signal averaging models.

At the same time, bidding behavior conveys less information than the theoretical benchmark. The information reflected in early dropouts of the JEA is partly obfuscated by heterogeneity in the bidding of early leavers. In agreement with the fact that it is relatively cheap to drive up the price in the JEA, spiteful bidders may stay longer in the JEA than in the AV, forcing cooperators to stay longer in the cases where they want to win. Such spiteful bidding by early leavers may neutralize the

[42] Note that Hoelzl and Rustichini (2005) find that people are underconfident in complicated tasks. Their result agrees with our finding that bidders place more weight to what others do in the strategically complicated common value setup.

revenue diminishing force of the Bayesian signal averaging heuristic. Our support for a spiteful motive resonates with some empirical findings in other auction environments (Andreoni et al. (2007), Bartling and Netzer (2016)).

In the OO, bidders choose how much information to reveal through their bids. Overall, bids in the OO convey as much information as those in the JEA. However, in the OO-format the available information is least well processed, and the price paid by the winner is the worst approximation of the common value among all three formats.

Instead, the OO activates some behavioral biases that enhance revenue. Bidders who suffer from endowment effects lose more money in these auctions. When they become the provisional winner, auction fever strikes and they become willing to submit higher bids than otherwise expected. In addition, the OO encourages bidders to submit jump bids. In contrast to the theoretical literature, jump bids do not enhance winners' expected profits. Jump bidders are more likely to win the auction, but they tend to lose money doing so.

Oral Outcry auctions may be popular not because they allow bidders to aggregate information. Instead, a more important rationale for using Oral Outcry auctions may be that they activate revenue enhancing behavioral biases.

Appendix

2.A eBay data, revenue predictions, and additional analyses

We start with an analysis of information effects on eBay. Then, we discuss revenue predictions for different parameterizations, and we present behavioral models. We present cursed equilibrium for the JEA and show the results for a horse race between different models in the AV and the JEA. We also present some robustness checks of the analyses in the main paper.

2.A.1 Information usage on eBay

eBay gives bidders access to a detailed bidding history during an auction. To investigate the effects of information use on eBay, we collected data from eBay-auctions between August 8 and September 27, 2019. We chose one of the most frequently auctioned cellphones in that moment, the Apple iPhone X, 64GB, with a total of 1194 phones. These phones vary considerably in the condition they are sold, with buyers potentially making inference on the phone's value for example based on pictures, descriptions, or the sellers' reviews. Crucially, the interested bidders can study others' bids, which may allow them to learn about a specific phone's value.

To explore this endogenous learning, we perform median splits of the data on a number of dimensions which might convey information during the auction, such as the interim price, the number of bids per bidder and the average increments between consecutive bids. We then study if median splits along these variables explain variation in the final price. Before performing the median splits, we regress the final price on a number of observable characteristics, such as the (exogenously set) length of the auction, the reserve price, the number of bids and the number of bidders, as well as the review count of the seller. By extracting the residuals obtained from these regressions, we factor out all variation that can be explained by these observable exogenous characteristics.[43]

[43]The described pattern is also found when directly comparing prices across the same median

In Figure 2.A.1, we plot the average residual for cellphones above and below the median for each of the three different splits of the data. First, we split the auctions based on the price of the cellphone when half of the total length of the auction elapsed. Second, we perform a split based on the average increment per bidder (given by the price at the end less the reserve price, divided by the number of bids submitted). This captures the degree of jump bidding observed. Third, we split on the number of bids per bidder.[44]

We observe that the average residuals of the iPhones are different depending on which half of the data they are categorized in. The effects are also quite sizable, as the average price of the phones is \$425.8, standard deviation of \$175.8. This implies that there is systematic variation in the prices of these phones which cannot be explained by the observable exogenous characteristics. Instead, this residual variation can be explained by the categories we perform the median split by, and these variables may capture information generated endogenously in the auction. This indicates that information revelation might matter.

Crucially, this type of observational data cannot be used to establish unambiguously that information revelation is taking place and what effect revealing information has. First of all, the direction of causality is not clear (e.g., are expensive phones attracting many bids, or do many bids increase the price?). More importantly, we cannot evaluate what information is processed without observing bidders' information sets and the underlying value of the good to be sold. Also, we are unable to determine the impact of information without providing a control condition where no information is being revealed. However, this is possible in our laboratory experiment.

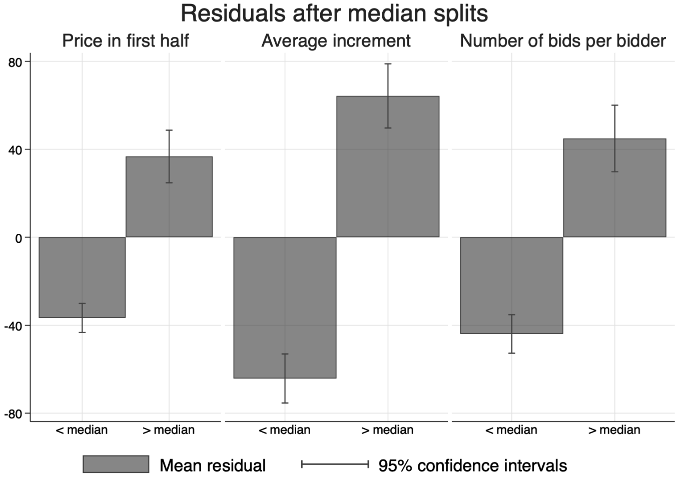
2.A.2 Revenue predictions for different parameterizations

In the choice of parameterizing the mean and variances of the values and signals for this experiment, we simulated revenues of the AV and the JEA to generate predictions of the symmetric Nash equilibrium. In Table 2.A.1, we report results of these simulations. For each parameterization, we draw 50,000 sets of signals according to the procedures of the draws for the experiment, then calculate average revenues based on all simulated bids. In Table 2.A.1, R_{AV} are revenues in the AV, R_{JEA} are revenues in the JEA. We simulate different parameterizations, for the full set of parameters $(\mu, \sigma_V, \sigma_\epsilon)$, which is the mean μ and standard deviation σ_V of the value distribution as well as the standard deviation σ_ϵ of the error distribution. Within each parameterization, we give mean revenues in the first row, and the standard deviation of the revenue in the second row. From the Table, it is clear that revenue

splits.

[44]As we perform the median split on these characteristics, we do not residualize the reserve price when performing the split on the price in the first half, and we do not residualize the number of bids and the number of bidders when performing the split on the number of bids per bidder.

Figure 2.A.1: Residuals obtained from regressions of final price



differences of the Nash equilibrium are quite small across specifications. Theoretical revenue differences for uniformly distributed values and errors are similarly low, the case studied by the previous literature.

Additionally, we calculated revenue differences for varying numbers of bidders. In Table 2.A.2, for the paramaterization used in this experiment, $(\mu, \sigma_V, \sigma_\epsilon) = (100, 25, 35)$, we state Nash equilibrium revenue differences for different numbers of bidders. Evidently, theoretical revenue differences between treatments are not driven by the size of our auctions.

Table 2.A.1: Revenue Nash predictions with varying parameters

$(\mu, \sigma_V, \sigma_\epsilon)$		R_{AV}	R_{JEA}
(50, 10, 12)	Mean	48.3555	48.7510
	Standard deviation	(8.3056)	(8.6844)
(100, 10, 12)	Mean	98.3877	98.7540
	Standard deviation	(8.4207)	(8.8667)
(100, 10, 30)	Mean	98.3874	98.5852
	Standard deviation	(5.3808)	(6.0436)
(100, 20, 20)	Mean	96.9314	97.6790
	Standard deviation	(17.6035)	(18.2750)
(100, 20, 30)	Mean	96.5708	97.4156
	Standard deviation	(15.1637)	(6.0436)
(100, 20, 40)	Mean	96.1720	97.2063
	Standard deviation	(12.7600)	(12.4809)
(100, 30, 20)	Mean	96.6016	97.5939
	Standard deviation	(26.8225)	(27.2350)
(100, 30, 30)	Mean	95.4797	96.9314
	Standard deviation	(23.5650)	(24.1363)
(100, 40, 40)	Mean	92.9095	95.5161
	Standard deviation	(26.4453)	(26.4859)
(200, 40, 40)	Mean	194.1535	195.6345
	Standard deviation	(35.1111)	(36.3651)

Table 2.A.2: Revenue predictions varying number of bidders, $(\mu, \sigma_V, \sigma_\epsilon) = (100, 25, 35)$

Number of bidders		R_{AV}	R_{JEA}
3	Mean	93.4861	94.2020
	SD	(18.0011)	(18.2828)
5	Mean	95.6073	96.9290
	SD	(18.2036)	(18.9653)
7	Mean	96.1953	97.7570
	SD	(18.1058)	(19.0956)
9	Mean	96.6706	98.4264
	SD	(17.7684)	(18.8915)
11	Mean	96.6956	98.5875
	SD	(17.6533)	(18.7593)

2.A.3 Naïve models

In this Section we discuss some behavioral models that have been discussed in the literature, to explain observed behavior in the AV and the JEA. In the AV, there are two principal behavioral models which might capture bidding behavior. First is the "bid signal"-heuristic, according to which bidders might just enter a bid equal to their own signal:

$$b(x_i) = x_i$$

In expectation, this will result in overbidding of the winning bidder, as the bidder neither includes information on the distribution of signals and values nor considers the informativeness of winning.

Second, somewhat more sophisticated bidders will incorporate information about the prior distribution of the value. In the "Bayesian bid signal"-heuristic, bidders still suffer from the Winner's curse, but bid the expected value of the good for sale, conditional on one's signal, as in Goeree and Offerman (2003b):[45]

$$b(x_i) = \mathbb{E}[V|x_i] = x_i - \mathbb{E}[\epsilon_i|x_i]$$

To explain behavior in the JEA, Levin et al. (1996) propose a "signal averaging rule", according to which bidders bid an equally weighted average of their own signal and the signals of their fellow bidders, revealed from the previous dropouts. This rule incorporates revealed information in a natural way.

Close to the bid-signal heuristic is the "symmetric signal averaging rule", introduced by Levin et al. (1996). Here, all bidders assume that all other bidders follow this rule as well. After k bidders dropped out, with the vector of revealed signals being $\mathbf{Y}_i$, this implies the following bid:

$$b_j(x_i, \mathbf{Y}_i) = \frac{1}{j} x_i + \frac{1}{j} \sum_{k=1}^{j-1} Y_{i}, (k)$$

This formulation can be rewritten to only depend on the last dropout price, for the vector of previous dropout prices $\mathbf{p}_{j-1}$, p_{j-1} being the $j-1$-th observed dropout:

$$b_j(x_i, \mathbf{p}_{j-1}) = x_i + \frac{j-1}{j} p_{j-1}$$

A variant of this rule is the "asymmetric signal averaging rule", according to which bidders assume that other dropouts are based on the heuristic of bidding equal

[45]Within the setup of our experiment, we can use that $\epsilon_i|x_i \sim N\left(\frac{\sigma_\epsilon^2(x_i-\mu)}{\sigma_\epsilon^2+\sigma_V^2}, \frac{\sigma_\epsilon^2\sigma_V^2}{\sigma_\epsilon^2+\sigma_V^2}\right)$. As derived in Goeree and Offerman (2003b): $b(x_i) = \frac{\sigma_V^2 x_i+\sigma_\epsilon^2\mu}{\sigma_V^2+\sigma_\epsilon^2}$

to signal. This would enable bidders to more easily include others' information. Additionally, it appears to be an intuitive rule given the information salient in the auction process. If bidders follow the asymmetric signal averaging rule, with p_{j-1} being the $j-1$-th dropout, bids are given by:

$$b_j(x_i, \mathbf{p}_{j-1}) = \frac{1}{j}x_i + \frac{1}{j}\sum_{k=1}^{j-1} p_k$$

Similar to the "Bayesian bid signal" heuristic, signal averaging rules can also incorporate information about the prior. According to the "Bayesian signal averaging rule", bidders apply Bayes rule in combination with the symmetric signal averaging rule. In this case, after $j-1$ observed dropouts, bidder i calculates the average of available signals $\bar{x}_i = \frac{1}{j}x_i + \frac{1}{j}\sum_{k=1}^{k-1} Y_{i,(5-k)}$:

$$b(\bar{x}_i) = \frac{\sigma_V^2 \bar{x}_i + \sigma_\epsilon^2 \mu}{\sigma_V^2 + \sigma_\epsilon^2}$$

While it is unlikely that a bidder that is sophisticated enough to apply Bayes rule correctly would rely on a signal averaging rule, Bayesian signal averaging is most of all useful in anchoring bidding to the prior, compared to standard signal averaging. Even if Bayes rule in itself is too sophisticated, it is also unlikely that bidders rely purely on averaging available signals and fully ignoring all information on the prior distribution of values.

2.A.4 Cursed equilibrium in the JEA

As shown by Eyster and Rabin (2005), the expected payoffs from winning in the χ-virtual common value auction is given by:

$$\pi(V, p) = (1-\chi)V + \chi\mathbb{E}[V|X_i = x_i] - p$$

for price p, compared to winners' payoff in Nash equilibrium of $\pi(V, p) = V - p$. We continue to analyze a game where χ is homogeneous across participants, as well as during the auction. This implies bidder's cursedness is not affected by observing other's bids. From Milgrom and Weber (1982), we know that a symmetric Bayes Nash equilibrium in the JEA is given by

$$b_j(x_i) = \mathbb{E}\Big[V|X_i = x_i, Y_{i,(1)} = x_i, \ldots, Y_{i,(5-j)} = x_i, p_1 = b_1\left(Y_{i,(4)}\right), \ldots$$
$$\ldots, p_{j-1} = b_{j-1}\left(Y_{i,(5-j+1)}\right)\Big]$$

This conditional expected value in a χ-virtual game is equal to

$$\begin{aligned}
&\mathbb{E}\Big[(1-\chi)V + \chi\mathbb{E}[V|X_i = x_i]\Big|X_i = x_i, Y_{i,(1)} = x_i, \dots, Y_{i,(5-j-1)} = x_i, p_1 = b_1\left(Y_{i,(4)}\right), \dots \\
&\qquad\qquad \dots, p_{j-1} = b_{j-1}\left(Y_{i,(5-j+1)}\right)\Big] \\
&= (1-\chi)\mathbb{E}[V|X_i = x_i, Y_{i,(1)} = x_i, \dots, Y_{i,(5-j-1)} = x_i, p_1 = b_1\left(Y_{i,(4)}\right), \dots, p_{j-1} = b_{j-1}\left(Y_{i,(5-j+1)}\right)] \\
&\qquad + \chi\mathbb{E}[V|X_i = x_i]
\end{aligned}$$

As Milgrom and Weber (1982) have shown that $b_j(x_i)$ is a Nash equilibrium in the original game, the expression above is a symmetric cursed equilibrium in a χ-virtual game, for $\chi \in [0, 1]$.

To employ cursed equilibrium, we need to estimate the additional parameter χ. This also provides a measure of the cursedness of our subjects.

We estimate for the AV:

$$\begin{aligned}
b(x_i) = &\underbrace{\left(x_i - \frac{\int_{-\infty}^{\infty} \epsilon\phi_V(x_i-\epsilon)\phi_\epsilon^2(\epsilon)\Phi_\epsilon^3(\epsilon)d\epsilon}{\int_{-\infty}^{\infty} \phi_V(x_i-\epsilon)\phi_\epsilon^2(\epsilon)\Phi_\epsilon^3(\epsilon)d\epsilon}\right)}_{=w_i} + \\
&+\chi\underbrace{\left(\frac{\sigma_\epsilon^2 x_i + \sigma_\epsilon^2\mu}{\sigma_V^2 + \sigma_\epsilon^2} + \frac{\int_{-\infty}^{\infty} \epsilon\phi_V(x_i-\epsilon)\phi_\epsilon^2(\epsilon)\Phi_\epsilon^3(\epsilon)d\epsilon}{\int_{-\infty}^{\infty} \phi_V(x_i-\epsilon)\phi_\epsilon^2(\epsilon)\Phi_\epsilon^3(\epsilon)d\epsilon}\right)}_{=z_i}
\end{aligned}$$

We simulate all terms using bidders' signals and then regress bids using OLS:

$$b(x_i) = \beta_1 w_i + \beta_2 z_i$$

In a constrained regression, we impose no constant and $\beta_1 = 1$. Then, $\beta_2 = \chi$. For the JEA, we proceed similarly. We first simulate Nash equilibrium bids, based on the inference of observed dropouts.[46] We also use OLS to estimate χ in the following equation:

$$b_k(x_i, \bar{x}_i) = \underbrace{\frac{5\bar{x}_i\sigma_V^2 + \mu\sigma_\epsilon^2}{5\sigma_V^2 + \sigma_\epsilon^2}}_{=w_i} + \chi\underbrace{\left(\frac{\sigma_V^2 x_i + \sigma_\epsilon^2\mu}{\sigma_V^2 + \sigma_\epsilon^2} - \frac{5\bar{x}_i\sigma_V^2 + \mu\sigma_\epsilon^2}{5\sigma_V^2 + \sigma_\epsilon^2}\right)}_{=z_i}$$

[46]Note that we do not use the theoretical, unobserved signals other bidders hold for simulations. These predictions differ from the Nash equilibrium predictions by not incorporating realized dropout prices, but these do require inferences bidders are not able to make given the observed dropouts in the laboratory.

We regress dropout prices on x_1, x_2:

$$b_k(x_i, \bar{x}_i) = \beta_1 w_i + \beta_2 z_i$$

Again using constrained regression with no constant and $\beta_1 = 1$, we obtain $\beta_2 = \chi$. In Table 2.A.3, we summarize the regression results. We estimate χ once for all pooled data and once for the fourth dropouts in (2) and (4), respectively.

The coefficient on z_i is $\hat{\chi}$, which turns out to be low in our sample. Recall that $\chi = 0$ corresponds to Nash equilibrium bidding, thus our bidding behavior appears to be close to this benchmark judged by the cursedness of the participant pool.

Table 2.A.3: Estimating χ

	(1) AV	(2) AV, d_4	(3) JEA	(4) JEA, d_4
w_i	1.000	1.000	1.000	1.000
	(.)	(.)	(.)	(.)
z_i (for $\hat{\chi}$)	0.058	0.985	-0.137	0.186
	(0.058)	(0.072)	(0.059)	(0.041)
Observations	2417	598	2453	599
Estimation	OLS	OLS	OLS	OLS

Notes: Standard errors in parentheses and clustered at the matching group level.

2.A.5 Horse race between models

To understand how bidding behavior can be characterized, we analyze how well individual bids can be predicted by the available models. For each bid in each round and based on the available information, such as the signals and observed dropouts, we simulate all models described previously. Then, we calculate the distance between each of the bids and all theoretical predictions, using the squared difference. Denote $\delta_{i,t,m}$ the distance of the bid by bidder i in round t, compared to model m. $b_{j;i,t}$ is the observed dropout price of bidder i in round t, dropping out at order j.[47] $b^m_{j;i,t}$ is the theoretically predicted dropout price by model m for this bid. The distance $\delta_{i,t,m}$ is given by:

$$\delta_{i,t,m} = \left(b_{j;i,t} - b^m_{j;i,t}\right)^2$$

After calculating each of the distances for all bids and models, we can determine which model fits individual bids best. Then, we calculate the average distance of all models across all bids. In other words, as a measure of fit, we state the mean squared error in predicting bids for each model.

[47] Here, we only consider bidders who actively choose to drop out.

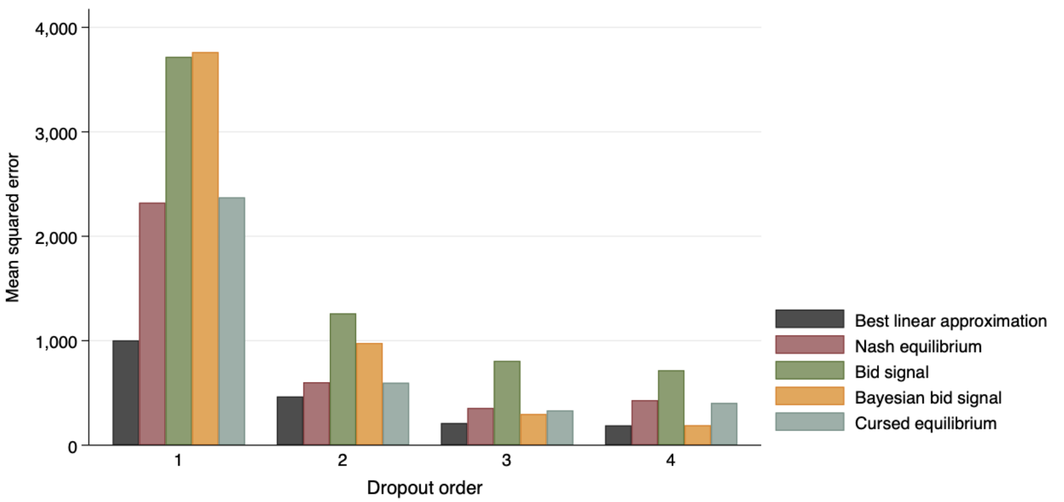
To allow for a comparison of the size of the error, we also provide a benchmark linear rule.[48] For this, we run regressions which use the identical available information as the models, which is the bidder's signal and dropouts in case they are observable. We then state the mean squared error of this prediction. By design, this minimizes this error within the class of linear models, which nests all models which are competing in this analysis.

In our analysis, we distinguish bids by dropout order. The first dropout order are all bidders who drop out first in an auction, and so forth. Note that the fourth dropout order is the most interesting, as these determine revenue.

Horse race for the AV

We start by comparing bidding behavior in the AV to the benchmarks. At this stage, we consider four models. We compare the Nash equilibrium benchmark and three naïve models: i) bidders exactly bid their signal, ii) Bayesian bid signal, where bidders suffer from the winner's curse, but do take the base rate into account, as in Goeree and Offerman (2003b), and iii) bidders in cursed equilibrium as proposed by Eyster and Rabin (2005), with an estimated $\hat{\chi}$ = .0578. Next to it, we provide the mean squared error of the linear benchmark at each dropout order, where only the private information signal is observable by bidders.

Figure 2.A.2: Mean squared error of model predictions in the AV



The first key insight is the fact that bidding behavior at early dropout orders is substantially less well predicted, as the mean squared error of the benchmark is

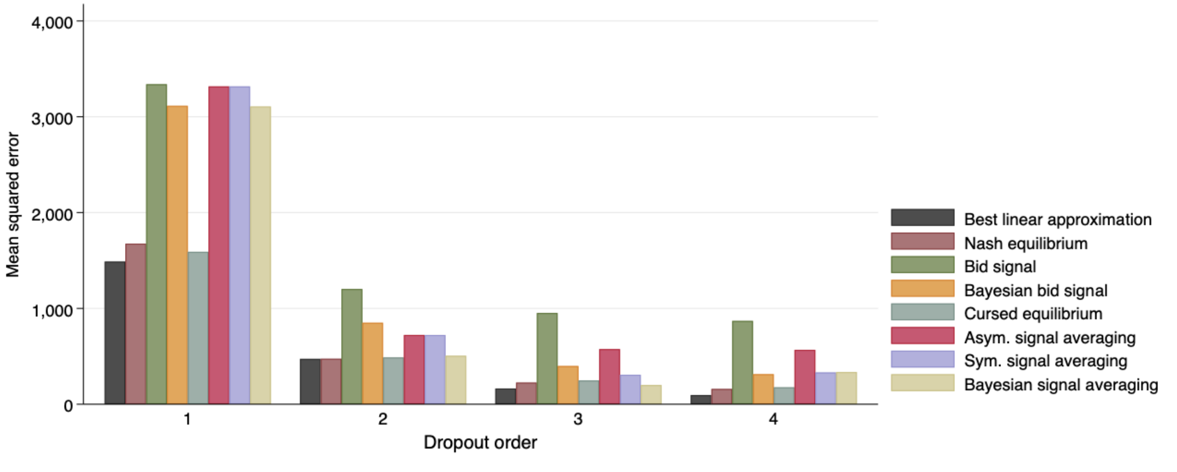
[48]Note that all models are in fact linear models. Derivations are available on request.

much larger for early dropout orders than for late dropout orders. This decrease in the error in dropout orders also holds for most other models considered. Second, especially for the later dropouts, Bayesian bid signal shows the lowest error, and comes very close to the benchmark prediction error.

Horse race for the JEA

We now continue this analysis for the JEA, using the identical classification procedure. We incorporate all models tested above.[49] Additionally, information revelation allows us to evaluate naïve models where bidders incorporate others' bids. For this, we test three signal averaging rules. In these rules, bidders are bidding the average of all signals available, both the private information signal as well as signals inferred from opponents' bidding behavior. The symmetric signal averaging rule, originally introduced by Levin et al. (1996), uses that bidders assume that also their opponents apply such a signal averaging rule. The Bayesian signal averaging rule is additionally applying information on the prior, similar to the difference between bid signal and Bayesian bid signal-rules for second-price auctions. The asymmetric signal averaging rule assumes that other bidders bid their signal, thus allows for straightforward computations. For the JEA, the best linear approximation incorporates all bids at earlier dropout orders, as these are observable when deciding on a bid.

Figure 2.A.3: Mean squared error of model predictions in the JEA



The main pattern observed in the AV carries over to the JEA: later bids can and in fact are predicted more precisely. Compared to the AV, the prediction error is

[49] For this auction format, we estimate $\hat{\chi} = -0.137$.

much lower in the JEA at late dropout orders, suggesting that bidding behavior is more predictable at this point (e.g., the best linear approximation for the fourth dropout shows a mean squared error of 189.5 in the AV and 96.1 in the JEA). At early dropout orders, there is however more noise in the JEA than in the AV. This might complicate matters for remaining bidders trying to estimate the value based on this revealed information in the JEA.

Interestingly, Nash equilibrium fits bidding behavior quite well, when comparing the mean squared error to the benchmark error of the regression.[50] Within the signal averaging rules, the Bayesian signal averaging rule performs best. Note that all signal averaging rules imply low intercepts in the linear bidding model, and we have presented evidence for substantial intercepts in the main text. This contributes to the high errors found for all signal-averaging rules.

2.A.5.1 Bid classification tables

Table 2.A.4 reports distances to predictions based on observed bidding.

Table 2.A.4: Classifying bids into models

	AV	JEA
	First dropout	
Nash	**2321.8**	*1675.4*
Bid signal	3718.3	3339.3
Bayesian bid signal	3763.3	3114.3
χ cursed	*2372.2*	**1590.5**
Sym. signal average		3316.7
Asym. signal average		3316.7
Bay. signal average		3106.6
Best linear approx.	1004.0	1489.3
	Second dropout	
Nash	*602.3*	**474.7**
Bid signal	1261.3	1202.6
Bayesian bid signal	978.8	850.3
χ cursed	**597.6**	*488.2*
Sym. signal average		722.3
Asym. signal average		722.9
Bay. signal average		505.6
Best linear approx.	465.9	473.5

	AV	JEA
	Third dropout	
Nash	356.3	225.8
Bid signal	807.5	952.6
Bayesian bid signal	**298.1**	398.6
χ cursed	333.3	248.1
Sym. signal average		307.8
Asym. signal average		575.2
Bay. signal average		**200.3**
Best linear approx.	212.2	163.9
	Fourth dropout	
Nash	431.5	**159.3**
Bid signal	717.7	869.7
Bayesian bid signal	**190.3**	313.5
χ cursed	404.0	*176.4*
Sym. signal average		331.6
Asym. signal average		567.4
Bay. signal average		335.6
Best linear approx.	189.5	96.1

Note: Average distance of observed bids to all considered models, by auction format and dropout order. Distances are squared distance from observed bid to bid predicted by each model. The best fitting model's distance is in bold, models within 10% of the best model's fit are italicized.

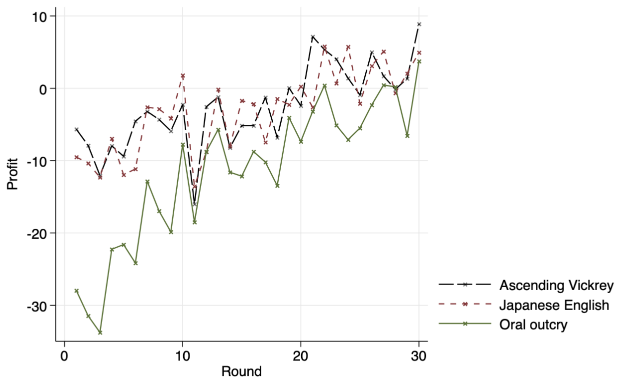
[50] Note that the simulated Nash equilibrium bids, as well as all other models incorporating observed dropouts, are based on inverting observed bids to retrieve the underlying signal. To do so, these rules make assumptions about how other bidders form their bids. This often leads to inferences about other bidders' signals which are incorrect, as other bidders did not, in fact, bid exactly as predicted by these models.

2.A.6 Learning effects

Information revelation in auctions potentially affects how bidders learn over time. In open auctions, this learning might also take place during the auction itself, and before information is revealed in sealed bid auctions, at the end of an auction.

In Figure 2.A.4, we plot the evolution of the winning bidders' profits over rounds, by auction format.

Figure 2.A.4: Evolution of profits over rounds by auction format



There is learning in the sense that profits increase over rounds. However, there are no meaningful differences in the evolution of profits between the JEA and the AV, learning in the OO is strongest in the sense of increases in profits over time. As we discuss in the main text, our results on revenue continue to hold in our auction data separately both in the first and last 15 rounds.

2.A.7 Estimations with experienced bidders

In the following, we present results of repeating estimations we report in the main text when only using the second half of our data, rounds 16 to 30.

In Table 2.3 in the main text, we study how available information correlates with bids. Table 2.A.5 repeats this anaylsis for rounds 16-30.

Across dropout orders, bidders appear to rely relatively less on public dropouts, and relatively more on their own private signal in late rounds. (8) shows that observed bids are more informative in late rounds than in the full data set. However, bidders still rely too strongly on the observed dropouts than what the empirical best response in (9) suggests.

In Table 2.4, we show that bids are more strongly correlated in the JEA than in the AV. Table 2.A.6 repeats this analysis for rounds 16-30. Results are in line with results in the full data set, apart from the coefficient on the interaction term of b_1 and JEA in the regression of b_2.

Table 2.A.5: Bidders' use of information in JEA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	b_1	b_2	b_3	b_4				V	$\hat{BR}$
	Observed	Observed	Observed	Observed	Nash	SA	BSA		
x	0.329	0.274	0.179	0.149	0.287	0.250	0.168	0.232	0.290
	(0.078)	(0.057)	(0.047)	(0.035)	(0.000)	(0.000)	(0.000)	(0.040)	(0.001)
b_1		0.341	0.038	0.029	0.100	0	0	-0.010	0.017
		(0.063)	(0.026)	(0.015)	(.)	(.)	(.)	(0.035)	(0.005)
b_2			0.537	-0.010	0.167	0	0	-0.011	0.047
			(0.089)	(0.019)	(.)	(.)	(.)	(0.057)	(0.008)
b_3				0.641	0.333	0.750	0.832	0.302	0.143
				(0.088)	(.)	(.)	(.)	(0.085)	(0.009)
t	-0.592	0.201	0.288	-0.063				0.379	-0.041
	(0.522)	(0.423)	(0.163)	(0.140)				(0.187)	(0.007)
Constant	35.745	34.694	23.235	26.580	11.265	0	0	41.700	50.207
	(17.143)	(10.104)	(5.317)	(4.495)	(.)	(.)	(.)	(6.651)	(0.597)
Observations	300	300	300	300				300	300
Adj. R^2	0.167	0.394	0.751	0.833				0.370	0.988
Rounds	16-30	16-30	16-30	16-30				16-30	16-30
Estimation	FE	FE	FE	FE				OLS	OLS

Notes: b_j: dropout price at order j; V: common value; x: own signal. (1) to (4) are fixed effects estimates (within estimation) of information use. Dependent variables (in columns) are dropout prices at each order, e.g. (1) are all bidders dropping out first in an auction. Regressors (in rows) are the available information at each dropout order, i.e., the signal x and the preceding dropout prices $\mathbf{b}_{j-1}$. (5) to (7) show how information is used in three canonical models, only for the fourth dropout. SA refers to the signal averaging-rule, BSA to the Bayesian signal averaging-rule. Note that these show how bids respond to earlier bids, where these bids are also calculated to follow the theoretical models. (8) shows how the price-setting bidder would have to use information to predict the common value after observing three dropouts. (9) shows how the bidder dropping out fourth would weigh information in an empirical best response. We provide adjusted R^2 of the original within-estimated model, as well as from estimating standard OLS where we include subject-specific absorbing indicators. The latter also includes fit obtained from subject fixed effects. Standard errors in parentheses, clustered at the matching group level.

Table 2.A.6: Comparing information use in the AV and the JEA, round 16-30

	b_2	b_3	b_4
b_{j-1}	0.316	0.268	0.325
	(0.050)	(0.026)	(0.039)
JEA × b_{j-1}	0.026	0.269	0.316
	(0.079)	(0.091)	(0.094)
Observations	599	599	599
Adjusted R^2	0.432	0.733	0.784

Notes: b_{j-1} denotes the just preceding dropout, e.g. is b_1 for b_2. JEA is a dummy equal one for JEA auctions. Other variables in regression omitted from table: all regressions include signal x, round t, all preceding dropouts (b_{j-k} for all $k \in \{1, \dots, j-1\}$), as well as all these variables interacted with the JEA-dummy and a constant. Standard errors in parentheses and clustered at the matching group level.

In Table 2.5, we study whether separately elicited characteristics of subjects correlate with the fixed effect we estimate from bidders' information use. Table 2.A.7 repeats this with experienced bidders.

Point estimates are mostly comparable to the analysis with the full data set in the main text. There are some estimates with larger standard errors, e.g. Imitator is no longer significant in (1). Point estimates for Imitator in the JEA in (2) turn negative, but remain not significant. The coefficient on SVO in the AV in (3) and the JEA in (4) turns positive and significant, comparable to the coefficient for early bids in (1). Note however that by restricting the dataset to the last 15 rounds, we will estimate the fixed effects much less precisely, as we on average only have 3 observations per individual to estimate those. In addition, for one bidder for b_1 & b_2, as well as for five bidders for b_3 & b_4, we cannot obtain a fixed effect any longer, as we don't have observations at these dropout orders for these bidders.

In Figure 2.5 in the main text we plot squared distances from the value to the prices in the auction and to some benchmarks, respectively. In Figure 2.A.5, we show this based on auction rounds 16 to 30. We observe some learning, as distances decrease compared to the analysis in the main text. This is strongest for the OO, where distances move closer to the no information benchmark, and bids in the empirical best response reveal more information than they do in the JEA.

Figure 2.A.5: Squared distance to common value, rounds 16 to 30

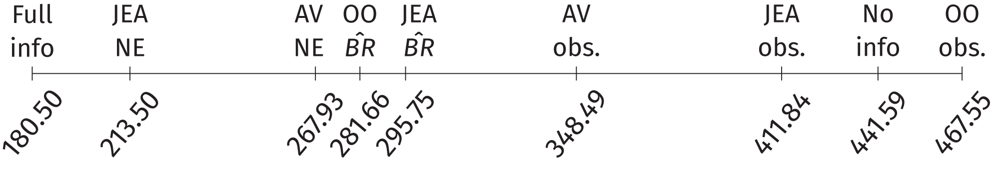


Table 2.A.7: Bidder fixed effects and their characteristics, rounds 16 to 30

	(1)	(2)	(3)	(4)
	Average bidder fixed effect			
	b_1 & b_2		b_3 & b_4	
	AV	JEA	AV	JEA
SVO	0.110	-0.222	0.100	0.117
	(0.043)	(0.213)	(0.034)	(0.030)
Imitator	4.469	-1.086	3.014	3.179
	(3.668)	(7.912)	(2.023)	(2.032)
Constant	0.894	7.993	-4.797	-3.264
	(1.847)	(3.317)	(1.295)	(1.949)
Observations	50	39	47	38
Adjusted R^2	-0.006	-0.013	0.035	0.141

Notes: Average fixed effects from regressing bids on available information for first and second vs. third and fourth dropout. SVO is a subject's social value orientation, in degrees. Imitator is a dummy variable equal one if a subject chose to retrieve social information when this contains no valuable information on the true state. Standard errors in parentheses, clustered at the matching group level.

In Table 2.6 in the main text, we investigate the effect of jump bids. In Table 2.A.8, we repeat this analysis for rounds 16 to 30, also constructing instruments only from experienced bids. Effects of jump bidding appear to be somewhat less pronounced in the second half of the data, but broadly similar.

Table 2.A.8: Effect of jump bids in the OO, rounds 16-30

	(1) Jump bid	(2) Pr(win)	(3) Profits	(4) Winners' profits
Total jump bid (IV)		0.341	-0.106	-0.232
		(0.097)	(0.054)	(0.067)
x	0.283	0.178	-0.118	-0.095
	(0.029)	(0.043)	(0.014)	(0.025)
t	-0.043		0.471	0.434
	(0.185)		(0.130)	(0.175)
V			0.587	0.630
			(0.059)	(0.071)
Constant	27.738	-14.532	-57.356	-49.753
	(7.274)	(2.465)	(4.318)	(8.201)
Observations	1309	1309	1309	300
Adjusted R^2	0.072	0.097	0.304	0.283
Estimation	OLS	2SLS	2SLS	2SLS
First-stage F-statistic		490.25	500.05	188.05
Hansen J-statistic, p-value		.800	.249	.620

Notes: Jump bid is the increment of a bid beyond the current price at the moment the bid was submitted. In 2SLS, we instrument the total jump bids and the maximum bid increment in other rounds. x is the submitting bidder's signal in round t. V represents the common value. Standard errors in parentheses, clustered at the matching group level.

2.A.8 Information usage in the AV and the JEA

In Section 2.6.2, we describe an empirical best response $\widehat{BR}$ in the JEA. It relies on estimated signals. Table 2.A.9 shows results of regressing signals on bids, which we in turn use to predict signals based on observable bids, where x_j refers to the signal of the bidder dropping out in j-th order in round t, and $\hat{x}_j$ refers to the predicted signal of the bidder dropping out in j-th order.

In Table 2.4, we show that bids are more strongly correlated in the JEA than in the AV, suggesting that information is actively used in the open format. Doing so controls for the correlation of unobservable dropouts in the AV, which arise as in these regressions bids are ordered. Table 2.A.10 shows the full regression results.

Table 2.A.9: Predicting signals with observed bids

	(1) x_1	(2) x_2	(3) x_3
d_j	0.205	0.567	1.009
	(0.064)	(0.089)	(0.087)
$\hat{x}_1$		0.048	0.081
		(0.191)	(0.226)
$\hat{x}_2$			-0.147
			(0.140)
t	0.270	0.177	0.166
	(0.202)	(0.175)	(0.140)
Constant	68.089	33.977	5.797
	(4.111)	(11.964)	(13.079)
Observations	600	600	600
Adjusted R^2	0.031	0.213	0.342
Rounds	1-30	1-30	1-30
Estimation	OLS	OLS	OLS
Session FE	Yes	Yes	Yes

Notes: Standard errors in parentheses and clustered at the matching group level.

Table 2.A.10: Comparing information use in the AV and the JEA

	b_1	b_2	b_3	b_4
x	0.247	0.297	0.242	0.227
	(0.0457)	(0.0216)	(0.0224)	(0.0298)
b_1		0.285	-0.00113	-0.0141
		(0.0309)	(0.0172)	(0.0209)
b_2			0.357	-0.0114
			(0.0319)	(0.0317)
b_3				0.465
				(0.0440)
t	-0.498	-0.0381	-0.126	-0.174
	(0.155)	(0.0872)	(0.0596)	(0.0341)
JEA × x	0.0464	-0.0296	-0.0704	-0.109
	(0.0718)	(0.0398)	(0.0342)	(0.0336)
JEA × b_1		0.0871	0.0244	0.0392
		(0.0463)	(0.0243)	(0.0253)
JEA × b_2			0.195	-0.0271
			(0.0533)	(0.0479)
JEA × b_3				0.244
				(0.0827)
JEA × t	0.181	-0.0844	0.0433	0.0991
	(0.315)	(0.141)	(0.0937)	(0.0455)
Constant	32.09	38.94	37.14	33.30
	(4.573)	(1.653)	(1.969)	(2.440)
Observations	1199	1199	1199	1199
Adjusted R^2	0.135	0.502	0.732	0.777
Estimation	FE	FE	FE	FE

Notes: Standard errors in parentheses and clustered at the matching group level.

2.A.9 Informational impact of dropouts

In this Section, we investigate the informational impact of earlier bids on subsequent bids. To do so, we first regress bids, by dropout order, on public information, and then predict residuals. As this estimation by design excludes all private information, for example a bidder's signal or bidders' idiosyncratic characteristics, this variation will be captured in the residual. Below, we reproduce the estimation used to predict residuals, we do use matching group fixed effects in this estimation.

Table 2.A.11: Residual estimations

	b_1	b_2	b_3
b_1		0.477	0.0221
		(0.0385)	(0.00955)
b_2			0.665
			(0.0435)
t	-0.555	-0.148	-0.0479
	(0.318)	(0.108)	(0.0949)
Constant	77.19	60.20	36.93
	(4.930)	(2.747)	(4.435)
Observations	600	600	600
Adjusted R^2	0.113	0.419	0.698
Fixed effects	matching group	matching group	matching group
Estimation	OLS	OLS	OLS

Notes: Standard errors in parentheses and clustered at the matching group level.

We then regress dropouts at later dropout orders on these residuals, results are reported in Table 2.A.12. Doing so, we can estimate the impact of information revealed in earlier bids on later bids, where we isolate the information contribution of each observed bid. For comparison, we repeat this exercise for Nash equilibrium and the Bayesian signal-averaging rule.[51]

In (1) to (3), we observe that the effect of a bidder's private information, captured by x is less than the public information, revealed through the dropouts. As in the analysis in the main text, we see that the just preceding dropout carries most weight in explaining bidding behavior. This does not lend support to bidders suffering from a strong correlation neglect, as we would expect higher coefficients on the impact of earlier residuals in that case (e.g., on e_1). Similarly, bidders' private information, x, is weighted less than in the benchmarks.

[51] In predicting corresponding residuals, we do not use matching group fixed effects nor do we control for round. For this estimation, note that the residuals are obtained from regressing simulated bids on simulated bids.

Table 2.A.12: Information effects captured by residuals

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Observed			Nash equilibrium			Bayesian signal-averaging	
	b_2	b_3	b_4	b_2	b_3	b_4	b_2	b_3	b_4
x	0.272	0.168	0.120	0.575	0.431	0.287	0.253	0.202	0.168
	(0.0220)	(0.0211)	(0.0180)	(0.00000261)	(0.00000411)	(0.00000549)	(0.00000220)	(0.00000475)	(0.00000568)
e_1	0.401	0.286	0.250	0.200	0.354	0.491	0.747	1.057	1.220
	(0.0392)	(0.0186)	(0.0134)	(0.00000342)	(0.00000524)	(0.00000515)	(0.00000366)	(0.0000154)	(0.0000112)
e_2		0.566	0.429		0.250	0.458		0.798	1.218
		(0.0455)	(0.0328)		(0.0000102)	(0.00000744)		(0.0000198)	(0.0000204)
e_3			0.719			0.333			0.832
			(0.0710)			(0.0000234)			(0.0000370)
Constant	60.04	78.92	90.86	37.08	49.18	63.64	65.16	69.01	74.32
	(2.940)	(2.711)	(2.151)	(0.000190)	(0.000386)	(0.000619)	(0.000157)	(0.000484)	(0.000610)
Observations	600	600	600	600	600	600	600	600	600
Adjusted R^2	0.439	0.713	0.772	1.000	1.000	1.000	1.000	1.000	1.000
Estimation	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS

Note: Standard errors in parentheses and clustered on matching group level.

2.A.10 Explaining heterogeneity in bidding

In Section 2.6.3, we study correlations of bidders' fixed effect with their separately elicited characteristics. A potential concern of this analysis is that these fixed effects are themselves estimated, and some might be more noisily estimated than others. To account for this, we study whether the results presented in the main text are robust to using Weighted Least Squares instead of OLS, where the weights are given by the inverse of the average variance of the estimate of each bidders' (averaged) fixed effect. This procedure ensures that particularly noisy fixed effects receive less weight in the regression. We present results in Table 2.A.13.

Table 2.A.13: Bidder fixed effects and their characteristics

	(1)	(2)	(3)	(4)
	Average bidder fixed effect			
	b_1 & b_2		b_3 & b_4	
	AV	JEA	AV	JEA
SVO	0.118	-0.231	0.037	0.045
	(0.024)	(0.037)	(0.024)	(0.037)
Imitator	5.364	4.840	5.606	1.080
	(1.030)	(1.695)	(1.030)	(1.695)
Constant	0.492	7.681	-3.719	-0.615
	(0.716)	(1.092)	(0.716)	(1.092)
Observations	50	40	50	40
Estimation	WLS	WLS	WLS	WLS

Notes: Average fixed effects from regressing bids on available information for first and second vs. third and fourth dropout; pooling data from the AV and the JEA. SVO is a subject's social value orientation, in degrees. Imitator is a dummy variable equal one if a subject chose to retrieve social information when this contains no valuable information on the true state. We use weighted least squares, with the weight given by the inverse average variance of the estimate of the bidder fixed effect, averaged at d_1 and d_2, and at d_3 and d_4. Standard errors in parentheses.

We observe that the point estimates presented in the main text carry over. In addition, some coefficients which are not significant in the main text are highly significant in this specification, e.g. the coefficient on SVO in JEA in (2).

2.A.11 Further results on jump bidding

In the main text, we report 2SLS estimations, where we instrument for the total jump bid with the maximum bid increment and average total jump bid of each bidder obtained in all other rounds.

In Table 2.A.14, we report results of the first stage. In addition, we report the Kleibergen-Paap F-statistic, which suggests that the instruments are relevant, and p-values for the Hansen J-statistic, which do not reject that the chosen instruments are valid. Columns (2), (3) and (4) show first-stage results for each of the corresponding second-stages in Table 2.6 in the main text.

Table 2.A.14: First stage for 2SLS estimation

	(2)	(3)	(4)
	Dependent variable: total jump bid		
Maximum increment in other rounds	0.163	0.163	0.151
	(0.032)	(0.032)	(0.060)
Mean total jump bid in other rounds	0.582	0.584	0.699
	(0.069)	(0.071)	(0.079)
x	0.279	0.302	0.175
	(0.028)	(0.029)	(0.046)
V		-0.076	0.106
		(0.018)	(0.046)
t		-0.163	-0.528
		(0.122)	(0.201)
Constant	-26.761	-19.123	-9.381
	(5.742)	(7.061)	(8.876)
Observations	2687	2687	600
F-statistic	96.4	98.7	143.3
Hansen J-test	.584	.582	.948

Notes: The dependent variable across all first-stage regressions is the total jump bid, given by the sum of bid increments beyond the current price within a round. As instruments, we use the maximum bid increment and the mean total jump bid for each bidder in all but the current round. x is the submitting bidder's signal in round t. V represents the common value. Standard errors in parentheses, clustered at the matching group level.

For robustness, we repeat the analysis in Table 2.6 using only the mean total jump bid in other rounds, and show results in Table 2.A.15.

Table 2.A.15: Effect of jump bids in the OO, one instrument

	(1)	(2)	(3)	(4)
	Jump bid	Pr(win)	Profits	Winners' profits
Total jump bid (IV)		0.343	-0.274	-0.315
		(0.087)	(0.134)	(0.151)
x	0.274	0.146	-0.064	-0.031
	(0.032)	(0.039)	(0.043)	(0.044)
t	-0.137		0.875	0.779
	(0.125)		(0.166)	(0.151)
V			0.624	0.634
			(0.046)	(0.066)
Constant	30.525	-12.128	-66.217	-58.989
	(5.937)	(2.756)	(6.987)	(10.231)
Observations	2687	2687	2687	600
Adjusted R^2	0.069	0.103	0.284	0.286
Estimation	OLS	2SLS	2SLS	2SLS
F-statistic		86.9	85.7	325.2

Notes: Jump bid is the increment of a bid beyond the current price at the moment the bid was submitted. In (1), we regress total jump bid on bidders' signals and round t. In (2) to (4), we use 2SLS, where we instrument using the average total jump bid in other rounds. (2) is the ex-post probability of winning, which is a dummy equal to one if a bidder wins the auction, 0 otherwise. Mean earnings are a participants' average earning across all auctions, winners' profits are the earnings for the auctions which a participant won. x is the submitting bidder's signal in round t. V represents the common value. Standard errors in parentheses, clustered at the matching group level.

2.A.12 Questionnaire results

In the questionnaire, we offered several reasons why bidders behaved as they did, asking participants how much they agree to a statement on a 7-point Likert scale.

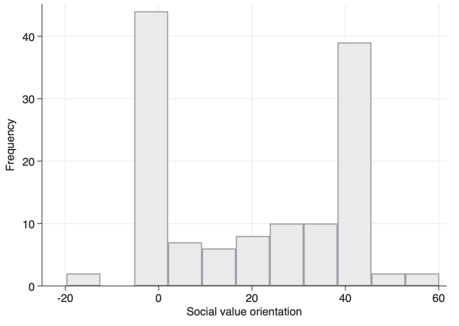
Below is the mean and standard deviation of how much people agree with a given statement. The scale is from 1 to 7, where 7 is fully agreeing, 4 is undecided.

Treatment	Statement	Mean	SD
AV	In auctions where I did not expect to win, I stayed in the auction longer to increase the price paid by the winner.	2.72	1.96
JEA	In auctions where I did not expect to win, I stayed in the auction longer to increase the price paid by the winner.	3.64	2.03
AV	In auction where I did not expect to win, I quit the auction sooner to decrease the price paid by the winner	2.98	2.00
JEA	In auction where I did not expect to win, I quit the auction sooner to decrease the price paid by the winner	3.08	1.83
JEA	When I observed other bidders leaving, I formed a more precise guess of the value of the item.	4.82	1.76
JEA	When I observed other bidders leaving, I also immediately left the auction, as I relied on the other bidders' guess of the value.	3.79	1.85
OO	All else being equal, I was more likely to enter a new bid if I have been the standing bidder for longer.	3.21	1.76
OO	All else being equal, I was willing to pay more for the item if I have been the standing bidder for longer.	3.26	1.82
OO	I tried to deter other bidders from bidding by entering a bid much higher than the current price.	4	2.12
OO	I tried to prevent other bidders from entering their desired bid by entering a bid much higher than the current price.	4.36	1.94
OO	I entered bids much higher than the current price because I thought this would allow me to pay a lower price for the item.	3.21	2.04
OO	I entered bids much higher than the current price because I was feeling impatient and wanted the auction to finish sooner.	3.15	1.95
OO	I entered bids much higher than the current price because I was becoming annoyed by being overbid by other participants.	3.26	2.11
OO	I entered bids much higher than the current price because it felt costly to decide on and enter new bids.	2.92	1.75

2.A.13 Circle test

We also elicited subjects' social value orientation. It is given as an angle. 0° is purely selfish (6 self, 0 other), whereas 45° is splitting equally between self and other (minimising inequality and maximising efficiency). Figure 2.A.6 gives a histogram of observed choices.

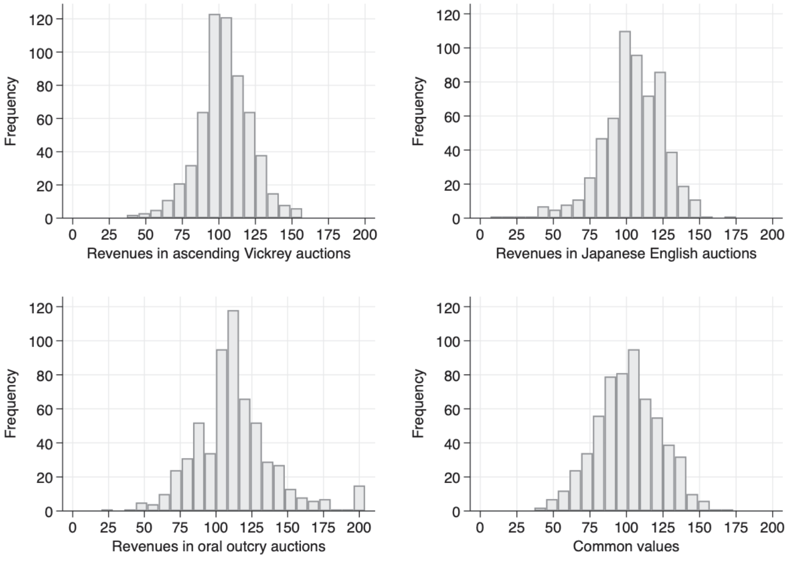
Figure 2.A.6: Angle in circle test



2.A.14 Histograms of auction revenues

In Figure 2.A.7, we plot histograms of the revenues in all three auction formats as well as a histogram of the common values drawn.

Figure 2.A.7: Histograms of the drawn common values and revenues.



2.B Instructions and screenshots of the experimental interface

In the following, we reproduce the instructions for participants as well as examples of the auction screens.

2.B.1 Experimental instructions

Page 1 Welcome!

Welcome to this experiment. Please read the following instructions carefully. You will also receive a handout with a summary. There is a pen and paper on your table, you can use these during the experiment. We ask that you do not communicate with other people during the experiment. Please refrain from verbally reacting to events that occur during the experiment. The use of mobile phones is not allowed. If you have any questions, or need assistance of any kind, at any time, please notify the experimenter with the CALL button on the wall to your left, the experimenter will then assist you privately.

Your earnings will depend on your decisions and may depend on other participants' decisions. Your earnings will be paid to you privately in cash at the end of today's session. All your earnings will be denoted in points. At the end of the experiment, each point that you earned will be exchanged for 25 eurocents.

Page 2 Decision and Payoffs

This experiment consists of 30 periods. In each period, you will be allocated randomly to a new group of five participants. Therefore, in each period you will be in a group with (most likely) different participants. You will never learn with whom you are in a group. At the end of the experiment, five periods will be randomly selected for payment. Your earnings will be the sum of the earnings in these five periods.

Description of the situation and possible earnings

In each period, an auction will take place. In each auction, a product of unknown value will be sold. In each period, you will be given a capital of 20 points. Any profits or losses you make in this period will be added to or subtracted from this capital.

Procedures

In each auction, each of the five participants (including you) can obtain the product. First, every participant indicates that he or she is ready, and, as soon as all participants indicate so, there will be a countdown of three seconds, after which the auction starts.

{JEA/AV: In the auction itself, the price will rise in increments of one point, starting at a price of 0. This will be indicated with a thermometer, where the level of the thermometer indicates the current price.

At any point while the price rises, you can decide to leave the auction. You do so by clicking on the “EXIT” button, indicating that you are not willing to buy the item and leave the auction for this period. For all remaining participants, the auction continues.

The auction stops after four of the five participants have pushed the “EXIT” button. The winner of the auction is the last participant remaining in the auction. The price the winner has to pay to buy the product is determined by the level of the thermometer when the fourth bidder has pushed the “EXIT” button. The price level at this point is called SELLING PRICE. The winner obtains the product and pays the SELLING PRICE. The earnings for the winner in the period are given by the value of the product minus the SELLING PRICE. These earnings are added to the capital of 20 points in this period. More details about how the value of the product is determined will follow. All participants who exited the auction will not obtain the product and will earn an amount equal to the capital of 20 points in this period.

{AV ONLY: During the auction, you will not observe how many participants remain in the auction. The price continues to rise as long as there are at least two participants in the auction including yourself.}{JEA ONLY: During the auction, you will be notified as soon as any other participant exits the auction. You will be shown at which price this other bidder left the auction, and there will be a pause of 4 seconds, in which the price will not be increasing. Afterwards, as long as there are at least two participants remaining in the auction, the price rises again.}

In the unlikely case in which multiple participants quit at the same moment and there is no bidder remaining in the auction afterwards, the program will randomly choose the person buying the item from all participants who were the last to exit and did so at the same time. The SELLING PRICE is then the level of the thermometer where these participants simultaneously pressed the button.

At the end of each period, the SELLING PRICE paid by the buyer will be shown to all participants within a group. The buyer will not literally receive a product. In addition to the capital for the period, he or she will receive an amount equal to the value of the product minus the selling price of the product (in points). The previously unknown value of the good will then be revealed to all bidders, as well as their earnings in points in this period. Afterwards, you will be matched with a new group of bidders and a new auction starts, with the same procedure.

Example: Suppose that the first 4 bidders who exit the auction do so at prices 40, 50, 70, 80. Further assume that the product’s value is 90 points. Then the last remaining bidder in the auction will receive the product and pay 80 points. His or her earnings from the auction will be 90-80=10, and the total earnings for the period will be 10 + 20, where 20 is the capital of the period. All other bidders will each earn the capital of 20 in that period.}

{OO: In the auction itself, participants will have the opportunity to enter maximum bids. A maximum bid tells the computer how much you maximally want to pay for the good. The computer will try to obtain the good as cheap as possible on your behalf, and at a price that is no higher than your maximum bid. If your maximum bid is the highest at some moment, then you are the current standing bidder. The standing bidder at the end of the auction obtains the product. This auction proceeds in bidding rounds in the following manner:

As soon as the auction starts, a 15 seconds countdown is initiated. Within these 15 seconds, each bidder can submit a maximum bid that is zero or higher. Whenever a maximum bid is submitted, the auction will be momentarily paused. The bidder who submitted the highest maximum bid so far will be recognized as the standing bidder. At the same time, the second highest maximum bid submitted up to this point will be the CURRENT PRICE for the good. The CURRENT PRICE will be displayed to all participants and a new bidding round immediately starts. Again, a countdown of 15 seconds is initiated, and bidders can submit new maximum bids. Any new maximum bid has to be higher than the CURRENT PRICE. The current standing bidder is notified that he is the standing bidder. He/she will only be able to submit a new maximum bid when he/she is no longer the standing bidder.

This procedure will then be repeated. As soon as new maximum bids above the CURRENT PRICE are submitted, there will be a brief pause, and afterwards a new CURRENT PRICE and standing bidder will be declared. During the bidding procedure, you will be able to see the last submitted maximum bid of each bidder (if a bidder submitted at least one maximum bid). Only the maximum bid of the current standing bidder is not revealed. Note that the bidder numbers do not enable you to identify bidders, as groups change over periods and these numbers are randomly reallocated.

Bidding will continue until no bidder in your group is willing to submit a maximum bid higher than the CURRENT PRICE, and the countdown elapses.

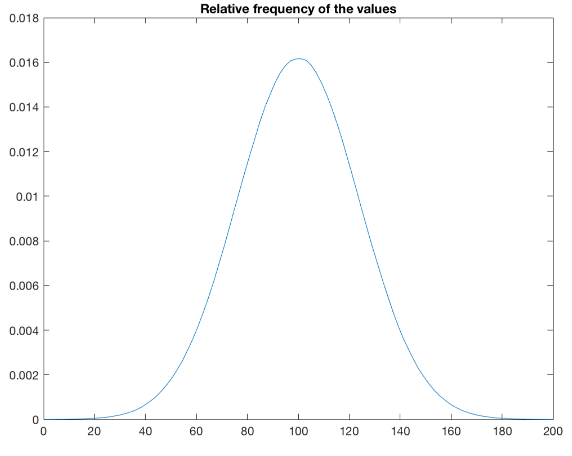
At the end of each period, so when a countdown elapses before any new bid is submitted, the earnings of the bidders are calculated as follows: The winner of the auction is the bidder who submitted the highest maximum bid, and he or she will pay a price equal to the CURRENT PRICE when bidding stopped. The buyer will not literally receive a product. He or she will receive an amount equal to the value of the product minus the CURRENT PRICE of the product (in points). This amount is added to the capital of 20 points in this period. All other bidders earn an amount equal to the capital of 20 points. The previously unknown value of the good will then be revealed to all bidders, as well as their earnings in points in this period. Afterwards, you will be matched with a new group of bidders and a new auction starts, with the same procedure. Notice that the winner of an auction can make a gain or a loss. A loss occurs if the price paid is higher than the value. Even though the final standing bidder pays a price equal to the second highest maximum bid, such bid may be high and result in a high price.}

In total, there will be 30 periods, and five randomly determined periods will be chosen to be paid out. Your earnings for the experiment will be equal to the sum of your earnings in these 5 periods.

Page 3 Value of the product and signals

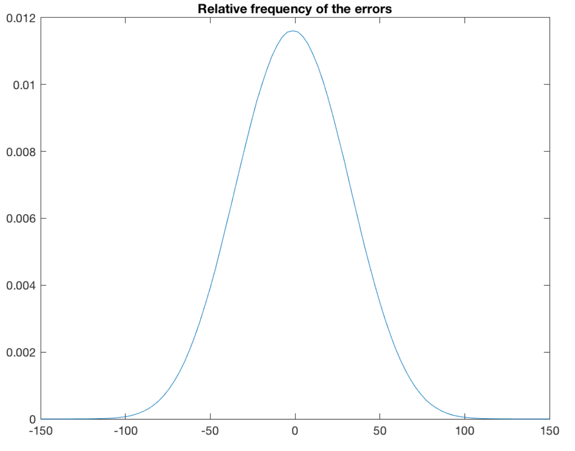
The value of the product is a random number which changes in each period. You cannot learn anything about subsequent value draws from previously observed values. Within the period the value is identical for all participants in the group. At the time of bidding, this value is unknown to all participants. Instead, each participant receives a signal which provides an imprecise indication of what the value may be. In the following, we will describe how the values and signals are determined in each period.

In each period, the value of the product will be randomly determined. The value can be any round number between 0 and 200. The figure below clarifies how frequently different values occur. You can see that values close to 100 occur most often (the frequency is highest when the value on the horizontals axis equals 100). Values below 100 occur as frequently as values above 100. Also, values below 50 occur as often as values above 150. You do not need to be familiar with such a distribution to participate in this experiment, and you will see some typical value draws on the next page.



The signals

Each participant will receive a (different) signal of the value. This signal gives a first indication of the value of the product in that period, although this is only imprecise information. In particular, the signal is the sum of the value and an error. The figure below shows how frequently different errors occur. You can see that errors close to 0 occur most often, and that errors below 0 occur as frequently as errors above 0.



The error (most likely) differs for every participant. Therefore, each participant in your group will (most likely) obtain a different signal of the value, even though the value of the product is the same for everyone. Signals higher than the value occur as frequently as signals lower than the value. Signals closer to the value are more likely than signals further away from the value. In this experiment, you will encounter only values and signals between 0 and 200.

Notice that each signal in a group is informative about the value of the product. If other bidders let their bidding depend on their signal, then their bidding will be informative about the value of the product.

Note that the signals will be newly determined in each period, therefore only the signals of this period are helpful for you to determine the value of the product for sale.

Payment

As mentioned before, out of the 30 periods, 5 will be randomly selected. You will receive the sum of the points that you earned in each of the 5 selected periods. In each period, every bidder receives a capital of 20 points. Then, any gains or losses a participant made in this period's auction are added to or subtracted from the capital. Notice that the buyer in a period can make a gain but also a loss. If the buyer pays a price higher than the value of the product, he or she makes a loss. Just like a profit is automatically added to the capital, a loss will automatically be subtracted.

Page 4 We will now illustrate in one particular example how the auction process works. We emphasize that this is only an example, and that these numbers are not relevant for the real auctions in which you will participate afterwards.

Example

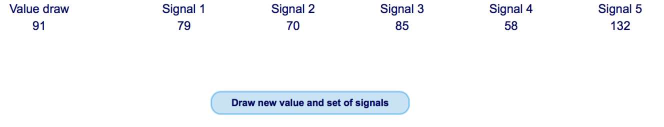
First, a value of the product is randomly determined, but not revealed to the participants. In our example, this value will be 121. Then, based on the value of 121, the signal for each participant will be drawn. The following signals are drawn: one bidder receives a signal of 60, one bidder a signal of 87, one bidder a signal of 126, one bidder a signal of 144 and the last bidder a signal of 175. Now the auction starts. {JEA/AV: The thermometer starts at 0, and rises continuously as soon as every participant indicated that he or she is ready and the countdown is initiated.

As the thermometer rises, bidders may decide to press the "EXIT" button and leave the auction. Imagine that the first participant exits at a price of 52, the second participant at a price of 77 and the third participant quits at a price of 109. Now, there still remain two bidders in the auction. {IN JEA: Each time a participant quits, all remaining participants will be notified about this, and will receive information about the price at which this participant chose to exit.} The thermometer will keep rising up to the point where the fourth bidder presses the "EXIT" button, for example at a price of 115. Then, the last remaining bidder buys the product at the selling price of 115. In this example, the product's value was 121 points. Therefore, the winner will earn 121-115=6 points in addition to his or her capital in this period, hence 6 + 20 = 26 points in total, if the period is selected for payment.} {OO: The countdown starts at 15 seconds, and is initiated as soon as every participant indicated that he or she is ready. Then, imagine that the first participant to enter a bid submits a maximum bid of 52. This bidder becomes the new standing bidder. As so far only one maximum bid has been submitted, the CURRENT PRICE will be 0, and this is shown to all bidders as soon as the next bidding round commences. The countdown is reset and starts immediately. Then, imagine a new maximum bid of 77 is submitted. As this is the current highest maximum bid, this bidder becomes the new standing bidder. The second highest maximum bid at this point is 52, and therefore 52 is the new CURRENT PRICE. Bidding continues in this fashion until the countdown elapses. For example, imagine that in the next rounds maximum bids of 109, 115 and 120 are, and in the next bidding round the countdown elapses. Then, the bidder who submitted the highest maximum bid (i.e. the bidder who bid 120) will win the auction. This bidder will pay the last CURRENT PRICE, which equals the second highest maximum bid (115 in this example). In this example, the product's value was 121 points. Therefore, the winner will earn 121-115=6 points in addition of his or her capital in this period, if the period will be selected for payment.}

Page 6 **Practice draws**

Now, you have the opportunity to see how typically values and corresponding signals are drawn. You can click on a button to draw new values and signals. Then, you will be shown a value and set of signals drawn according to the same procedure as those in the experiment. In the experiment, you will not be able to observe the value draw, but instead you receive one of the imprecise signals of the value. Five signals corresponding to this value are shown to you next to the value draw. When you click on the button again, a new value and corresponding set of signals will be drawn, you can repeat this as often as you like. Note that these example values and signals are not informative about the draws you will actually face in the experiment.

Figure 2.B.1: Screenshot of the practice draws from the instructions.



When you have tried a number of times, please continue to the practice questions on the next page.

2.B.2 Additional elicitations

For the last 14 sessions, we added two additional measures, elicited after the auctions concluded. Below are instructions for both tasks.

2.B.2.1 Imitation, adapted from Goeree and Yariv (2015)

Part 2

In this part of the experiment, you make an individual decision. The amount you earn depends only on your choices and your choices do not affect the earnings of other participants.

Guessing the urn

In this task, you have to guess which one of two possible urns has been selected. It is equally likely that you face a red or a blue urn. These urns contain red and blue balls as follows:

- Red urn: 7 red balls and 3 blue balls
- Blue urn: 7 blue balls and 3 red balls

Information

For your decision, you have to choose to receive one of two types of information:

- Draw: The color of one randomly selected ball drawn from your urn will be shown to you.
- History: The choices of three participants from previous sessions of this experiment will be revealed to you. These three participants faced the same urn as you do, but did not receive any of the two types of information you can choose between (neither Draw nor History).

Task

After you receive this information, you have to guess which of the two urns has been selected.

Payoff

You will earn 4 points if you guess correctly which urn has been chosen.

When you continue, it will be randomly determined whether you face the red or the blue urn. In the next screen, you first choose the type of information you would like to receive, then you have to enter your guess which urn you are facing.

2.B.2.2 Circle test, adapted from Linde and Sonnemans (2012)

Part 3

For this part of the experiment, you have been matched with one other randomly selected participant, called OTHER. Your subsequent decision will be anonymous, no participant will know with whom they have been matched. In the end, either your or OTHER's decision will be implemented.

Choice

In this part you have to choose between combinations of earnings for yourself and the OTHER. All possible combinations are represented on a circle. You can click on any point on the circle. Which point you choose determines how much money you and the OTHER earn. You can enter this choice on the next page.

Earnings

The axes in the circle represent how much money you and the OTHER earn when you choose a certain point on the circle. The horizontal axis shows how much you earn: the more to the right, the more you will earn. The vertical axis shows how much the OTHER will earn: the more to the top, the more the OTHER earns. The distribution can also imply negative earnings for you and/or the OTHER. Points on the circle left of the middle imply negative earnings for you, points below the middle imply negative earnings for the OTHER. When you click on a point on the circle the corresponding combination of earnings, in cents, will be displayed in the table to the right of the circle. You can try different points by clicking on the circle using your mouse. Your choice will only become final when you click on the "send" button.

Payoff

The OTHER is presented with the same choice situation. At the end of the experiment, either your decision or the decision of the OTHER will be paid. This will be determined by a random draw, your decision is as likely to be chosen as the decision of the OTHER. This draw is not affected by the choices you or others make.

2.B.3 Screenshots of the interface

In the following, some screenshots of the screens of auction participants for all three treatments:

Figure 2.B.2: Screenshot from a ascending Vickrey auction.

Figure 2.B.3: Screenshot from a Japanese English auction.

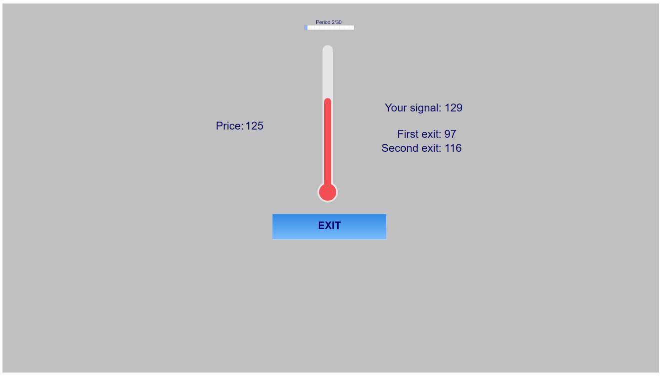
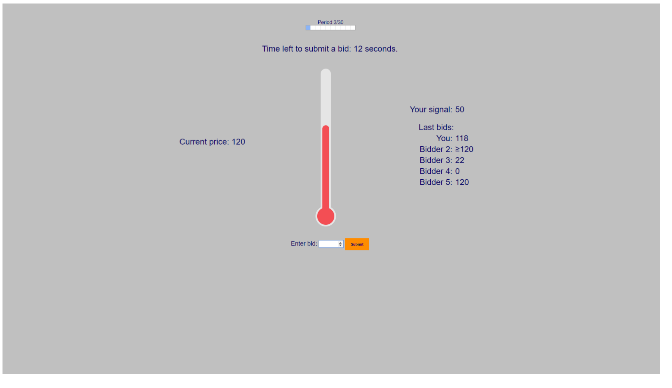


Figure 2.B.4: Screenshot from an oral outcry auction.



CHAPTER 3

Morals in multi-unit markets

This chapter is based on Ziegler, Romagnoli, and Offerman (2022).

3.1 Introduction

People's morals may easily take a back seat in markets. Consider the market for air travel. Passengers may think that if they refrain from buying a cheap ticket, someone else could take their place, leaving total emissions unchanged. This reasoning, the so-called replacement logic, may explain why frequent flying also occurs among environmentally conscious consumers (Barr, Shaw, and Coles, 2011). Airlines themselves may justify their offering of flights by arguing that a competitor will offer an additional flight if they decide to withdraw a connection. At the same time, the choices of a minority of consumers can have a disproportionate impact on aggregate outcomes. Gössling and Humpe (2020) find that in the US, 12% of adults account for 68% of all trips. As a result, aggregate behavior may not reflect the average person's concern for environmental damages. Anecdotal evidence suggests that resorting to the replacement logic to excuse morally questionable behavior as well as the disproportional activity of few irresponsible actors are features common to several morally questionable or highly polluting markets, such as the opioids market, the shipping industry and weapons trade.[1]

Recent laboratory experiments have investigated the extent to which morals are eroded in single-unit markets, which are markets where each participant is restricted to trade at most one unit. In a seminal paper, Falk and Szech (2013) find that while 45.9% of subjects are willing to kill a mouse for € 10 in individual decision-making, 75.9% do so in single-unit markets. In the multi-lateral bargaining setting, they also find a decline in prices as a result of competition, which the authors interpret as further evidence for moral erosion. However, key results of Falk and Szech (2013) are contested. Market prices can decline also without moral erosion (Sutter et al., 2020). Moreover, while Falk and Szech (2013) compare a single decision in individual decision-making with repeated decisions in a market, Bartling et al. (2023) show that the partial erosion in markets disappears under repetition of both environments. So far, the evidence that people's morals are eroded in markets is inconclusive.[2]

[1] In the opioids market, a spokes-woman for McKesson, which was the largest distributor in the US from 2006 to 2012, stated: "Any suggestion that McKesson influenced the volume of opioids prescribed or consumed in this country would reflect a misunderstanding of our role as a distributor" (`https://apnews.com/98963bb70e0f462295ccc02fe9c68e71`). In contrast, also in this market single firms can be responsible for a significant share of overall harm: Purdue Pharma's marketing campaign for OxyContin increased sales and the associated overdose deaths (Alpert, Evans, Lieber, and Powell, 2022). In 2017, the number of Americans dying from an overdose of opioids (47,600) surpassed the number dying from car accidents (Scholl, Seth, Kariisa, Wilson, and Baldwin, 2019). Empirically, Vuillemey (2020) documents an erosion of standards in the shipping industry, where jurisdictions compete to register additional ships by relaxing regulatory requirements. In the market for weapon trading, both UK prime minister Tony Blair (in 2002) and British Secretary of State Boris Johnson (in 2016) made the argument that they could stop the defense industry operating in their country, but that then someone else would step in to supply the arms that they supplied (Bartling and Özdemir, 2023; Falk, Neuber, and Szech, 2020).

[2] In Section 3.2, we position our paper more precisely in the literature.

In our view, many real world markets are poorly approximated by the single-unit markets studied so far. In addition, these markets inhibit forces which may contribute to a strong erosion of morals, which may have lead to an underestimation of the effect of markets on morals so far. We focus on more realistic multi-unit markets[3] and distinguish between two forces that may drive erosion: (i) *market selection* and (ii) *replacement logic*.

For market selection, we assume market participants trade whenever the material profits exceed their moral costs associated with causing negative externalities. Multi-unit markets remove individuals' constraint to trade at most once. This allows the less-moral participants to capture a larger share of the market, as they can also trade units associated with low profits. Trade stops when even the participants least concerned about the externalities are no longer willing to trade. Market selection then implies that outcomes in multi-unit markets are predominantly determined by the least moral traders, as the abstention of the more-moral traders no longer restricts the exchange of additional units. This effect is further enhanced when preferences are characterized by diminishing marginal moral costs for the negative externality, as trading repeatedly generates an additional competitive advantage for the least moral traders.

Further, in single-unit markets, traders possess substantial market power. For each pair of active traders, at least one of them is pivotal: The total quantity traded would be reduced if this trader refrains from trading. This reduces the scope for the replacement logic. According to this principle, participants may decide to trade as they realize that their individual actions do not affect aggregate outcomes. They then feel justified in trading and reaping the profits for themselves (Sobel, 2007). In the multi-unit markets we study, no trader is pivotal. Thus, traders on both sides of the markets can excuse their trading with the argument that if they had not traded, someone else would have taken advantage of the opportunity.

The existing literature inferred people's moral deterioration by comparing their choices in individual tasks to their trading behavior in markets. As a consequence, this literature could not distinguish between norm erosion and an erosion of norm compliance. For policy applications, it is important to understand the reason behind a possible shift to more selfish behavior in markets. If people's norms are not affected while norm compliance deteriorates – i.e., if even the traders themselves regard their behavior in the market as inappropriate – this could lead to a stronger case for government interventions that reduce the extent of moral erosion in markets.

In this paper, we employ a laboratory experiment to investigate how the erosion of morals depends on the ability of traders to affect aggregate market outcomes. We measure morality as participants' valuations of donations for measles vaccines

[3] Also within experimental economics, markets were extensively studied in multi-unit rather than single-unit settings (e.g. Smith, 1962; Ketcham, Smith, and Williams, 1984; Plott, 1983).

to UNICEF. Consistent with our participants' perception, we call the decision to cancel a donation to UNICEF, in exchange for money to one-self, immoral.[4] We then measure how participants' evaluations for the same donation change in markets, where the choice to trade increases money to self while producing a negative externality in the form of a cancelled donation. We explore how these evaluations change across a set of multi-unit markets which vary traders' pivotality for aggregate outcomes. Our main contribution is threefold. First, we investigate whether market outcomes reflect participants' concerns towards causing negative externalities and the extent to which this is affected by individuals' market power. Second, we identify whether the moral erosion is due to a shift of norms or a deterioration of norm compliance. Third, we disentangle how much moral erosion is due to either market selection and the replacement logic.

Our experiment is based on four main between-subject treatments: Three multilateral market treatments and, as in the previous literature, an individual decision-making control treatment, MPL. In this treatment, we employ multiple price lists to elicit participants' reservation value for avoiding canceling a donation for measles vaccines. We repeat individual decision-making in MPL as often as we repeat all markets. This allows us to control for a potential erosive effect of repetition. In addition to the separate MPL treatment, we also use the individual decision-making task to elicit individual preferences at the start of all market treatments. In all treatments, we also directly measure people's perceptions of the norm about canceling these donations in exchange for money.

Across our market treatments, we vary how many units each market participant can trade. Our first market, treatment SINGLE, is a single-unit market. This treatment is comparable to the markets studied in the current literature and connects the main market treatments of interest, MULTI and FULL, to the previous literature. MULTI is a scaled-up version of SINGLE, where instead of one unit, three units per participant can be traded in each market period. In MULTI, each trader is similarly pivotal as in SINGLE. In FULL, we remove pivotality, as each trader is now able to serve the entire market by herself. This activates both the replacement logic as well as the market selection effect. In all market treatments, we use a common supply and demand schedule. With this schedule, costs and values are equalized across all traders, i.e., they only change in the aggregate quantity exchanged by all traders. Two of the benefits of the common schedule are the following. First, it models features that are typical of markets with negative externalities we want to represent in the laboratory, such as the ones for weapons and flights. In these markets, these common cost and value components are very salient. Second, it allows us to study behavior of the traders holding constant monetary gains from

[4] Using the elicitation method by Krupka and Weber (2013), we find that 666 out of 781 participants rate taking € 1 as a payment to one-self instead of donating € 1.50 to UNICEF in an individual decision-making task as "(somewhat) socially inappropriate".

all trades. By doing so, only differences in morality affect the willingness to engage in trading.

We provide conclusive evidence for a partial erosion of morals in single-unit markets when comparing SINGLE and MPL. Our main interest is in the comparison of different market treatments. Erosion in SINGLE and MULTI is comparable. Strikingly, we detect a full erosion of morals in FULL. Trading in this unrestricted multi-unit market is statistically indistinguishable from selfish competitive equilibrium, consistent with participants completely disregarding that their trading causes negative externalities.

Next to documenting that multi-unit unrestricted (FULL) markets fully erode morals, we find that this deterioration is due to an erosion in norm compliance. While we find some evidence for generalized norm erosion in markets compared to individual decision-making tasks, remarkably, norms are eroded to a similar extent across all market designs: On average, trading is considered approximately equally socially inappropriate in all market treatments. However, in contrast to the unchanged norms, morals are eroded to a much larger extent in unrestricted multi-unit markets because of a deterioration in norm compliance. Norm compliance starts to deteriorate in MULTI compared to SINGLE and entirely breaks down in FULL. In this treatment, norms are fully ineffective, as fully selfish trade emerges.

We further show that the deterioration of morals and norm compliance is largely driven by the widespread use of the replacement logic. In FULL, 83% of market participants attempt to trade units yielding minuscule gains and comparatively large negative externalities whereas only 16% of participants in SINGLE and 32% in MULTI attempt to trade at these same monetary terms.

Additional treatments allow us to provide direct evidence for the two mechanisms. To shed light on the selection argument, we include a treatment similar to FULL in which we divide participants on the basis of their individual decision-making preferences in either a homogeneous group or a heterogeneous group. In the homogeneous group, subjects know that they are matched with traders who, just like them, are close to the median moral preference, which should substantially reduce the scope for market selection. Even under these circumstances, the market exhibits the same degree of erosion documented in the FULL treatment. We infer that, when the replacement logic is available, market selection does not contribute to an erosion of morals. To shed direct light on the replacement excuse, we include treatments similar to FULL and MULTI in which we elicit subjects' beliefs about whether they are pivotal. In agreement with the replacement excuse, we observe that subjects believe to be more likely to be replaced in FULL than in MULTI and are more active when they think that they are more replaceable.

A further noteworthy result is that we find evidence for biased social learning in markets. After the markets, we elicit participants' beliefs about the median subjects' morals exhibited in individual decision-making at the start of the experiment.

We find that multi-unit markets lead to strongly biased social learning. Subjects in these markets are overly pessimistic about their fellow traders' morals. This points to subjects' beliefs not accounting well for how much market selection and, especially, the replacement logic may impact the observable outcomes in such markets.

In the following, we start by positioning our paper in the related literature. We then describe the experimental design and present the novel features of the experimental markets in detail. We continue by presenting our hypotheses and by describing our results. We conclude by discussing the implications of our findings.

3.2 Related literature

In this section, we discuss how our paper contributes to the literature on erosion of norms in markets and the literature on erosion in other interactions. Following Samuelson and Nordhaus (2005, p. 26), we define a market as a mechanism through which buyers and sellers interact to determine prices and exchange goods and services. In a market, traders affect each others' outcomes when they compete to buy and sell valuable products or services. According to this definition, the decisions that people individually make when they trade off money and a negative external effect in an individual price list are not considered *market* decisions. In an individual price list, there is no competition for a scarce good, and people's decisions do not affect other traders' outcomes.[5]

We start with the related literature on moral erosion in markets. The paper by Falk and Szech (2013) inspired a follow-up literature that investigates how different market forces affect traders' morals. Bartling, Weber, and Yao (2015) show that fair and unfair products can co-exist in a market and that it is not necessarily the case that unfair products crowd out fair products. They find only a modest role of erosion. In their Swiss sample, consumers make the fair choice on average 14 percentage points more often in the individual decision-making task than in the market, and the difference is not consistently significant across all specifications (in their Chinese sample they find slightly more erosion).[6] Other papers have investigated the role played by other factors on moral erosion, such as anonymity, market framing, joint decision-making or relative share of buyers versus sellers affect traders' morals in markets (Kirchler, Huber, Stefan, and Sutter, 2016; Irlenbusch and Saxler, 2019; Sutter et al., 2020). Engelmann, Friedrichsen, and Kübler (2018) show that the morality of behavior in laboratory markets correlates with the

[5]Our finding that subjects find trading less socially inappropriate in markets than in individual decision-making reveals that markets and individual decision-making do not only differ technically, but also in the minds of our subjects.

[6]Bartling et al. (2015)'s findings are robust to different specifications of the externalities (Bartling, Valero, and Weber, 2019).

type of choice they are intended to capture outside of the laboratory. All these papers exclusively focus on single-unit markets that de-activate the selection effect and the replacement excuse. Instead, the forces they focus on are active across all our market treatments, so are held constant in the comparison between market treatments we are focusing on. All these studies also do not independently elicit subjects' perceptions of norms, so they cannot distinguish between norm erosion and the erosion of norm compliance.

Besides Bartling et al. (2015), there are also some other papers that study specific market structures that allow markets to partially sustain pro-social behavior. Schneider, Brun, and Weber (2020) document an endogenously arising wage premium, and associated sorting, for morally questionable tasks. Other examples in which competition and pro-social behavior can be mutually reinforcing are provided by Byambadalai, Ma, and Wiesen (2019) and van Leeuwen, Offerman, and Schram (2020). In a large non-student sample, Riehm, Fugger, Gillen, Gretschko, and Werner (2022) highlight the importance of norms in these types of markets: Traders prefer to condition their decisions on others' entry and punishment opportunities for immoral trading are frequently used. Ockenfels, Werner, and Edenhofer (2020) and Herweg and Schmidt (2022) compare (experimentally the former, theoretically the latter) taxes and cap-and-trade schemes to regulate moral markets with negative externalities.

Our conjecture that market selection can be an important force is based on a literature that shows that there is substantial heterogeneity in people's social preferences (Offerman, Sonnemans, and Schram, 1996; Fischbacher, Gächter, and Fehr, 2001; Burlando and Guala, 2005). Falk, Becker, Dohmen, Enke, Huffman, and Sunde (2018) document heterogeneity in social preferences within and across many countries. Given that the most immoral traders are the ones who may determine how much is traded in a market, heterogeneity can furnish selfish aggregate outcomes.

Our paper also contributes to a literature that investigates how the replacement logic and diffusion of pivotality affect behavior in non-market games. Dana, Weber, and Kuang (2007) show that a diffused responsibility for moral outcomes erodes moral behavior in dictator games. Grossman (2014) demonstrates that this effect survives when subjects have to actively seek to remain ignorant. In an individual decision-making context, Falk and Szech (2014) find that almost a third of their subjects pay for a diffused notion of being pivotal for a questionable moral outcome. Serra-Garcia and Szech (2022) study how the demand for moral ignorance depends on monetary incentives. They find that the demand for ignorance does not respond to social norm messages. Exley (2016) demonstrates that uncertainty about the impact of a charity may serve as an excuse not to give. Falk et al. (2020) find support for the replacement logic in committee decisions. A string of papers study diffusion of pivotality in ultimatum games with proposer or responder competition. Roth, Prasnikar, Okuno-Fujiwara, and Zamir (1991); Prasnikar and

Roth (1992); Fischbacher, Fong, and Fehr (2009) find that the side with competition receives almost nothing of the endowment.[7]

There are also studies that find only limited support for the replacement logic. Bartling and Özdemir (2023) find that subjects do not employ the replacement excuse if a social norm exists that classifies the selfish action as immoral. In a voting context, Brütt, Schram, and Sonnemans (2020) find mixed evidence for the effect of decreased pivotality.

An important contribution of Behavioral Economics is to study how findings from stylized, simple settings generalize to market settings (e.g. List, 2003; Enke and Zimmermann, 2019; Enke, Graeber, and Oprea, 2022). In this light, our paper studies the generalizability of the replacement logic to markets. Compared to the previous stylized settings, we can study the importance of the replacement logic in a market environment where competing forces are active. These can be previously studied forces that erode morals already in single-unit markets, as well as the market selection effect we introduce in multi-unit markets. Our findings show that the replacement logic substantially increases the erosion of morals in markets, beyond the erosion in single-unit markets. Lastly, insofar as normative judgments are context-specific, our paper sheds novel light on how norms and norm compliance are shaped in market contexts by the availability of the replacement logic argument. In particular, we find a full erosion of morals driven by the replacement logic, against the prevailing norm.

3.3 Experimental design

The experiment consisted of three main parts.

Parts 1 and 3 were identical to each other and the same in all treatments. In these parts, subjects faced an individual decision-making task which elicited their willingness to accept (WTA) to cancel donations towards UNICEF for varying stakes. In Section 3.3.3, we give more details on the donation opportunity. We employed multiple price lists where subjects chose between varying amounts of money and donations to UNICEF. Monetary amounts ranged between € 0 and twice the monetary amount of the donation under consideration, with a total of 21 steps in each list. Each subject faced separate price lists for 1, 2, 3, 5, 7, 10 and 15 units of donation, in increasing order. We restricted participants to switch at most once in each price list. In our analysis, we set a subject's moral costs equal to the payment at which the subject switched.[8] We set the moral costs of subjects who never choose

[7] There is also theoretical work on the replacement logic. Besides Sobel (2007), the papers of Huck and Konrad (2005), Grossman and Van Der Weele (2017), and Rothenhäusler, Schweizer, and Szech (2018) have theoretically studied diffused notions of pivotality.

[8] We do this to match behavior in the markets, where we can only infer that a subject's moral costs is at most equal to the profit margin of a submitted or accepted offer.

to cancel a donation equal to the upper bound of the multiple price list.

Part 2 varied in the four between-subject treatments. In our control treatment (individual decision-making, or MPL), part 2 presented a repetition of the task of part 1 for four times. In the three market treatments, four market periods were implemented. Part 3 repeated the individual decision-making task of part 1 in each treatment.

3.3.1 Markets

We implemented two-sided posted offer markets characterized by common supply and demand schedules. We here explain these features and the rationale behind them.

3.3.1.1 Two-sided posted offer markets

We implemented the market as a two-sided posted offer market with induced values and costs. Each market consisted of five buyers and five sellers interacting repeatedly and anonymously. Buyers posted bids, sellers asks, and all traders could accept an offer of the other market side. If accepted, a trade was implemented at the price of the accepted offer. The buyer received a payment corresponding to the induced value minus the price and the seller received a payment equal to the price minus the induced costs. For every unit traded, a donation to UNICEF which costs approximately € 1.50 was cancelled.

Buyers and sellers moved in turns, trading unit by unit. In each market period, one side of the market – i.e., the buyers or the sellers – was randomly determined to move first. The starting side had the opportunity to submit offers to the second movers within a time constraint of 14 seconds. We restricted all offer submissions to yield non-negative profits for both market sides. Afterwards, the second movers could either accept the most favorable standing offer, or decide to submit a counter offer. A counter offer had to improve upon all preexisting offers. If no trader accepted an offer, the most favorable counter offer was presented to the original starting side, and traders could again decide whether to accept the most favorable offer or improve upon the best offer they had submitted so far.

If both market sides did not accept or submit an improved offer at least twice, the market period ended and no further units could be traded. Participants were shown a reminder of this feature after neither side had been active once. Whenever an offer was accepted and the 14 seconds time limit had elapsed for all traders currently moving, the trade was implemented for the two agreeing traders. If more than one trader accepted an offer, or if multiple offers were equally favorable, one randomly determined buyer and one randomly determined seller traded, irrespective of the exact time at which an offer was made or accepted.

After a unit had been traded, all pre-existing offers were removed and the previous second-movers were first to propose new offers for the subsequent unit. These design features have three key advantages: (i) the responding market side has most bargaining power, as they only observe the most favorable offer of the proposers, therefore we obtain relatively tight bounds on the profits proposers deem acceptable; (ii) subjects have 14 seconds to decide, which gives participants sufficient time to think and simultaneously generates observations on the willingness to trade for all active traders (and not only the fastest to react). This goes beyond what is normally observed in a traditional double auction where trade is implemented immediately after agreement. Notice further that the posted offer element fits the product markets that we target, whereas standard double auction rules are more representative of financial markets.

To ensure that the negative externalities were salient, each time when participants traded a unit and at the conclusion of a market period, traders were reminded about the consequences of their trading for the charity.

3.3.1.2 The common schedule

In our markets, we use a common schedule. In a common schedule, a seller's cost for supplying a unit and a buyer's value for buying a unit depend on the total quantity already traded in the market, while they are held constant between traders. As a consequence, costs and values depend on the timing of when the trade happens, compared to the other trades in the market. In the common schedule of our paper, for any trader, profit margins of early trades are larger than profit margins of later trades. In contrast, in a private schedule, each trader's costs and values depend only on the quantity traded by themselves, and they differ across traders.

Our motivation for choosing a common schedule is threefold. First, a common schedule captures essential features of the markets that we target. While real-world market schedules have both private and common elements, we think that in markets with negative external effects common elements are often particularly salient. Consider for instance the market for weapons. In a war, the buyers of weapons benefit much more from guns that they are able to secure early in the conflict than guns that they obtain later, while at every moment the strategic advantage the weapons afford are first-order similar across potential buyers. Likewise, in the short run, there is only a limited number of factories in the world that produce for instance AK-47 guns, and a trader who acquires these guns early may do so at lower costs than a trader who does it later when the factories are closer to their capacity constraints. Thus, in the market for weapons, the willingness to pay for the products and the costs of the products depend to a large extent on the timing of the trade. Similar common schedule features characterize other important markets with negative external effects. In the aviation market, airlines lease a substantial part of the aircrafts. This feature represents a strong common cost element for airlines in this market. Consumers may prefer to fly to interesting places before they become less attractive for everyone due to overtourism. In the market for illegal construction permits, constructors will prefer to acquire early permits which allow them to choose the best spots to build their resorts. Corrupt officials will find it easier to hand out early permits before public opposition becomes organized.[9]

Second, such a schedule has the advantage of providing a clean interpretation of trading data: For each unit traded, all buyers (sellers) face the same values (costs). Because they compete on even ground from a monetary perspective, a differential propensity to trade can be ascribed to a difference in their moral costs.

Third, equalizing the monetary terms across participants after each trade ensures that traders remain fully replaceable with each other. This means that both the replacement logic argument and market selection have the same opportunity to arise, irrespective of traders' earlier behavior. In contrast, with a private schedule, participants who had refrained from trading gain a competitive advantage, which inhibits both forces.

Opportunities to replace other traders can also occur in markets with private schedules. Here, the shape and slope of the private schedules affect the size of the maximal potential impact for moral erosion that can be produced by the replacement logic argument and market selection. In Appendix Section 3.A.11, we provide a few examples of private schedules that can trigger replacement thinking.

3.3.1.3 Main market treatments

We ran three main market treatments: SINGLE, MULTI and FULL. In the single-unit market treatment, SINGLE, each trader is restricted to trade at most one unit, so up to five units could be traded in the entire market. This treatment allows for most market forces of erosion considered in the current literature.

The multi-unit market, MULTI, was implemented identically to SINGLE, with the exception that each trader could trade up to three units. This implies that in each market, 15 units could be traded. We also scaled up induced values and costs exactly proportionally. Doing so, MULTI only differs from SINGLE in the scale of an otherwise identical market.

We allowed each trader to cater to the entire market in the unrestricted market, FULL. Treatment FULL was identical to MULTI apart from one key aspect: We removed the capacity constraints of each trader. This means that each participant

[9] For some background on these markets, see `https://www.theguardian.com/world/2001/jul/09/armstrade.iantraynor`; "Mid-life aircraft trading patterns and the impact of lessors". Flightglobal, 7 March 2017; `https://www.theguardian.com/world/2020/jan/25/overtourism-in-europe-historic-cities-sparks-backlash`; `https://www.phnompenhpost.com/national/apsara-raises-concerns-over-illegal-construction-angkor`.

was able to trade up to 15 units and thus serve the entire market.

In all treatments, costs and values each trader faces were identical (as a consequence of the common schedule) and known to all traders. In Figure 3.1, we plot the costs and values we induced using the common schedule in treatment SINGLE on the left and treatments MULTI and FULL on the right. The first units were designed such that trade is efficient: The surplus available to traders is larger than the associated costs to UNICEF by trading these units (surpluses of € 3.80 and € 2.40 compared to a cost of donating of € 1.50). Profitability decreased progressively in subsequent units where market participants could split € 0.60, € 0.40 and € 0.20.

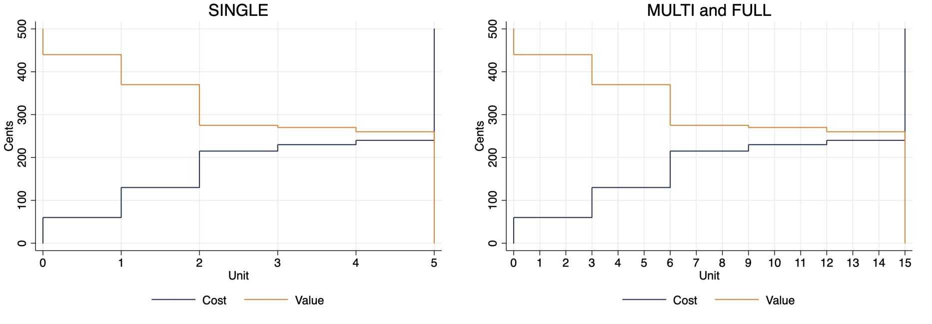


Figure 3.1: Induced common costs and values

In each market treatment, traders first participated in a practice market where no externality was present, to make them familiar with the market environment. Afterwards, we implemented four market periods in which every trade caused an externality through the cancelled donations.

Subjects' trading in the practice market without externalities allows us to see if our design features lead to different market outcomes than previously established in the literature. Across all groups, all units were traded in the practice market period. Therefore, our trading institution produces standard results for experimental markets in the absence of externalities. Lower trading volumes can be cleanly attributed to the introduction of negative externalities. Moreover, as a control market, we ran a double auction with a private schedule. We report on this treatment in Appendix Section 3.A.11.

3.3.1.4 Other treatments

We included some follow-up treatments that allow us to further investigate the mechanisms behind our main results. To provide direct evidence on the selection effect, we ran two additional FULL markets differentially activating market selection. On the basis of participants' moral costs elicited in individual decision-making in part 1, we formed groups either consisting of the middle two quartiles (so, participants close to the median preference) or of the first and fourth quartiles. The latter, HET, fully activates market selection as participants are very heterogeneous in their preference for the external effect. The former, HOM, generates homogeneous market groups, where market selection has less scope to affect outcomes. To ensure that participants hold correct beliefs about their fellow traders' morals, we informed participants both at the start of part 1 and part 2 of the group formation procedure, in part 2 they also learned which type of group they belonged to.[10]

To shed direct light on the replacement logic, we included three treatments in which we directly elicited beliefs about other traders' activity in markets. Treatments B-MULTI and B-FULL replicate MULTI and FULL with additional belief elicitations about the trading of unit 10, 12, 13 and 15. Just before trading of these units started, traders reported their non-incentivized beliefs about the probability that the next unit will be traded, either with or without their participation. In addition, we elicited the (cognitively less demanding) prediction of how many of the other traders will attempt to trade the next unit. This last prediction was incentivized: If and only if participants predicted this number correctly, they would earn a bonus of € 1.50. Next to the treatments with direct belief elicitation in the markets, we ran a treatment with spectators, SPEC. The spectators were not directly involved in any market transaction. Instead, they followed the series of screens and received the identical information of a randomly matched participant from B-FULL and were asked to report their own beliefs in the same fashion. Comparing B-FULL and SPEC allows us to test for self-serving belief reports in B-FULL.

3.3.2 Additional elicitations

In all treatments, we included additional measurements of subjects' views and attitudes after part 3. We elicited: (i) beliefs about the median trader's WTA to cancel donations; (ii) norms about behavior in individual decision-making and markets; (iii) risk preferences. For the beliefs, subjects were asked to fill in a multiple price list reporting what they "think the average participant did" in the first list of part 1. If their belief matched the choice of the median participant, they received € 1. To elicit subjects' perception of the norms for canceling donations in either individual decision-making or the market, we followed the procedure by Krupka and Weber (2013) and asked subjects to state whether scenarios described to them were considered "socially appropriate" and "consistent with moral or proper social behavior" on a 4-point scale from "very socially inappropriate", to "somewhat socially (in)appropriate" and "very socially appropriate". For one randomly picked

[10]This information was processed well, as beliefs about the median participants' morals are more precise in HOM (average absolute error of 38.8) than in HET (average absolute error of 69.8), the difference is statistically significant (MWU, 8 observations per treatment, p-value=.003).

scenario, subjects received € 2 if their choice matched the modal choice in their session. Among the scenarios described were "[Individual] 1 chooses to receive 1 Euro instead of making a donation of 4 doses of measles vaccine to UNICEF" and "[Individual] 2 decides to accept an offer which allows him to earn 1 EURO". For the full list of scenarios, see the Appendix Section 3.A.5. We also elicited risk attitudes using the method introduced by Eckel and Grossman (2002).

3.3.3 Experimental procedures

For the treatments MPL, SINGLE, MULTI and FULL, the computerized laboratory experiment was run in 28 sessions in September and October 2019, at the CREED laboratory of the University of Amsterdam. We preregistered the experiment (Offerman, Romagnoli, and Ziegler, 2019b). In total, 381 subjects participated. 47% were women, with an average age of 21. We had 100 participants per market treatment and 81 participants in MPL. Sessions lasted on average 1.5 hours, with average payments of € 19 per subject, besides payments to UNICEF.

We conducted the follow-up treatments from October 2021 to January 2022. These were pre-registered separately (Offerman, Romagnoli, and Ziegler, 2021). In total, 441 participated in the new sessions. Out of those, 208 participants were recruited from the pool at the CREED laboratory at the University of Amsterdam. The remaining 233 participants were recruited from the pool at the CentERlab at Tilburg University. Treatments were balanced in the composition of participants from Amsterdam and Tilburg (between 63% and 69% of participants were from Tilburg), apart from treatment PRIV, which was fully ran in Amsterdam. We did so as only data from PRIV was directly compared to the original treatments, which were also only ran in Amsterdam. All treatments consisted of 80 participants, apart from SPEC with 41 participants. 55% of participants were women, with an average age of 21. Sessions lasted on average 1.7 hours, with average payments of € 20.4. In Appendix Section 3.A.2, we show that participant characteristics are balanced across all treatments.

Subjects knew that they were paid for only one randomly selected part from the first three parts. All subjects within a session were paid for the same part. If individual decision-making was selected, one decision from one of the multiple price lists was randomly chosen and paid for each subject. If one of the markets was selected, the sum of earnings in two out of the four market periods and the practice market was paid. Additionally, subjects received a show-up fee of € 7, all earnings from the three additional elicitations at the end of the experiment as well as an unannounced lump-sum payment of € 9 if the markets were selected for payment, to guarantee sufficient minimum earnings.

Subjects read the computerized instructions at their own pace and separately for each part of the experiment (see Appendix Section 3.C). They also received handouts with summaries of the instructions. Subjects were required to complete a set of test questions before they could proceed. Subjects were paid in cash and in private at the end of the experiment.

In the experiment, several choices affected donations to UNICEF. As in Kirchler et al. (2016) and Sutter et al. (2020), donations were intended for measles vaccine. We used a text of UNICEF to inform subjects about the consequences of measles.[11] One dose of measles vaccine through UNICEF costs approximately € 0.375, and two doses are required to vaccinate one person. In the experiment, one unit was chosen to consist of four doses, corresponding to a donation of € 1.50. This amount was communicated to subjects in the instructions and the handout.[12] In the instructions, subjects were presented with sample receipts of such a donation to UNICEF.[13] At the end of each experimental session, the donation was immediately implemented by the experimenter. Subjects were presented with the UNICEF receipt for their session (i) immediately in the experimental interface, jointly with their experimental earnings; (ii) when receiving their earnings in cash; (iii) via email if subjects so desired. These emails were collected on separate handouts and thus could not be linked to specific subjects or choices in the experiment. Subjects were made aware of this procedure at the start of the experiment. In total, approximately € 2111 (€ 889 in 2019 and € 1222 in 2021/22) was donated to UNICEF as a result of subjects' choices.

3.4 Hypotheses

In this section, we elaborate on the hypotheses behind the main contributions of this paper, namely (i) the role played by market power in eroding morals in markets; (ii) the distinction between norm erosion and the erosion of norm compliance; and (iii) the separation of the role played by the replacement logic vis-à-vis market selection. These hypotheses, preregistered in (Offerman et al., 2019b) and (Offerman et al., 2021), are summarized and motivated below.

[11] "Measles are highly infectious and very often deadly. Each day hundreds of children become victims of this disease. The survivors often suffer consequences for their whole life, like blindness or brain damages. This, even though protecting the children would be so easy. Measles kills more than 160,000 children worldwide each year."

[12] This particular donation was only available in packs of 40 doses, excess donations were made over to UNICEF as a generic donation, which subjects were aware of and could verify as well.

[13] At the time of the sessions in 2019, this donation is available at `https://market.unicef.org.uk/inspired-gifts/measles-vaccines-to-protect-20-children/S359163X/`, which we also communicated to subjects. In 2021/22, we instead donated to UNICEF in Austria, `https://unicef.at/shop/produkte/`. Costs per dose were approximately constant and all procedures were kept identical otherwise.

The erosion of morals in single-unit markets

We start by exploring the erosion of morals in single-unit markets by comparing our treatment SINGLE to individual decision-making elicitations in MPL. In doing so, we study the treatment effects from prior literature in our experimental setting. Falk and Szech (2013) report limited erosion of morals in single-unit markets. Bartling et al. (2015) find limited erosion in most specifications. Bartling et al. (2023) fails to reject this hypothesis. Our first hypothesis is thus:

H1. There is no erosion of morals in single-unit markets.

The erosion of morals in multi-unit markets with market power (MULTI)

The following hypothesis bridges our multi-unit markets to the current literature, which studied single-unit markets. Treatment MULTI is a scaled-up version of SINGLE. In both treatments, a single trader can trade up to 1/5th of the maximal market size and retains full pivotality, in that she can unilaterally decide to reduce the maximum aggregate quantity by not trading her units.

H2A. Compared to single-unit markets (SINGLE), there is no additional erosion in restricted multi-unit markets (MULTI).

The erosion of morals in multi-unit markets without market power (FULL)

While MULTI serves as a benchmark treatment for the introduction of multi-unit trading, the next hypothesis is the key hypothesis in our paper. Here, we focus on unrestricted multi-unit markets with treatment FULL. Between MULTI and FULL, the market structure remains identical, apart from removing individual traders' capacity constraints, so each trader can serve the entire market.

H2B. Unrestricted multi-unit markets (FULL) do not lead to more moral erosion than restricted multi-unit markets (MULTI).

Norm erosion and erosion of norm compliance

Our next hypothesis is concerned with the question of whether differences in the degree of moral erosion across treatments are due to changes in norms or in the degree of norm compliance.

H3. Norms are (A) not eroded in markets in comparison to individual decision-making and (B) not differentially affected by the specific market institution.

H3 is also a key hypothesis of our paper. Our independent measures for subjects' norms allow us to distinguish between norm erosion and the erosion of norm compliance. Previous literature highlighted the importance of norms for the availability of the replacement logic (Bartling and Özdemir, 2023).

The mechanisms behind moral erosion in unrestricted markets: Market selection versus replacement logic

Our remaining hypotheses are concerned with investigating the relative role played by the two mechanisms of market selection and replacement logic in the erosion of morals that we expect to detect in treatment FULL. We here provide a definition of both forces.

Market selection. According to this mechanism, traders compare the material profit from trading to the moral costs that they incur from imposing the associated externality. Each trader continues to trade until their own moral costs no longer justify the monetary returns. As trade progresses, the profit margins get smaller, justifying trade for an ever smaller number of traders, i.e., those for whom moral costs are lowest. The final units will be traded by the traders with the lowest moral costs within their market. Additionally, a potential decrease in the least moral traders' marginal moral costs further increases the quantity traded.

The replacement logic. The replacement logic is a mechanism based on the following strategic thinking: Traders ask themselves whether their trading will affect the aggregate quantity traded in the market, assuming that other traders behave as if they are selfish (thus willing to trade all units available to them). If under this assumption their own behavior would not impact the aggregate volume traded, then this motive convinces them to trade irrespective of their own moral costs.

Notice that the belief of other traders behaving selfishly will be correct not only when other traders *are* actually selfish (i.e., genuinely unconcerned with the negative externality), but also when other moral traders *act* selfishly because they themselves apply replacement logic thinking, in a self-fulfilling cycle. Because traders can always replace each other in the unrestricted FULL market, the application of the replacement logic could lead to full trade and thus a full erosion of morals in this treatment. In the case of SINGLE or MULTI, traders' unilateral withdrawal from trade diminishes the aggregate quantity. This remains to hold even when all other traders act selfishly. Therefore, traders conclude that their behavior will matter for the aggregate outcome and not trade units where moral costs exceed their profits. Notice that this view of the replacement logic is similar in spirit to Falk et al. (2020).

Our hypotheses regarding the mechanisms of moral erosion are thus:

H4. Any erosion of morals in FULL compared to MULTI is not driven by market selection.

H5. Any erosion of morals in FULL compared to MULTI is not driven by the replacement logic.

3.5 Results

In this section, we present the results of the experiment. For all market outcomes, we perform tests on the basis of averages of matching-group data, which yields 10 observations for each market treatment SINGLE, MULTI and FULL (10 groups with 10 participants each per treatment), as well as 8 observations for HOM, HET, B-MULTI, and B-FULL. MPL and SPEC feature no interaction, with 81 and 41 observations, respectively. For all tests on the individual level, for which participants do not interact, we study individual level data. To construct the confidence intervals in the graphs, we used a bootstrap procedure. We do this to correct for floor and ceiling effects of proportions close to 0% or 100%.[14]

3.5.1 Morals in individual decision making

In the individual decision-making task, the moral costs connected to causing the negative externality are quite substantial, with an average evaluation of € 1.42 for a € 1.50 donation to UNICEF.

Two factors contribute to a potential effect of market selection in multi-unit markets: (i) initial heterogeneity in how traders value donations, and (ii) decreasing marginal moral costs in traders' preferences for causing the negative externality.

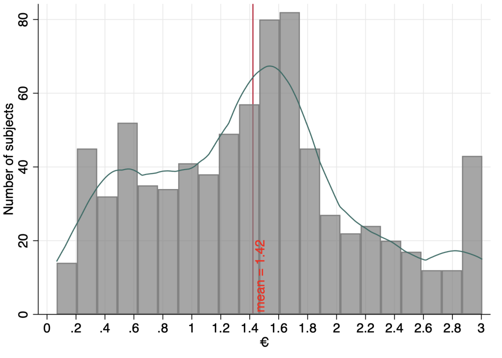
On the basis of individual decision-making data, we verify that these two factors can play a role. From the choice data for units 1, 2, 3, 5, 7, 10 and 15 we calculate the average per-unit valuation of a € 1.50 donation to UNICEF.

In Figure 3.1, we provide a histogram of the per-unit moral costs of all subjects in part 1 of the experiment, averaged at the subject level. We show the minimum payment that a subject requires to be willing to cancel a donation to UNICEF. Evidently, there is substantial heterogeneity in how subjects value the opportunity to donate to UNICEF. A minority of subjects hardly cares about donating to UNICEF. There is also a remarkable share of subjects whose moral costs are estimated to be above € 1.50, implying that they value donating *more* than the corresponding monetary value.[15]

[14] In the bias-corrected confidence intervals that we plot, we introduce clustering at the matching group level (the market group for market treatments and the participant for MPL or SPEC) and use 10,000 replications.

[15] Bénabou, Falk, Henkel, and Tirole (2020) show that elicited moral costs can be affected by the method of elicitation, when using either direct elicitation or multiple price lists, since image motives are affected differently by these methods. In our experiment, we keep the elicitation method constant across treatments. In our data, we find only few "observationally deontological" subjects, those who never cancel a donation across all price lists, as only 28 out of 781 subjects do so across part 1, compared to 26% of subjects who do not cancel the donation for any monetary amount in Bénabou et al. (2020).

Figure 3.1: Heterogeneity in valuations of donations



Notes: Histogram of subjects' average moral costs for cancelling a donation with a value of € 1.5. For each subject, we use the switching points from all multiple price lists for cancelling donations in part 1. Kernel density is displayed in green, the mean in red.

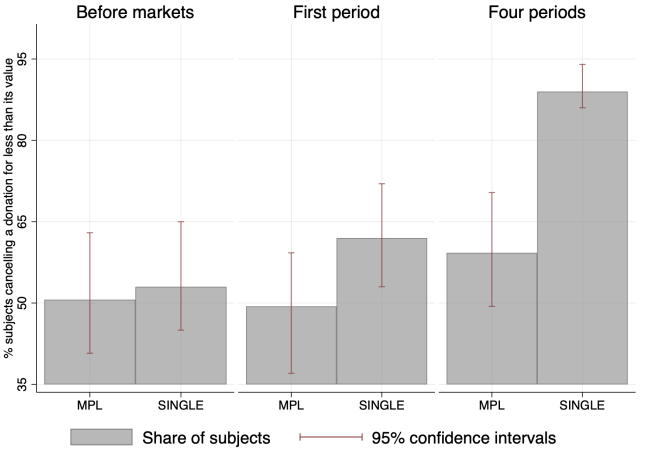
We also detect decreasing marginal moral costs and provide an analysis in Appendix Section 3.A.3. Given these data, there is a clear possibility for market selection to play an important role.

3.5.2 Moral erosion in markets

In this section, we investigate whether market behavior and outcomes display moral erosion. Whether moral erosion is due to norm erosion or an erosion of norm compliance is the topic of the next section. We start with measuring erosion in single-unit markets, as in Falk and Szech (2013). We compare individual-level decisions to cancel donations across individual decision-making and single unit-markets. In Figure 3.2, we plot the share of subjects who cancel a donation in exchange for € 1.50 (i.e., its value) or less in different environments and at different stages of the experiment. In the first two bars, we plot the share of subjects who cancel the first unit of donation for a payment of at most € 1.50 in individual decision-making in part 1. These treatments are balanced in this dimension. The following two groups of bars compare behavior of these participants either in repeated individual decision-making in MPL or in markets in SINGLE. For the markets, we study whether a trader concluded a trade for which she was paid at most € 1.50. This is the comparison that speaks to the literature on erosion in single-unit markets. In the middle panel, we compare behavior in the first period in part 2. We observe that there appears to be an erosion of morals in markets. In the right panel, we use the entire four periods of the experiment and plot the share of participants who at least once cancelled a donation for at most € 1.50 in part 2.

Table 3.1 reproduces estimation result of the corresponding effect. The depen-

Figure 3.2: Cancellation of donations between environments and treatments



Notes: Share of participants who cancelled a donation for at most its value (€ 1.50) in individual decision-making and in trades in the market. The left panel shows cancellation rates in part 1 of the experiment and the middle panel plots cancellation rates in the first period of part 2. The right panel displays the share of participants who, in the four periods of part 2, at least once cancelled a donation.

dent variable is a dummy variable equal to one if a participant cancelled a donation for at most its value either (i) in period 1 of part 2 or (ii) at least once in periods 1-4 of part 2. Models (1) and (2) suggest that there is erosion through repetition, as in Bartling et al. (2023): more participants cancel a donation in the entire part 2 than only in its first period. In our setup, we do find evidence for an erosion in markets: models (3) and (4) suggest that more participants cancel a donation in SINGLE than in the corresponding time interval in MPL. Model (5) confirms that this is particularly strong when testing for erosion in the pooled data of part 2, compared to only the first period.[16] Summarizing, we find evidence for both a partial erosion of morals in markets as well as erosion when measured by a subject cancelling a donation at least once in a repeated task, compared to a non-repeated measurement.

[16]A more conservative approach would be to halve the moral costs in the market as a result of shared responsibility. The effect of erosion in SINGLE in models (4) and (5) is robust to defining erosion within markets as the decision to cancel a donation for a payment of € 0.75 or less. For example, the estimate on SINGLE corresponding to (4) is .247 (p-value < .001). In Section 3.5.3, we also present direct evidence on norm erosion between individual decision-making and markets.

Table 3.1: Erosion in single-unit markets and through repetition

	(1) MPL	(2) SINGLE	(3) MPL & SINGLE	(4)	(5)
			Period 1	Period 1-4	Pooled data
Period 1-4	0.099***	0.270***			0.099***
	(0.033)	(0.052)			(0.033)
SINGLE			0.126*	0.297***	0.126*
			(0.074)	(0.059)	(0.075)
SINGLE × Period 1-4					0.171***
					(0.060)
Constant	0.494***	0.620***	0.494***	0.593***	0.494***
	(0.056)	(0.051)	(0.056)	(0.055)	(0.056)
Observations	162	200	181	181	362

Note: Dependent variable is a dummy equal to one if a subject cancelled a donation for a payment of at most its value (€ 1.50) either in SINGLE or in MPL. Period 1-4 is a dummy variable equal to one if the choice is measured as occurring at least once in period 1 to 4 in part 2 of the experiment, the omitted category is cancellation in period 1. SINGLE is a dummy equal to one if the choice occurred in treatment SINGLE, with the omitted category MPL. Standard errors, clustered on subject level for MPL and matching group level for SINGLE, are presented in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Result 1. *We reject hypothesis H1, and find partial erosion of morals in single-unit markets.*

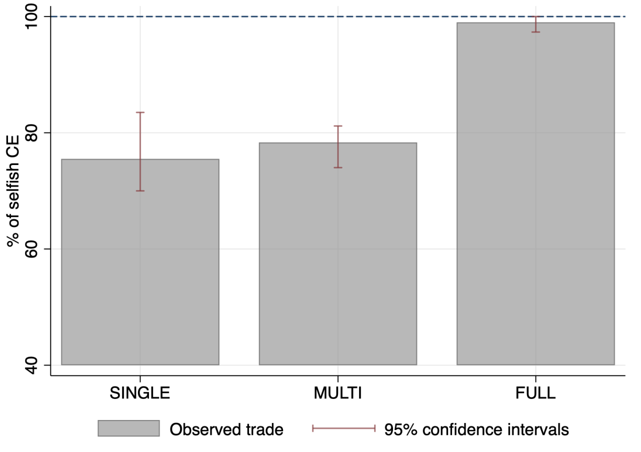
Our key hypotheses are on behavior in multi-unit markets. We want to establish whether there is an erosion in these markets, *in excess* of the erosion we find in single-unit markets. To measure erosion, we will focus on aggregate quantities traded. Higher quantities imply larger negative externalities, so they are a natural measure of the overall effect of the market structure on the morality of trading outcomes. In addition, we can exploit that our markets featured decreasing gains from trade, while damages to UNICEF are kept constant at € 1.50 per unit traded. Thus, the trading of larger volumes also implies that traders are willing to accept lower trading margins, which directly ties to the measure of moral erosion commonly used in the literature.

In Figure 3.3, we plot the observed market quantities. All quantities are relative to the selfish competitive equilibrium outcome, according to which 5 units are traded in SINGLE, and 15 units in MULTI and FULL.

The bars show traded quantities relative to the competitive equilibrium across the three treatments. SINGLE and MULTI show similar traded quantities, consistent with a comparable amount of erosion in these markets. In contrast, we observe that market outcomes in FULL are fully selfish. Traded quantities exceed quantities in other market treatments, indicating substantially stronger erosion in FULL.

Erosion appears to be particularly strong in FULL if the shrinking gain of surplus of the additional units is taken into account. Induced gains from trade are decreasing at higher quantities, while damages stay constant. Below 40%, trading

Figure 3.3: Market outcomes



Notes: Average quantities relative to selfish competitive equilibrium. Trading units below 40% is efficient (gains from trade exceeds the externality). Compared to the negative externality of € 1.50 per unit, each unit between 40% and 60% yields gains from trade of € 0.60, each unit between 60% and 80% yields € 0.40 and each unit between 80% and 100% yields € 0.20.

is efficient, as the damage to UNICEF is less than the associated payments to market participants. An increase of trade from 40% to 60% leads to additional negative externalities of € 4.50, while traders receive € 1.80. A further increase from 80% to 100% again yields damages of € 4.50, however traders only receive the meagre total payments of € 0.60.[17]

In Table 3.2, we summarize market quantities relative to the selfish competitive equilibrium quantities together with *p*-values of Mann-Whitney U-tests (10 observations per treatment) of quantity comparisons between treatments.[18]

Result 2. *We detect full erosion of morals in unrestricted multi-unit markets (FULL). Erosion in MULTI is similar to erosion in SINGLE.*

We also included an additional control treatment in which we implemented a standard double auction with a private schedule, with a multi-unit design and a scope for replacement similar to MULTI. In this treatment, we assigned values and costs in such a way that the aggregate supply and demand coincides with MULTI.

[17] This result is also supported by using part 1 data to predict market outcomes under the assumption that moral costs are not changing in a market environment. When we compute the moral competitive equilibrium, we find ample scope for market selection and erosion of morals in FULL. We provide details in Appendix Section 3.A.1.

[18] These treatment differences also arise when regressing quantities on treatment indicators, with and without controlling for period indicators, moral costs (average, median and minimum within matching group), as well as risk measures; see the Appendix Section 3.A.4 for results.

Table 3.2: Treatment effects

		SINGLE	MULTI	FULL
Quantity in %		75.5	78.3	99
p-values	vs. SINGLE	-	.378	.0005
	vs. MULTI	-	-	.0001

Notes: Average quantities relative to selfish competitive equilibrium. Mann-Whitney U-tests, on matching group averages, 10 observations per treatment.

We report on these results in Appendix Section 3.A.11. The main takeaway from this treatment is that morals are eroded to an approximately similar extent as in MULTI.

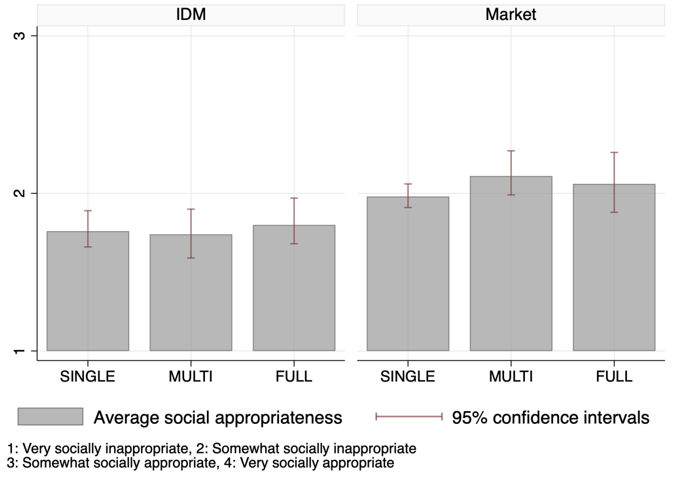
3.5.3 Norms and norm compliance

The preceding section presented evidence for a complete erosion of morals only in FULL markets. An important question is whether this change can be attributed to a change in norms or whether it is the result of an erosion of norm compliance. Did traders feel that cancelling donations in exchange for minuscule profits in FULL was "consistent with moral or proper social behavior"?

To this end, we elicited subjects' norms in individual decision-making tasks and experimental markets after the markets took place, using the method proposed by Krupka and Weber (2013). Subjects were incentivized to report what they believed was their session's modal answer on a 4-point scale from "very socially inappropriate" (indexed 1), to "very socially appropriate" (indexed 4) in response to scenarios in which a participant in an experiment chose to cancel donations of € 1.50 when paid € 1 either in individual decision-making or in an experimental market.

In Figure 3.4, we display the mean answers to two (otherwise identical) questions regarding the social appropriateness of canceling a € 1.5 donation in exchange for € 1 in individual decision-making (left panel), and in a market (right panel). We observe that across all market treatments and both environments, cancelling such donation is rated on average at best as "somewhat socially inappropriate". Thus, there does exist a clear norm that cancelling donations and trading is not appropriate. This norm particularly contradicts the rather frenzied trading behavior observed in FULL.

Figure 3.4: Norms in individual decision-making and in markets



Notes: Average norm in response to cancelling one donation of €1.50 when paid €1 in individual decision-making (left panel) and in the market (right panel). A rating of 2 corresponds to "somewhat socially inappropriate".

In accordance with even single-unit markets eroding morals, causing an externality in a market is perceived as less inappropriate as the same choice in individual decision-making (Wilcoxon signed-rank, 300 observations, p-value< .001).

Somewhat surprisingly, differences in elicited norms do not map one to one to differences in behavior between market treatments. In particular, the more selfish behavior in FULL is not supported by a further erosion of the norm compared to the other market treatments.[19] We cannot reject equality of norms in markets comparing SINGLE and MULTI (MWU, 100 observations per treatment, p-value=.238) and between MULTI and FULL (MWU, 100 observations per treatment, p-value=.705).[20] We report additional descriptive statistics for other scenarios in the Appendix Section 3.A.5, which yield similar conclusions.

Even though norms do not further erode in FULL compared to the other treatments, we see a complete break-down of norm compliance. When traders can take advantage of trading opportunities foregone by other traders, norms take a back seat in participants' decision making. In the next section we shed light on

[19]We find no evidence of excuse-driven norm reports, see Appendix Section 3.A.10.

[20]Results are similar when regressing subjects' norms (2 elicitations for 781 subjects, so 1562 observations) on treatment fixed effects, a dummy for the market scenario and interactions of this dummy with the treatment fixed effects, clustering standard errors on the matching group. Significant is the dummy for the market scenario (p-value = 0.02), but none of the interactions is significant (all p-values > .1). This confirms that there is not a specific treatment effect on norms in markets.

the question whether the complete breakdown of norm compliance is caused by market selection or the replacement logic.

Result 3. *We reject hypothesis 3A. Traders find cancelling a donation less inappropriate in markets than in individual decision-making. We do not reject hypothesis 3B. Norms are not differentially affected by market treatments. The finding that market outcomes are most selfish in FULL is caused by a breakdown of norm compliance.*

3.5.4 Mechanisms: Market selection versus replacement logic

A crucial question is the mechanism behind the full erosion of morals in FULL. In this section we aim at providing direct evidence for each of these forces separately.

In a first step in distinguishing the two mechanisms, we study which traders are active in the market. Under market selection, only the least moral participants trade the last units, while all other participants abstain. In contrast, the replacement logic can be used by any trader and is most powerful if many traders become active. We thus study which traders are active in submitting or accepting offers for the final units, the least profitable units which yield gains from trade of € 0.20. To evaluate which type of trader is active we split the sample into those with below- and above-median moral costs in part 1. If market selection drives erosion in FULL, we would expect that few very immoral traders are active. If in turn the replacement logic is active, we expect many active traders, and there need not be a correlation between individual activity and the valuations in individual decision-making.

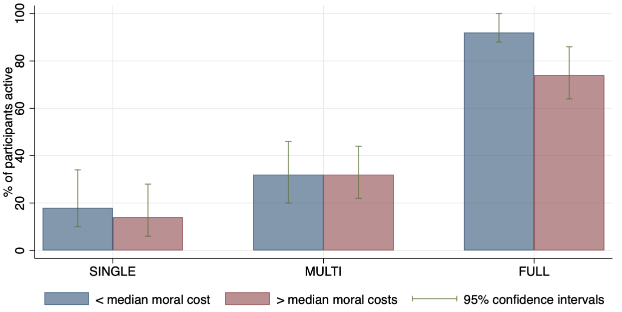
In Figure 3.5, we plot the share of traders who are active at least once at these least profitable units. We see that in both SINGLE and MULTI, both groups of market participants are similarly active. However, the share of active participants is much higher in FULL, where 94% of traders with below-median moral costs are active, but even 72% of traders with above-median moral costs are active. The difference between the above- and below-median group is significant only in FULL (MWU, 10 observations per above- or below-median group per treatment, p-value=.023).[21] This is however not robust to using a regression, see Appendix Section 3.A.6.

This points to only a minor role for market selection. Traders with above-median moral costs are hardly less active than traders with below-median moral costs. This evidence hints at a major role for replacement thinking. A large share of participants across are actively trading when the replacement logic is available, providing justification for the trading of others.[22]

[21] Differences in other treatments move in the expected direction for earlier units with larger associated gains from trade, e.g. in MULTI 78% of above-median participants are active for units 10 to 12, while 92% of below-median participants are active.

[22] In the Appendix Section 3.A.6, we provide further evidence in line with this analysis. While

Figure 3.5: Share of traders active at the least profitable units



Notes: Share of traders who submit or accept an offer at the final units, which yield gains from trade of € 0.20 in exchange for an externality of € 1.50. Median splits are based on moral costs within the matching group.

A set of follow-up treatments distinguish between these two forces more directly. To study the role of market selection in FULL, we compared homogeneous groups that consist of traders close to the median preference for canceling donations (HOM) to heterogeneous groups that include the traders on both extremes (HET). The main interest is in comparing outcomes in the HOM groups to the original FULL treatments as well as to HET. If market selection drives the erosion of morals in FULL, limiting its scope in HOM would lead to less erosion compared to the erosion in HET and FULL.

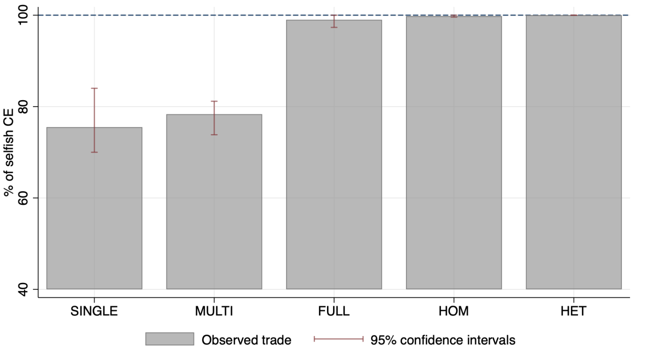
In Appendix Section 3.A.2, we show that the participants in these two groups are balanced across other characteristics we observe. Yet, crucially, participants in HOM are more homogeneous than those in HET. Therefore, this treatment successfully manipulates the potential for market selection to drive outcomes, while other characteristics are not affected.

In Figure 3.6, we present average quantities traded, relative to the selfish competitive equilibrium. Strikingly, market outcomes are similarly selfish in HOM, HET and FULL. Average quantities in HOM are not statistically distinguishable between HOM and FULL (MWU, 8 observations in HOM and 10 in FULL, p-value=.632) as well as between HOM and HET (MWU, 8 observations per treatment, p-value=.317). This indicates that even when limiting the scope of market selection, the replacement logic is sufficient to produce fully selfish market outcomes.

Result 4. *We do not reject Hypothesis 4. Both more and less moral traders are*

traders in SINGLE and MULTI submit or accept less than 1.4 offers on average, traders in FULL engage in 8.2 actions per trader. In addition, we show that a similar picture emerges for the traders who revealed to not use consequentialist reasoning in individual decision-making, since they declined to cancel donations even when paid more than the monetary value of these donations.

Figure 3.6: Market outcomes: HOM and HET



Notes: Average quantities relative to selfish competitive equilibrium. The trading of units below 40% is efficient (gains from trade exceeds the externality). Compared to the negative externality of € 1.50 per unit, each unit between 40% and 60% yields gains from trade of € 0.60, each unit between 60% and 80% yields € 0.40 and each unit between 80% and 100% yields € 0.20.

active. Market selection does not contribute to the complete erosion of morals in FULL.

B-MULTI, B-FULL and SPEC allow us to shed direct light on the replacement logic. In these treatments, we directly elicited beliefs about others' activity in the trading of unit 10, 12, 13 and 15 just before trading of these units started. In the pre-registration, we announced that we will focus on the non-incentivized measure if the two measures correlate. Unfortunately, the two measures do not correlate. Within B-MULTI, the Spearman correlation between non-incentivized reports for the statement "What is the probability that whether or not the next unit is traded depends on your behavior?" and the incentivized report for the statement "How many participants other than you will attempt to trade this unit?" is -0.016 (p-value=.718, 500 observations). The same correlation in B-FULL is -0.003 (p-value=.910, 1280 observations).[23] In the main text, we therefore focus on the simpler incentivized measure.[24] Results for the non-incentivized measures are presented in the Appendix Section 3.A.7 and are in line with these results unless otherwise noted. To avoid selection issues in treatment comparisons, and as pre-

[23]This analysis assumes independence of observations, even though e.g. the same participant reports multiple beliefs. The conclusions are robust to using participant-level averages or regressions with standard errors clustered on a matching group level.

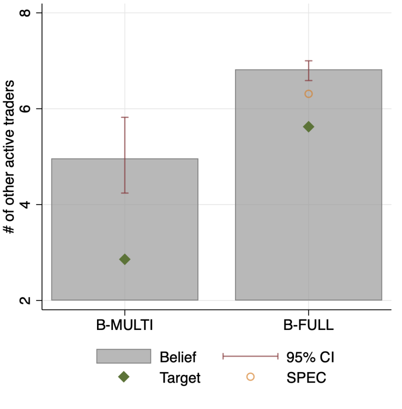
[24]Other reasons to focus on the incentivized measure are that: (i) it correlates more strongly with the underlying true values; and (ii) while we do not find a correlation between incentivized and unincentivized measures for traders in B-MULTI and in B-FULL, we do find the expected correlation for spectators. The latter suggests that we may have been asking too much of our traders, and that they may have decided to focus on the incentivized questions. See Appendix Section 3.A.7 for details.

registered, we use only beliefs for which 13 out of 16 groups are observed – i.e., have continued to trade to the corresponding unit. This only allows us to compare data for unit 10. For subsequent units, beliefs in B-MULTI are only available for a self-selected sample, as already at unit 12 only 40.6% of groups had continued to trade.[25]

We use belief data for two purposes. First, we test whether our treatments induced differences in beliefs on others' activity. If the replacement logic drives the enhanced trading in FULL, we would expect that participants believe that more traders are active in FULL than in MULTI. Second, we check whether within-subject correlations between actions and beliefs are in line with replacement logic thinking, which implies that participants who believe to be more replaceable are those who are more active.

In Figure 3.7, we report the average number of other traders believed to be active in the trading of unit 10, including the corresponding target in the data. Traders in B-FULL believe that more other traders will be active than traders in B-MULTI do, consistent with replacement logic thinking. The difference between these two treatments is significant, with a p-value of .002 (MWU, 8 observations per treatment).

Figure 3.7: Beliefs about other traders' activity



Notes: Number of other traders believed to be active (grey bar), actual number of others active (the target; green diamond) and belief of spectators (orange circle).

Figure 3.7 also presents the target for these reports, based on the actual trading behavior of the other traders. Consistent with the beliefs, we observe more activity in B-FULL than in B-MULTI already at unit 10. Lastly, we show the corresponding re-

[25]Treatments B-FULL and B-MULTI also allow us to investigate the robustness of the original results. In Appendix Section 3.A.8 we reproduce the other analysis presented in the main text including the new treatments. Results are qualitatively in line with the original treatments.

ports for the spectators, in SPEC. Directionally, this data is in line with self-serving reports, but differences between spectators' beliefs and traders' beliefs are minor and not significant (MWU, 8 observations in B-FULL and 41 in SPEC, p-value=.393).

This data can also be used to test whether traders who believe to be more replaceable are those traders who trade most frequently. In Table 3.3, we regress the decision to be active at the last units in the market, those with gains from trade of € 0.20, on participants' beliefs about others' activity. As we do not compare data across treatments, we now use the full data set. We observe that both in B-MULTI and B-FULL, participants who expect others to be more active are more inclined to trade themselves, again consistent with the replacement logic.[26]

Table 3.3: Beliefs and activity

	(1) B-MULTI	(2) B-FULL
# active traders	0.053***	0.069***
	(0.011)	(0.013)
Average moral cost	-0.019	-0.111***
	(0.023)	(0.025)
Period	0.001	-0.028*
	(0.039)	(0.013)
Constant	0.076	0.391**
	(0.119)	(0.121)
Unit FE	Yes	Yes
Observations	500	1280

Note: Dependent variable is a dummy equal to one if a subject submitted or accepted an offer for units with gains from trade of € 0.20. Average moral costs are the average moral costs for a participant, based on averaging per-unit moral costs based on part 1 individual decision-making. Standard errors clustered on matching group level in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Result 5. *We reject Hypothesis 5. Fully selfish market outcomes in unrestricted multi-unit markets are driven by the replacement logic.*

3.5.5 Effects of market exposure

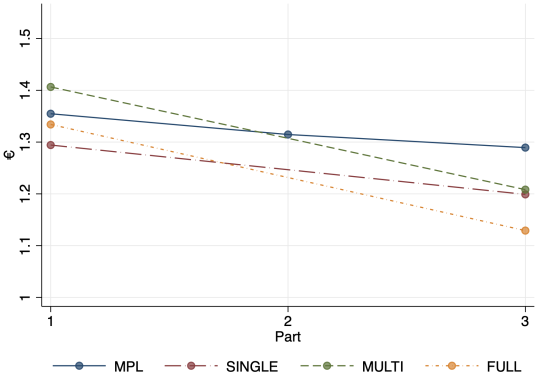
Our experimental design also allows us to test whether morals are eroded within an identical decision environment, as participants faced identical individual decision-making tasks in parts 1 and 3. Treatment MPL allows us to study whether repetition by itself is eroding morals. Comparing this erosion to erosion after experiencing markets in treatments SINGLE, MULTI and FULL allows us to determine whether the

[26] This is the only beliefs analysis that does not generalize when we use the unincentivized belief report (see Table 3.A.8 in Appendix Section 3.A.7).

erosion in markets has an effect outside the immediate market environment. In addition, we can evaluate whether specific market features lead to stronger erosion outside the market.

In Figure 3.8, we plot the average elicited moral costs per treatment, by parts. In treatment MPL, we elicit moral costs in parts 1, 2 and 3. In the market treatments, we use individual decision-making only in parts 1 and 3.

Figure 3.8: Persistence of erosion



Notes: Average per-unit valuations in individual decision-making, for € 1.50 donations, by part. In part 2, only MPL employs individual decision-making.

We observe that moral costs are decreasing over time. In MPL, average per-unit moral costs in part 3 decrease by 6.5 cents (relative to a donation of € 1.50), compared to the moral costs in part 1. This change slightly increases in the markets, in SINGLE it amounts to 9.5 cents. In the multi-unit markets MULTI and FULL, erosion is most drastic, with decreases of moral costs of 19.8 cents and 20.5 cents, respectively, after market exposure. This decrease is significant across all market treatments (Wilcoxon signed-rank tests, 100 observations per market treatment, 81 in MPL, *p*-values of MPL=.108, SINGLE=.002, MULTI=.000, FULL=.000). Comparing the decrease between treatments, we do not find significant differences between MPL and SINGLE (Mann-Whitney U-test, 81/100 observations, *p*-value=.289). We find that multi-unit markets in turn show somewhat stronger erosion, as the decrease in MULTI compared to SINGLE is significant (MWU, 100 observations per treatment, *p*-value=.008), while the decrease between MULTI and FULL is similar (MWU, 100 observations per treatment, *p*-value=.799). This indicates that, surprisingly, erosion of morals does seem to persist outside of markets, especially so in multi-unit markets. Repetition seems to contribute to erosion as well, but its role appears to be less pronounced than that of multi-unit market exposure.

We further investigate how trading experience in our experimental markets affects traders' perceptions about the morality of other traders. For this, we elicited subjects beliefs about the median moral costs of canceling a donation in individual decision-making. Subjects were paid a bonus of € 1 if they correctly estimated the median participant's choices in the first multiple price list, for the first unit in the first part of the experiment, within their session.

In Figure 3.9, we report for each main treatment the mean difference between predicted and actual moral cost of the median trader in the left panel, together with the absolute prediction error in the right panel. Observing their fellow peers does not help participants to improve their estimate: the absolute error is not decreasing in the markets compared to MPL. Also, there do not appear to be strong differences between the market treatments.

Interestingly, the direction of the error changes systematically between treatments.[27] If anything, participants in MPL slightly overestimate how much the median participant values a donation to UNICEF. While there is a slight decrease in SINGLE, the multi-unit markets MULTI and FULL lead to systematic errors: participants strongly underestimate how much their participants care about donations for the measles vaccine.[28]

Summarizing, there is biased social learning in the sense that participants believe that their peers are more selfish than they truly are. Participants do not sufficiently take into account that other traders' behavior in the market is to a large extent shaped by market forces. This is also consistent with multi-unit markets complicating inference about the moral costs of traders who are less active. Market participants observe frequent trading, but do not comprehend that this may not reflect the preferences of an average participant outside of the market.

3.6 Discussion

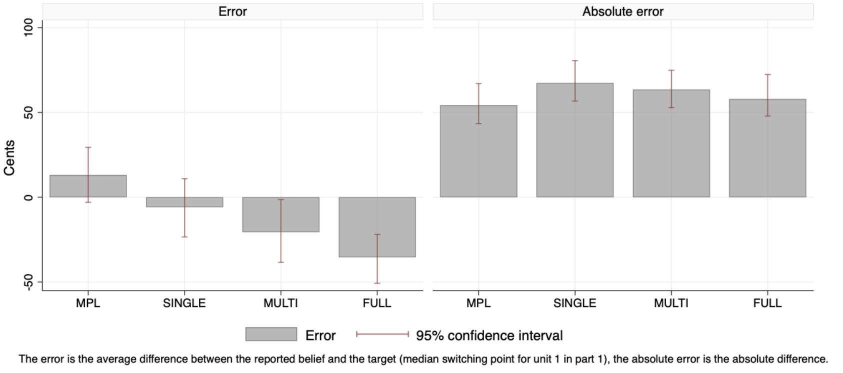
In this paper, we study market forces that can lead to a widespread erosion of morals and selfish market outcomes. As market power is reduced by allowing traders to take advantage of trading opportunities foregone by other traders, we show that aggregate outcomes as well as the behavior of a large share of market participants change dramatically.

Our paper provides conclusive evidence that markets can erode morals. We start by documenting that markets which retain pivotality of individual traders lead to a partial erosion of morals, as we observe more participants cancelling

[27] Regressing subjects' absolute errors on treatment fixed effect shows insignificant dummies (781 observations, clustering standard errors on matching group level; all p-values > .1 for SINGLE, MULTI and FULL). Regressing the error on treatment fixed effects shows differences in fixed effects, compared to the MPL baseline, for SINGLE (estimate of -19.0, p-value=.117), MULTI (estimate of -33.7, p-value=.009) and FULL (estimate of -48.5, p-value < .001).

[28] We find no evidence of excuse-driven belief reports, see Appendix Section 3.A.9.

Figure 3.9: Errors in beliefs about median subject's moral cost



Notes: Average error in estimating the session's median subject's moral cost for canceling one unit of donation of € 1.50 in part 1 of the experiment. The left panel displays the average difference between prediction and target, the right panel the absolute distance between prediction and target.

donations in markets than in individual decision-making. These results support Falk and Szech (2013)'s conclusion that single-unit markets partially erode morals.

We then expand the analysis of markets by introducing multi-unit trading and removing pivotality. These changes lead to a full erosion of morals. Participants appear to entirely disregard their moral concerns towards preventing negative externalities in these markets. Meanwhile, they are willing to forgo substantial amounts of money before and after markets in an individual decision-making task.

We further investigate the relative role played by market selection and the replacement logic in deteriorating market outcomes. We show that there is substantial heterogeneity in our traders' preferences for canceling donations, which leaves substantial scope for the selection effect to play a role. However, in our markets we find that less moral traders are hardly more active than more moral traders. Moreover, when we create homogeneous groups of traders who know that their preferences for the negative external effect are close to the median preference, we continue to see that all units in the market are traded. We conclude that the selection effect plays at most a minor role in our data. In contrast, and in agreement with the replacement logic, we find that (i) subjects become more active in trading when they are more convinced that their behavior does not have an impact on the aggregate outcome and (ii) subjects expect that their own behavior has less consequences for outcomes in FULL than in MULTI. Furthermore, our subjects' beliefs are hardly biased in a self-serving direction, instead they correctly predict that many participants are trading.

It is particularly interesting and worrisome to see the extent to which replacement thinking can deteriorate market outcomes. Absent pivotality, large shares of subjects engage in frenzied trade of units which cause large damages compared to the available gains from trade: 83% of subjects are willing to trade when they can share gains from trade of € 0.2, whereas only 9% of these same subjects are willing to cancel the first donation when each is paid € 0.2 in individual decision-making, averaged on part 1 and 3-data.

Strikingly, this frenzied trading contrasts with the prevailing norm. Even though we observe some deterioration in subjects' norms in markets compared to individual decision-making, we do not see that norms are further eroded when pivotality of trading in markets disappears. Still, norm compliance is completely eroded when subjects can be replaced when others refrain from trading. This led to widespread frustration among subjects, some of whom spontaneously wrote down their thoughts after the experiment. One subject commented: "The level of selfishness displayed on market 2 has almost made me cry during the experiment. Today, my faith in humanity has taken a giant blow".

Our findings suggest implications for policy. Because selection effects hardly play a role, efforts to restrain the more immoral players in a market may not affect market outcomes as long as these immoral players can be replaced by others. For instance, we think that it is doubtful that the recent dissolution of Purdue Pharma will solve the crisis in the opioids market. Instead, it may be more promising to pursue measures that restore or create pivotality in the market. One way to accomplish this would be to individually constrain traders in the quantities that they can trade. The treatment that implements this shows much less moral erosion. Further, because even the traders themselves normatively disapprove of the outcomes in the unrestrained markets, we expect that there may be support for measures that restore pivotality. As an alternative to individual capacity constraints, externalities can be mitigated by introducing taxes on the relevant behavior (Plott, 1983). On the other hand, aggregate quotas (i.e., *cap-and-trade* systems) can crowd-out moral behavior as they remove pivotality and make traders replaceable in the acquisition of the permits (Herweg and Schmidt, 2022).

The large erosion of morals we detect has also implications for our understanding of markets as aggregators of preferences. Using market outcomes to infer individuals' preferences regarding damages to third parties is complicated by key market design features. Simultaneously, obtaining a precise measurement of moral preferences in one environment may not be particularly useful to understand behavior in other environments. Participants can behave very selfishly and quite generously depending on specific features of the market structure. A poor understanding of the forces that apply in a given environment might fundamentally lead to a misrepresentation of individuals' preferences. In this sense, markets may not aggregate preferences in a straightforward way. Aside from concerning economists attempting to estimate preferences, this inference problem affects market participants themselves: Our subjects strongly underestimate how much their peers care about the donation to UNICEF after having participated in multi-unit markets. This brings up another potential danger of inference from market outcomes: We might be systematically underestimating by how much fellow members of our society would actually want to prevent the externalities they cause.

Appendix

3.A Additional analyses

In this appendix, we provide additional analyses of the data.

3.A.1 Predicting moral costs and moral competitive equilibria

In addition to the analysis presented in the main text, we can use individual decision-making data to predict outcomes in markets. For this, we proceed in two steps. First, we explain how we fit a moral cost curve to individual decision-making data. Second, we can use predicted moral costs to simulate market outcomes under the assumption that moral costs are not affected by moving to markets, to predict a moral competitive equilibrium.

3.A.1.1 Moral cost curves

We begin by fitting a moral cost curve to individual decision-making data. Denote $\Theta_i(q)$ the total moral costs subject i incurs for cancelling q units of donation. We use the moral costs we had elicited for $q \in \{1, 2, 3, 5, 7, 10, 15\}$ to estimate α_i, β_i in i's moral cost curve using OLS, where $\epsilon_{j,q}$ is an individual-unit error:

$$\Theta_i(q) = \alpha_i q + \beta_i q^2 + \epsilon_{j,q}$$

After estimating the above equation, we can use $\hat{\alpha}_i, \hat{\beta}_i$ to predict moral costs for any quantity $q \in \{1, 2, \dots, 15\}$ and each individual i. This predicts total moral costs $\hat{\Theta}_i(q)$, so the total moral costs for cancelling q units of donation. However, we often are interested in per-unit marginal moral costs $\theta_i(q)$, for unit q. These are the moral costs for cancelling an additional q-th unit of donation, after having cancelled $q - 1$ units earlier. So, we want to decompose predicted total moral costs $\hat{\Theta}_i(q)$ into a sum of q per-unit, marginal moral costs $\theta_i(j)$: $\hat{\Theta}_i(q) = \sum_{j=1}^{q} \hat{\theta}_i(j)$. To obtain per-unit moral costs $\hat{\theta}_i(q)$ for unit q, we use the predicted total moral costs for unit q, $\hat{\Theta}_i(q)$, and unit $q-1$, $\hat{\Theta}_i(q-1)$, and take their difference. By repeating this exercise for all $q \in \{1, 2, \dots, 15\}$, we obtain all per-unit moral costs for all units for all individuals.

3.A.1.2 Moral competitive equilibria

In the market treatments, we use the moral cost curves to predict market outcomes under the assumption that markets do not erode morals. That is, we use traders' estimated moral costs $\hat{\theta_i(q)}$ and predict how many units we would expect to be traded if $\hat{\theta_i(q)}$ is *not* affected by moving to our market setup, given the market rules of the treatment subjects are participating in. Since the literature finds repetition to be a force behind erosion (Bartling et al., 2023), we correct moral costs in each period by the average erosion we find in the corresponding repetition of the treatment MPL. We estimate moral costs each period, and rescale estimated moral costs for each market period with the average erosion found in MPL.

For this, starting from the first unit, we randomly draw a buyer *b*, with marginal moral costs to trade an additional unit of $\hat{\theta_b(q)}$, and a seller *s*, with marginal moral costs of $\hat{\theta_s(q')}$. If the sum of the two moral costs do not exceed the available gains from trade, given by the difference in induced values and costs, this pair of traders is designated to trade. Afterwards, we proceed to the next unit, and repeat the procedure. If the marginal moral costs of the pair (b, s) exceed the gains from trade, we attempt to find 200 times a pair for whom trading is feasible. In drawing random pairs, we keep track of the number of units previously traded, which may affect marginal moral costs or individual capacity constraints. At the point where no further pair can be found, the predicted quantity is the last unit which can be traded. Our predictions are the average outcome of 10,000 simulations, to account for differences in drawing random buyer-seller pairs.[29]

To be precise, we take *p* as the price agreed between one buyer and one seller. For unit Q to be traded, $v(Q)$ are induced values, $c(Q)$ induced costs, which are common across all traders at this unit. $\hat{\theta_i(q)}$ are estimated marginal moral costs for trader *i* to cancel a *q*-th unit of donation.

For heterogeneous moral competitive equilibria, we take the following steps, in each market period, where the simulation proceeds sequentially unit by unit:

1. We record individually traded quantities at every step, keeping track of which traders are constrained by capacity constraints (in SINGLE and MULTI) and what the predicted marginal moral costs to trade one more unit are for each trader *i*: $\hat{\theta_i(q)}$.
2. First, we verify whether any trade made in the experiment is consistent for both this buyer-seller pair we observe. That is, profits are larger than the

[29] In order to focus on the most relevant equilibria, we keep those trades observed in the experiment which are consistent with traders' moral costs in our simulations.

moral costs if $\theta_s(\hat{q}) \leq p - c(Q)$ for seller s and $\theta_s(\hat{q}) \leq v(Q) - p$ for buyer b. We keep all trades which are consistent for this buyer and seller. By doing so, we keep those equilibria which are closest to observed trading behavior.

3. Second, we verify whether additional units can be traded, beyond the number of units kept in step 2. For each additional unit, we draw at most 200 times a random pair of buyer b and seller s. In drawing random traders, we incorporate that our market picked one buyer and one seller randomly among those who submitted an equally favorable offer and among those who accepted the offer in question. We check whether for a candidate pair of traders, their moral costs allow them to trade one more unit, compared to the available gains from trade. That is, we verify that the sum of marginal moral cost is at most the difference in induced values and costs: $\theta_s(\hat{q}) + \theta_b(\hat{q}) \leq v(Q) - c(Q)$. If moral costs satisfy this equation, the two traders can agree on a price p at which they are both willing to trade. For the first randomly drawn pair of traders for whom the equation is satisfied, we designate these two to trade the Q-th unit, and continue to the $Q + 1$-th unit. We continue to simulate additional units, up to the point where for all 200 randomly drawn pairs of traders, marginal moral costs are prohibitively high: $\theta_s(\hat{q})+\theta_b(\hat{q}) > v(Q)-c(Q)$. At this point, trading stops, and the predicted quantity is the last unit which could be traded.
4. For each market and period, we repeat this procedure 10,000 times, as the order in which trader pairs are drawn potentially affects outcomes. Predictions shown are averages across all simulations and periods.

For homogeneous moral competitive equilibria, we adapt the above procedure only in the predicted marginal moral costs $\theta_i(\hat{q})$: for each market group, we use the median trader's moral costs for the first unit as the moral costs for all traders and all units. We thus remove both initial heterogeneity within a market and the decreasing marginal moral costs from estimated moral costs. We then perform the above procedure, which again yields a predicted quantity to be traded.

We call the outcome of this exercise the "competitive equilibrium with moral costs" or "moral competitive equilibrium". Note that this exercise is only possible in a design such as ours, where we observe participants both in individual decision-making and in a market environment. This exercise is meaningful, as we observe full trade in the first practice market period, which is incentivized but features no externalities, across all treatments featuring our market institution. This is consistent with the standard competitive equilibrium arising in the absence of negative externalities. Any decrease in trading volume can thus cleanly be interpreted as traders' concern for preventing the negative externality.

The benchmark of the moral competitive equilibrium allows us to: (i) disentangle whether observed market outcomes can be reconciled with the preferences of market participants or whether markets do erode morals; (ii) carry out counterfactual simulations to highlight the role of market selection. Regarding (i), we compare the degree of moral erosion by ranking the extent to which observed quantities exceed predicted quantities in the moral competitive equilibrium between treatments. This is of particular interest in treatment FULL: due to market selection, the least moral traders can determine quantities by themselves. If preferences are heterogeneous, or additionally if marginal moral costs are strongly decreasing for some of the traders, predicted quantities in the moral competitive equilibrium are higher in FULL than in MULTI or SINGLE. Under market selection, aggregate market outcomes in FULL appear to be more selfish than we would expect on the basis of homogeneous traders having the same median preferences. However, this does not imply that moral costs have eroded in markets, it just represents the fact that the traders least concerned with causing an externality are setting quantities. These traders might not be representative of the median trader. In the analysis, we will use each trader's estimated moral costs to verify whether her trading behavior is consistent with her stance outside of markets.

The possibility to run counterfactual simulations, in (ii), provides another important advantage of the moral competitive equilibrium. In predicting quantities, we use the estimated moral cost curve. By comparing outcomes in the heterogeneous to the homogeneous moral competitive equilibrium, we measure of how severe market selection is in this benchmark, or, how well markets are predicted to reflect the preferences of an average market participant.

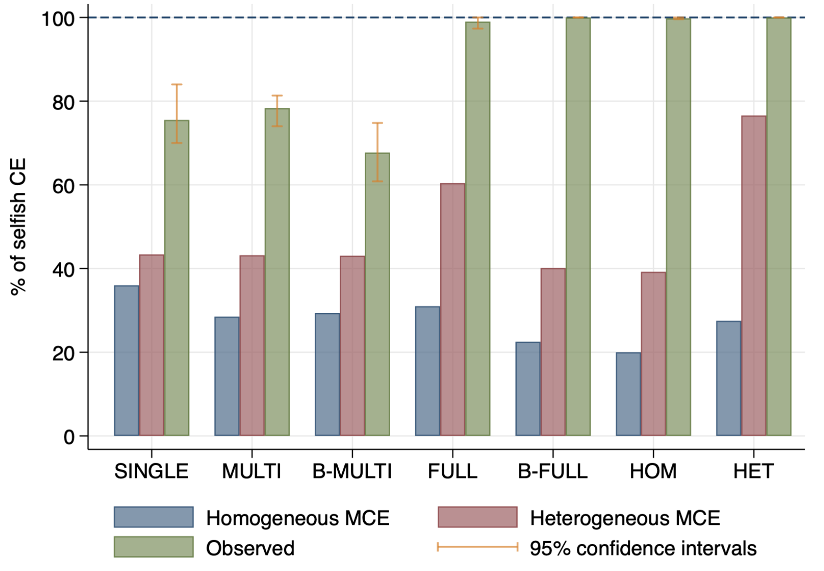
If market quantities exceed the predictions of the moral competitive equilibrium, markets do erode morals in the sense that traders care less about the externality they cause in a market than outside of a market.

In Figure 3.A.1, we present the results for this exercise. The first bar for each treatment, in grey, shows the predicted quantity in the competitive equilibrium with homogeneous and constant moral cost. For each market, we use the median trader's moral costs for the first unit to simulate how many units will be traded on average. Average quantities are between 28.5% and 36%. These differences between treatments are purely driven by initial heterogeneity of subjects, and are not related to underlying market features.[30] As it turns out, our subjects valued donations to UNICEF somewhat higher in FULL and MULTI than in SINGLE.

The second bar, in red, shows predicted quantities given the heterogeneous moral costs of market participants. These are higher quantities in all treatments than in the homogeneous moral competitive equilibrium. As expected, the differences are largest in FULL. The difference between the two equilibria can be attributed to market selection: the least moral traders in FULL are no longer con-

[30] Note that average moral costs between treatments are quite similar. However, for this exercise, we rely on distributions of the median, where we continue to see some variability between treatments.

Figure 3.A.1: Market outcomes and competitive equilibria



Notes: Average quantities relative to selfish competitive equilibrium for two moral competitive equilibrium ("MCE") benchmarks and observed quantities. MCE use participants' moral costs elicited in individual decision-making to predict market quantities. Heterogeneous MCE are based on actual moral costs, homogeneous MCE are based on the median trader's moral cost for the first unit within the matching group.

strained, thus they can expand the size of the market. This market force increases quantities by 29.4 percentage points. In MULTI and SINGLE participants' heterogeneity has a smaller impact on traded quantities. Whereas in SINGLE the increase is only 7.4 percentage points, this increases to 14.7 percentage points in MULTI.

The third bar, in green, shows observed quantities across the three treatments. We see that there is erosion of moral costs in all treatments. We observe partial erosion of morals in SINGLE and MULTI. In FULL, market outcomes are fully selfish. Compared to the competitive equilibrium with heterogeneous moral costs, quantities increase in SINGLE. They increase stronger in MULTI, and by even more in FULL.

Moral erosion in FULL is particularly strong, even though differences between observed and heterogeneous moral competitive equilibria might appear not too different between MULTI and FULL at first sight in Figure 3.A.1. Erosion is much stronger in FULL as additional units traded are causing larger negative externalities the more units have already been traded, relative to the potential gains from trade. This is the case as the induced gains from trade are decreasing at higher quantities, while damages stay constant. Below 40%, trading is efficient, as the damage to UNICEF is less than the associated payments to market participants. An increase from 40% to 60% leads to additional negative externalities of € 4.50, whereas traders receive € 1.80. The gains relative to damages to UNICEF decrease further, and an increase from 80% to 100% also yields damages of € 4.50, however traders only receive total payments of € 0.60. To quantify the size of the erosion, we summarize how many additional units compared to the moral competitive equilibrium benchmark are traded in each treatment in Table 3.A.1. We also show what damages to the donation traders are willing to accept for an additional payment of € 1 per additional unit that is traded. Damages, and the associated erosion of moral costs, are highest in FULL.

Table 3.A.1: The size of erosion in markets

	SINGLE	MULTI	FULL	B-MULTI	B-FULL	HOM	HET
Normalized units	4.8	5.3	5.8	3.7	9.0	9.1	3.5
Damage per € 1 gain per unit	3.2	3.2	4.9	2.9	4.6	4.5	6.0

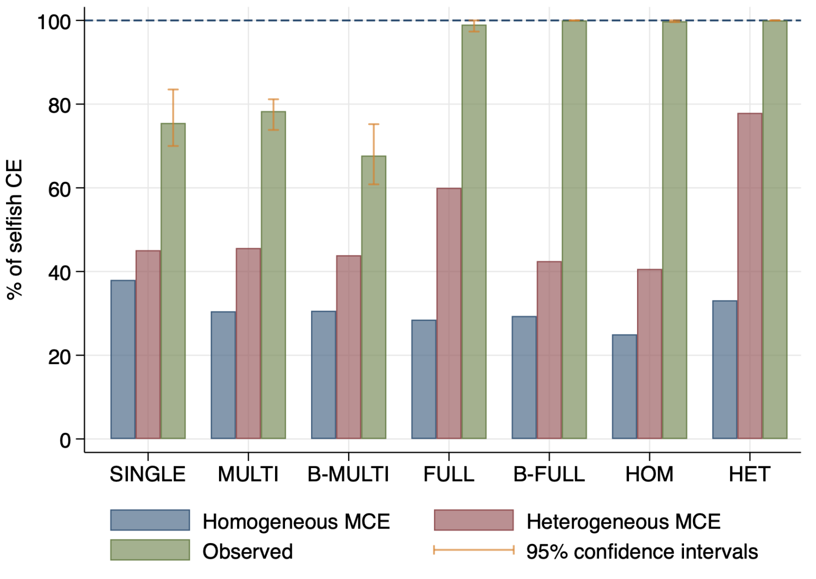
Notes: Number of units traded beyond heterogeneous moral competitive equilibrium as well as damages to UNICEF on average per additional unit, normalized across treatments. Damage per unit is fixed at € 1.50, gains from trade vary between € 0.20 and € 3.80.

The results of the moral competitive equilibrium exercise hinge on assumptions on the moral cost curve we use to fit individual decision-making data. In the following, we provide results on two exercises to test the robustness of the above conclusion. First, we use a linear moral cost curve to fit data from individual decision-making. This assumes marginal moral costs to be constant. Figure 3.A.2 presents the results of this exercise. The results are in line with the findings when allowing for non-linear moral cost curves.

Second, we repeat the procedure assuming that moral costs are halved when moving from individual decision-making to markets. This can account for the fact that decisions in markets always involve two participants and these trades generate payoff for two participants. Figure 3.A.3 presents the results. We continue to observe erosion compared to this benchmark. In particular, predicted average quantities in the moral competitive equilibria with halved moral costs are 11.1 units, an increase from the 9.1 units in the baseline simulation. This quantity still falls substantially short of the observed traded quantities of 14.9 units.[31]

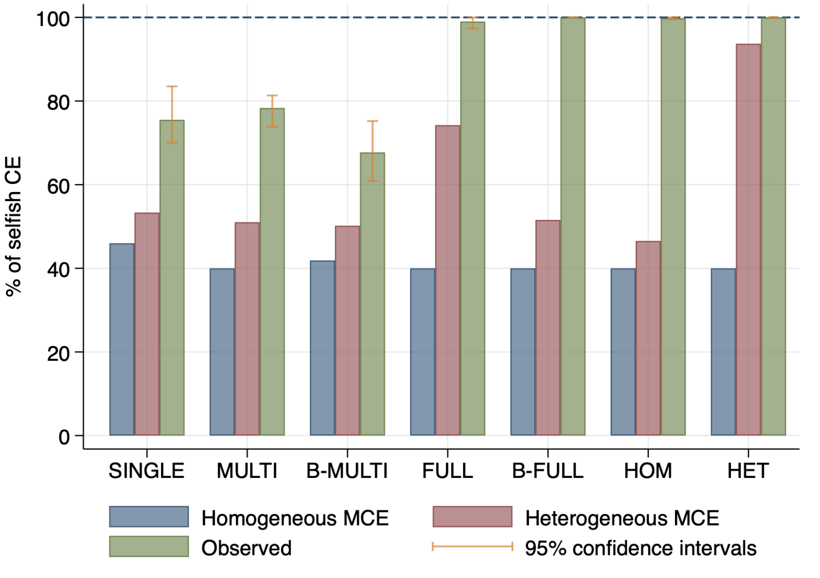
[31]We pre-registered a second method of evaluating moral erosion, that was based on the information conveyed in traders' offers. The results of this analysis are qualitatively similar to the results reported in this section. Details will be sent on request.

Figure 3.A.2: Market outcomes and competitive equilibria: Linear moral costs



Notes: Average quantities relative to selfish competitive equilibrium for two moral competitive equilibrium ("MCE") benchmarks and observed quantities. MCE use participants' moral costs elicited in individual decision-making to predict market quantities, using linear cost curves to estimate moral costs. Heterogeneous MCE are based on actual moral costs, homogeneous MCE are based on the median trader's moral cost for the first unit within the matching group.

Figure 3.A.3: Market outcomes and competitive equilibria: Halved moral costs



Notes: Average quantities relative to selfish competitive equilibrium for two moral competitive equilibrium (“MCE”) benchmarks and observed quantities. MCE use participants’ moral costs elicited in individual decision-making to predict market quantities, which are divided by two. Heterogeneous MCE are based on actual moral costs, homogeneous MCE are based on the median trader’s moral cost for the first unit within the matching group.

3.A.2 Balancing

In Table 3.A.2, we show covariate balance across treatments. Of particular interest is the comparison between HOM and HET. Apart from the intended manipulation of heterogeneity, these treatments are balanced. Note that the data for the treatments MPL, SINGLE, MULTI and FULL was collected first at the CREED laboratory in Amsterdam, in September and October 2019. The data for the remaining treatments was collected October 2021 to January 2022. Sessions were ran both at the CREED laboratory in Amsterdam as well as at the CentERlab of Tilburg University.

Table 3.A.2: Balancing table

	Age	% women	% international	Switching point part 1	Risk
MPL	21.6	42	86	9.9	3.5
SINGLE	20.6	46	80	9.5	3.6
MULTI	20.7	52	81	10.3	3.5
FULL	21.7	45	76	9.8	3.6
B-MULTI	21.1	48	74	10.4	3.6
B-FULL	21.3	51	78	11.1	3.4
HOM	21.5	56	75	10.9	3.6
HET	21.5	63	76	11.5	3.4
SPEC	21.6	59	80	-	-
HOM vs. HET (p-values)	.973	.424	.856	.520	.455
Kruskal-Wallis (p-values)	.248	.453	.964	.219	.961

Notes: Average characteristic by treatment. Switching point part 1 and Risk were not elicited for the SPEC treatment. In the second-last row we report p-values of a t-test comparing HOM with HET, 80 observations per treatment. In the last row we report p-values of a Kruskal-Wallis test, comparing equality across all treatments.

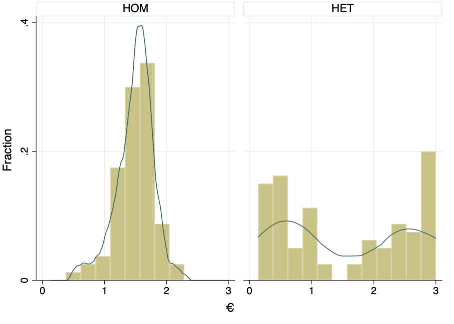
In Figure 3.A.4, we show histograms of the average marginal moral costs for HOM and HET, using part 1 data. We do observe that generating homogeneous groups in HOM and heterogeneous groups in HET was successful.

3.A.3 Are marginal moral costs decreasing?

In Figure 3.A.5, we provide evidence of decreasing marginal moral costs. We plot the average valuation implied by choice data in individual decision-making, averaged on the unit level. At larger stakes, subjects need to be paid less, averaged per unit, such that they are willing to cancel a donation. The effect is quite strong: For the first unit, subjects on average reported moral costs of € 1.68, this decreases to € 1.27 for the fifteenth unit.

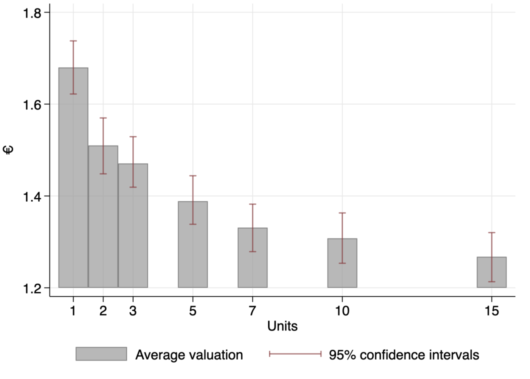
The decreasing pattern of marginal moral costs is statistically significant. In an OLS regression with subjects fixed effects, we allow for changes in marginal moral costs as a function of the size of the donation (Unit). The results, presented in

Figure 3.A.4: Histograms of average marginal moral costs



Notes: Histograms and kernel densitites of average moral costs based on the elicited valuations, with a value of € 1.5 each.

Figure 3.A.5: Decreasing marginal moral costs



Notes: Average moral costs based on the elicited valuations, with a value of € 1.5 each.

Table 3.A.3, show that this variable is empirically important. The estimate on Unit is negative and significant.

Table 3.A.3: Evidence for decreasing marginal moral costs

	Marginal moral costs
Unit	-.0248***
	(.00166)
Constant	1.575***
	(.0101)
Observations	5,467
# of subjects	781
Subject FE	Yes
Adjusted R^2	0.814

Note: Dependent variable is average per-unit valuation elicited in individual decision-making, in Euros. Unit captures the unit number, from 1, 2, 3, 5, 7, 10, 15. Subject fixed effects control for level differences in valuations across subjects. Standard errors clustered on matching group level in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.A.4 Robustness of treatment effects

In Table 3.A.4, we regress the quantities on treatment indicators to verify robustness of our main results. Each market outcome provides one observation. Quantities in FULL differ significantly both in (1) and when including controls in (2).

Table 3.A.4: Treatment effects

	(1)	(2)
	Relative quantity traded	
(1 if MULTI)	2.833	4.040
	(3.980)	(3.256)
(1 if FULL)	23.500***	23.315***
	(3.604)	(3.506)
(1 if Period=2)		-4.667**
		(1.795)
(1 if Period=3)		-7.111***
		(2.020)
(1 if Period=4)		-4.889**
		(2.153)
Mean moral cost		0.040
		(0.080)
Median moral cost		-0.098
		(0.065)
Minimum moral cost		0.108
		(0.100)
Mean risk measure		3.834
		(3.842)
Constant	75.500***	69.044***
	(3.512)	(12.992)
Observations	120	120
Adjusted R^2	0.508	0.555

Note: Dependent variable is observed quantity relative to selfish competitive equilibrium. Mean, median and minimum moral cost are the mean, median and minimum of marginal moral costs, averaged on a subject level, in part 1 within a matching group. Mean risk is the average chosen lottery in the risk task per matching group. Standard errors clustered on matching group level in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.A.5 Norms

We elicited norms using the method introduced by Krupka and Weber (2013). We described seven different scenarios in the experiment, where subjects evaluated whether they deemed the behavior as "socially appropriate" and "consistent with moral or proper social behavior" on a 4-point scale from "very socially inappropriate", to "somewhat socially (in)appropriate" and "very socially appropriate". In particular, we described four scenarios involving individual decision-making as well as three scenarios in an experimental market. In Appendix B, we reproduce the full instructions and interface.

Scenarios 1 to 4 mirror the individual decision-making task in the experiment, where Individual 1 makes the following choices (as a reminder, 4 doses cost approximately € 1.5.):

1. "1 chooses to receive 1 Euro instead of making a donation of 4 doses of measles vaccine to UNICEF."
2. "1 chooses to receive 2 Euro instead of making a donation of 4 doses of measles vaccine to UNICEF."
3. "1 chooses to receive 3 Euro instead of making a donation of 12 doses of measles vaccine to UNICEF."
4. "1 chooses to receive 6 Euro instead of making a donation of 12 doses of measles vaccine to UNICEF."

Three scenarios with Individual 2 mirror the experimental markets, where trading canceled a donation of four doses of measles vaccine.

5. "2 decides to accept an offer which allows him to earn 1 Euro."
6. "2 decides to accept an offer which allows him to earn 2 Euro."
7. "2 makes an offer in the market. If a trade is concluded based on this offer, 2 would earn 1 Euro."

In addition to the data presented in the main text, below are histograms of the responses of subjects for all scenarios across the four treatments.

Figure 3.A.6: Norms in individual decision-making

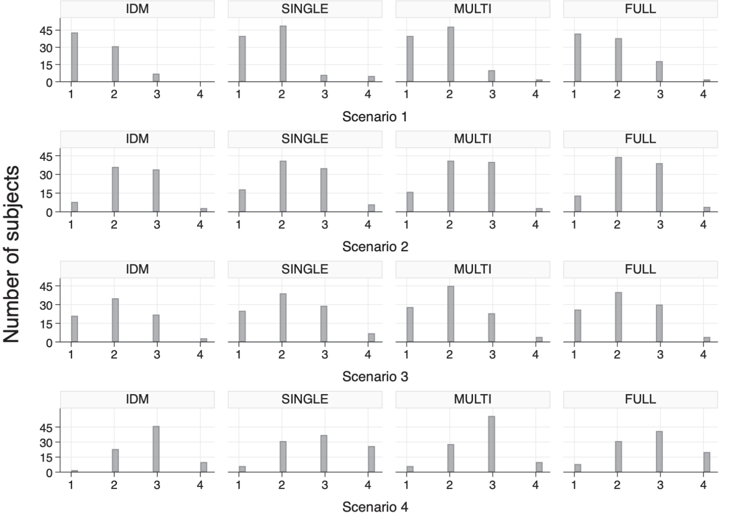
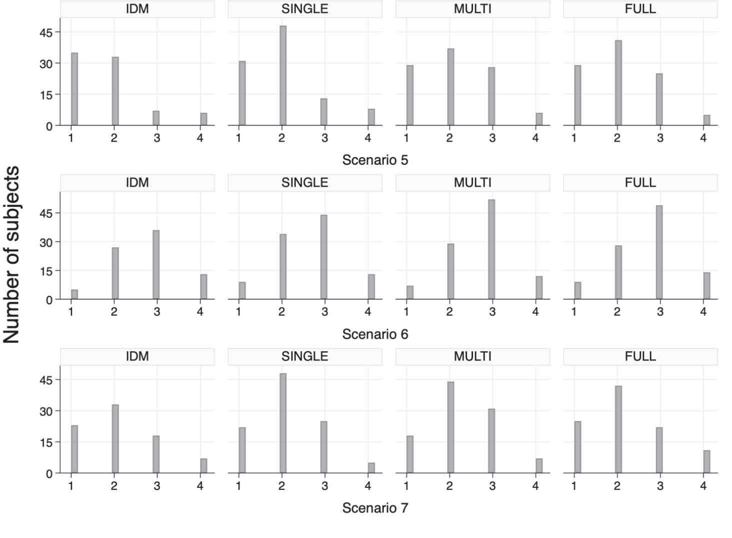


Figure 3.A.7: Norms in markets

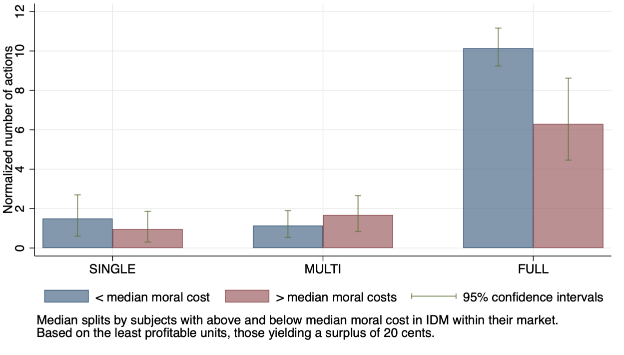


3.A.6 The replacement logic: Intensive margins and deontological subjects

In the main text, we show that a large share of subjects engage in trading when these units only yield € 0.20 for two market participants, in exchange for causing a damage of € 1.50 to UNICEF.

In Figure 3.A.8, we also show the intensive margin of this phenomenon: how many offers and acceptances do we observe from traders? To normalize the number of actions per trader across treatments, so to account for the smaller total market size in SINGLE, we multiply the observed number of actions in SINGLE by 3. We again observe that erosion due to replacement logic appears to matter most. Frequent trading of both types of traders is observed in FULL, with 8.2 actions per trader observed on average, whereas in SINGLE and MULTI only 1.2 and 1.4 actions per traders are observed on average.

Figure 3.A.8: Number of acceptances and offers at the least profitable units



Notes: Average number of offer submissions or acceptances per trader at the final units, which yield gains from trade of € 0.20 in exchange for an externality of € 1.50. Median splits based on predicted moral costs within matching group.

The difference between the above and below median group is only significant for FULL (MWU, 10 observations per group, p-value=.003). Table 3.A.5 repeats the analysis using a regression. Model (1) repeats the analysis from the main text, regressing a dummy equal one if a subject was active for the last units in the markets on treatment dummies, a dummy equal one if a subject had above median moral costs and their interactions. We confirm that participants are more active in FULL. However, the interaction for above median participants in FULL is not significant. Model (2) uses the dependent variable from the appendix analysis, counting the number of acceptances and offers per participant. This analysis is robust to this specification, where above median participants are significantly less frequently trading.

Table 3.A.5: Replacement logic or market selection?

	(1) Active	(2) Nr. actions
MULTI	0.140	-0.160
	(0.090)	(0.660)
FULL	0.740***	9.160***
	(0.069)	(0.789)
Above median	-0.040	-0.300
	(0.049)	(0.572)
MULTI × Above median	0.040	0.680
	(0.076)	(0.897)
FULL × Above median	-0.140	-4.340***
	(0.097)	(1.411)
Constant	0.180***	1.380**
	(0.061)	(0.530)
Observations	300	300

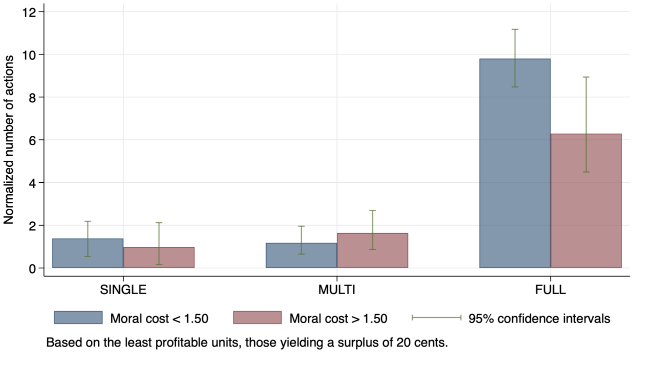
Note: Dependent variable is a dummy equal to one if a subject made or accepted an offer at least once for the units with gains from trade of € 0.20 in (1), or how many offers or acceptances a subject made for these units in (2). Above median is a dummy equal one if a subject has above median moral costs. MULTI (FULL) is a dummy equal to one if the choice occurred in treatment MULTI (FULL), with the omitted category SINGLE. Standard errors, clustered on matching group level, are presented in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In this analysis, part of the traders with above median moral costs are potentially consequentialistic subjects, who can use the replacement logic: given that the donation will in any case not go through, it may be legitimate to trade.

Interestingly, this activity also carries over to subjects who likely do not use consequentialistic reasoning. In the first part of the experiment, we have a subset of participants who report moral costs above the corresponding value of the donation. This set of participants decided to forgo a higher payment in order not to cancel the donation, which they could have instead donated to UNICEF themselves. Approximately 34% of subjects report such preferences.[32]

In Figure 3.A.9, we show what share of traders are active at the least profitable units in the markets, splitting them into subjects with moral costs below and above € 1.50. While in SINGLE and MULTI, these subjects rarely are active, they are very active in FULL. For these subjects, it appears to be the case that their morals were eroded. This is the case as for these subjects, the replacement logic is hardly a justification to trade.

Figure 3.A.9: Replacement logic in non-consequentialistic subjects



Notes: Average number of offer submissions or acceptances per trader at the final units, which yield gains from trade of € 0.20 in exchange for an externality of € 1.50. Splits based on average predicted moral costs above and below € 1.50, the cost of the donation.

In Table 3.A.6, we show correlates of an indicator capturing whether a subject was active at the least profitable units, those with available gains from trade of € 0.20. All statements in quotation marks are statements from the questionnaire, rated from 1 to 7 whether subjects agreed with a given statement. What appears to matter are (1) initial moral costs of subjects, (2) leaning politically to the right,

[32] Note that this is unlikely to be driven by misunderstanding: regressing subjects' moral costs, or equivalently a dummy equal one if they report moral costs above € 1.50, on the number of attempts this subject required to complete the practice questions for part 1 shows an insignificant correlation. Results are also similar when splitting subjects at even higher moral costs, such as at € 1.70 or € 2, which implies transaction costs are also unlikely to explain these results.

(3) using a statement modeled to fit the replacement logic: "I decided to trade in market 2 because I realized the units I traded would have been traded by others in any case.". In (2), we report average marginal effects of the logistic regression in (1), as well as OLS estimates in (3).

Table 3.A.6: Who uses the replacement logic?

	(1)	(2)	(3)
Dep. variable:		(1 if active at last units)	
Change in moral cost from part 1 to 3	0.267	0.036	0.033
	(0.531)	(0.071)	(0.069)
Moral cost in part 1	-0.829***	-0.111***	-0.094**
	(0.279)	(0.037)	(0.035)
(1 if male)	-0.914***	-0.123***	-0.111**
	(0.311)	(0.040)	(0.042)
(1 if international student)	0.168	0.023	0.016
	(0.378)	(0.050)	(0.053)
Risk measure	-0.013	-0.002	-0.003
	(0.098)	(0.013)	(0.015)
Belief about median subject's moral cost	0.068*	0.009*	0.009
	(0.040)	(0.005)	(0.006)
Norm in ind. dec.-making	0.219	0.029	0.024
	(0.306)	(0.041)	(0.044)
Norm in market	-0.175	-0.023	-0.022
	(0.218)	(0.030)	(0.033)
"I believe the donations for measles vaccines to UNICEF are helpful."	0.221	0.030	0.030
	(0.186)	(0.025)	(0.019)
"I believe measles vaccines save lifes."	-0.021	-0.003	-0.001
	(0.157)	(0.021)	(0.021)
"When making a moral decision, I try to always follow a rule, instead of evaluating the consequences of each particular option every time."	-0.177	-0.024	-0.024
	(0.111)	(0.015)	(0.016)
"When deciding on whether I should trade in market 2, I studied at what profits other traders were willing to trade."	0.065	0.009	-0.000
	(0.136)	(0.018)	(0.018)
"I decided to trade in market 2 because I realized the units I traded would have been traded by others in any case."	0.338***	0.045***	0.047***
	(0.099)	(0.012)	(0.012)
"How competitive are you?" (1 not competitive, 7 very; Buser, Niederle, and Oosterbeek (2020))	-0.052	-0.007	-0.008
	(0.148)	(0.020)	(0.021)
"Where do you see yourself in the left-right political spectrum?" (1 left, 7 right)	0.325**	0.044***	0.039**
	(0.128)	(0.016)	(0.015)
(1 if MULTI)	0.987**	0.142**	0.146**
	(0.435)	(0.061)	(0.070)
(1 if FULL)	4.434***	0.694***	0.686***
	(0.511)	(0.047)	(0.053)
Constant	-5.227***		-0.195
	(1.792)		(0.219)
Study fixed effect	yes	yes	yes
Observations	273	273	278
Estimation	Logit	Avg. ME	OLS

Note: Dependent variable is a dummy equal one if a subject submitted or accepted an offer at least once for units with gains from trade of € 0.20. Change in moral cost is defined as moral costs in part 3 less moral costs in part 1 in Euro. Standard errors clustered on matching group level in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.A.7 Non-incentivized belief measures

In the main text, we analyze the incentivized belief measure. In this subsection, we provide some additional analysis and replicate the main analysis with our second belief measure, which is non-incentivized.

The non-incentivized belief measure consisted of three questions, each eliciting the three potential scenarios for the trade of the upcoming unit. The questions were:

1. What is the probability that whatever you do, the next unit will be traded?
2. What is the probability that whether or not the next unit is traded depends on your behavior?
3. What is the probability that whatever you do, the next unit will not be traded?

Participants received a payment of 300 cents irrespective of the correctness of their reports. Traders are pivotal only in the second scenario, thus the probability of being pivotal is measured by the likelihood ascribed to scenario 2. Sometimes the analysis requires the probability of being replaced. This is the probability of not being pivotal conditional on trade happening (i.e., the chance to have at least one other trader active on the own side of the market, conditional on the other side being active), and formally calculated by the probability of scenario 1 divided by the sum of the probabilities of scenarios 1 and 2.

For the incentivized question, we asked participants to report the following: "How many participants other than you will attempt to trade this unit?". When correctly reporting the number of active traders, they received a payment of 150 cents.

In the main text, we focus on the incentivized measure as this correlates more strongly with the underlying true values. The Spearman correlation coefficients between the predicted and actual number of active traders averaged on a subject-level is 0.422. The same correlation between traders belief to be pivotal and the realized event to have actually been pivotal is 0.181. When bootstrapping the difference in test statistics this difference is significant with a p-value of .021 (160 observations, 1000 repetitions). The same pattern arises when calculating correlations treating each report as an independent observation. The correlation coefficient for the incentivized measure is 0.239, for the non-incentivized measure it is 0.120. The difference is significant with a p-value of .001 (1780 observations, 1000 repetitions).

First, Table 3.A.7 presents Spearman correlation coefficients of the two measurements of the beliefs. All data are based on the four market periods with externalities. The first row uses individual report-level data, the second row presents correlations between averages on a participant level. Both in B-MULTI and B-FULL there are no detectable correlations. We do find the expected correlation for SPEC, which suggests that eliciting beliefs while simultaneously trading in markets inhibited this correlation. We think that we asked too much of our subjects in B-MULTI and B-FULL, which made them focus on the incentivized questions and pay less attention to the unincentivized ones.

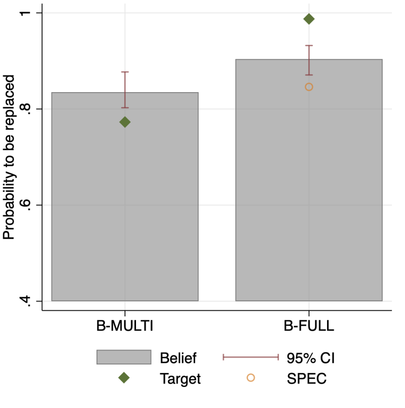
Table 3.A.7: Correlation between belief measures

	SPEC	B-MULTI	B-FULL
All data	-.278 (.000)	-.016 (.718)	-.003 (.910)
Participant averages	-.447 (.003)	-.131 (.247)	.011 (.923)

Notes: Spearman correlation coefficients between the incentivized and non-incentivized belief measure, *p*-values in parentheses.

In Figure 3.A.10, we report the non-incentivized belief of the probability to be replaced across different treatments. These are calculated as the belief to have at least one other trader active on the own side of the market, conditional on the other side being active. Conclusions are in line with the analysis in the main text. The beliefs are significantly different between B-MULTI and B-FULL (MWU, 8 observations per treatment, *p*-value=.0209), while they do not differ between B-FULL and SPEC (MWU, 8 observations for B-FULL and 41 for SPEC, *p*-value=.704).

Figure 3.A.10: Unincentivized beliefs about own probability to be replaced



Notes: Probability to be replaced (grey bar), actual probability to be replaced (green diamond) and belief of spectators (orange circle).

Table 3.A.8 presents the results of an analysis in which we regress a dummy indicator of being active in the market on the unincentivized belief to be replaced. Surprisingly, participants who believe to be more replaceable are less likely to

trade. Possibly, some of our participants may have become confused about the questions that we were asking and may have thought that if they planned not to trade the subsequent unit, it will not be traded even if they were allowed to trade.

Table 3.A.8: Beliefs and activity

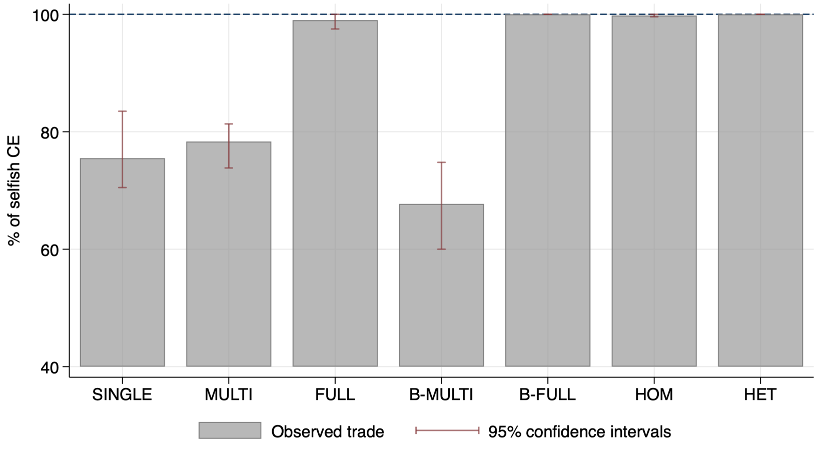
	(1) B-MULTI	(2) B-FULL
Prob. to be replaced	-0.141*	-0.687***
	(0.066)	(0.092)
Average moral cost	-0.001	-0.001***
	(0.000)	(0.000)
Period	-0.031	-0.022
	(0.030)	(0.015)
Constant	0.589***	1.485***
	(0.073)	(0.097)
Unit FE	yes	yes
Observations	466	1279
Adjusted R^2	0.025	0.124

Note: Dependent variable is a dummy equal to one if a subject submitted or accepted an offer at least once for units with gains from trade of € 0.20. Average moral costs are the average moral costs for a participant, based on average estimated per-unit moral costs based on part 1 individual decision-making. Standard errors clustered on matching group level are presented in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

3.A.8 Main analysis in HOM, HET, B-MULTI and B-FULL

In Figure 3.A.11 and Table 3.A.9, we report market outcomes across all treatments. Quantities across all treatments using FULL market rules (FULL, B-FULL, HOM and HET) are all fully selfish and statistically indistinguishable. Quantities in B-MULTI are slightly below those in MULTI, suggesting that additionally eliciting beliefs in this treatment leads to slightly more moral behavior.

Figure 3.A.11: Market outcomes



Notes: Average quantities relative to selfish competitive equilibrium. Trading units below 40% is efficient (gains from trade exceeds the externality). Compared to the negative externality of € 1.50 per unit, each unit between 40% and 60% yields gains from trade of € 0.60, each unit between 60% and 80% yields € 0.40 and each unit between 80% and 100% yields € 0.20.

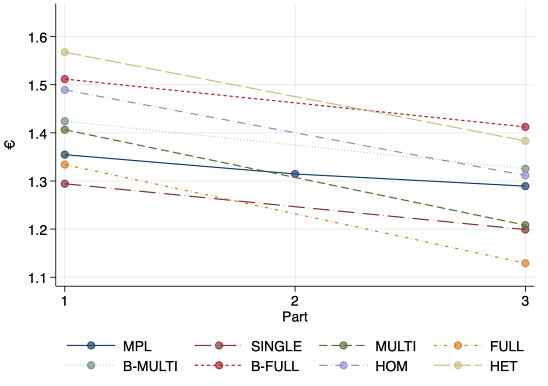
Table 3.A.9: Treatment effects

		SINGLE	MULTI	FULL	B-MULTI	B-FULL	HOM	HET
Quantity in %		75.5	78.3	99	67.7	100	99.8	100
p-values	vs. SINGLE	-	.378	.0005	.0899	.0006	.0009	.0006
	vs. MULTI	-	-	.0001	.0308	.0002	.0002	.0002
	vs. FULL	-	-	-	.0002	.1931	.6318	.1931
	vs. B-MULTI	-	-	-	-	.0003	.0004	.0003
	vs. B-FULL	-	-	-	-	. -	.3173	1.000
	vs. HOM	-	-	-	-	. -	-	.3173

Notes: Average quantities relative to selfish competitive equilibrium. Mann-Whitney U-tests, on matching group averages, 10 observations per treatment.

Figure 3.A.12 reports data on erosion across parts across all treatments, complementing Figure 3.8.

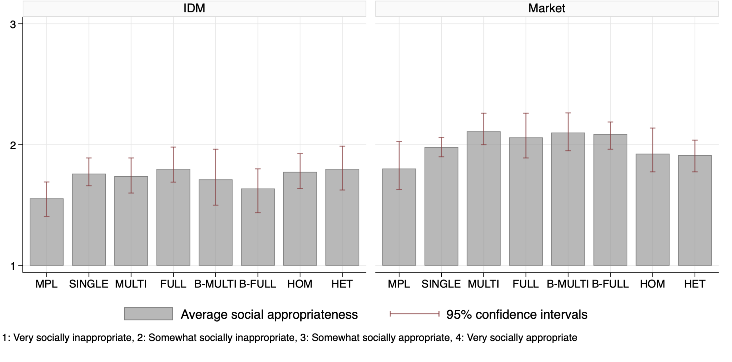
Figure 3.A.12: Persistence of erosion



Notes: Average per-unit valuations in individual decision-making, for € 1.50 donations, by part. In part 2, only MPL employs individual decision-making.

Figure 3.A.13 reports data on norms across all treatments, complementing Figure 3.4.

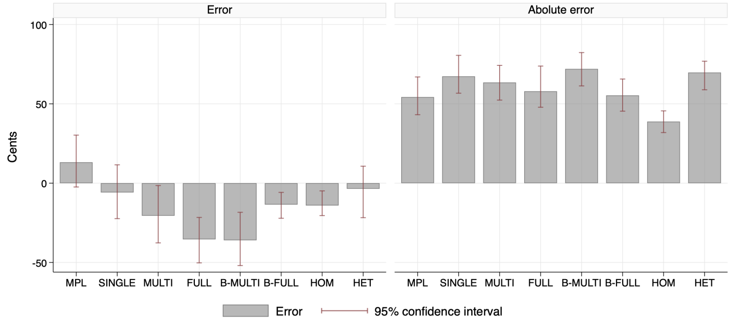
Figure 3.A.13: Norms in individual decision-making and in markets



Notes: Average norm in response to cancelling one donation of € 1.50 when paid € 1 in individual decision-making (left panel) and in the experimental market (right panel). A rating of 2 corresponds to "somewhat socially inappropriate".

Figure 3.A.14 reports data on beliefs across all treatments, complementing Figure 3.9.

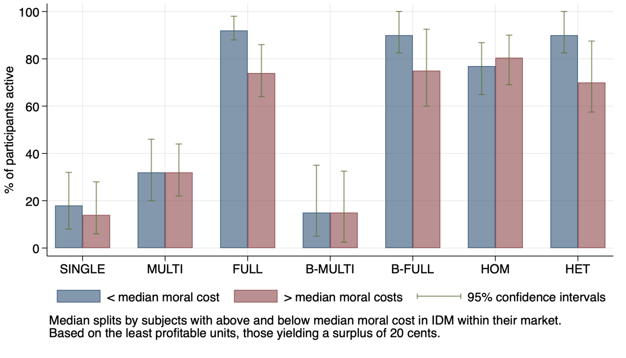
Figure 3.A.14: Errors in beliefs about median subject's moral cost



Notes: Average error in estimating the session's median subject's moral cost for cancelling one unit of donation of € 1.50 in part 1 of the experiment. In grey the absolute distance between prediction and target, in red the difference between prediction and target.

Figure 3.A.15 reports data on norms across all treatments, complementing Figure 3.5.

Figure 3.A.15: Share of traders active at the least profitable units



Notes: Share of traders who submit or accept an offer at the final units, which yield gains from trade of € 0.20 in exchange for an externality of € 1.50. Median splits based on predicted moral costs within matching group.

3.A.9 Beliefs as excuses

In the main text, we document that participants hold biased beliefs about others' morals outside markets. One potential concern may be that subjects report beliefs in order to provide an excuse for their own selfish behavior in the markets. These excuses may be needed most in treatments MULTI and FULL, where we also find that subjects are most pessimistic about their peers' morals.

To verify whether this might be driving our results, we report average beliefs, by treatment, for those traders who likely need the excuse the most: those traders who we observe to be active at the least profitable units, those yielding profits of € 0.20. In Table 3.A.10, we see that there are no meaningful patterns that would support such excuse-driven reporting of beliefs. Similarly, regressing beliefs (or errors in beliefs) on a dummy variable equal to one if a trader was active at the least profitable units, with treatment fixed effects, yields insignificant, and for that matter positive, coefficients on the dummy variable capturing the need for an excuse. Therefore, it is unlikely that our findings on beliefs can be explained by participants' need to provide justification for their own selfish behavior.

Table 3.A.10: Average beliefs for (in)active traders at last units

	SINGLE	MULTI	FULL
Inactive	10.20	10.07	8.24
Active	10.25	10.56	8.84

Notes: Average belief of median participant's switching point in the multiple price list for the first unit (11 corresponds to indifference between payments to self and UNICEF). Split by whether the subject was active at the final units, those with gains from trade of € 0.20.

3.A.10 Norms as excuses

As for the beliefs, it may be the case that subjects report perceived norms to excuse their behavior in part 2 of the experiment. We report the same analysis for the elicited norms in Table 3.A.11.

Table 3.A.11: Average norm report for (in)active traders at last units

	SINGLE		MULTI		FULL	
Norm in	IDM	Market	IDM	Market	IDM	Market
Inactive	1.76	1.99	1.68	2.01	1.65	2.12
Active	1.75	1.94	1.88	2.31	1.83	2.05

Notes: Average norm report for cancelling donations of € 1.50 in return for a payment of € 1 in individual decision-making (MPL) or with in an experimental market (Market). 2 corresponds to "somewhat socially inappropriate". Split by whether the subject was active at the final units, those with gains from trade of € 0.20.

There are no systematic patterns which suggest that norms are reported self-servingly. This is confirmed by regression evidence, similar to the analysis for beliefs. Regressing the reported norm in the experimental market on a dummy variable equal to one if a trader was active at the least profitable units, with treatment fixed effects, yields insignificant coefficients on the dummy variable (p-value=.469) and on the treatment fixed effects (p-value=.222 for MULTI, p-value=.923 for FULL).

3.A.11 Morals in a double auction with a private schedule

Our experimental markets use a two-sided posted offer institution and a common schedule. In this section, we compare outcomes of our treatment MULTI to an additional control treatment PRIV, in which we implement a standard double auction with a private schedule.

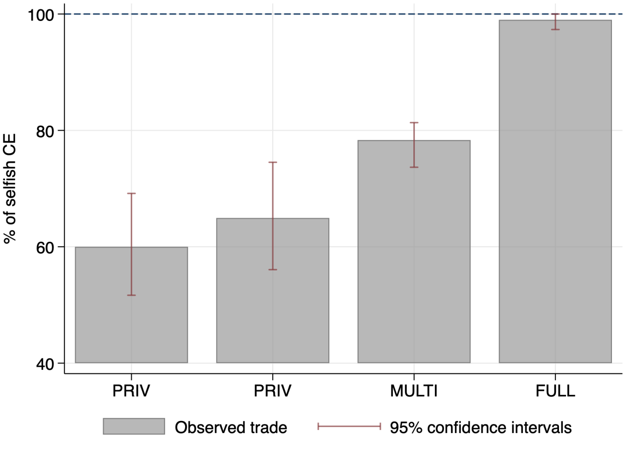
In the double auction, buyers and sellers submitted offers simultaneously and units were traded if participants agreed on a price (if a bid exceeded an ask). Each market period ran for eight minutes. In this market, traders face a private cost/value schedule. For comparability, we mapped the schedule we use in our multi-unit markets to a private schedule. For this, every trader received three randomly drawn values or costs from the common cost or value schedule, redrawn every market period. This way aggregate costs and values were kept identical, and each participant was restricted to trade at most three units. In agreement with standard double auction procedures, participants did not know other traders' costs or values. This is a further difference with our two-sided posted offer market.

All other elements of our experiment were kept identical. In a first market period, participants could trade in a market without externalities. In this market, quantities were slightly below the competitive equilibrium quantity, at 13.89 units on average. Before the start of the last four periods, participants learned that per trade donations of € 1.50 to UNICEF were cancelled.

In Figure 3.A.16, we plot the traded amounts in PRIV, compared to MULTI and FULL. Quantities in PRIV are the first two bars. The first bar normalizes quantities relative to the selfish competitive equilibrium, the second bar relative to the quantity traded in the market period without externalities. For MULTI and FULL, presented in the in the third and fourth bar, these normalizations are identical.

Quantities in PRIV are below quantities in MULTI and FULL. Comparing the normalization relative to selfish competitive equilibrium, quantities in PRIV are significantly different from MULTI (MWU, 8 observations in PRIV and 10 in MULTI, p-value=.003) and from FULL (MWU, 8 observations in PRIV and 10 in MULTI, p-value<.001). However, they are still consistent with a partial erosion of morals, as trade continued beyond 40%, which implies that units where the damage to UNICEF exceeds the associated gains from trade have been traded.

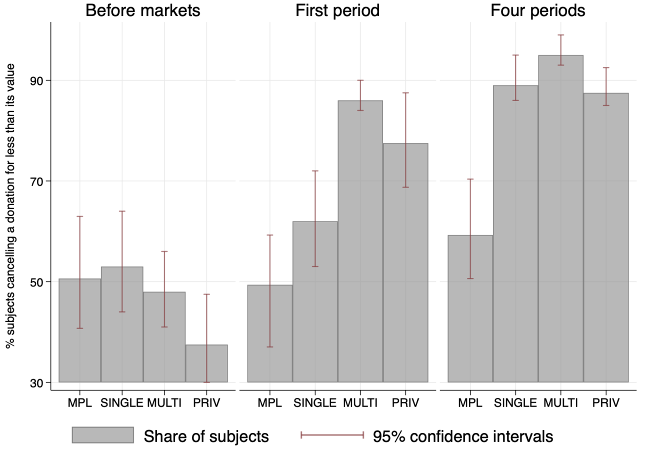
Figure 3.A.16: Market outcomes in PRIV



Notes: Average quantities traded relative to selfish competitive equilibrium (first, third and fourth bar) and relative to quantities traded in markets without externalities (second bar).

An important question is to which extent morals are eroded in the standard double auction, and how this compares to the two-sided posted offer market. To investigate this question, we repeat the analysis of how many traders trade a unit for a payment of less than € 1.50. In Figure 3.A.17, we present the results, including the treatment MULTI for comparison. In this comparison we observe erosion in PRIV, at least as large as the erosion in SINGLE, and slightly below the erosion detected in MULTI.

Figure 3.A.17: Cancellation of donations between environments and treatments



Notes: Share of participants who cancelled a donation for at most its value (€ 1.50) in individual decision-making and in implemented trades in the market. The left panel shows shows cancellation rates in part 1 of the experiment and the middle panel plots cancellation rates in the first period of part 2. The right panel displays the share of participants who, in the four periods of part 2, at least once cancelled a donation.

In Table 3.A.12, we repeat the analysis from the main text comparing the three market treatments SINGLE, MULTI and PRIV. When we pool the data, we see that at the start there is similar erosion in PRIV as in MULTI, while there is more erosion in PRIV than in SINGLE. Then over time the difference with SINGLE disappears.

Table 3.A.12: Erosion in markets and through repetition

	(1) SINGLE	(2) MULTI	(3) PRIV	(4)	(5)	(6)
				SINGLE, MULTI & PRIV		
				Period 1	Period 1-4	Pooled data
Period 1-4	0.270***	0.090***	0.100*			0.100**
	(0.052)	(0.023)	(0.046)			(0.044)
SINGLE				-0.155**	0.015	-0.155**
				(0.071)	(0.033)	(0.071)
MULTI				0.085	0.075**	0.085
				(0.053)	(0.029)	(0.053)
SINGLE × Period 1-4						0.170**
						(0.067)
MULTI × Period 1-4						-0.010
						(0.050)
Constant	0.620***	0.860***	0.775***	0.775***	0.875***	0.775***
	(0.051)	(0.016)	(0.053)	(0.050)	(0.024)	(0.050)
Observations	200	200	160	280	280	560

Note: Dependent variable is a dummy equal to one if a subject cancelled a donation for a payment of at most its value (€ 1.50). Period 1-4 is a dummy variable equal to one if the choice is measured as occurring at least once in period 1 to 4 in part 2 of the experiment, the omitted category is cancellation in period 1. SINGLE (MULTI) is a dummy equal to one if the choice occurred in treatment SINGLE (MULTI), with the omitted category PRIV. Standard errors, clustered on matching group level, are presented in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The possibility of replacement thinking is not limited to markets with a common schedule. Here, we discuss two ways in which replacement excuses can be introduced in markets with private schedules. The first possibility arises when aggregate supply and demand are horizontally extended beyond the competitive equilibrium, as is illustrated in Figure 3.A.18. This schedule results when, compared to the PRIV schedule, we allow traders on both sides of the market to trade additional units at a price close to the competitive equilibrium price. If the buyers' values for these additional units are slightly below the competitive equilibrium price, while the sellers' costs are slightly above, then the competitive equilibrium is not affected. This schedule allows traders to take full advantage of trading opportunities foregone by others. Assuming that other traders are selfish, traders would anticipate that their trading decisions will not matter for the aggregate outcome, and replacement thinking will excuse their trading. Note that this type of schedule does not feature elements of markets we want to capture, for example traders' costs and values do no longer depend on others' trading.

The second possibility is to add traders on both sides, with similar cost and demand schedules. Combined with a restriction for aggregate trade not to exceed the original competitive quantity, traders will again recognize that the replacement excuse applies, and feel free to trade as much as they can.

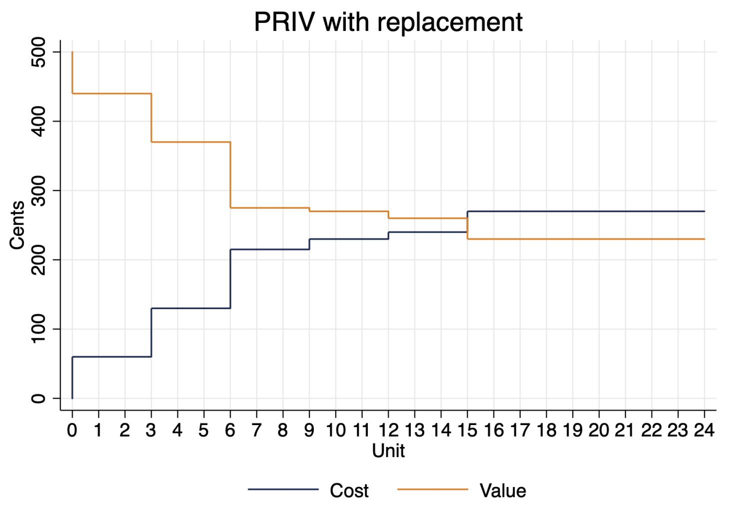


Figure 3.A.18: Cost and value schedule with replacement for PRIV

3.B Experimental interface

Below is an example screenshot from the experimental markets.

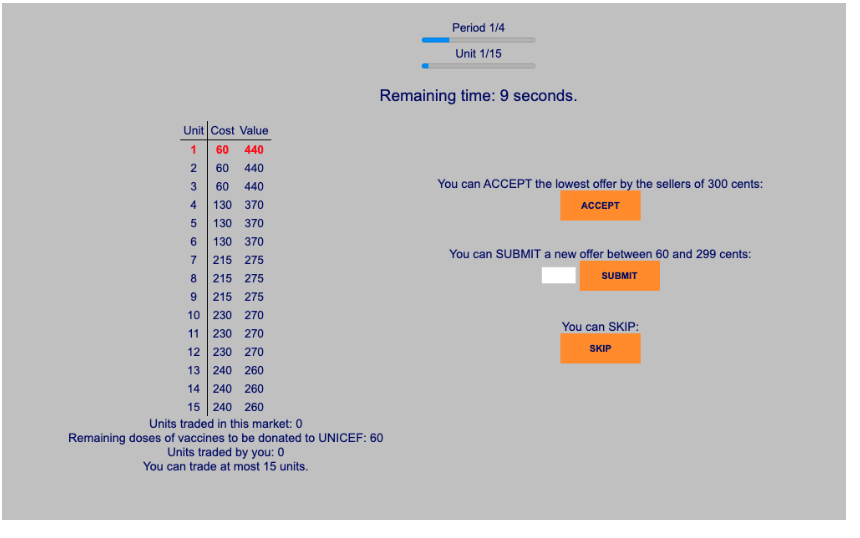


Figure 3.B.1: Experimental market interface

After each unit, traders received the following feedback:

You accepted an offer. You earn 140 cents.

The price agreed upon was 300. The buyer earns 440-300=140 cents, the seller 300-60=240 cents.

Four doses of measles vaccines will not be donated to UNICEF because of this trade. If no more units are traded in this market, 56 doses of measles vaccines will be donated to UNICEF.

In this market period, you have traded 1 unit and earned 140 cents so far.

Figure 3.B.2: Feedback after each trade

Further, after the end of each market period, participants received this feedback:

The market for this trading period has closed, no more units can be traded.

In this period, you traded one unit. You earn 140 cents if this period will be randomly selected for payment.

In the entire market, 1 unit was traded.

If this period will be randomly selected for payment, 56 doses of vaccine will be donated to UNICEF. Donations for 4 doses have been cancelled in this market.

Figure 3.B.3: Feedback after a market period

Therefore, traders were reminded of the negative externality that was caused by trading continuously within the market, after each unit and at the end of each market period.

3.C Instructions part 1

3.C.1 Page 1

Welcome!

Welcome to this experiment. Please read the following instructions carefully.

Please do not communicate with other people and refrain from verbally reacting to events that occur during the experiment. The use of mobile phones or laptops is not allowed.

There are pen and paper on your table, you can use these during the experiment. We will also distribute a handout with some key facts about this experiment later.

If you have any questions, or need assistance at any time, please notify the experimenter by raising your hand. The experimenter will assist you privately.

3.C.2 Page 2

General information

This experiment consists of multiple parts. {NOT in HOM/HET Your decisions in one part will not affect any of your choices or potential earnings in other parts. You will receive instructions for each part separately.}

For your participation in this experiment, you will be paid 7 Euro. Additionally, you can earn money by your decisions in this experiment. These earnings will depend on your decisions and may depend on other participants' decisions. **One out of the first three parts will be randomly selected to be paid to you.** Additionally, you will be paid for three short tasks at the end of the experiment. Your earnings will be paid to you privately in cash at the end of today's session. **All your earnings will be denoted in cents** (100 Cents = 1 Euro).

{HOM/HET Your decisions in part 1 will affect with whom you will interact in a later part of this experiment. Like the other participants, you will either be assigned to a group of participants that made quite similar choices as the average participant did in part 1, or to a group in which participants behaved quite differently from the average participant.}

3.C.3 Page 3

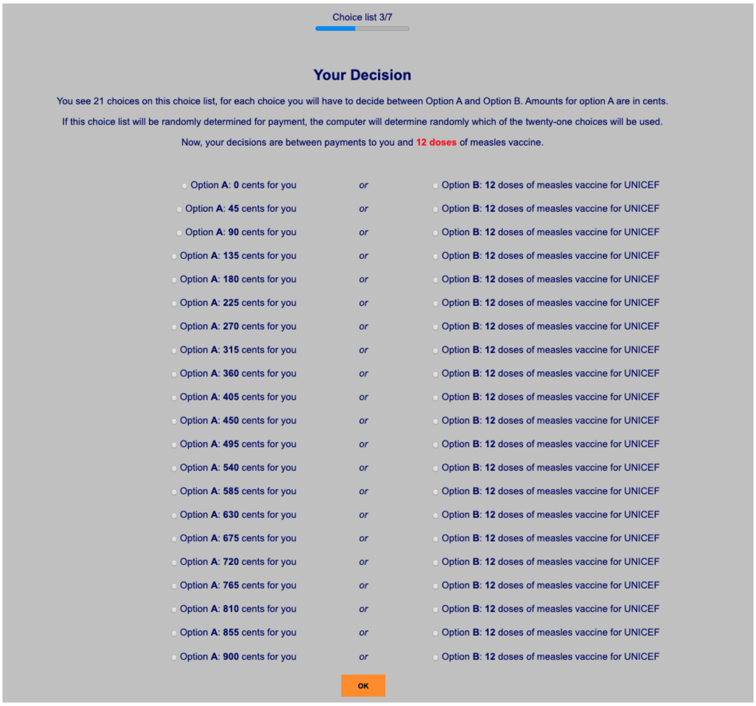
Part 1

In this first part, you will repeatedly choose between two options, A and B:

- **A**: This option will pay a certain amount of money to you.
- **B**: This option will donate a certain amount of money to UNICEF. With this donation, UNICEF will buy measles vaccines. With two doses of this vaccine, one child can be vaccinated against measles (details on the donation follow below).

A list of repeated choices between A and B on one screen is called a choice list.

Below is an example of a choice list. In this example, you choose between varying amounts of money paid to you on the left (option A) and 12 doses of measles vaccine on the right (option B). A donation of four doses costs approximately 1.5 Euro. Even though you will be asked to make multiple decisions, at most one of them will affect your earnings.



In the second example screenshot below, you see another choice list. Here, you choose between varying amounts of money paid to you in option A and *28* doses of measles vaccines in option B. Note that also the available payments in option A vary across choice lists.

You will face several choice lists like the one in the screenshots above. On each list, two things change. First, the number of doses of vaccine donated to UNICEF change, which are 12 and 28 in the two examples given here. Second, the available payments in option A change.

Within each choice list, only option A changes between choices. As you scroll down the list, the amount of measles vaccines donated to UNICEF stays the same. The money that would be paid to you if you choose not to donate to UNICEF is increasing on each choice list. To simplify the decision, as soon as you click on one choice, the computer will pre-fill subsequent choices automatically. If for a particular choice you chose A (money to you), then all choices on the choice list below this choice of A pay even more money to you than that choice while option B does not change. Then, the computer will pre-fill A for all these choices below. Similarly, if for a particular choice you chose B (donation to UNICEF), then all choices above this choice pay even less money to you than that choice, so then the computer will pre-fill B for all these choices above. Until you click on OK, you can always change your decision. The pre-filled choices will adjust automatically while you change your decision.

3.C.4 Page 4

Payment

If this part is randomly selected for payment, one of your decisions from this part will be randomly selected to be paid out at the end of the experiment. For the decision to be paid out, first, one of the seven choice lists you faced will randomly be chosen, with each choice list being equally likely. Second, within this chosen list, one decision will be randomly chosen for payment, with each decision being equally likely. If you chose option A for this decision, you will be paid the number of cents indicated for this choice. If you instead chose option B, the specified number of doses of measles vaccines will be donated to UNICEF at the end of the experiment.

Details about the donation Two doses of measles vaccine are sufficient to vaccinate one child (see the next page for more details) and can be bought with a donation of approximately 75 cents. Depending on your choice in the selected period a certain amount of money is donated to UNICEF by the experimenters. We will show you a donation receipt by UNICEF at the end of this experiment, right after we transferred the announced donation. As an example, below we show you how such a receipt for a previous donation looks like. A confirmation of the donation to UNICEF can also be sent to you via email, to allow you to verify the correctness of the statements made here. To do so, you can write your email address on the form on your table, which will be collected after the experiment. Your email address will not be linked to any other data in this experiment.

As UNICEF only allows us to donate in bundles of 40 doses, any excess donations in your session will be paid to UNICEF as a direct transfer, and this transfer will be included in the receipt we show you.

[DONATION RECEIPT EXAMPLES]

3.C.5 Page 5

Information on the measles vaccines

Measles are highly infectious and very often deadly. Each day hundreds of children become victims of this disease. The survivors often suffer consequences for their whole life, like blindness or brain damages. This, even though protecting the children would be easy. Measles kills more than 160,000 children worldwide each year.

Measles are extremely infectious and spread especially fast when many people live densely together, as in refugee camps. Especially with weakened children the disease often ends deadly or leads to lasting physical or mental damages. Measles are one of the main causes for blindness among children and often become critical when no medical help is available. This, even though measles vaccination offers quick, reliable, and cheap protection. UNICEF conducts major vaccination campaigns, especially after natural disasters and in other emergency situations, to prevent the spreading of the disease. With a measles vaccination you do not only protect the children, but you also reduce the risk for all who get in contact with them.

Source text on measles vaccines by UNICEF: `https://unicef.at/shop/index.php/gesundheit-und-schutz/masern-impfstoff.htm` and `https://market.unicef.org.uk/inspired-gifts/measles-vaccines-to-protect-20-children/S359163X/`

Source pictures: `https://market.unicef.org.uk/inspired-gifts/measles-vaccines-to-protect-20-children/S359163X/`

3.C.6 Practice questions (page 6)

NOTE: ALL NUMBERS IN THE QUESTIONS ARE ARBITRARILY CHOSEN, AND ARE NOT RELEVANT FOR THE EXPERIMENT.

Please answer the following questions:

1. {NOT in HOM/HET Your decisions in other parts do not affect your earnings in this part. Also, your choices in this part do not affect your earnings in other parts. [TRUE/false]}

2. The following choice has been randomly selected for payment:

   In this choice, you have chosen option A, as indicated. How much will be paid to you? [FREE FORM: 360] cents

   How many doses of measles vaccines will be donated to UNICEF? [FREE FORM: 0] doses

3. Now, the following choice has been randomly selected:

   In this choice, you have chosen option B, as indicated. How much will be paid to you? [FREE FORM: 0] cents

   How many doses of measles vaccines will be donated to UNICEF? [FREE FORM: 28] doses

4. At the end of this experiment, the promised donations will immediately be transferred by the experimenter. You can verify this with the receipt from UNICEF. [TRUE/false]

3.C.7 Page 7

End of instructions

You have reached the end of the instructions. When you are ready for the experiment, please push the button READY. When all participants have pushed READY, the experiment will start.

If you still have questions, please raise your hand, and the experimenter will assist you!

3.D Instructions part 2

{FOR IDM: Repeated instructions of part 1: These are the identical instructions as those you saw at the start of the experiment [see above]}

{FOR MARKET TREATMENTS [ONLY SELLER INSTRUCTIONS ARE REPRODUCED, BUYERS APPROPRIATELY ADJUSTED]:

3.D.1 Page 1

Market instructions

{NOT in HOM/HET: In this part of the experiment you will repeatedly trade in a market. In the market, 5 sellers can trade with 5 buyers. You will be a SELLER in the entire experiment. You will trade in two markets, market 1 and market 2, which proceed according to similar rules. After market 1 is completed, you will receive additional instructions for market 2.} {HOM/HET: In this part of the experiment you will repeatedly trade in a market. In the market, 5 sellers can trade with 5 buyers. As explained in part 1, you will either be assigned to a group of participants that made quite similar choices as the average participant did in part 1, or to a group in which participants behaved quite differently from the average participant. You will participate in a group in which participants' choices are similar/different from the average participant.

You will be a BUYER in the entire experiment. You will trade in two markets, market 1 and market 2, which proceed according to similar rules. After market 1 is completed, you will receive additional instructions for market 2.}

Market 1

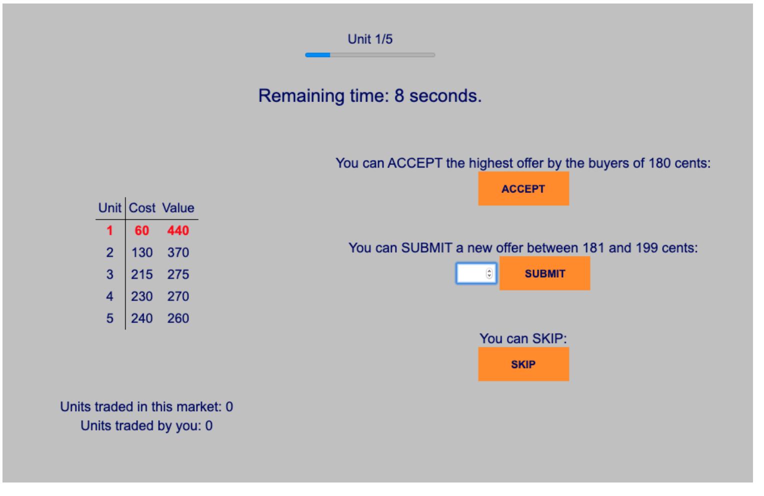
Trading profits

In market 1, a total of 5 units can be traded. Each trader can trade at most {FULL: 5 units; MULTI: 2 units; SINGLE: 1 unit}. Trading will proceed unit by unit. For each unit, one buyer and one seller can conclude a trade by agreeing on a price for that unit.

If bought, each unit has a certain cost to the seller. This will be denoted in cents. Similarly, each unit sold will have a value in cents to the buyer. Earnings for the buyer and seller for concluding a trade are:

- The seller earns the difference between the price and the cost for this unit: PRICE-COST
- The buyer earns the difference between the value and the price for this unit: VALUE-PRICE

These costs and values are presented during the market, as in the screenshot below. In this example, the first unit is being traded, which is highlighted by the red first line in the table.



Example in the screenshot: The buyer has a value given by 440, the seller has a cost given by 60. You, as the seller, and one of the buyers agree on a price of 180. Then,

- You get: PRICE - COST = 180 - 60 = 120 cents.
- The buyers get: VALUE - PRICE = 440 - 180 = 260 cents.

In the screenshot, notice that the cost of the seller and the value of the buyer change with the unit transacted (e.g. for the first unit the cost for the seller is 60 cents and the value for the buyer is 440 cents, for the third unit the cost is 215 cents and the value is 275 cents and so on). **However, for each unit, they are the same for all buyers or sellers.** Costs and values only depend on the number of units traded up to that point in the entire market by any of the traders. That is, they do not depend on the number of units you yourself have traded previously.

3.D.2 Page 2

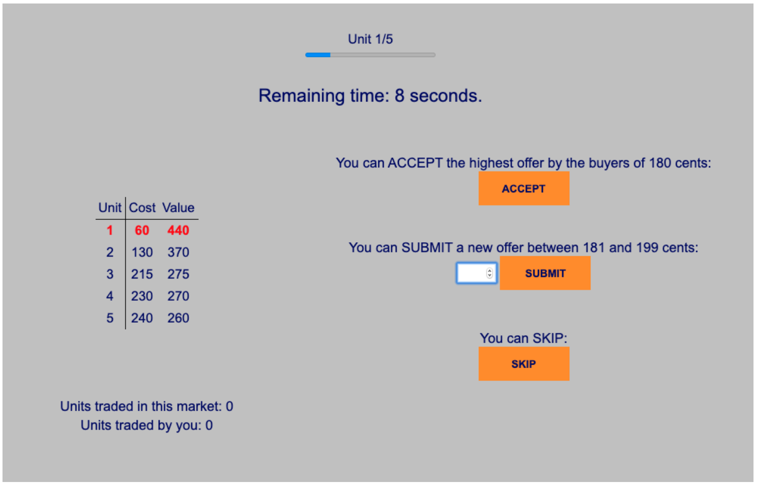
Trading protocol

To agree on a price, the side of the sellers and the side of the buyers submit and accept offers sequentially. This means that first one side of the market decides ("the active side"), afterwards this side will wait and the other side of the market decides. If trading continues, the first side of the market is allowed to decide again, and so forth.

While your side (the sellers' side) is active in the market, you have three available choices:

1. SUBMIT: Submit an offer to the buyers
2. ACCEPT: Accept an offer of the buyers
3. SKIP

You can see all three options available in the screenshot below:



Each of the options works according to these rules:

1. ACCEPT:
   - You will see the highest price offered by any of the buyers.

- You can accept this highest offer. If you do so, a trade for one unit is concluded, the profits are calculated as explained before.
- If multiple sellers accept an offer, or if multiple offers are equally good, it will be randomly chosen which of the traders who wanted to can conclude this trade.
- Afterwards, trading of the next unit can begin, old offers are removed and new ones can be made.

2. SUBMIT:

   - You can submit a new offer, which will be presented to the buyers as soon as they become active.
   - A new offer has to improve upon previous offers. This means that a new offer needs to be above the lowest offer submitted by any of the other sellers. A new offer cannot be above the buyers' values, or below the highest offer by the buyers.

3. SKIP:

   - If you skip, you immediately move to the waiting screen.
   - As soon as all sellers are on the waiting screen, the buyers become active and can submit new offers or accept the lowest offer of the buyers. Clicking on skip can speed up the market. However, you will no longer be able to submit or accept an offer at that moment.
   - If you do not submit or accept an offer within the trading time of 14 seconds, you will skip automatically.

End of trading

Trading ends if all available units are sold in the market.

Also, if no trader on both sides of the market chooses SUBMIT or ACCEPT, a warning sign will be shown. Then, each trader on both market sides can once again SUBMIT, ACCEPT or SKIP. If again no trader on either of the two sides chooses SUBMIT or ACCEPT, the market ends for this and all subsequent units. This means that you will not be allowed to trade additional units after this happens.

3.D.3 Page 3

Additional details

- At the start of the market for the first unit, it is randomly determined whether the side of the buyers or the side of the sellers first becomes active. For the next units, the active side for making the first offer is alternated.

- On the top of the trading screen you always see the remaining trading time. We will also show how many units you have traded. The specific moment at which you submit or accept offers does not matter, as long as you submit or accept within the 14 second trading time.
- No trader knows with whom in the room he or she has traded. That means that your anonymity is ensured.

Reminders

- At each moment, you can choose only one of the three options (SUBMIT, ACCEPT and SKIP). If trading continues and your side of the market becomes active again, you can again choose between these options.
- A maximum of five units can be traded in market 1; after the 5th unit is sold the market ends. Each trader can trade at most {FULL: 5 units; MULTI: 2 units; SINGLE: 1 unit}.
- Each unit is traded by one buyer and one seller, all other traders get a payoff of zero for that unit.

Payment

If this part and this market is selected for payment, for each trade a participant concluded, his or her payment is calculated with the rules described above. That is, for each unit, the seller will be paid the difference between the price and the cost for this unit. The buyer will be paid the difference between the value for this unit and the agreed upon price. {FULL/MULTI: The earnings for this market are then given by the sum of earnings for all units traded by each participant.}

In part 2, there will be a total of 5 markets. 2 out of the 5 markets will be randomly selected to be paid.

3.D.4 Predictions {ONLY IN B-FULL/B-MULTI}

{ONLY IN B-FULL/B-MULTI: While you are trading, you will occasionally be asked to predict how future trading will proceed. At these moments, you will be asked what you think will happen in the trading of the next unit.

Each time, we will ask you to predict four things.

The first three predictions concern probabilities of whether trading will occur for the next unit. For each next unit, there are three possible events:

1. Whatever you do, the unit will be traded. This means that even if you do not participate in trading, the unit will be traded by the others.

2. Your behavior will determine whether the unit is traded or not. This means that if you do not participate, the unit will not be traded, while if you do participate, the unit will be traded.

3. Whatever you do, the next unit will not be traded. This means that even if you do try to trade the next unit, this will not happen because the buyers are not participating.

We will ask you for the probabilities that each of these events occurs. These probabilities are your predictions of how likely it is that each possible event will happen. A higher probability means that an event is more likely to happen. As a probability, your predictions can be between 0% (will not occur) and 100% (will certainly occur). As the three events above include all possible scenarios in which this experiment progresses, the probabilities you report across 1. to 3. need to add up to 100%.

The last prediction concerns the number of sellers and buyers who will be active for the next unit. We will ask you to predict how many buyers and sellers other than you within your market will attempt to trade the next unit. By this we mean the total number of participants other than you who will either submit an offer to and/or accept an offer. Your predictions can be between 0 participants (no other participant will be active) and 9 participants (all other participants will be active).

At the end of the experiment, if this part is selected for payment, you will be paid for a set of predictions for one unit in one period of the markets. This will be another period than the period for which your trading determines your earnings. For the first three predictions, you will receive 300 cents. For the fourth prediction, you will receive an additional payment of 150 cents if you correctly predict how many participants other than you attempt to trade the next unit.}

3.D.5 Practice questions (page 4)

NOTE: ALL NUMBERS IN THE QUESTIONS ARE ARBITRARILY CHOSEN, AND ARE NOT RELEVANT FOR THE EXPERIMENT.

Please answer the following questions:

1. Each seller pays the same costs as any of the other sellers to supply any unit, and each buyer values any unit equally as any of the other buyers. [TRUE/false]

2. If no buyer or seller submits an improved offer twice, the market for this period will end and no more units can be traded. [TRUE/false]

3. {ONLY IN B-FULL/B-MULTI How much will you earn if you correctly predict how many other participants will attempt to trade the next unit? [FREE FORM: 150] cents}

4. We will ask you several questions about the scenario below. Note that the behavior in this scenario is randomly determined, only for the purpose of asking these questions.

   (a) The first unit is being traded in this market. This unit costs 60 cents to any of the sellers, and has a value of 440 cents to any of the buyers. The buyers were randomly selected to first submit offers.

   (b) Buyer B1 decides to submit a price of 140 cents to the buyers and buyer B2 submits a price offer of 200 cents. The trading time of 14 seconds expires without any other buyer submitting an offer.

   (c) Now the sellers become active. As buyer B2's offer is the highest offer, the sellers will only see buyer B2's offer of 200 cents.

   (d) However, none of the sellers decides to accept this offer. Instead, seller S1 submits a new offer. This offer needs to be higher than 200 cents, as otherwise accepting buyer B2's offer is more favorable to seller S1. Seller S1 submits a new offer of 260 cents. Again, the trading time of 14 seconds expires without any other seller submitting or accepting an offer.

   (e) Now, the buyers become active again. Seeing seller S1's offer of 260 cents, buyer B3 decides to accept this offer. The trading time of 14 seconds expires without any other buyer accepting this offer. This means that the first unit has been traded.

   (f) Afterwards, bargaining about the second unit begins.

   How many cents does buyer B1 earn from the first unit? [FREE FORM: 0] cents

   How many cents does buyer B2 earn from the first unit? [FREE FORM: 0] cents

   How many cents does buyer B3 earn from the first unit? [FREE FORM: 180] cents

   How many cents does seller S1 earn from the first unit? [FREE FORM: 200] cents

[MARKET 1 TAKES PLACE, AFTERWARDS INSTRUCTIONS FOR MARKET 2 (with externality) FOLLOW]

3.D.6 Page 1

This concludes market 1. Now, trading in **market 2** begins.

Generally, the same rules apply in this market. We will therefore highlight here only the differences between the two markets:

- Trading behavior in this market determines an amount of money that will be donated to UNICEF, in addition to your own earnings. The number of units successfully traded in the market is used to calculate how many doses of measles vaccines will be donated to UNICEF. The maximum number of doses donated to UNICEF in one market period is {FULL/MULTI: 60 doses; SINGLE: 20 doses}. The more units are traded in the market, the less will be donated to UNICEF: for each unit that is traded in market 2 that is selected for payment, 4 doses of measles vaccines will be subtracted from the donation to UNICEF, which cost approximately 1.5 Euro. Recall that with two doses of measles vaccine, one child can be protected. UNICEF will be paid the donation amount at the end of the experiment. The following table summarizes how the number of traded units in the market translates into the number of MEASLE DONATIONS. For example, if at the end of the market, zero units have been traded, then a total of {FULL/MULTI: 60 doses; SINGLE: 20 doses} are donated to UNICEF for this market. If at the end of the market 3 units have been traded then in total {FULL/MULTI: 48 doses; SINGLE: 8 doses} doses are donated. Donations to UNICEF are only affected by the overall number of units traded in the market and not by whom these units are traded.

Final number of units traded and number of doses: [TREATMENT-SPECIFIC TABLE WITH COST/VALUES]

- Each unit traded has a VALUE and a COST according to the table below. These costs and values (in cents) will be the same in all markets of this experiment. [TABLE HERE, STATING NUMBER OF TRADED UNITS AND CORRESPONDING DONATIONS]

- While market 1 only lasted for 1 period, you will now be trading in a sequence of 4 market periods. Each market period is conducted in the same way. Your choices in one period have no consequences on any other period.

- {FULL/MULTI: While in market 1 a maximum of 5 units could be traded, now the maximum number of units tradeable in each market period is 15.; SINGLE: As in market 1, a maximum of 5 units can be traded.} Just like in market 1, fewer than {FULL/MULTI: 15; SINGLE: 5} units will be traded if the traders no longer SUBMIT or ACCEPT after the warning sign. Moreover, each trader can trade at most {FULL: 15 units; MULTI: 3 units; SINGLE: 1 unit}. {MULTI/SINGLE: This means that if you decide not to trade one unit that you are allowed to trade, you reduce the number of units that can be traded by one, which would also reduce the corresponding damage to the donation to UNICEF.}

Payment

If this part is selected for payment, two of the market results are randomly selected for payment. It is equally likely that each one of the 4 market periods of market 2 or the one period in market 1 is selected for payment. Payment for participants are then calculated according to the same rules as in market 1.

If a market period of market 2 is selected, the trades in the selected period also determine the amount donated to UNICEF. At the end of the experiment, the experimenter will transfer this amount.

3.D.7 Page 2

[REPEATED INFORMATION ON UNICEF, SEE INSTRUCTIONS FOR PART 1]

3.D.8 Practice questions (Page 3)

NOTE: ALL NUMBERS IN THE QUESTIONS ARE ARBITRARILY CHOSEN, AND ARE NOT RELEVANT FOR THE EXPERIMENT.

Please answer the following questions:

1. If this part is selected for payment, two market results are randomly selected for payment. These can be market 1 or one of the market periods of market 2. {FULL/MULTI: Each trader earns the sum of cents generated by all of his or her trades} [TRUE/false]

2. For each unit that is traded, how many doses of measles vaccines will be subtracted from the donation to UNICEF? [FREE FORM: 4] doses

3. We will ask you several questions about the scenario below. Note that the behavior in this scenario is randomly determined, only for the purpose of asking these questions.

   (a) The first unit is being traded in the market. This first unit costs 60 cents to any of the sellers, and has a value of 440 cents to any of the buyers. The sellers are first to submit offers.

   (b) Seller S1 decides to submit a price of 290 cents to the buyers. Also, seller S2 submits a price offer, of 310 cents. The trading time of 14 seconds expires without any other seller submitting an offer.

   (c) Now the buyers become active. As seller S1's offer is the lowest offer, the buyers will only see seller S1's offer of 290 cents.

   (d) Buyer B1 and buyer B2 decide to accept this offer.

(e) It is randomly determined that buyer B2 buys the first unit. This means that the first unit has been traded and that 4 fewer doses of measles vaccines will be donated to UNICEF.

(f) Afterwards, bargaining about the second unit begins.

How many cents does seller S1 earn from the first unit? [FREE FORM: 230] cents

How many cents does seller S2 earn from the first unit? [FREE FORM: 0] cents

How many cents does buyer B1 earn from the first unit? [FREE FORM: 0] cents

How many cents does buyer B2 earn from the first unit? [FREE FORM: 150] cents

}

3.E Instructions part 3

3.E.1 Page 1

Part 3

You will now face a set of choices identical to the choices at the start of the experiment. As before, you have several choice lists, where each choice asks you to choose between points for yourself or varying doses of measles vaccine donated to UNICEF.

This part is conducted identically to the first part, and you will also be paid according to the same rules. On the next page, we reproduce the instructions from the start of the experiment in case you want to review them again.

Note that your earnings from your decisions in this part are not depending on any decision you have made up to now, or on any of your decisions you will make in the following set of questions.

3.E.2 Page 2

[SEE ABOVE FOR INSTRUCTIONS]

3.F Instructions for the three additional tasks

This is the end of the main parts of this experiment. In the remainder you will be able to make some additional money for three short tasks.

3.F.1 Instructions part 4 (belief elicitation)

Now, think of all other subjects who participate in this session today. The first task everyone completed in this experiment was a choice list where you could choose between an amount for yourselves and a donation of 4 doses of measles vaccines donated to UNICEF.

What do you think other participants chose on average in this choice list?

Please fill out this choice list **as you think the average participant did** in their first choice list. If your choice matches what the average participant did, you will earn an additional bonus of 100 cents.

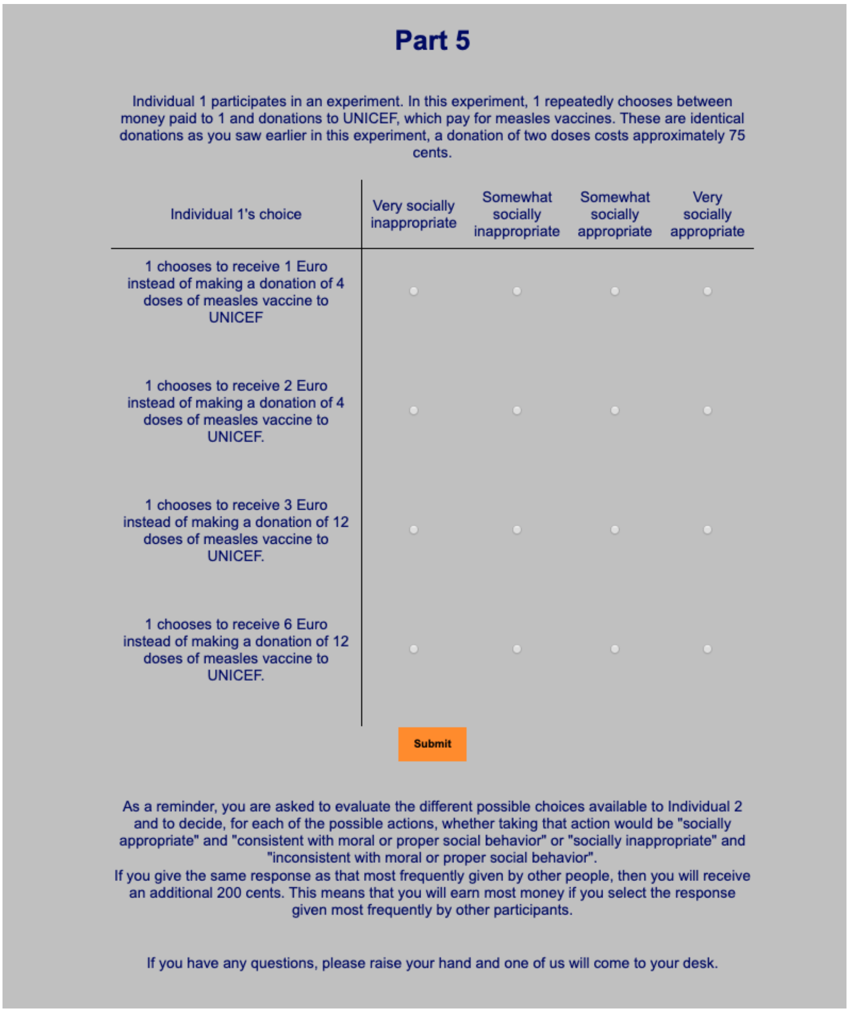
3.F.2 Instructions part 5 (norms)

On the following screens, you will read descriptions of a series of situations. These descriptions correspond to situations in which one person, Individual 1, must make a decision.

After you read the description of the decision, you will be asked to evaluate the different possible choices available to the person and to decide, for each of the possible actions, whether taking that action would be “socially appropriate” and “consistent with moral or proper social behavior” or “socially inappropriate” and “inconsistent with moral or proper social behavior”. By socially appropriate, we mean behavior that most people agree is the “correct” or “ethical” thing to do. Another way to think about what we mean is that if Individual 1 were to select a socially inappropriate choice, then someone else might be angry at Individual 1 for doing so.

In each of your responses, we would like you to answer as truthfully as possible, based on your opinions of what constitutes socially appropriate or socially inappropriate behavior.

At the end of the experiment today, we will randomly select one of the situations. For this situation, we will also randomly select one of the possible choices that Individual 1 could make. Thus, we will select both a situation and one possible choice at random. For the choice selected, we will determine which response was selected by most people participating in this experiment right now. If you give the same response as that most frequently given by other people, then you will receive an additional 200 cents. This means that you will earn most money if you select the response given most frequently by other participants.



Part 5

Individual 1 participates in an experiment. In this experiment, 1 repeatedly chooses between money paid to 1 and donations to UNICEF, which pay for measles vaccines. These are identical donations as you saw earlier in this experiment, a donation of two doses costs approximately 75 cents.

Individual 1's choice	Very socially inappropriate	Somewhat socially inappropriate	Somewhat socially appropriate	Very socially appropriate
1 chooses to receive 1 Euro instead of making a donation of 4 doses of measles vaccine to UNICEF				
1 chooses to receive 2 Euro instead of making a donation of 4 doses of measles vaccine to UNICEF.				
1 chooses to receive 3 Euro instead of making a donation of 12 doses of measles vaccine to UNICEF.				
1 chooses to receive 6 Euro instead of making a donation of 12 doses of measles vaccine to UNICEF.				

Submit

As a reminder, you are asked to evaluate the different possible choices available to Individual 2 and to decide, for each of the possible actions, whether taking that action would be "socially appropriate" and "consistent with moral or proper social behavior" or "socially inappropriate" and "inconsistent with moral or proper social behavior".
If you give the same response as that most frequently given by other people, then you will receive an additional 200 cents. This means that you will earn most money if you select the response given most frequently by other participants.

If you have any questions, please raise your hand and one of us will come to your desk.

Part 5

Individual 2 also participates in an experiment. In this experiment, 2 is a seller in a market. 2 negotiates with other participants, and can earn money for every trade. However, **whenever a trade is made in this market, a donation with four doses of measles vaccines for UNICEF is cancelled.** A donation of four doses costs approximately 1.5 Euro.

2 can now make different offers or accept others' offers in this market, see the options below.

Individual 2's choice	Very socially inappropriate	Somewhat socially inappropriate	Somewhat socially appropriate	Very socially appropriate
2 decides to accept an offer which allows him to earn 1 Euro.				
2 decides to accept an offer which allows him to earn 2 Euro.				
2 makes an offer in the market. If a trade is concluded based on this offer, 2 would earn 1 Euro.				

Submit

As a reminder, you are asked to evaluate the different possible choices available to Individual 2 and to decide, for each of the possible actions, whether taking that action would be "socially appropriate" and "consistent with moral or proper social behavior" or "socially inappropriate" and "inconsistent with moral or proper social behavior".
If you give the same response as that most frequently given by other people, then you will receive an additional 200 cents. This means that you will earn most money if you select the response given most frequently by other participants.

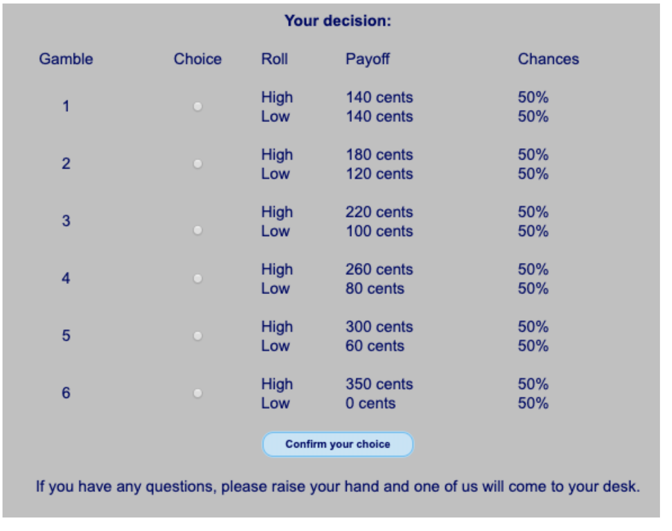
If you have any questions, please raise your hand and one of us will come to your desk.

3.F.3 Instructions part 6 (risk aversion)

For this part, you choose one gamble you would like to play from among six different gambles. The six different gambles are listed below. You must select one and only one of these gambles.

Each gamble has two possible outcomes (Roll Low or Roll High). For every gamble, each Roll has a 50% chance of occurring. At the end of the study, it will be randomly determined which event will occur.

For example, if you select Gamble 4 and Roll High occurs, you will be paid 260 cents. If Roll Low occurs, you will be paid 80 cents.



Your decision:

Gamble	Choice	Roll	Payoff	Chances
1		High Low	140 cents 140 cents	50% 50%
2		High Low	180 cents 120 cents	50% 50%
3		High Low	220 cents 100 cents	50% 50%
4		High Low	260 cents 80 cents	50% 50%
5		High Low	300 cents 60 cents	50% 50%
6		High Low	350 cents 0 cents	50% 50%

Confirm your choice

If you have any questions, please raise your hand and one of us will come to your desk.

3.G Instructions for SPEC

Participants read instructions as their counterpart in B-FULL did. They did not see tasks related individual decision-making or any of the tasks at the end of the experiment (beliefs, risk aversion, norms).

The instructions started with:

3.G.1 Page 1: Your task

In this experiment, you will first read instructions similar to the instructions that participants received in earlier sessions of this experiment. However, you will *not* participate in that experiment, and will not be able to make the choices described in these instructions.

It is important that you read and understand these instructions well, as your earnings depend on decisions in these earlier sessions.

After you have read the instructions, you will make predictions about choices these earlier participants made. Your earnings in this experiment will depend on the accuracy of your predictions about choices these earlier participants made.

3.G.2 Instructions and quiz of their matched participant (see B-FULL)

3.G.3 Page 2: Your prediction task

We now describe the choices that you will make in this experiment. You will follow one particular participant in an earlier session of this experiment. We will call this person your matched participant. We will show you the market interaction that your matched participant experienced in the experiment. Within a dark grey box, you will see a screen identical to the screen your matched participant observed. There is also some basic explanation on how to interpret this screen in the markets.

We will ask you to make predictions about the choices of *other* participants in the session of your matched participant.

Each time, we will ask you to predict four things.

The first three predictions concern probabilities of whether trading will occur for the next unit. For each next unit, there are three possible events:

1. Whatever your matched participant does, the unit will be traded. This means that even if your matched participant does not participate in trading, the unit will be traded by the others.

2. Your matched participant's behavior will determine whether the unit is traded or not. This means that if your matched participant does not participate, the unit will not be traded, while if your matched participant does participate, the unit will be traded.

3. Whatever your matched participant does, the next unit will not be traded. This means that even if your matched participant does try to trade the next unit, this will not happen because the buyers/sellers are not participating.

We will ask you for the probabilities that each of these events occurs. These probabilities are your predictions of how likely it is that each possible event will happen. A higher probability means that an event is more likely to happen. As a probability, your predictions can be between 0% (will not occur) and 100% (will certainly occur). As the three events above include all possible scenarios in which this experiment progresses, the probabilities you report across 1. to 3. need to add up to 100%.

The last prediction concerns the number of sellers and buyers who will be active for the next unit. We will ask you to predict how many buyers and sellers other than your matched participant within your matched participant's market will attempt to trade the next unit. By this we mean the total number of participants other than your matched participant who will either submit an offer to and/or accept an offer. Your predictions can be between 0 participants (no other participant will be active) and 9 participants (all other participants will be active).

At the end of the experiment, you will be paid for three randomly selected sets of predictions. Within each set of predictions, for the first three predictions, you will receive 300 cents. For the fourth prediction, you will receive an additional payment of 150 cents if you correctly predict how many participants other than you attempt to trade the next unit.

3.G.4 Page 3: Practice questions

Please answer the following questions:

1. In this experiment, you will only make predictions about the choices of other participants in an earlier experiment. [TRUE/false]
2. You will observe the market interaction your matched participant observed in this earlier experiment. [TRUE/false]
3. You will also trade in a market similar to the market your matched participant participated in. [true/FALSE]
4. How much will you earn if you correctly predict how many participants other than your matched participant will attempt to trade the next unit? [FREEFORM: 150] cents

3.H Double auction: PRIV

All participants received instructions on the individual decision-making task as in the other market treatments. Below we reproduce all instructions for part 2.

3.H.1 Page 1: Market instructions

In this part of the experiment you will repeatedly trade in a market. In the market, 5 sellers can trade with 5 buyers. You will be a SELLER in the entire experiment. You will trade in two markets, market 1 and market 2, which proceed according to similar rules. After market 1 is completed, you will receive additional instructions for market 2.

3.H.1.1 Market 1

Trading profits

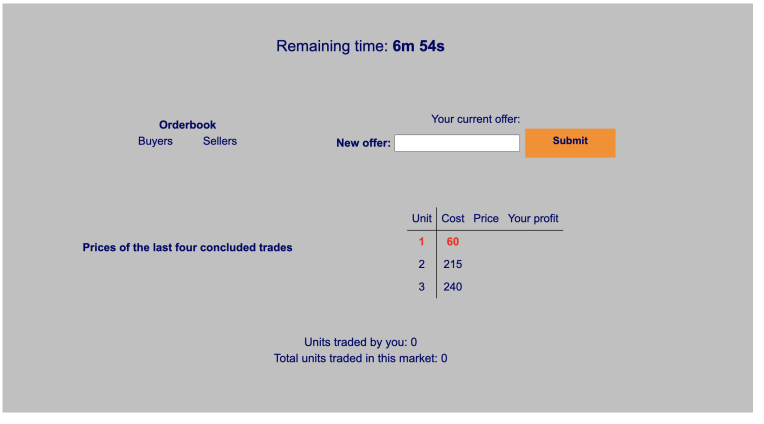
In market 1, up to 15 units can be traded. To trade one unit, one buyer and one seller can conclude a trade by agreeing on a price for that unit.

If sold, each unit has a certain cost. This will be denoted in cents. Earnings for concluding a trade are:

- You earn the difference between the price and the cost for this unit: PRICE-COST

You will only know your costs in the market. You will not know other sellers' costs. In general, all participants have different costs. It is randomly determined which participants has what costs.

In the example in the screenshot below, the seller can trade their first unit, which is highlighted by the red first line in the table.



Example in the screenshot: Imagine you have a cost given by 60. You, as the seller, and one of the buyers agree on a price of 260. Then,

- You get: PRICE - COST = 260 - 60 = 200 cents.

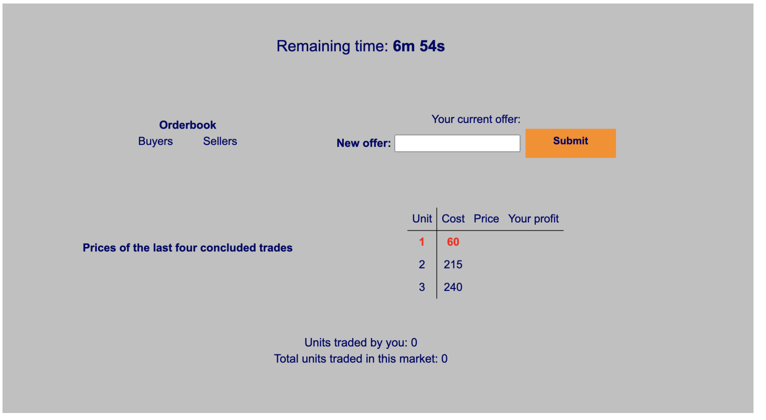
In the screenshot, notice that the cost of the seller change with the unit transacted (e.g. for the first unit the cost for the seller is 60 cents, for the third unit the cost is 240 cents).

3.H.2 Page 2: Market instructions

Trading protocol

To agree on a price, all participants submit offers simultaneously. At any time during the trading, you can submit an offer to the buyers.

You can see how you can submit an offer in the screenshot below:



If you submit an offer, one of two things will happen:

1. A trade happens immediately:

   - If your new offer is at a price *below or equal to* the best current offer of the buyers, you will trade.
   - You will trade with the buyer who made this offer. The price will be the price offered by this buyer.
   - Your offer is removed and new ones can be made.

2. Your offer enters the order book:

   - If your new offer is at a price above the best current offer of the buyers, you will not trade yet.
   - Instead, your offer will enter the order book. In the order book, all current offers of all buyers and sellers are collected and shown to all buyers and sellers.
   - If your offer is the lowest current offer among all sellers, and a buyer makes an offer *above* your price, you will trade with this buyer at a price equal to your offer.
   - You can always decide to adjust your offer.

After you have traded a unit, you can submit new offers to trade additional units. Note that your cost for trading the next unit may be different.

End of trading

Trading ends automatically after 8 minutes.

3.H.3 Page 3: Market instructions

Additional details

- On the top of the trading screen you always see the remaining trading time.
- We will also inform you about how many units you have traded, as well as the price and profits for these units.
- You will see all offers currently in the order book. The most favorable offers are ranked highest in the order book, which are the highest offered price by the buyers and the lowest offered price by the sellers.
- You also see the prices of the last four concluded trades.
- Note that offers are executed at the time they arrive.
- No trader knows with whom in the room he or she has traded. That means that your anonymity is ensured.

Payment

If this part and this market is selected for payment, for each trade a participant concluded, his or her payment is calculated with the rules described earlier. That is, for each unit, you will be paid the difference between the price and the cost for this unit. In this market, the earnings for each participant are then given by the sum of earnings for all units traded by the participant.

In part 2, there will be a total of 5 market periods (1 period for market 1 and 4 periods for market 2). 2 out of the 5 market periods will be randomly selected to be paid.

3.H.4 Page 4: Practice questions

NOTE: ALL NUMBERS IN THE QUESTIONS ARE ARBITRARILY CHOSEN, AND ARE NOT RELEVANT FOR THE EXPERIMENT.

Please answer the following questions:

1. Each seller pays the same costs as any of the other sellers to supply each unit. true/FALSE
2. Trading automatically ends after 8 minutes. TRUE/false
3. We will ask you several questions about the scenario below. Note that the behavior in this scenario is randomly determined, only for the purpose of asking these questions.

- The market begins, and the two sellers as well as one buyer submit an offer.
- First, the two sellers submit an offer. Seller S1 has a cost of 60 cents and submits an offer with a price of 200 cents. Seller S2 has a cost of 130 cents and submits an offer with a price of 190 cents.
- All market participants see the two offers in the order book. As seller S2'S offer is more favorable, it will be shown first in the order book.
- Next, one buyer submits an offer: Buyer B1 submits an offer with a price of 210 cents.
- As B1's offer is higher than S2's offer, B1 immediately trades with S2. They will trade at the price offered by S2.
- Both B1's and S2's offers are removed from the order book and trading can continue.

Please calculate the earnings of S1 and S2 at this point in the market:

How many cents does seller S1 earn? [FREEFORM: 0] cents

How many cents does seller S2 earn? [FREEFORM: 60] cents

[MARKET 1 TAKES PLACE, AFTERWARDS INSTRUCTIONS FOR MAR- KET 2 (with externality) FOLLOW]

3.H.5 Page 1: Part 2

This concludes market 1. Now, trading in **market 2** begins.

Generally, the same rules apply in this market. We will therefore highlight here only the differences between the two markets:

- Trading behavior in this market determines an amount of money that will be donated to UNICEF, in addition to your own earnings. The number of units successfully traded in the market is used to calculate how many doses of measles vaccines will be donated to UNICEF. The maximum number of doses donated to UNICEF in one market period is 60. The more units are traded in the market, the less will be donated to UNICEF: for each unit that is traded in market 2 that is selected for payment, 4 doses of measles vaccines will be subtracted from the donation to UNICEF, which cost approximately 1.5 Euro. Recall that with two doses of measles vaccine, one child can be protected. UNICEF will be paid the donation amount at the end of the experiment. The following table summarizes how the number of traded units in the market translates into the number of MEASLE DONATIONS. For example, if at the end of the market, zero units have been traded, then a total of 60 doses are donated to UNICEF for this market. If at the end of the market 3 units have been traded then in total 48 doses are donated. Donations to UNICEF are only affected by the overall number of units traded in the market and not by whom these units are traded.

- Final number of units traded and number of doses: [TABLE WITH NUMBER OF UNITS AND CORRESPONDING DONATION AMOUNTS, from 15 units traded (0 donations) to 0 units traded (60 units donation).]
- While market 1 only lasted for 1 period, you will now be trading in a sequence of 4 market periods. Each market period is conducted in the same way. Your choices in one period have no consequences on any other period.

Payment

If this part is selected for payment, two of the market results are randomly selected for payment. It is equally likely that each one of the 4 market periods of market 2 or the one period in market 1 is selected for payment. Payment for participants are then calculated according to the same rules as in market 1.

If a market period of market 2 is selected, the trades in the selected period also determine the amount donated to UNICEF. At the end of the experiment, the experimenter will transfer this amount.

3.H.6 Page 2

[REPEATED INFORMATION ON UNICEF, SEE INSTRUCTIONS FOR PART 1]

3.H.7 Page 3: Practice questions

NOTE: ALL NUMBERS IN THE QUESTIONS ARE ARBITRARILY CHOSEN, AND ARE NOT RELEVANT FOR THE EXPERIMENT.

Please answer the following questions:

1. If this part is selected for payment, two market results are randomly selected for payment. These can be market 1 or one of the market periods of market 2. Each trader earns the sum of cents generated by all of his or her trades. TRUE/false
2. For each unit that is traded, how many doses of measles vaccines will be subtracted from the donation to UNICEF? [FREEFORM: 4] doses
3. We will ask you a question about the scenario below. Note that the behavior in this scenario is randomly determined, only for the purpose of asking the question.

- The market begins, and one seller as well as one buyer submit an offer.
- First, one seller submits an offer. Seller S1 has a cost of 130 cents and submits an offer with a price of 300 cents.
- All market participants see the offer in the order book.
- Next, one buyer submits an offer: Buyer B1 submits an offer of 350 cents.
- As B1's offer is higher than S1's offer, B1 immediately trades with S1. They will trade at the price offered by S1.
- As one unit was traded, four doses of measles vaccine are *not* donated to UNICEF.
- Both B1's and S1's offers are removed from the order book. Trading can continue.

Please calculate the earnings of S1 at this point in the market:

How many cents does seller S1 earn? [FREEFORM: 170] cents

CHAPTER 4

Persuading an audience: Testing information design in the laboratory

This chapter is based on Ziegler (2023).

4.1 Introduction

Senders frequently speak to an audience of multiple receivers. For example, governments communicate with their citizens, and the leadership in private organizations addresses their employees or customer base. I focus on the sender's key choice of communication channel. The sender may employ *public* announcements, in which information is jointly revealed to all receivers. Alternatively, the sender may rely on *private* messages to individual receivers. In practice, senders often employ public communication strategies to convince their audiences to take a desired action. For instance, governments hold public press conferences, and central banks ensure that market participants can access their communication.[1] In other settings, private messages can be advantageous—for example, when route-planning services such as Google Maps or Waze recommend routes to their customers.[2] Miscoordinated routes minimize average travel times by reducing congestion; to ensure that drivers stay on their designated paths, recommendations to others are kept private.

Using a laboratory experiment, I provide the first empirical evidence on whether choosing the right communication channel helps a sender persuade her audience and what role the audience members' strategic interaction plays in that decision. As in Bayesian persuasion (Kamenica and Gentzkow, 2011), the sender can access superior information about the state of the world. In contrast to Bayesian persuasion, the sender communicates with an audience of multiple receivers. The presence of other receivers in the audience may affect how persuasive different communication channels are. Theoretically, the receivers' strategic interaction determines whether private signals or public announcements are a more effective tool of persuasion, a prediction from the literature on information design (for example, Bergemann and Morris, 2019). This strategic interaction is essential, because a receiver's optimal action frequently depends on other receivers' actions, as in the examples above. This interdependence implies that the sender may benefit from tailoring messages to the audience's strategic interaction.

In particular, I introduce coordination and miscoordination motives into the audience members' strategic interaction. To capture coordination motives, the receivers' strategic environment features strategic complementarities. Each receiver's incentive to choose an action increases in the number of other receivers choosing that action. With these complementarities, public messages are predicted to improve persuasion. A public message encourages all receivers to choose an identical action. Common actions reinforce incentives to select that action,

[1] For example, central banks may want to tailor their communication to be commonly understood by the general public in order, for example, to anchor inflation expectations (Haldane and McMahon, 2018; Binder, 2017; Bholat, Broughton, Parker, Ter Meer, and Walczak, 2018; Haldane, Macaulay, and McMahon, 2021).

[2] See Das, Kamenica, and Mirka (2017) for a theoretical analysis.

and observing everyone's recommended actions increases incentives to choose the favored action by minimizing strategic uncertainty. To capture miscoordination motives, the receivers' environment features strategic substitutes; that is, each receiver's incentive decreases in others' choice of the same action. In this environment, private messages are predicted to perform better. Each receiver is encouraged to take a potentially different action and does not observe other receivers' messages.[3] By miscoordinating actions and withholding information about the state from some receivers, persuasion can induce the favored action more frequently.

To create exogenous variation in the communication channel and strategic environment, I study persuasion in a laboratory experiment. In the field, researchers only observe the receivers' response to the communication channels that the senders select, which are often public. By contrast, the laboratory setting allows me to vary the communication strategies exogenously, to control each receiver's information set, and to hold constant other features that affect a sender's persuasiveness, such as her reputation. I can measure the importance of the channel choice, which allows me to evaluate whether any practical constraints on the channel choice, such as a legal requirement to use public communication or an inability to ensure public dissemination of signals, limit a sender's persuasiveness. In addition, in the field, it is difficult to assess whether the success of persuasion is affected by the strategic environment. Many communication settings feature interacting audience members, but settings and audiences vary not only in strategic interaction.

The experiment is designed to test the theoretical rationale for using either channel, which I can disentangle from other forces driving receivers' responses. These other forces turn out to be important, as using public signals increases persuasion for reasons not yet captured theoretically. Public communication results in less noisy behavior because it has a simpler, symmetric structure, and it is advantageous given receivers' aversion to differential treatment through private signals.

I employ two experiments that build on an investment game introduced by Bergemann and Morris (2019). In that investment game, the receivers choose whether to invest without knowing whether the state of the world is good or bad. A receiver wants to match the state. In addition, a receiver's payoff depends on the choice of the other receiver, creating room for strategic complements or substitutes. Without information beyond the prior, investment is not profitable for receivers. Investment is attractive in the good state, yet receivers, on average, make a loss when investing without additional information about the state. This creates scope for persuasion. I assume that the sender wants to persuade receivers to in-

[3] The strategic tension between the sender's and the receivers' interests means that private signals cannot be revealed publicly. If both receivers have access to the private information revealed to each other, the sender can no longer exploit her superior information about the state. I discuss this feature in more detail in Section 4.3.

vest, irrespective of the state. As in Bayesian persuasion (Kamenica and Gentzkow, 2011), the sender communicates by committing to an information structure. The signals are action recommendations that reveal information about the state and others' signals. The receivers need to judge whether they can trust a sender's recommendation. This decision depends not only on their own inference but also on their beliefs about others' information processing and decisions, which I elicit in the laboratory.

In the first experiment, I focus on receiver behavior. Computerized senders recommend actions to two participants in the role of receivers. I vary three treatment dimensions. First, I vary whether the game features strategic complements or substitutes. Second, I vary whether the information structure uses public or private signals. Third, I vary how aggressively the sender persuades the receivers by varying how often they receive a recommendation to invest in the bad state. Higher probabilities of this recommendation decrease expected gains from following recommendations. Formally, this varies whether an information structure satisfies obedience constraints, which measure whether a receiver can best respond by following recommendations. I test three levels of aggressiveness, where expected payoffs from following are held constant at each level, and two levels satisfy obedience constraints for risk-neutral receivers. By comparing following rates in the three levels, I test whether obedience is predictive of behavior. These constraints are widely used theoretically, but, to the best of my knowledge, this paper is the first to test them empirically.

Comparing public and private communication, I find that a channel's persuasiveness depends on the strategic environment, but I also observe surprising deviations from the predictions. In particular, I find that public structures perform well in a broader sense than expected. I observe the theoretically predicted advantage of public structures in settings with strategic complements. The empirical benefit even exceeds the theoretically predicted wedge. With strategic substitutes—a setting in which private signals are predicted to enhance receivers' persuasion—both public and private platforms perform equally well. Empirically, receivers are less willing to follow private than public recommendations. Interestingly, they anticipate this effect, as they believe other receivers follow public recommendations more frequently than private ones. Senders thus benefit in ways not captured by existing theory from using public signals, providing a justification for the frequent use of public communication in practice.

Two mechanisms drive the empirical superiority of public signals. First, the receivers' behavior exhibits more variance than predicted in response to private signals. Therefore, there is less additional unintended variation with public signals. The noise specific to private structures adds uncertainty about others' behavior beyond what is deliberately introduced by the sender and beyond what is optimal to persuade receivers. Hence, the receivers' best response is to follow private recommendations less often, which decreases persuasion. The additional noise with private signals is consistent with their complexity. Only with private signals do the receivers have to reason through the uncertainty about which recommendations others have received. As a second mechanism, I show that whether the signals are public or private affects the receivers' reaction to experiencing *bad advice*. Here, bad advice is defined as the recommendation to invest in the bad state against the receivers' interest. With private signals, bad advice is sent to only one receiver, while the sender recommends that the other receiver not invest, a form of differential treatment. In contrast, both receivers receive a common (bad) recommendation with public signals. I find that only receivers who receive bad advice with private signals subsequently reduce their investment. This pattern is consistent with receivers disliking this differential treatment.

While not capturing the benefits of public persuasion, theory otherwise predicts behavior well. Depending on the information structure's aggressiveness, 78% to 90% of recommendations that theory predicts will be followed are indeed followed. In contrast, recommendations that cannot always be followed in equilibrium are followed only in 66% of periods. Therefore, the obedience constraints organize receivers' behavior well. Using data on beliefs, I show that the decisions to follow are consistent with the theoretically predicted mechanism: Receivers update their beliefs reasonably well, and signals are processed close to the theoretically predicted way. Beliefs about the state show some conservative updating but evolve in line with Bayesian predictions. Furthermore, participants have a good understanding of the average response of other receivers to different signals. Even more striking is that given receivers' beliefs, their decisions are close to their best response, especially so with public persuasion.

In a second experiment, computerized senders are replaced by human senders. The senders are incentivized to maximize receivers' investment and choose among the same information structures that were exogenously assigned in the first experiment. They choose between different levels of aggressiveness in persuasion and between public or private signals. Between subjects, I vary whether the receivers' game features strategic substitutes or complements.

This experiment allows me to test whether receiver behavior in response to endogenous choices by participants is different from that under computerized communication strategies. This is an important distinction, as senders deliberately choose communication strategies in practice and their intentions may affect receivers' responses. For example, motives such as reciprocity or an aversion to being deceived may change receiver behavior, which in turn may affect the sender's optimal communication strategy. Empirically, I find little evidence for changes in receiver behavior across the two experiments. Recommendations are followed slightly less often, but this change is similar across both games and all information structures.

This experiment also allows me to study how participants in the role of senders persuade. This is important for two reasons. First, it means I can test whether senders adapt their choice to the receiver's strategic environment: do senders use public signals more frequently with strategic complements and private ones with substitutes? Second, it means I can assess whether senders foresee and react to the empirical superiority of persuasion with public signals.

I find that senders employ public signals in 55% of periods. Crucially, they respond to the receivers' strategic interaction: they use public signals more frequently in games with strategic complements than in games with strategic substitutes. Senders apparently exploit both the theoretically predicted benefit specific to each game (as they use public signals more frequently in settings with strategic complements) and the empirical advantage of public signals (as they use public signals more frequently when pooling data across the two settings). Senders' beliefs indicate that they anticipate that receivers respond to a change in communication strategies. However, they underestimate how strongly receivers react to changes in communication strategies, which leads them to not fully capitalize on the potential gains from public signals.

In the experiment, senders persuade quite forcefully. The senders' median choice is the sender-optimal structure, which maximizes their own self-interested payoffs at the receivers' expense; it is just obedient for risk-neutral receivers to trust these signals. If anything, senders err by being even more aggressive than what theory predicts will maximize their self-interested payoffs. While senders believe that more aggressive persuasion leads receivers to implement the sender's desired action less frequently, they do not fully account for the strength of the receivers' response. This aggressiveness in this complex environment, in which senders communicate by committing to an information structure, contrasts with findings from settings with more direct communication, such as cheap-talk games. In cheap-talk games, senders typically communicate more truthfully than if they were motivated purely by self-interest (Blume, Lai, and Lim, 2020; Abeler, Nosenzo, and Raymond, 2019).

In sum, I provide the first empirical evidence on the persuasion of audiences as modeled in the theoretical literature on information design. Along many dimensions, the behavior in the laboratory is consistent with the theoretical predictions. For example, in my empirical test of the theoretical concept of obedience constraints, choices are close to the predictions. Crucially, I find empirically that public messages help senders to persuade their audience in ways not yet captured theoretically. The messages' persuasiveness can be attributed to their simplicity, leading to less noisy behavior, and their equal treatment of receivers. Senders take advantage of the superiority of public signals.

In the following, I start by positioning the paper in relation to the literature. Section 4.3 describes the theoretical background, the theoretically motivated hypotheses, and the experimental design. Section 4.4 presents the results, and Section 4.5 concludes.

4.2 Relation to the literature

This study builds on a setup introduced by Bergemann and Morris (2019) within the literature on information design. Information design generalizes Bayesian persuasion (Kamenica and Gentzkow, 2011) to multiple receivers. In the laboratory, I test whether a sender can leverage strategic uncertainty by choosing an appropriate communication channel to enhance persuasion. Bergemann and Morris derive this insight on the channel choice in the investment game used in this experiment. Relatedly, a large theoretical literature compares public and private signals as well as different types of strategic interaction. For example, Angeletos and Pavan (2007) study welfare, Ely (2017) bank runs, Arieli and Babichenko (2019) information disclosure as in advertising, and Inostroza and Pavan (2021) stress tests. Taneva (2019) studies designer-optimal information design. Mathevet, Perego, and Taneva (2020) introduce an investment game and study adversarial equilibrium selection, and Taneva and Wiseman (2022) consider strategically ignorant receivers.

More abstractly, Bergemann and Morris (2016) introduce Bayes correlated equilibria.[4] These equilibria are widely used theoretically—for example, for informationally robust auction design (Bergemann, Brooks, and Morris, 2019; Brooks and Du, 2021). They build on obedience constraints, which require that receivers' best response is to follow recommendations. I am the first to study whether these constraints capture receiver behavior empirically. I focus on the question whether receivers' empirical response depends on specific information structures—for example, whether their response depends on the publicness of a signal.

Several strands of experimental literature are related to this study. First, a small recent literature tests Bayesian persuasion in the laboratory, such as Frechette, Lizzeri, and Perego (2022), Aristidou, Coricelli, and Vostroknutov (2019), and Au and Li (2018). These papers test persuasion of a single receiver, whereas I focus on games with multiple interacting receivers.

More closely related are two other strands of literature. The first studies other models of strategic information transmission experimentally, usually using cheap-

[4] Bayes correlated equilibria generalize correlated equilibria (Aumann, 1987) to games of incomplete information, see Forges (1993) for similar generalizations. Correlated equilibria have been tested in the laboratory—for example, by Van Huyck, Gillette, and Battalio (1992); Brandts and Holt (1992); Moreno and Wooders (1998); Cason and Sharma (2007); Duffy and Feltovich (2010); Bone, Drouvelis, and Ray (2013); Anbarci, Feltovich, and Gürdal (2018); Kurz, Orland, and Posadzy (2018); Friedman, Rabanal, Rud, and Zhao (2022). A connected line studies information transmission through mediators in the laboratory (Casella, Friedman, and Archila, 2020; Blume, Lai, and Lim, 2023). Unlike this literature, I study a sender that can not only correlate agents' play, but crucially has access to information about the uncertain state of the world, which she can use to persuade.

talk games (Crawford and Sobel, 1982). This literature focuses on when information about the state of the world is transmitted to and trusted by receivers. It typically finds overcommunication (see Blume et al. (2020) for a recent survey). In contrast to this large literature, I study the understudied setting with multiple interacting receivers, and I show that this interaction matters for a sender's optimal communication.[5] Theoretical work on communication with audiences began with Farrell and Gibbons (1989). Instead of capturing strategic interaction on the receivers' side, this literature focuses on receivers that differ in their degree of preference misalignment. The presence of multiple receivers may lead the sender to communicate less or more truthfully than in cheap-talk games with a single receiver, depending on whether the message is private or public. In experimental tests of this work, communication is more truthful with public signals (Battaglini and Makarov, 2014; Drugov, Hernán-González, Kujal, and Troya-Martinez, 2021).[6] A small literature on microtargetting study messages that target heterogeneity between receivers, compared to public messages common to all voters (van Gils, Müller, and Prüfer, 2022; Tappin, Wittenberg, Hewitt, Berinsky, and Rand, 2022).

Within this literature, more closely related are two papers that capture some elements of audience interaction. However, neither one captures how a sender can enhance persuasion by choosing channels optimally, nor do they systematically investigate audience members' strategic interaction. Agranov and Schotter (2013) study an announcement game in which a player in the role of the government can choose to reveal information about the state to its citizen-players. The authors focus both on what information about the state is revealed when the preference misalignment between the government and its receivers varies and on which natural language is used.[7] Cooper, Hamman, and Weber (2020) consider a cheap-talk game in which a leader encourages followers to choose an action. Both papers fix the strategic interaction of the audience members. In contrast, I show that both anticipating the receivers' interaction and communicating publicly can be beneficial to a sender. I contribute empirical evidence on why public messages are prevalent in practice, whereas theoretically the benefits of these public messages are limited to games of strategic complements.

The second closely related strand experimentally studies strategic uncertainty within global games (Carlsson and van Damme, 1993; Morris and Shin, 1998, 2002).

[5]A related literature compares behavior between games of strategic complements and substitutes (Fehr and Tyran, 2008; Potters and Suetens, 2009; Embrey, Mengel, and Peeters, 2019; Mermer, Müller, and Suetens, 2021).

[6]The only paper that investigates public signals in the field is Kapoor and Magesan (2014), who find that when public information generated from traffic light countdowns is observable by all participants, it increases accidents.

[7]Conceptually related is work on language barriers. Introducing uncertainty about others' ability to understand messages may impede the efficiency of communication (Blume and Board, 2013; Blume, 2018; Giovannoni and Xiong, 2019), mirroring the importance of common knowledge about others' signals to enhance persuasion with strategic complementarities.

Players in a game of strategic complements can receive private or common signals about the state of the world. These signals do not contain direct information about others' actions.[8] In contrast, I study a sender that attempts to coordinate agents' actions, which is a feature of many sender-audience interactions, such as governments' rhetorical interactions with their citizens. Furthermore, by comparing games of strategic complements and substitutes, I disentangle forces captured theoretically from behavioral forces that change behavior between public and private signals. Importantly, I focus on communication setups and ask whether a sender can exploit the audience members' interaction to persuade them. In experiments, behavior in the two types of information structures is more similar than theoretically predicted (Heinemann, Nagel, and Ockenfels, 2004, 2009; Cabrales, Nagel, and Armenter, 2007).[9] Trevino (2020) studies financial contagion between linked financial markets and finds that biases enhance contagion through traders' social learning, compared to contagion based purely on fundamentals. Avoyan (2022) allows agents in a global game to communicate, and Szkup and Trevino (2021) study information acquisition in global games.

4.3 Theoretical setup and experiment

In the laboratory experiment, I use an investment game devised by Bergemann and Morris (2019).[10] Here, I summarize key aspects of the theory underlying the experiment.

In this game, two firms simultaneously choose an action: to invest or not invest. Payoffs depend on both firms' actions. In addition, payoffs depend on the state of the world: $\theta \in \{\text{good}, \text{bad}\}$. Firms share the common prior of $Pr(\theta = \text{good}) = \frac{1}{2}$. Table 4.1 summarizes payoffs in the symmetric game, in which firm 1 is the row player and firm 2 the column player.

Here, x captures the payoff from investment in the good state, with $0 < x < 1$. ϵ characterizes the strategic interaction of the firms. When $\epsilon > 0$, the firms face strategic complements: their payoffs from investing compared to not investing are increasing if the second firm also invests. $\epsilon < 0$ implies strategic substitutes: payoffs from one firm's investment are decreasing in the second firm's investment.

In the experiment, I compare firms' behavior in a game with strategic comple-

[8] Related to this is the literature on sunspot equilibria, in which a sunspot realization serves as a correlation device. Coordination rates are higher than in the literature on correlated equilibria (Duffy and Fisher, 2005). Contrary to what theory predicts, both public and sufficiently correlated private signals generate sunspot equilibria (Fehr, Heinemann, and Llorente-Saguer, 2019).

[9] Cornand and Heinemann (2008) study theoretically to what extent signals in global games are optimally public. Experimentally, participants place a larger weight on a public signal over a private signal with stronger coordination incentives (Cornand and Heinemann, 2014).

[10] See Taneva (2019) on how to solve information design problems with common priors, as in this paper.

Table 4.1: Investment game

θ = good		Firm 2	
		invest	not invest
Firm 1	invest	$x+\epsilon$, $x+\epsilon$	x, 0
	not invest	0, x	0, 0

θ = bad		Firm 2	
		invest	not invest
Firm 1	invest	-1+ϵ, -1+ϵ	-1, 0
	not invest	0, -1	0, 0

ments to a game with strategic substitutes. Section 4.3.1 describes the parameterization and other details of how the game is implemented in the experiment.

Sender. In addition to the two firms, this setup includes a sender (or information designer) who commits to an information structure. Conditional on the state realization, she sends a signal—in particular, a recommendation to firms to either invest or not invest. The probability that she makes a particular recommendation may depend on the state, as in typical persuasion games. Additionally, it can depend on the recommendation the other firm receives. This allows the information designer to (mis)coordinate the firms' actions.

To study persuasion setups, I focus on senders that want to maximize receivers' investment across all states. In doing so, and in assuming that the sender is committed to an information structure, I connect to the literature on Bayesian persuasion (Kamenica and Gentzkow, 2011) and information design (Bergemann and Morris, 2019; Taneva, 2019). My main interest is in the receivers' strategic interaction and how this interaction affects the sender's optimal choice of channel; these are strategic elements that are also present with other communication protocols.

In the first experiment, the sender is computerized and the choice of information structure is a treatment variable. Receivers have no information on the sender's intentions. In the second experiment, participants in the role of senders are explicitly incentivized to maximize investment. They receive a payoff for each receiver that chooses to invest. The goal and payoff structure are known to the receivers. The sender can persuade the receivers to invest by committing to an information structure.

Information structures. Table 4.2 presents the notation for general information structures in this setup. Each cell gives the probability that, conditional on a given state, the row-column combination of action recommendations is sent to the firms. $p_\theta - r_\theta$ is the probability that each firm receives a separate recommendation to invest in state θ, and r_θ is the probability that both firms receive a simultaneous recommendation to invest in state θ.

For a sender, it is optimal to always recommend investment to both firms in the good state and thus to set $r_{\text{good}} = p_{\text{good}} = 1$. Investment is always profitable in the

Table 4.2: General information structures

θ = good	invest	not invest
invest	r_{good}	$p_{good} - r_{good}$
not invest	$p_{good} - r_{good}$	$1 + r_{good} - 2p_{good}$

θ = bad	invest	not invest
invest	r_{bad}	$p_{bad} - r_{bad}$
not invest	$p_{bad} - r_{bad}$	$1 + r_{bad} - 2p_{bad}$

good state. By maximizing investment in this state, the sender generates positive expected payoffs for receivers. This enables her to also sometimes recommend investment in the bad state, counterbalancing the gains in the good state with some expected losses in the bad state. This increases expected investment, as with the persuasion trade-off in Kamenica and Gentzkow (2011).

My focus, however, is on how the information structure's publicness affects persuasion. The information designer may use a public information structure by setting $r_{bad} = p_{bad}$ and $r_{good} = p_{good}$. In doing so, all firms always receive identical recommendations; messages are perfectly coordinated. Perfectly coordinating the signals generates common knowledge in the sense that both receivers know that they have received identical recommendations and have identical knowledge about the state. In the experiment, the receivers use the information structure to infer this perfect correlation. In practice, when persuading receivers to take an action, revealing information in a public announcement generates exactly the required common knowledge: all receivers are aware that this action has been recommended to each receiver.

Alternatively, the designer may use a private information structure. For example, she can set $r_{bad} = 0$ and $p_{bad} > 0$ in the bad state. Based only on the recommendation one firm received, this firm cannot infer with certainty what recommendation the other firm received. With a private information structure, firms' actions can be miscoordinated when the firms follow recommendations, as sometimes one firm invests while the other firm does not. Private signals feature two components: firms receive different signals and do not observe the other firm's signal. The definition of the private signals considered in this experiment, in which $r_{good} = 1$ and $p_{bad} - r_{bad} > 0$, clarifies why each receiver's private signal cannot be revealed to both receivers. Conditional on the state being bad, each firm receives the recommendation to invest with probability $p_{bad} - r_{bad}$. In that case, the other firm then receives the recommendation not to invest. If these two recommendations were revealed to both receivers, they would learn that the state is bad. In the bad state, the receiver can no longer best respond by investing. Therefore, when private signals are publicly revealed, the sender can no longer persuade receivers to invest in the bad state. The misaligned interests in the bad state between sender and receiver require that private signals remain private.[11]

[11]It might also not be in the receivers' own interest to reveal signals truthfully. Conditional on considering investing, a receiver wants the second receiver not to invest in games of strategic

Besides coordinating or miscoordinating firms' actions, a signal also transmits information about the state of the world, which a receiver can use to form a Bayesian posterior. Assume that a sender always recommends investment in the good state ($r_{\text{good}} = 1 = p_{\text{good}}$) and uses public signals that recommend investment with a probability of 50% in the bad state ($r_{\text{bad}} = p_{\text{bad}} = 0.5$). Conditional on receiving the recommendation to invest, the sender believes that the state is good with $Pr(\theta = \text{good}|\text{invest}) = \frac{Pr(\text{invest}|\theta{=}\text{good})Pr(\theta{=}\text{good})}{Pr(\text{invest}|\theta{=}\text{good})Pr(\theta{=}\text{good})+Pr(\text{invest}|\theta{=}\text{bad})Pr(\theta{=}\text{bad})} = \frac{.5}{.5+.25} = \frac{2}{3}$. Therefore, the firm learns that the state is more likely good than it believed before receiving the recommendation to invest. Given the new posterior, investment may now be profitable.

Obedience. Obedience constraints capture the degree to which a firm can trust an information designer and implement the recommended action.[12] Consider a firm receiving the recommendation to invest. It can use this recommendation to infer information about the state and about the action recommended to the second firm. By choosing the probabilities for each action recommendation appropriately, the information designer can ensure that firms' best response is to follow her recommendations. Following a recommendation is obedient if taking the recommended action is a best response; in that case, the Bayes Nash equilibrium is for both firms to follow. Knowing what is obedient allows the information designer to anticipate receivers' responses to different information structures. Then she can optimize over structures knowing firms' responses.

When a risk-neutral firm receives the recommendation to invest, obedience holds iff

$$\frac{1}{2}\underbrace{\left(r_{\text{bad}}(-1+\epsilon) + (p_{\text{bad}} - r_{\text{bad}})(-1)\right)}_{\text{Investment in the bad state}} + \frac{1}{2}\underbrace{\left(r_{\text{good}}(x+\epsilon) + (p_{\text{good}} - r_{\text{good}})x\right)}_{\text{Investment in the good state}} \geq 0 \qquad (4.1)$$

To verify obedience, receivers first use Bayes' rule (for compactness, I cancel out common terms in Equation 4.1). The right-hand side equals 0, as the payoffs from no investment are normalized to zero.

Theoretically, all obedient information structures capture the set of Bayes correlated equilibria (Bergemann and Morris, 2016). In this experiment, I determine whether this representation corresponds to game play in the laboratory or whether some equilibria are easier or more difficult to induce than others.

For each information structure, games of strategic substitutes feature a unique equilibrium, while games of strategic complements generally feature two equilibria. I discuss equilibria for the parameters and information structures in the

substitutes and wants the second receiver to always invest in games of strategic complements. In the experiment, information cannot be shared.

[12] For a formal definition following Bergemann and Morris (2016), see Appendix Section 4.A.

experiment in Section 4.3.2.

4.3.1 Experimental implementation of the investment game

In the laboratory experiment, players face either strategic complements or substitutes. In addition, they face (i) either private or public information structures, and (ii) different information structures, which vary their expected payoffs from following recommendations. In the first experiment, these two characteristics are varied exogenously. In the second experiment, they are chosen by another participant in the role of the sender.

The games are parameterized and normalized such that all payoffs are non-negative. All payoffs are denoted in points, which are exchanged at a rate of one point for five cents.

Table 4.3 presents the payoffs when the firms face strategic complements. As in the general example, investing is profitable only when the good state materializes. Firms face strategic complements, as the firms receive higher payoffs when both firms simultaneously invest. For example, if firm 1 invests in the good state, its payoff increases from 180 points to 210 points if firm 2 switches from not investing to investing.

Table 4.3: Game with strategic complements

θ = good		Firm 2		θ = bad		Firm 2	
		invest	not invest			invest	not invest
Firm 1	invest	210, 210	180, 170	Firm 1	invest	100, 100	70, 170
	not invest	170, 180	170, 170		not invest	170, 70	170, 170

Table 4.4 presents the payoffs when the firms face strategic substitutes. As in the game with strategic complements, investment is only profitable in the good state. In contrast to that game, firms prefer that the other firm does not invest: firm 1's payoff from investing decreases when Firm 2 switches from not investing to investing.

Table 4.4: Game with strategic substitutes

θ = good		Firm 2		θ = bad		Firm 2	
		invest	not invest			invest	not invest
Firm 1	invest	210, 210	260, 170	Firm 1	invest	20, 20	70, 170
	not invest	170, 260	170, 170		not invest	170, 70	170, 170

Both states are equally likely ($Pr(\theta = \text{good}) = 0.5$). Without information beyond this prior, firms would not be willing to invest in this game, as expected profits from investing are negative. The information designer can persuade firms to invest by conditioning signals on the state.

This experiment's primary interest is in understanding how players respond to different information structures. To this end, players face different exogenously designed information structures in the first experiment. Here, the role of the information designer is computerized. The structures themselves are revealed to participants. Across all information structures, all players always receive the recommendation to invest in the good state ($r_{\text{good}} = p_{\text{good}} = 1$). Players then either face private ($r_{\text{bad}} = 0$) or public information structures ($r_{\text{bad}} = p_{\text{bad}}$). In the first experiment, this is varied between subject.

For each class of information structures (private or public), each player faces three different information structures. They vary players' expected payoffs from following recommendations. Two of the information structures are obedient for risk-neutral players. *Optimal* structures yield close to the highest possible investment frequencies and thus are optimal for an information designer maximizing investment. If both firms follow the recommendations, their expected gains are barely positive, with fewer than five points for each firm. *Low* structures feature a less frequent recommendation to invest in the bad state. This decrease in frequency increases expected gains from following the recommendations to at least 22 points per firm and leads to a comparatively low level of investment. Unlike the *optimal* structures, *low* structures are also obedient for moderately risk-averse receivers.

Finally, *high* structures frequently feature the recommendation to invest in the bad state. These structures are not obedient, as they too frequently feature the recommendation to invest. If both firms follow these recommendations, they expect to lose more than five points.

Table 4.5 presents parameters and the receivers' probabilities of investing in the Bayes Nash equilibrium with maximal investment.

Table 4.5: Treatment table: Information structures

	Complements				Substitutes			
	Public		Private		Public		Private	
	r_{bad}	Pr(invest)	$p_{\text{bad}} - r_{\text{bad}}$	Pr(invest)	r_{bad}	Pr(invest)	$p_{\text{bad}} - r_{\text{bad}}$	Pr(invest)
High	71%	0%	48%	0%	32%	58%	48%	62%
Optimal	48%	74%	34%	67%	23%	62%	34%	67%
Low	19%	60%	14%	57%	10%	55%	14%	57%

Notes: Treatment parameters within the information structures (r_{bad}, $p_{\text{bad}} - r_{\text{bad}}$) and the probability each firm will invest in the equilibrium with maximal following (Pr(invest)). The left panel shows parameters for games of strategic complements, the right panel for games of strategic substitutes. Across all information structures, firms always receive the recommendation to invest in the good state ($r_{\text{good}} = p_{\text{good}} = 1$). r_{bad} is the probability that firms receive the joint recommendation to invest in the bad state. $p_{\text{bad}} - r_{\text{bad}}$ is the probability that only one firm receives the recommendation to invest, while the other receives the recommendation not to invest, in the bad state. With public structures, only common signals are used: $r_{\text{bad}} > 0$, while $p_{\text{bad}} - r_{\text{bad}} = 0$. With private structures, firms never receive the common recommendation to invest in the bad state: $r_{\text{bad}} = 0$, while $p_{\text{bad}} - r_{\text{bad}} > 0$. Within each level of obedience—*high*, *optimal*, and *low*—I fix the expected profits from following recommendations, assuming that the other receiver follows. *Optimal* and *low* are obedient for risk-neutral receivers.

Fixing the level of obedience, I set parameters such that the private information structures are identical between games of strategic complements and substitutes. For example, at the *optimal* level, each firm receives the private recommendation to invest in the bad state with a probability of 34% in both games. When following, this leads to identical expected profits across the two games.

The strategic advantage of public structures in games of complements and the advantage of private structures in games of substitutes become evident in the difference between public and private structures within each level for each game. Within each level of obedience, I fix expected profits from following the recommendations and then calculate the implied probability of recommending joint investment to both firms. In games of complements, this is a higher probability than was the case with private structures. For example, at the *optimal* level, both firms receive the recommendation to invest in the bad state with a probability of 48%, instead of the 34% with private structures. Crucially, in both public and private structures at the *optimal* level, firms expect to gain about five points if both firms follow. In games of substitutes, the probability of investment with public signals is lower than the probability with private signals. Again at the *optimal* level, firms receive the public signal to invest in the bad state with a probability of 23%, while they receive a private signal to invest in the bad state with a probability of 34%.

By fixing expected payoffs from following within each level (*low*, *optimal*, or *high*), play across the different structures (public versus private) becomes comparable. Signals are not equally informative across public and private signals, as the probability of the recommendation to invest in the bad state is changing.

In the second experiment, participants take on the role of the information designer. They receive a fixed payoff of 90 points each period and earn an additional 100 points for each receiver that chooses to invest. The senders choose among the six information structures that are used in the first experiment. Their choice thus entails two dimensions: Should they use a public or private information structure to persuade receivers? And which of the three levels of obedience should they use to maximize investment? After choosing a structure, the choice is revealed to participants jointly with the computer-generated signal.

4.3.2 Equilibria: Characterization and multiplicity

Conditional on choosing a particular information structure, these games generally feature two equilibria for games of strategic complements and one equilibrium for the games of strategic substitutes.

In the case of strategic substitutes, following an obedient information structure (*low* or *optimal*) constitutes the unique Bayes Nash equilibrium for risk-neutral receivers. If a structure is not obedient, only a mixed-strategy equilibrium survives, in which both receivers only probabilistically follow the recommendation to invest.

In the case of strategic complements, one Bayes Nash equilibrium for obedient structures is to follow recommendations. Therefore, as with strategic substitutes, *low* and potentially *optimal* information structures feature an equilibrium with following receivers. In the second Bayes Nash equilibrium, both receivers never invest, thus do not follow recommendations to invest. If one receiver does not follow the recommendation with sufficient likelihood, the equilibrium with full following is not attainable with complements. This is the case because only simultaneous investment by both receivers generates the complementary payoffs, $\epsilon = 30$ points. Crucially, this payoff is anticipated by the sender in calculating obedience, and receivers might no longer expect to gain from following recommendation if this payoff is not realized. This introduces another reason to potentially choose *low* structures: if receivers believe that others' best respond only noisily, it may no longer be a best response to follow in *optimal* information structures even for risk-neutral receivers. It is of theoretical interest in the literature on information design which of these equilibria prevails; for example, Mathevet et al. (2020) discuss sender-adversarial equilibrium selection. In the case of non-obedient information structures, the games of strategic complements feature only the equilibrium of not following.

When analyzing the experimental data, I use the equilibrium with the highest investment as a benchmark and compare data to this benchmark. This is the sender-preferred equilibrium and the unique equilibrium in games of substitutes. This equilibrium turns out to be a closer fit to the data than the alternative equilibrium with no investment in games of complements.

4.3.3 Theoretical predictions

In the first experiment, I test two dimensions central to the theory. First, I study the strategic advantage of public (private) structures in cases with strategic complements (substitutes).

Prediction 1. *Private structures induce more investment than public structures with strategic substitutes. Public structures induce more investment than private structures with strategic complements.*

The setup in this experiment captures the above predictions, which are typical in the information design literature. Table 4.5 illustrates the advantage of either public or private structures with the parameters of this experiment, within each level of obedience. With strategic complements, investments can be maximized with public signals; with strategic substitutes, private signals induce more investments than public signals.

Second, I test whether obedience captures empirical responses to information structures. Based on the expected profits, following is expected to be strongest in *low* levels. Following in *optimal* levels is equal to or lower as in *low* levels.

The ranking of *low* and *optimal* depends on receivers' risk aversion: risk-neutral receivers follow in *optimal* structures; however, sufficiently risk-averse receivers follow only in *low* structures. The least amount of following is expected in *high*, levels, in which the choice to always follow does not constitute a best response.

Prediction 2. *The frequency of following recommendations is characterized by the following ranking:*

$$low \geq optimal > high$$

Theoretically, the information designer anticipates the receivers' responses across different information structures. She can use these responses to choose structures advantageous to herself. However, empirically, play may differ. As a first step, players need to update their beliefs and comprehend that the information released in the recommendation is valuable. As a second step, players must choose accordingly and understand that following obedient information structures is profitable. What makes this setup particularly interesting is the inferences players make about others' behavior. Obedience relies on the common knowledge of players following recommendations.

In the second experiment, I focus on the information designers' choices. To maximize their own expected payoffs, if senders assume that the receivers are risk neutral, they can choose the information structure that maximizes receivers' expected investment. The first way they can do so is by exploiting the channel that theoretically enhances persuasion in each game.

Prediction 3. *In games of strategic complements, information designers choose public structures more often than they do in games of strategic substitutes.*

Second, payoff-maximizing senders choose the level of obedience that maximizes the level of investment conditional on receivers following:

Prediction 4. *Information designers choose structures according to the following ranking:*

$$optimal > low > high$$

4.3.4 Experimental design

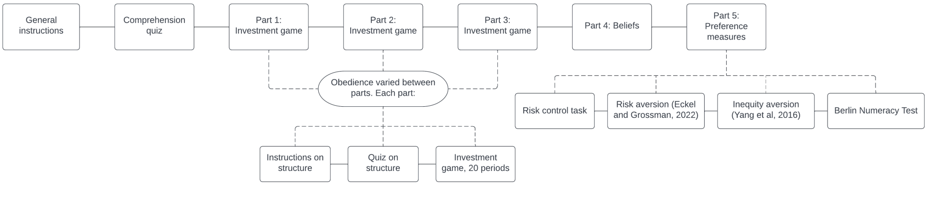
The experiment closely follows the theoretical setup, except that in the laboratory, the game is framed as two workers' decision to work or not work, not two firms' decision whether to invest. Each player's payoffs depend on their own decision and the decision of their coworker. In the first experiment, a computerized manager recommends actions, while in the second experiment this role is played by a participant. Information structures are implemented as a recommendation plan, according to which the workers receive recommendations. The state in the investment game is implemented as the randomly determined difficulty of the project, which is called difficult or easy.

At the moment that receivers decide, the screen summarizes the recommendation they received, the game, and the recommendation plan. After their decision, the state and the recommendations are revealed, participants learn their and their coworker's payoff and, in the second experiment, the manager's payoff. In addition, they learn what payoff they would have received if they had chosen the alternative action. In the second experiment, the sender's decision screen summarizes, for each available information structure, how frequently receivers in their matching group invested and followed recommendations in earlier periods.

First experiment. In the first experiment, I vary two between-subject treatment dimensions: (i) whether the strategic interaction of the receivers features complements or substitutes and (ii) whether the information structure that receivers face uses public or private signals.

Participants first receive general instructions on the investment game and have to pass a comprehension quiz. The investment game is played in three parts, with 20 periods per part. In each of these parts, players face one of the three levels *low*, *optimal*, and *high*. This treatment dimension, the third, varies within subject and with a counterbalanced order. At the beginning of each part, players first receive specific instructions for the new information structure and a comprehension quiz. Figure 4.1 shows the timeline of this experiment. Participants are allocated to matching groups of six participants, with random rematching every period.

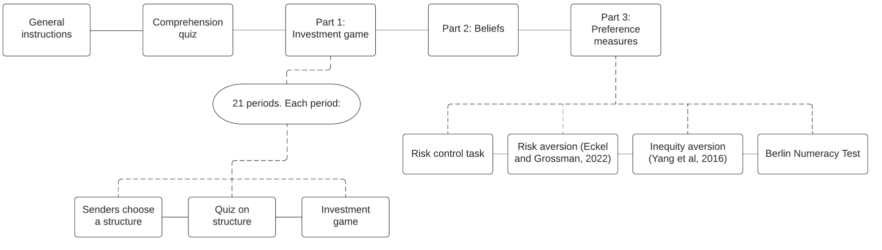
Figure 4.1: Timeline in the first experiment



Second experiment. In the second experiment, participants again first receive general instructions. For receivers, these are instructions similar to the first experiment, but they include some additional instructions on the senders' choice set and incentives. For senders, these instructions fully describe their own and receivers' decisions. Both senders and receivers have to pass a comprehension quiz afterward. During the experiment, senders also receive information about the receivers' responses to the information structures the senders in their matching group chose earlier. In this experiment, I vary only one between-subject treatment dimension: whether the receivers' strategic interaction features strategic complements or substitutes. To persuade receivers, the senders choose among the six different information structures that are varied exogenously in the first experiment. As in the first experiment, the information structure is revealed to the receivers.

This investment game is played only in one part, with 21 periods. Each period, receivers also have to answer one randomly selected question from a comprehension quiz similar to the quiz in the first experiment. Figure 4.2 shows the timeline of this experiment. Participants are allocated to matching groups of nine participants, with three senders and six receivers, with random rematching every period.

Figure 4.2: Timeline in the second experiment



Additional elicitations. The experiment concludes with measurements of beliefs and participants' characteristics. In both experiments, I elicit participants' beliefs induced by the information structures. I elicit beliefs about whether the state is good and whether the other participant decides to invest—once for other participants that receive the recommendation to invest, and once for other participants that receive the recommendation not to invest. Participants predict in how many of 10 randomly drawn decisions the state was good and in how many decisions others invested, conditional on those participants having received the recommendation to invest or not invest. In the first experiment, this generates a set of 12 reports, 4 for each of the three levels of obedience. Out of the 12 reports, 1 is randomly drawn to be paid out. In the second experiment, beliefs for all six structures (combinations of public versus private and the three levels of obedience) are elicited and again 1 report is randomly paid out. If their report matches the actual value for 10 randomly selected instances, they receive a payment of 40 points in both experiments.

Second, I elicit participants' choices in an individual decision-making transformation of the investment game. The transformation strips away the strategic aspect of the game. By comparing choices between the two environments, we learn about the importance of these strategic aspects. Within each level of obedience, all structures and games generate equal expected payoffs. However, structures and games differ in their riskiness. In particular, the payoffs from investment in the bad state differ between public and private structures as well as between games. The probabilities of the bad state, conditional on receiving the recommendation to invest, counterbalance the difference in the payoffs. This preserves expected payoffs but affects the variances of payoffs. For example, joint investment with complements pays 100 points with public signals, while separate investment with private signals only pays 70 points. This payoff difference is offset by recommending investment in the bad state with a probability of 48% with public signals but only 34% with private signals. To measure whether individuals change their behavior in individual decision-making in agreement with the patterns I observe in the investment game, I introduce an individual control task. To generate this task, I use the investment game and associated information structure. Then, I assume that the second receiver follows recommendations, which removes the strategic element of the game. I compute expected payoffs from following a recommendation to invest for the game and for all information structures that each participant faces in the experiment, once conditional on the bad state materializing and once conditional on the good state materializing. The required probabilities of either state occurring are defined by the Bayesian posterior for the good and bad state materializing, conditional on the recommendation to invest. With the implied posterior probability, the good state materializes or the bad state materializes. In the experiment, the decision is framed as a lottery choice. The participants can choose a safe payoff, calibrated to match the payoff from no investment in the investment game. Alternatively, they can choose a risky payoff. This leads to a gain corresponding to the expected profit from investment in the good state, with the Bayesian posterior of the good state occurring when investment is recommended. With the remaining probability, this leads to a loss corresponding to the expected loss from investment in the bad state.

Third, I elicit risk preferences using the Eckel and Grossman (2002) task. Fourth, I elicit the parameters of the Fehr and Schmidt (1999) model for inequity aversion using the task in Yang, Onderstal, and Schram (2016). Fifth, participants' skills in understanding statistical information and risk are measured using the Berlin numeracy test (Cokely, Galesic, Schulz, Ghazal, and Garcia-Retamero, 2012).[13] Screenshots of all instructions are presented in Appendix Section 4.C.

Experimental procedures. Both experiments, hypotheses and all analyses are preregistered at the AEA RCT registry (Ziegler, 2021, 2022). Experimental payments are exchanged at a rate of one point for five cents. In Appendix Section 4.B.1, I provide balancing tables for both experiments. Treatments across all experiments are balanced, apart from Aheadness aversion in the second experiment (*p*-value=0.097). Controlling for this measure does not affect the results.

[13] In the second experiment, only the first and third questions are used.

The first experiment was conducted in March 2021. Due to COVID-19 restrictions, the experiment was conducted online using a standard laboratory sample. The participants were recruited from the traditional subject pools of CREED at the University of Amsterdam in the Netherlands and MELESSA at LMU Munich in Germany, with the participants at MELESSA using ORSEE (Greiner, 2015). Both laboratories frequently conducted online experiments at that time, and protocols for running them online were in place. Besides the computerized experiment, participants were required to join a Zoom meeting with the experimenter. Participants were anonymized in the meeting and could only chat with the experimenter. This allowed close monitoring of potential problems, and participants could ask questions as in regular laboratory sessions. To verify their identity, participants either received a personalized link (at MELESSA) or had to verify their identity by taking pictures of themselves and their student ID using their webcams. Images were stored separately and deleted immediately after the sessions. Payments were implemented using bank transfers. Participants recorded their IBAN (and never their names or any other personal information) either in separate surveys (LimeSurvey at MELESSA) or in separate parts of the experimental software (at CREED). Almost all participants finished the experiment: out of 432 participants, only 1 participant dropped out (because of technical problems). This participant made 48 out of 60 decisions in the first three parts.

In the first experiment, payments were given for two randomly selected periods, each from a different randomly selected part. In total, 432 participants joined for 1 of 18 sessions, 288 of them being registered at CREED. Each session consisted of three to five matching groups, with six participants per matching group. The average age was 22.7 years. 249 out of the 432 participants were women; average earnings were 26.3 euros; and sessions took on average 82 minutes.

The second experiment was computerized and conducted in person in August and September 2022, in the laboratories of CREED in Amsterdam and MELESSA in Munich. In total, 360 participants joined for 1 of 22 sessions, 225 of them being registered at CREED. Participants received payments from two randomly selected periods. They were paid out in cash in all sessions apart from three sessions at MELESSA, which used the same payment procedure as the first experiment. Each session consisted of one to four matching groups, with nine participants per matching group. The average age was 22.6 years. 202 out of the 360 participants were women; average earnings were 26.9 Euros; and sessions took on average 99 minutes.

4.4 Results

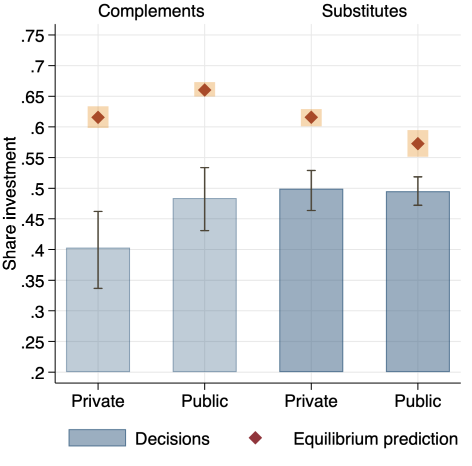
This section presents the results of the experiment. I focus on the first experiment first. The data from that experiment allows me to study receivers' behavior in different games and in response to exogenously assigned information structures.

4.4.1 Investments

The experiment was set up to measure whether receivers can be persuaded to invest. The measure of investment share is shown in Figure 4.1. Unless otherwise noted, all figures compare data on the two obedient levels (*low* and *optimal*) to ease interpretation, as this holds constant the existence of an equilibrium with full following. For regressions, I pool all data. Results are robust to using either approach.

The red diamonds illustrate equilibrium predictions. For strategic complements, theory predicts higher investment in public than in private structures. For strategic substitutes, theory predicts higher investment in private structures than in public ones.[14]

Figure 4.1: Investment decisions



Notes: Average frequency of investment by treatment, bars indicate observed choices and red diamonds choices in the Bayes Nash equilibria with the highest investment. The figure only uses data from *low* and *optimal* structures. Bars and shaded areas indicate 95% bootstrapped confidence intervals.

Overall, investment rates are substantial, with an average investment of 47% across all treatments. Absent information beyond the prior, for both separate and joint investment, investing would not be profitable, as participants would expect

[14]The theoretical treatment effects shown in Figure 4.1 are comparatively small because these data are averaged across obedience levels (*low* and *optimal*). For the *low* level, theoretical differences are relatively small, while I chose parameters to generate large treatment differences for *optimal* information structures. For example, the theoretically predicted interaction effect of public versus private signals interacted with the game is 14.9 percentage points with *optimal* structures (see Appendix Section 4.B.2). I discuss parameter choices in more detail in Appendix Section 4.A.1.

to lose between 5 and 55 points. Therefore, the appropriate benchmark to evaluate whether persuasion succeeded is no investment. This benchmark is also consistent with the individual risk measurement discussed in Section 4.4.4.

Participants frequently invest when receiving the recommendation to do so. The high investment rate suggests that the participants trust the signals they receive and trust their fellow participants to make the same inference as themselves. This can be interpreted as a mark of successful information design, as persuasion frequently succeeds.

Trusting others to follow is most crucial in games of strategic complements. In these environments, investing is only profitable if other receivers are also investing. Yet receivers invest in only 44% of these cases. In contrast, in games of strategic substitutes, when others do not follow, it reinforces the incentive to invest. Consistent with this difference in strategic incentives, average investment frequencies increase to 50% in games of substitutes.

Nevertheless, even though always following is an equilibrium for risk-neutral receivers, overall investment is still below the predicted investment. Two forces contribute to this finding. First, participants' beliefs exhibit some conservatism in updating about the probability that the state is good when receiving a recommendation to invest, which decreases expected profits from investment. This feature is discussed in more details in Section 4.4.3. In addition, these predictions assume risk neutral receivers. However, empirically, many participants exhibit risk aversion in the two control tasks at the end of the experiment. Using estimates of risk aversion from these tasks in the equilibrium prediction captures that empirically, investment rates are lower, and partially even predict lower investment than observed. I discuss this exercise in Appendix Section 4.B.3.

These data are also informative about equilibrium selection in games of strategic complements. For the two obedient structures in these games, investment is predicted in 64% of cases in the equilibrium of maximal following. Thus, empirically, investment frequencies come closer to the equilibrium with maximal investment, and inducing this equilibrium is frequently successful. In addition, the difference between predicted and observed investment is to a large extent driven by the fact that only some receivers within each group are not willing to invest when they receive the recommendation to invest. If instead the equilibrium without investment drove the behavior of some groups and thereby explained the difference between predicted and observed investment, we would expect to see some groups with very low average investment and some with high average investment. However, even at the *optimal* level, we observe low investment, coded as average investment in at most 3 of the 20 periods, for only 4% of groups. This rareness is inconsistent with the possibility that a non-investment equilibrium is prevalent for some groups. While the alternative equilibrium without investment exists, this does not appear to limit the sender-optimal equilibrium's attainability.

Table 4.1 presents estimation results of the treatment effect. All columns compare investment behavior in the data (columns (1), (3), (5), and (7)) to the predicted behavior in the Bayes Nash equilibria with maximal investment (columns (2), (4), and (6)). To generate the equilibrium data, I use the recommendation draws from the experiment, and impose equilibrium following. Columns (1) and (2) compare data only within games of strategic substitutes, columns (3) and (4) only within games of complements. The key specifications are columns (5) and (6), which pool all data.[15] These specifications allow for a difference-in-differences interpretation between games and information structures. Column (7) only uses data from obedient information structures (r_{bad} and $p_{\text{bad}} - r_{\text{bad}}$ at *low* and *optimal* levels).

Strikingly, the comparative statics for public and private information structures reveal a surprising pattern and an advantage of public information structures in the data across all strategic environments. Private structures perform no better with strategic substitutes than public ones (coefficient of -0.009 on Public; p-value=0.643; column (1)). This contrasts with the equilibrium prediction of higher investment with private signals (coefficient of -0.043; column (2)). In games of strategic complements, public structures increase investments by 9 percentage points (p-value=0.034; column (3)). This is in line with the theoretical prediction that public signals perform well with strategic complements. However, the empirical treatment effect exceeds the theoretically predicted benefit of just 3 percentage points (column (4)).

Column (5) documents the interaction effect—moving from private to public signals and from games of substitutes to games of complements—which is the main effect of theoretical interest. Investment increases by 10 percentage points (coefficient on Public × Complements; p-value=0.035; column (5)) when using public compared to private signals and when moving between games. Again, this slightly exceeds the theoretically predicted increase of 7 percentage points (column (6)).

To show that investments increase with public structures compared to the empirical predictions, I interact models (5) and (6) and show estimates in Appendix Section 4.B.2. Across both strategic environments, the empirical advantage exceeds the predicted advantage by 3 percentage points (p-value=0.080). This effect does not differ between strategic environments (p-value=0.604). At the *optimal* level, at which parameters are chosen to maximize power, the difference between the empirical and the predicted effect of public structures increases to 6 percentage points (p-value=0.024), while it is again similar between environments

[15]The negative coefficient in column (6) on Complements is driven by *high* information structures. In that case, following one's recommendation does not constitute a Bayes Nash equilibrium. With complements, this implies no investment. A mixed-strategy equilibrium with partial investment arises with substitutes, where recommendations are followed only probabilistically. The coefficient is not significant in *optimal* and *high* structures, as the private structures across these two games are designed to be identical and recommend investment equally often. The maximal-investment Bayes Nash equilibria have both players always following these recommendations. Therefore, they induce equal investment.

(p-value=0.421).

Summarizing, I find both evidence for the game-specific advantage of public signals in games of strategic complements and evidence for the general advantage of public signals. For the latter, I find that public structures do not perform worse than private structures even with strategic substitutes. This suggests that, in practice, public messages appear to possess inherent advantages when persuading receivers.

In Appendix Section 4.B.2, I reproduce Figure 4.1 separately for all levels of obedience. As preregistered, I show that the analysis of Table 4.1 is robust to including controls, to using logistic regressions, and is similar over time in Appendix Section 4.B.4. This also holds when only studying part-one data, where all treatment dimensions, including the level of obedience, were assigned between subject.

Table 4.1: Treatment effects: Investment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Substitutes		Complements		Diff-in-Diff		
	Data	NE	Data	NE	Data	NE	Data
Public	-0.009	-0.043***	0.087**	0.030***	-0.009	-0.043***	-0.004
	(0.019)	(0.011)	(0.039)	(0.007)	(0.020)	(0.011)	(0.022)
Complements					-0.108***	-0.211***	-0.096***
					(0.032)	(0.009)	(0.034)
Public × Complements					0.096**	0.073***	0.085*
					(0.045)	(0.013)	(0.045)
(1 if level=*optimal*)	-0.009	0.082***	-0.040*	0.122***	-0.025*	0.102***	-0.024*
	(0.018)	(0.014)	(0.021)	(0.016)	(0.014)	(0.011)	(0.014)
(1 if level=*high*)	-0.038*	0.058***	-0.073***	-0.577***	-0.055***	-0.260***	
	(0.020)	(0.013)	(0.020)	(0.011)	(0.014)	(0.039)	
Constant	0.514***	0.574***	0.445***	0.562***	0.533***	0.674***	0.508***
	(0.022)	(0.010)	(0.038)	(0.013)	(0.022)	(0.014)	(0.022)
Period trend, part and lab FE	Yes	No	Yes	No	Yes	No	Yes
Only obedient signals	No	No	No	No	No	No	Yes
Observations	12960	12960	12948	12948	25908	25908	17268
# clusters	36	36	36	36	72	72	72
# participants	216		216		432		432

Notes: The table reports OLS estimates and includes all data. The dependent variable is a dummy variable equal to 1 if the participant decided to invest (Data) or was predicted to invest in the Bayes Nash equilibrium with maximal investment (NE). Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. (1 if level=*optimal*) and (1 if level=*high*) are dummy variables equal to 1 if the information structures used *optimal* or *high* probabilities to persuade receivers to invest, relative to the omitted category *low*, respectively. Standard errors in parentheses, clustered at the matching-group level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Result 1. *Public information structures induce higher investments than private structures with strategic complements, more than theoretically predicted. In contrast, private information structures do not induce higher investment than public structures with strategic substitutes, contrary to theoretical predictions.*

The regression results in Table 4.1 also reveal how investment changes in *high* and *optimal* information structures compare to those in *low* structures. Consistent with the theoretical prediction that *high* information structures are not obedient, we observe less investment in this treatment. This effect is, however, smaller than theoretically predicted, especially for games of strategic complements. This implies that receivers partially trust recommendations they should not trust in equilibrium. Investment decreases by 4 percentage points (p-value=0.071; column (1)) when receivers face a *high* structure with strategic substitutes. For this game, investments are even predicted to increase in equilibrium for *high* structures (coefficient of 6% for *high* structures; column (2)), which highlights the empirical importance of persuading not too aggressively. With strategic complements, investment decreases by 7 percentage points (p-value<0.001; column (3)) when receivers face a *high* structure, consistent with the conjecture that when others do not follow, it reduces the incentive for own investment.

In addition, *optimal* structures do not increase investment compared to *low* structures. This runs contrary to theoretical predictions when assuming risk-neutral receivers, as we expected an increase in investment (coefficient on *optimal* levels in columns (2) and (4)). Empirically, however, there is no significant effect for strategic substitutes (p-value=0.625; column (1)). For strategic complements, investment even decreases by 4 percentage points (p-value=0.068; column (3)). Some receivers are only willing to invest when substantial informational rents from following are available, consistent with some receivers' risk averseness. The next section discusses the following frequencies and obedience in more detail.

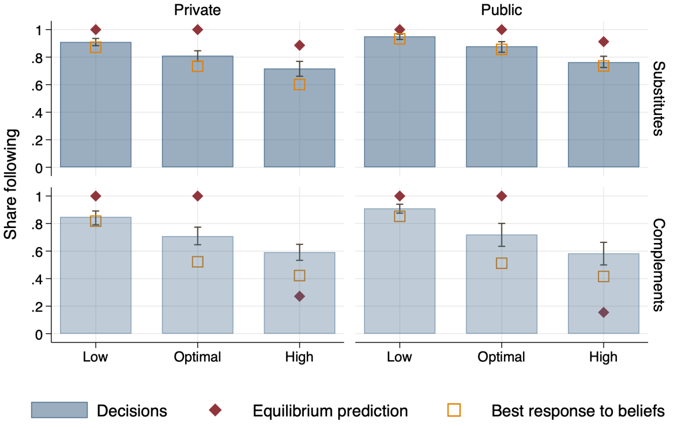
4.4.2 Following behavior

Participants face the critical decision of whether to trust and follow a recommendation. The investment behavior presented in Section 4.4.1 compounds two factors. First, how often is a recommendation to invest sent to receivers? Second, how often is this recommendation followed? As the former factor varies between information structures, focusing on the following behavior allows for a clean measure of receivers' responses to information structures.

Figure 4.2 presents average following behavior, differentiated by game, publicness, and information structure level. Following behavior is coded such that it is equal to 1 whenever a recommendation is followed (investing after the recommendation to invest, not investing after the recommendation not to invest), and zero otherwise. Table 4.2 reports accompanying regressions. Columns (1) and (3) use data, while columns (2) and (4) repeat the same analysis for predicted behavior in the equilibrium with maximal following. Columns (1) and (2) use data only from the obedient information structures (*low* and *optimal*), while columns (3) and (4) also use data from *high* structures. Column (2) reflects the equilibrium feature that all recommendations are followed in the equilibrium with maximal investment for obedient structures, as the estimate on the constant is one and there are no changes across treatment conditions.

Five facts emerge. First, receivers respond to the level of the information struc-

Figure 4.2: Following rates



Notes: Average frequency of following a recommendation by treatment and by the level of the information structure. The variable is a dummy equal to 1 if a recommendation was followed (investment after the recommendation to invest, or no investment after the recommendation not to invest). Bars indicate observed choices, diamonds indicate the following rate in the equilibrium with the highest following, and squares are empirical best responses based on participants' separately elicited beliefs. Error bars indicate 95% bootstrapped confidence intervals.

ture precisely as expected. Most following occurs with the strongest incentive to follow in *low* structures. The constant of 93% in column (1) indicates that in the baseline level (*low*), following is very prevalent and is close to the full following predicted in equilibrium in column (2). We observe intermediate levels of following for intermediate incentives in *optimal* structures. Compared to the omitted category *low*, following decreases by 13 percentage points in *optimal* structures (p-value<0.001; column (1)). Risk-neutral receivers are expected to respond equally to *optimal* and *low* structures; see column (2). Behavior in the laboratory is more nuanced, consistent with at least some risk-averse receivers. Last, there is the least following with the weakest incentives in *high* structures, with following rates 24 percentage points lower (p-value<0.001; column (3)).

Second, across most treatments, observed following is lower than in the Bayes Nash equilibria with the highest investment. For example, and not surprisingly, we can reject the null that there is full following for obedient structures (H_0: Constant=1; p-value=0.003; column (1)). Nevertheless, behavior is frequently in line with the sender-preferred equilibrium instead of the equilibrium with no investment, so we can reject the null that following in games of complements is canceling out the high baseline following in games of substitutes (H_0: Complements=Constant; p-value<0.001; column (1)). In addition, we observe more following than predicted in *high* levels. While following is predicted to decrease following by 44 percentage points with *high* levels (column (4)), following is observed to decrease by only 24 percentage points (p-value<0.001; column (3)). Therefore, some

Table 4.2: Treatment effects: Following

	(1) Data	(2) NE	(3) Data	(4) NE
Public	0.054***	0.000	0.052**	0.009***
	(0.018)	(.)	(0.021)	(0.003)
Complements	-0.082***	0.000	-0.096***	-0.205***
	(0.029)	(.)	(0.030)	(0.004)
Public × Complements	-0.018	0.000	-0.030	-0.048***
	(0.039)	(.)	(0.043)	(0.005)
(1 if level=*optimal*)	-0.125***	0.000	-0.125***	-0.000
	(0.015)	(.)	(0.015)	(0.000)
(1 if level=*high*)			-0.241***	-0.444***
			(0.015)	(0.041)
Constant	0.934***	1.000	0.956***	1.110***
	(0.022)	(.)	(0.024)	(0.014)
Level of obedience	*low* & *optimal*		*low*, *optimal* & *high*	
Period trend, part and lab FE	Yes	No	Yes	No
Observations	17268	17268	25908	25908
# clusters	72	72	72	72
# participants	432		432	

Notes: The table reports OLS estimates and includes all data. The dependent variable is a dummy variable equal to 1 if the participant decided to follow a recommendation (invest after recommended to invest, or not invest after recommended not to invest) (Data) or was predicted to follow in the Bayes Nash equilibrium with maximal investment (NE). Columns (1) and (2) use data only from obedient structures, while columns (3) and (4) pool all data. Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. (1 if level=*optimal*) and (1 if level=*high*) are dummy variables equal to 1 if the information structures used *optimal* or *high* probabilities to persuade receivers to invest, relative to the omitted category *low*, respectively. Standard errors in parentheses, clustered at the matching-group level, * p < 0.10, ** p < 0.05, *** p < 0.01.

receivers continue to follow the recommendation to invest in *high* levels.

Third, participants are more likely to follow recommendations from public information structures. Theoretically, this is surprising. The structures were designed to induce equal following in equilibrium for the obedient levels; see column (2). However, empirically, participants appear to trust private recommendations less than public ones, as following increases by 5 percentage points in public structures (p-value=0.004; column (1)). This feature drives the two key deviations from predicted investments reported in Section 4.4.1. First, with games of strategic substitutes, the higher following of public signals leads to similar investment rates across private and public signals. While private structures are more likely to recommend investment in the bad state, receivers' decreased following almost exactly cancels out this advantage. Second, with games of strategic complements, the increased following of public signals leads to the higher-than-predicted investment with public signals. Theoretically predicted effects are slightly different when including *high* structures in columns (3) and (4).[16] Nevertheless, the same pattern arises, as public recommendations are followed more frequently than theoretically predicted. In Section 4.4.4, I disentangle potential drivers of this effect.

[16]This arises because the mixed equilibria for *high* levels in games of strategic substitutes feature slightly different following probabilities across public and private structures. In addition, recommendations not to invest are predicted to be followed in games of strategic complements but are sent at different frequencies for public and private structures.

Fourth, games of strategic substitutes generate higher following behavior than games of strategic complements. Following frequencies decrease by 8 percentage points with complements (p-value=0.006; column (1)). This is in line with the conjecture that receivers anticipate the noisy behavior of fellow receivers; as in equilibrium there is no change between games (column (2)). In games of strategic substitutes, when other receivers do not invest when they receive the recommendation to invest, it increases incentives to follow recommendations to invest. When the other receiver does not invest, it generates larger payoffs for receivers driven by the gains from investing in the good state. In contrast, receivers in games of strategic complements need the other receiver to invest to make their own investment profitable, especially in the *optimal* structure. Given that other receivers are not always following recommendations, following all recommendations is no longer a best response for receivers with complements.

Fifth, and most strikingly, behavior overall is remarkably close to the behavior in a best response to participants' beliefs. For this best response, I use beliefs about the state and about others' behavior conditional on each recommendation, described in Section 4.4.3. These beliefs were elicited only at the end of the experiment, so they represent the beliefs of experienced participants. Based on these beliefs, I predict which recommendations should be followed by payoff-maximizing risk-neutral receivers. To do so, I predict expected profits of following recommendations given each receiver's beliefs, and I predict they follow recommendations if the expected profit exceeds the no-investment payoff of 170 points. Since behavior is close to this best response, participants apparently understand this game well. When accounting for their beliefs about the play of others, which may differ from behavior in the Nash-equilibrium benchmark, as well as when accounting for their potential non-Bayesian inference about the state, participants behave close to what standard theory would predict. In addition, behavior is closer to the best response in public structures, as reported in Appendix Section 4.B.5. This indicates that play is particularly sophisticated when participants face public signals, but less so when facing private signals. The closeness of behavior to the best responses is a mark of success of information design: we can use standard models to predict behavior. The next step is to investigate the induced beliefs.

Result 2. *Receivers respond to incentives to follow recommendations as theoretically predicted, and behavior is close to best responses to beliefs. In contrast to theoretical predictions, public information structures generate more following than equivalent private information structures. Consistent with theoretical predictions and moderate risk aversion, the frequency of following recommendations is characterized by the following ranking:*

$$low > optimal > high$$

Appendix Section 4.B.5 reports additional analyses. The results reported in Table 4.2 are similar when estimating the models using data only from recommendations to invest. In addition, they are robust to including additional controls. As expected, more risk-averse participants follow recommendations less. No characteristics other than gender correlate with following behavior.

4.4.3 Beliefs

In the first experiment, a computerized sender attempts to persuade receivers to invest by changing their beliefs. So far, we have observed that receivers' behavior changes. In the following, I present data on elicited beliefs for each between-subject treatment to measure whether the change in behavior is consistent with the changes in beliefs.

Theoretically, information on the state is inferred using Bayes' rule. In addition, in the equilibrium with maximal following, others are predicted to follow recommendations if they are obedient. In the experiment, participants reported beliefs at the very end after making all choices in the investment game. All beliefs presented here are conditional on having received the recommendation to invest.[17]

In the left panel of Figure 4.3, I show the average belief about the response of other participants to the recommendation to invest. The red diamond represents the observed following behavior. We observe that participants predict others' following behavior remarkably well.

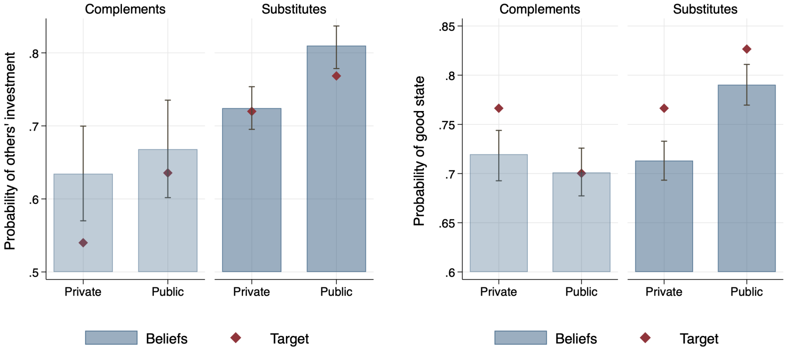
In Table 4.3, I regress errors and squared errors in beliefs on treatments. The squared errors are informative about the presence of a prediction error. The errors are informative about the direction of this error, if present. Column (1) use the distance between the target and the reported belief about others' following a recommendation to invest and column (2) the squared distance. Prediction errors are larger for games of complements, in which receivers overestimate others' investment by 9 percentage points (p-value=0.025 in column (1); p-value=0.014 in column (2)). Errors also increase for *high* structures: receivers overestimate that others will invest by 14 percentage points (p-value<0.001 in column (1); p-value=0.002 in column (2)). Crucially, receivers predict others' following in public and private structures equally well (p-value=0.138; column (2)). The main deviation of behavior from theoretical predictions, the advantage of public signals, is also present in this belief channel.

In the right panel of Figure 4.3, I show the average belief that the state is good conditional on receiving the recommendation to invest. The red diamonds indicate the Bayesian posterior. In Table 4.3, I again regress errors and squared errors in beliefs on treatments, where column (3) uses the distance between the Bayesian

[17] In Appendix Section 4.B.6, I present averages for each level and for the recommendation not to invest.

posterior and the reported belief that the state is good after receiving the recommendation to invest, and column (4) uses the squared distance. Participants are generally slightly more pessimistic than predicted, so they under-respond to good news. This is reflected in the constant, in which they underestimate the odds that the state is good by 8 percentage points (p-value<0.001 in columns (3) and (4)). Otherwise, they only overestimate how likely the state is to be good in *high* structures, compared to *low* structures, by 3 percentage points (p-value<0.001 in column (3); p-value=0.009 in column (4)).

Figure 4.3: Beliefs about the state and others' following behavior



Notes: Left panel: average reported belief that other participants invest, conditional on receiving the recommendation to invest, by treatment. Right panel: average reported belief that the state is good, conditional on receiving the recommendation to invest by treatment. This figure pools data from all levels of obedience. Bars indicate observed choices, diamonds indicate the observed target in the data, and error bars indicate 95% bootstrapped confidence intervals.

Result 3. *Beliefs evolve in line with Bayesian updating about the state and about the play of other receivers. Participants predict others' following behavior well and expect public structures to induce higher following.*

4.4.4 Mechanisms: Explaining the advantage of public structures

Contrary to theoretical predictions, participants are more willing to follow public signals. In addition, participants correctly believe that others do the same. In this section, I investigate mechanisms that may explain this advantage of public structures.

First, I study whether the advantage of public structures is still present in an individual control task that mirrors the game but removes the strategic interaction. The advantage of public structures can stem from two sources. First, it may result from the strategic uncertainty in the interaction with the other receiver. Second,

Table 4.3: Errors in beliefs

	(1)	(2)	(3)	(4)
	Others' following		State is good	
	Error	Error^2	Error	Error^2
Public	-0.038	-0.014	-0.018	-0.002
	(0.032)	(0.010)	(0.015)	(0.005)
Complements	-0.091**	0.032**	-0.011	0.004
	(0.040)	(0.013)	(0.016)	(0.004)
Public × Complements	0.102*	0.004	-0.027	0.001
	(0.054)	(0.016)	(0.023)	(0.006)
(1 if level=*optimal*)	-0.048**	0.003	-0.029***	-0.008***
	(0.020)	(0.008)	(0.007)	(0.003)
(1 if level=*high*)	-0.135***	0.029***	-0.054***	-0.004
	(0.022)	(0.009)	(0.009)	(0.003)
Constant	0.030	0.063***	0.084***	0.025***
	(0.033)	(0.009)	(0.012)	(0.004)
Part and lab FE	Yes	Yes	Yes	Yes
Observations	1293	1293	1293	1293

Notes: The table reports OLS estimates and includes all data. The dependent variables are errors in beliefs (target - belief) in columns (1) and (3) and squared errors in beliefs ((target - belief)^2) in columns (2) and (4). Columns (1) and (2) use the belief about others' investment after they receive the recommendation to invest. Columns (3) and (4) use the belief about the state being good after others receive the recommendation to invest. Public and Complements are dummy variables equal to 1 if the belief was reported for facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. (1 if level=*optimal*) and (1 if level=*high*) are dummy variables equal to 1 if the information structures used *optimal* or *high* probabilities to persuade receivers to invest, relative to the omitted category *low*, respectively. Beliefs were not elicited for one participant that dropped out earlier. Standard errors in parentheses, clustered at the matching-group level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

different information structures may differ in their riskiness, even when stripped from the game.

I first investigate the second possibility: differences in riskiness. Within the investment game, all structures are calibrated such that a risk-neutral receiver is equally willing to follow within each level of obedience (*low*, *optimal*, or *high*), but differences in riskiness may contribute to differences in choices observed in the investment game.

To obtain an individual control task for each structure and game, I remove strategic uncertainty about others' behavior by assuming that others follow their recommendations. Participants choose to either take a risky lottery, corresponding to following a recommendation to invest, or take the safe payoff, corresponding to not investing. The risky lottery is calibrated to match the expected payoffs and probabilities of the investment game and the associated information structure. Section 4.3.4 explains the task in more details.

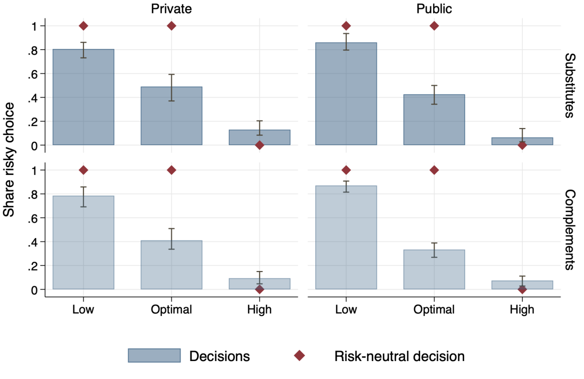
Each participant makes three choices in this task, corresponding to the three information structures they face in the main parts of the experiment. Risk-neutral participants would accept the lotteries associated with the *low* and *optimal* structures and reject the lottery associated with the *high* structures. In Figure 4.4, I present the average share of participants who accept the risky choice. The red diamonds indicate the choices a risk-neutral participant makes. Table 4.4 presents the corresponding regressions of the decision to accept the risky lottery on treatment indicators.

The data indicate that the majority of participants are risk averse: while 86% accept the lottery corresponding to the *low* structures (coefficient on the constant, because *low* is the omitted category; p-value<0.001), as expected gains decrease, take-up of the lottery decreases: by 42% for the *optimal* lottery (p-value<0.001), which has an expected value just above the safe payoff, and by 74% for the *high* lottery (p-value<0.001) compared to the *low* lottery's take-up.

Crucially, there are no systematic differences between treatments (Public: p-value=0.528; Complements: p-value=0.279; Public×Complements: p-value=0.679). While behavior in the game indicates that participants are more likely to follow public signals, this increase in following is not present in this individual control task. Any change in behavior we see between these treatments is driven by the strategic interaction in the game and not by any differences in the riskiness of the structures.

Consistent with this finding, I do not detect significant correlations between following and a standard risk-preference measure (Eckel and Grossman, 2002), the treatment variables and their interactions (see Appendix Section 4.B.7).

Figure 4.4: Control lottery task



Notes: Bars indicate observed choices, diamonds indicate the risk-neutral choice, and error bars indicate 95% bootstrapped confidence intervals.

Result 4. *Differences in riskiness cannot explain the higher following in public information structures. Such structures' advantage only arises when receivers strategically interact.*

I also measure inequity aversion (Fehr and Schmidt, 1999), and I show in Appendix Section 4.B.8 that it does not explain the higher following in public information structures. To summarize, I find that differences in play are not driven by features unrelated to the strategic nature of the game—namely, the game's inher-

Table 4.4: Control lottery choice

Public	-0.025	(0.039)
Complements	-0.045	(0.042)
Public × Complements	0.021	(0.050)
(1 if level=*optimal*)	-0.415***	(0.025)
(1 if level=*high*)	-0.740***	(0.024)
Constant	0.860***	(0.032)
Observations: 1293, # clusters: 72, # participants: 431		

Notes: The table reports OLS estimates and includes all data. The dependent variable is a dummy variable equal to 1 if a participant chose to take up the risky lottery, corresponding to following a recommendation to invest. Public and Complements are dummy variables equal to 1 if the lottery decision was made capturing a public information structure, with the omitted category being a private structure, or capturing a game with strategic complements, with the omitted category being a game of strategic substitutes. (1 if level=*optimal*) and (1 if level=*high*) are dummy variables equal to 1 if the information structures used *optimal* or *high* probabilities to persuade receivers to invest, relative to the omitted category *low*, respectively. Choices were not elicited for one participant that dropped out earlier. Standard errors in parentheses, clustered at the matching-group level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

ent riskiness and inequalities in payoffs. In the following, I investigate two mechanisms that take the game's strategic component into account.[18]

As a first mechanism, I study whether private and public structures produce differences in the noisiness of behavior. Higher additional uncertainty about others' actions is detrimental to investment, as participants can no longer best respond by following recommendations. If public or private structures induce different degrees of noisiness, it may be desirable for a sender to rely more frequently on the less noisy environment to persuade receivers.

There are good reasons to expect that private structures generate more noisy behavior. One reason is that they require more complex strategic reasoning. Public signals generate common knowledge about others' signals. The symmetric decision structure with public signals may help receivers arrive at their best response and lead them to expect that others do so as well. In contrast, by design, private signals introduce uncertainty about others' signals. A corresponding increase in difficulty is consistent with Martínez-Marquina et al. (2019) finding that uncertainty—in this case about others' signals—contributes to failures of contingent reasoning. Similarly, Oprea (2020) finds that having to consider additional states—in this case the potential state of miscoordinated action recommendation, with one recommendation to invest and one not to invest—is perceived as complex and costly to process. In line with these findings, the number of errors in the quizzes associated with the information structures is significantly lower for public structures.[19] These quizzes directly measure their understanding for example, of what signals the second participant would receive if they themselves received a particular signal.

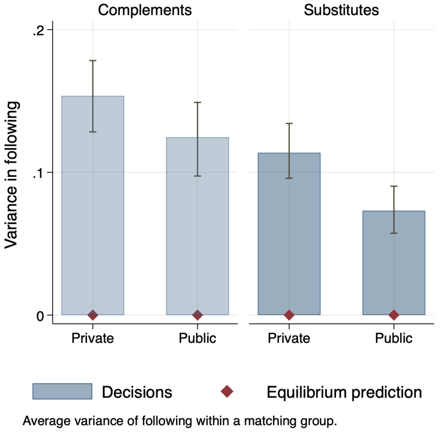
To document this mechanism, I begin by studying differences in the variance in the behavior between treatments. In Figure 4.5, I plot the average variance in the following behavior for *low* and *optimal* levels, calculated for each group and part separately. This provides a measure of how uncertain a participant is about the fol-

[18]The following analysis in this section is exploratory and was not preregistered.

[19]In a regression of the number of errors on treatment dummies, the coefficient on public is negative (-1.18, compared to a control average of 6.68) and significant (p-value=0.012, 431 observations, clustering standard errors on the matching-group level; all other coefficients are not significant at conventional levels).

lowing decisions of participants within their matching group. Theoretically, there is zero variance in following behavior, as all obedient signals are always followed in equilibrium. Empirically, however, the more complex private signals generate noisier behavior than public signals: public signals decrease the standard deviation in following by 0.055 (coefficient on Public; *p*-value=0.009; Table 4.5; column (1)).

Figure 4.5: Variance in following behavior



Notes: Average variance of following behavior, calculated on a matching group-part level. The figure only uses data from *low* and *optimal* structures. Bars indicate observed choices; red diamonds indicate the equilibrium predictions; error bars indicate 95% bootstrapped confidence intervals.

As explained, this increase in the variance in following behavior is detrimental to receivers' incentives to follow recommendations. Higher uncertainty about others' play implies that following is less frequently a best response. This was already documented in Section 4.4.2, as Figure 4.2 showed that the best response to receivers' beliefs implies lower following rates for private than for public structures. Even receivers' beliefs reflect the noisier behavior: there is more variance in beliefs about others' following a recommendation to invest for private than for public signals (see Appendix Section 4.B.9).

Next, I show how this variance in behavior correlates with the treatment effects I find. Within each treatment, I split groups into those showing above- and below-median variance. I interact treatment indicators with a dummy variable capturing whether a group has above-median variance within each treatment in columns (2) and (3) in Table 4.5. In column (2), I focus on the decision to invest. I find the theoretically predicted advantage of private structures with strategic substitutes for the low variance groups (coefficient of 5 percentage points on Public; *p*-value<0.001). This effect, however, reverses for the high-variance groups, for which public structures induce higher investment than private ones (coefficient of 8 percentage points on Public × High variance; p-value=0.003). These two counteracting effects produce the nonsignificant treatment effect of public structures documented in Table 4.1. In column (3), we see that high-variance groups follow recommendations less frequently (coefficient of 12 percentage points on High variance; p-value<0.001). Here, public structures prove beneficial, as they generate higher following rates for highly noisy groups (coefficient of 6 percentage points on Public × High variance; p-value=0.031). The noisy response to private signals thus indeed explains the superiority of public signals.

Table 4.5: Variance and heterogeneous effects

	(1) SD(following)	(2) Investment	(3) Following
Public	-0.055***	-0.049***	0.022
	(0.020)	(0.014)	(0.016)
Complements	0.053**	-0.074***	-0.073***
	(0.020)	(0.027)	(0.023)
Public × Complements	0.012	0.118***	-0.007
	(0.031)	(0.043)	(0.041)
High variance		-0.112***	-0.119***
		(0.022)	(0.020)
Public × High variance		0.081***	0.059**
		(0.026)	(0.027)
Complements × High variance		-0.050	-0.031
		(0.046)	(0.039)
Public × Complements × High variance		-0.062	-0.063
		(0.067)	(0.061)
(1 if level=*optimal*)		-0.025*	-0.126***
		(0.014)	(0.015)
(1 if level=*high*)		-0.055***	-0.241***
		(0.014)	(0.015)
Period		-0.003***	-0.003***
		(0.001)	(0.001)
Constant	0.378***	0.597***	1.024***
	(0.018)	(0.021)	(0.021)
Part and lab FE	Yes	Yes	Yes
Observations	216	25908	25908
# clusters	72	72	72
# participants	-	432	432

Notes: The table reports OLS estimates and includes all data. In column (1), the dependent variable is the standard deviation of the following behavior, calculated for each group and part separately. There are 72 groups making decisions across three parts each, which results in 216 observations. The dependent variables in columns (2) and (3) are the decision to invest and to follow a recommendation, respectively. High variance is a dummy variable equal to 1 if the average standard deviation of the matching group (calculated as in (1)) is above the median within each treatment. Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. (1 if level=*optimal*) and (1 if level=*high*) are dummy variables equal to 1 if the information structures used *optimal* or *high* probabilities to persuade receivers to invest, relative to the omitted category *low*, respectively. Standard errors in parentheses, clustered at the matching-group level, * p < 0.10, ** p < 0.05, *** p < 0.01.

Result 5. *Private signals induce noisier behavior than public signals. The increased uncertainty lowers receivers' incentives to follow private signals, which decreases the sginals' persuasiveness.*

As a second mechanism, I show that participants' behavior is consistent with them disliking the differential treatment private structures produce. With private structures, at most one of the participants receives bad advice at any moment. Here, bad advice is the recommendation to invest even though the bad state materialized. If followed, this advice generates a loss for the receiver. In contrast, with public signals, both participants receive such a recommendation and simultaneously suffer losses when following it. Therefore, only participants with private structures can experience being the sole receiver losing out after trusting the sender. I show that participants' behavior is affected by being the only loss-making participant in games with private signals; when both make a loss with public signals, it does not change their behavior.

To study this mechanism, I focus on participants' response to having received bad advice in the past and how the response depends on whether they face public or private signals. I split participants into those who receive bad advice in the first period in which they face a new information structure and those who do not receive such advice. Then, I assess whether the behavior of these two groups differs in all subsequent periods when they face this structure.

In Table 4.6, I regress the decision to invest or to follow a recommendation on treatment dummies, a dummy for having received bad advice, and the interaction of the two. Participants who received bad advice are less likely to invest or follow in all future periods. Having received bad advice reduces investments by 12 percentage points (p-value<0.001; column (1)). Bad advice also decreases following by 13 percentage points (p-value<0.001; column (2)). However, this is solely driven by those participants who face private signals, as the interaction effect for public signals with bad advice almost exactly cancels out this baseline effect. Investment increases for public signals by exactly the 11 percentage-point loss measured for those having received bad advice (p-value=0.020; column (1)). Following increases by 11 percentage points (p-value=0.007; column (2)) for participants with public signals with bad advice, compared to those with private signals and bad advice. Column (3) shows that the effects on following are robust to including additional controls.[20]

Result 6. *Bad advice in private structures, but not in public structures, decreases investment in later periods after receivers experienced differential treatment.*

In Appendix Section 4.B.11, I perform a back-of-the-envelope calculation of the relative contributions of the two mechanisms to the superiority of public signals. On average, public signals lead to 4 percentage points higher investment, which is an effect that is not predicted theoretically. When decomposing this effect, about

[20] In Appendix Section 4.B.10, I show that the pattern is similar when using other ways of measuring whether participants received bad advice, such as how often a participant overall received bad advice when facing an information structure. In addition, I show that the pattern is driven by those participants that receive bad advice, and not by participants that receive different recommendation than their matched participant, so not by participants that receive good advice while their matched participant receives bad advice. This finding is not consistent with alternative explanations for this pattern, such as conformism or a preference to always receive the same recommendations.

Table 4.6: Bad advice and future following

	(1) Investment	(2) Following	(3) Following
Public	-0.027	0.035	0.046**
	(0.021)	(0.021)	(0.022)
Complements	-0.110***	-0.098***	-0.090***
	(0.033)	(0.031)	(0.031)
Public × Complements	0.109**	-0.018	-0.031
	(0.047)	(0.045)	(0.045)
Bad advice	-0.122***	-0.128***	-0.123***
	(0.035)	(0.036)	(0.036)
Public × Bad advice	0.111**	0.125***	0.113**
	(0.047)	(0.045)	(0.045)
Complements × Bad advice	-0.009	0.009	-0.010
	(0.051)	(0.057)	(0.054)
Public × Complements × Bad advice	-0.033	-0.055	-0.028
	(0.066)	(0.069)	(0.067)
Constant	0.535***	0.963***	0.948***
	(0.022)	(0.024)	(0.081)
Period trend, part & lab FE	Yes	Yes	Yes
Additional controls	No	No	Yes
Observations	24612	24612	24510

Notes: The table reports OLS estimates and includes all data after period one in each part. Column (3) uses fewer observations, as some additional controls are not available for all participants. The dependent variables are the decision to invest or the decision to follow a recommendation. Bad advice is a dummy variable equal to 1 if a participant received a recommendation to invest when the state was bad in period 1 of the corresponding information structure. Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. The additional controls are participants' Fehr and Schmidt (1999) preferences, risk aversion, numeracy score, and demographic variables. Standard errors in parentheses, clustered at the matching-group level,* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

61% can be attributed to the complexity of private signals, with the remainder attributed to participants that had received bad advice decreasing their following.

4.4.5 Human senders

In the second experiment, I explore whether receivers respond differently to human senders and how participants approach the sender's problem. Participants in the role of sender are incentivized to maximize receivers' investment, a fact that receivers are aware of. This may change receiver behavior compared to the first experiment. For example, receivers may no longer be willing to follow recommendations of a sender that too frequently attempts to deceive them into investing. These receivers may exhibit an aversion to being deceived to invest in the bad state of the world, beyond what is justified by the strategic skepticism in the game. Furthermore, while the first experiment provides a good indication of how to persuade audiences, it is unclear whether real senders are capable of optimally adjusting their persuasion to their audience. Apart from introducing human senders, the second experiment mirrors the first as closely as possible. The senders can choose among the six information structures exogenously assigned in the first experiment: either using public or private signals, and using one of the three levels of obedience (*low*, *optimal*, or *high*).

Differences in receivers' behavior. I begin by comparing receivers' behavior between the first and second experiment using data on beliefs.[21] Direct choice data in the second experiment is less informative for two reasons. First, data on receivers' choices are only available for the structures senders choose.[22] Second, the senders likely particularly rely on structures that are successful for their group of receivers, but these structures may be heterogeneous across groups. This means that we observe receivers' choices in a selected distribution of structures.[23] To account for the selection in choice data, at the end of the experiment I elicit beliefs for the full set of potential structures. As I elicited the same beliefs in both experiments, I can compare data from experiments with and without participants in the role of senders. Within the second experiment, I can also compare senders' and receivers' beliefs separately.

Figure 4.6 shows receivers' belief, across the two experiments, about other receivers' following behavior after they received the recommendation to invest. The red diamonds represent the observed following behavior within each experiment. The left panel reproduces data from Figure 4.3 on the receivers' beliefs in the first experiment. The middle panel shows the receivers' beliefs elicited in the second experiment, and the right panel the senders' beliefs. Table 4.7 presents estimation results of the corresponding effects. I regress the belief that others invest after receiving the recommendation to invest on features of the information structure (public versus private, information-structure level) for three samples. In column (1), I use receivers' beliefs from the first experiment. In column (2), I use receivers' beliefs from the second experiment. Column (3) uses beliefs of the senders from the second experiment. Column (4) pools data from both experiments and both roles to estimate interaction effects.

Most behavioral patterns are robust across both experiments and roles. Between the first and second experiments, receivers believe that following behavior decreases somewhat: from 80% (coefficient on the constant; p-value<0.001; column (1)) to 73% (coefficient on the constant; p-value<0.001; column (2); interaction effect in column (4): p-value<0.001). Senders, in turn, predict following rates 21 percentage points lower (coefficient on Second exp., senders; p-value<0.001).

In addition, senders predict that receivers' changes in behavior in response to different structures are smaller than the response predicted by receivers. In that

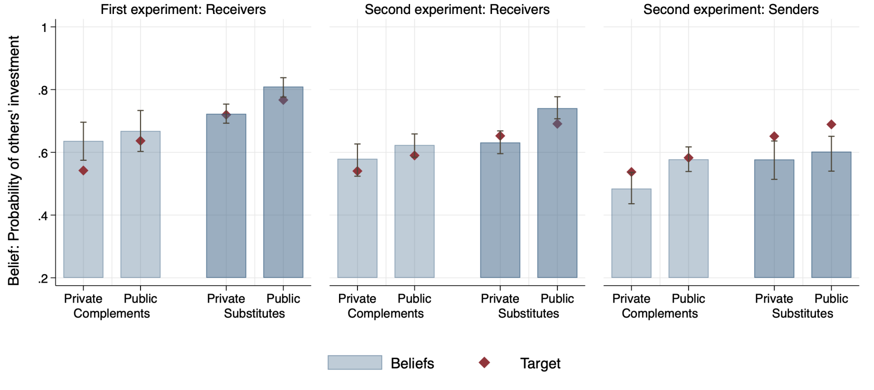
[21]Nevertheless, Appendix Table 4.B.18 shows that the main result is also present in the choice data: public structures increase investment. In the second experiment, this effect is similar in both strategic environments.

[22]For example, 6 out of the 40 groups in the experiment did not encounter all structures, as none of the senders in these groups exploited their whole choice set during the entire experiment.

[23]Consistent with this form of selection, the distribution of chosen structures is quite imbalanced. Of the 63 total possible choices of information structures for each matching group (three senders per matching group, 21 periods), 32 of the 40 groups faced at least one information structure fewer than five times. Simultaneously, in 17 of the 40 groups just one information structure accounted for more than half of the receivers' choices (so, for more than 32 sender choices).

sense, senders underestimate receivers' responses. For example, they believe that receivers respond to higher levels less than receivers believe other receivers respond. For *high* compared to *low* levels, receivers predict a decrease in following rates of 17 percentage points (p-value<0.001; column (2)), while senders only predict a decrease of 6 percentage points (p-value=0.026; column (3); interaction effect in column (4): p-value<0.001). Senders do not anticipate that receivers expect more following with public signals with strategic substitutes (p-value=0.400; column (3)), while receivers predict a decrease in the following rate of 11 percentage points (p-value<0.001; column (2); interaction effect in column (4): p-value=0.093). Thus, while senders partially anticipate the advantage of public structures, they underappreciate that receivers believe that public structures increase following behavior.

Figure 4.6: Beliefs across all experiments



Notes: Average reported belief that other receivers invest, conditional on them receiving the recommendation to invest, by treatment and role. Left panel: receivers in the first experiment. Middle panel: receivers in the second experiment. Right panel: senders in the second experiment. Bars indicate observed beliefs, diamonds indicate the observed target in the data, and error bars indicate 95% bootstrapped confidence intervals.

In Appendix Table 4.B.16, I show that beliefs about the state are also comparably updated across both experiments and roles. Across both experiments, receivers update as expected by becoming more pessimistic about the state with *optimal* and *high* structures. Again, senders underestimate the extent to which receivers believe others are more pessimistic.

Senders choosing public or private structures. Now I turn to senders' choices of information structures. Figure 4.7 shows the share of public structures used. Table 4.8 presents estimation results of the corresponding treatment effect. In column (1), I regress the decision to use a public structure on a treatment indicator. Senders on average choose public structures slightly more often than private

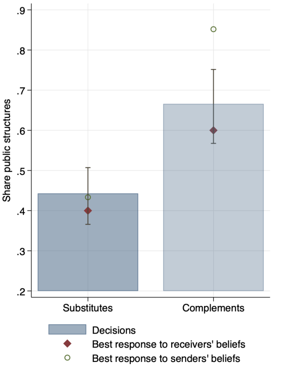
Table 4.7: Beliefs about others' following across all experiments

	(1)	(2)	(3)	(4)
	Belief: Probability others invest			
Public	0.087***	0.109***	0.025	0.087***
	(0.023)	(0.016)	(0.029)	(0.022)
Complements	-0.087**	-0.052	-0.091**	-0.086**
	(0.034)	(0.032)	(0.042)	(0.034)
Public × Complements	-0.055	-0.066***	0.068*	-0.055
	(0.050)	(0.023)	(0.035)	(0.050)
(1 if level=*optimal*)	-0.117***	-0.122***	-0.035*	-0.117***
	(0.010)	(0.013)	(0.018)	(0.010)
(1 if level=*high*)	-0.175***	-0.166***	-0.056**	-0.175***
	(0.014)	(0.015)	(0.024)	(0.014)
Second exp., receivers				-0.095***
				(0.026)
Second exp., senders				-0.214***
				(0.038)
Public × Second exp., receivers				0.023
				(0.027)
Public × Second exp., senders				-0.062*
				(0.036)
Complements × Second exp., receivers				0.036
				(0.047)
Complements × Second exp., senders				-0.005
				(0.054)
Public × Complements × Second exp., receivers				-0.011
				(0.055)
Public × Complements × Second exp., senders				0.123**
				(0.061)
(1 if level=*optimal*) × Second exp., receivers				-0.006
				(0.016)
(1 if level=*optimal*) × Second exp., senders				0.081***
				(0.021)
(1 if level=*high*) × Second exp., receivers				0.010
				(0.020)
(1 if level=*high*) × Second exp., senders				0.119***
				(0.028)
Constant	0.803***	0.726***	0.592***	0.811***
	(0.018)	(0.024)	(0.042)	(0.017)
Experiment	First	Second	Second	Both
Role	Receivers	Receivers	Senders	Both
Lab FE	Yes	Yes	Yes	Yes
Observations	1293	1440	720	3453
# clusters	72	40	40	112
# participants	431	240	120	791

Notes: The table reports OLS estimates. The dependent variable is the reported belief that other receivers invest after receiving the recommendation to invest. Column (1) uses data from the first experiment, with only receivers. Columns (2) and (3) use data from the second experiment. (2) are the receivers, (3) the senders. Column (4) pools data from both experiments and both roles. Public and Complements are the treatment indicators. Public and Complements are dummy variables equal to 1 if the belief was reported for facing a public information, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. (1 if level=*optimal*) and (1 if level=*high*) are dummy variables equal to 1 if the information structures used *optimal* or *high* probabilities to persuade receivers to invest, relative to the omitted category *low*, respectively. Second exp., receivers and Second exp., senders are dummies equal one if the belief is measured in the second experiment, for receivers and senders, respectively. The omitted category is the receivers in the first experiment. In column (1), beliefs were not elicited for one participant that dropped out earlier. Standard errors in parentheses, clustered at the matching-group level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

ones, in 55% of periods. Importantly, they respond to the receivers' interaction in making their own choice. They choose public structures 53% more frequently with strategic complements compared to substitutes (p-value<0.001; column (1)), consistently with the theoretical prediction.

Figure 4.7: Senders' choices of public and private signals



Notes: Share of public structures chosen by senders. Bars indicate observed choices; error bars indicate 95% bootstrapped confidence intervals. Red diamonds indicate the average best response to receivers' beliefs, green circles the best response to senders' own beliefs about receiver behavior.

In addition to senders' choices, I show best responses to receivers' and senders' beliefs in Figure 4.7. The best responses indicate what share of public structures would have maximized senders' payoffs when using beliefs to predict receivers' behavior.[24] The preceding analysis in this section revealed that compared to receivers' beliefs, senders believe that receivers do not respond strongly to changes in the information structure. Therefore, best responses to either senders' or receivers' beliefs may differ. Receivers may understand their own decision situation reasonably well. Senders, in contrast, are required to predict receivers' responses while simultaneously deciding on an optimal structure. A difference in the best responses to senders' and receivers' beliefs reveals the extent to which differences in beliefs affect the best response.

Senders' decisions match a best response to receivers' beliefs quite closely. This indicates that senders' choices are reasonably close to choosing structures that maximize their own payoffs, and they are optimal based on expected receiver behavior. Here, the best response to receivers' beliefs is likely the most infor-

[24]For the best response to receivers' beliefs, I first calculate each receiver's best response to recommendations, based on each receiver's own beliefs about the state and others' following behavior. I aggregate these best responses by calculating the average best response within a matching group. Using this exercise, I obtain predicted investments for each of the possible information structures. I define a sender's best response to receivers' beliefs to be the information structure that maximizes investment, given predicted receiver behavior. The best responses to receivers' beliefs always exist. However, they do not exist for 40 of 120 best responses to senders' beliefs, as these senders hold beliefs that do not generate investment under any information structure.

mative, as choices and beliefs in the first experiment revealed that participants' beliefs are reasonably accurate; thus, these beliefs give a good indication what investment behavior senders could have expected.

The best response to senders' own beliefs indicates that, if anything, they use public structures less frequently than expected. Clearly, senders anticipate both that public signals are generally more persuasive and that they are particularly valuable in games of strategic complements.

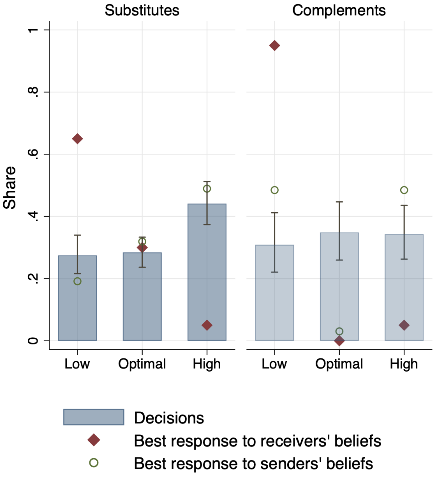
Result 7. *Senders on average use public signals slightly more often than private signals, and as predicted they use public signals more often when the receivers' strategic environment features strategic complements.*

Senders' choice of level. In Figure 4.8, I show how frequently senders choose each possible level of information structure. Senders are relatively aggressive in persuading receivers to invest frequently; the median choice in both games is the *optimal* structure. This structure recommends investment as often as possible while ensuring that risk-neutral receivers continue to best respond by following. However, this level also means that receivers' payoffs are quite low, while senders' payoffs are high if these recommendations are followed. In addition, senders surprisingly frequently employ *high* structures. In columns (2) and (4) in Table 4.8, I compare how much more frequently senders choose *optimal* instead of *low* structures. We can see that at the beginning of the experiment, senders on average are 18 percentage points more likely to choose *optimal* structures (coefficient on the constant; p-value=0.004; column (2)). However, senders over time learn to choose *low* structures more often (-1.5 percentage points per period, p-value=0.001; column (2)), which encourages investment. There is no significant difference in baseline choices between games of complements and substitutes (p-value=0.758; column (4)). However, in games of substitutes, senders are 17 percentage points more likely to choose *high* instead of *optimal* structures (coefficient on the constant; p-value=0.021; column (5)). In games of complements, senders are equally likely to choose either level (coefficient of -17% on Complements; p-value=0.082; column (5)). In Appendix Table 4.B.15, I repeat this analysis separately for the first third and last two-thirds of the data to study learning. Senders use public structures more frequently across both games as they gain experience and learn to avoid high levels in games of strategic complements.

A large majority of senders apparently understand that a too high level is not optimal, as receivers are no longer incentivized to follow. Yet, on average, they choose *high* levels, which reduce receivers' expected profits from following but increase their own profits if receivers do follow. Somewhat surprisingly, they are more aggressive than the best response to receivers' beliefs indicates. The senders would have generated higher investment by reducing their aggressiveness, as receivers would be more likely to follow recommendations. In addition, their own beliefs indicate that senders again underestimate the degree to which choosing a more aggressive persuasion strategy will affect receivers' choices, judged by the gap between the best responses to receivers' and senders' beliefs.

The aggressiveness in communication contrasts with typical findings in the earlier literature on cheap-talk experiments, in which senders typically overcommunicate relative to equilibrium predictions (Blume et al., 2020). Instead, communicating by committing to an information structure moves predictions closer to self-interested behavior. One reason may be that senders only deceive their receivers probabilistically, as uncertainty remains about which signals participants receive even conditional on the bad state materializing. This is in line with the literature on how uncertainty in choices diffuses participants' perceived responsibility for selfish choices (Falk and Szech, 2014; Exley, 2016).

Figure 4.8: Senders' choice of level



Notes: Share of periods in which senders choose *low*, *optimal*, or *high* information structures. Bars indicate observed choices; error bars indicate 95% bootstrapped confidence intervals. Red diamonds indicate the average best response to receivers' beliefs, green circles the best response to senders' own beliefs about receiver behavior.

Result 8. *Senders persuade aggressively. In games of substitutes, they choose structures according to the following ranking:*

$$high > optimal = low$$

In games of complements, they choose structures according to the following ranking:

$$high = optimal = low$$

Table 4.8: Senders: Treatment effects

	(1) Public	(2) *Optimal* vs. *low*	(3) *High* vs. *optimal*	(4) *Optimal* vs. *low*	(5) *High* vs. *optimal*
Complements	0.222***			0.031	-0.166*
	(0.062)			(0.099)	(0.093)
Period	0.003	-0.015***	0.001	-0.015***	0.001
	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Constant	0.415***	0.184***	0.085	0.168**	0.171**
	(0.060)	(0.061)	(0.056)	(0.076)	(0.071)
Lab FE	Yes	Yes	Yes	Yes	Yes
Observations	2520	2520	2520	2520	2520
# clusters	40	40	40	40	40
# participants	120	120	120	120	120

Notes: The table reports OLS estimates. In column (1), the dependent variable is a dummy equal to 1 if the sender chose a public structure. In columns (2) and (4), the dependent variable is a difference in level shares: the share of *optimal* structures minus the share of *low* structures. In columns (3) and (5), the dependent variable is a difference in level shares: the share of *high* structures minus the share of *optimal* structures. Complements is the treatment indicator. This is a dummy variable equal to 1 if the decision was made facing a game with strategic complements, with the omitted category being a game of strategic substitutes. Period is a linear period trend. Standard errors in parentheses, clustered at the matching-group level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.5 Conclusion

In this paper, I studied the optimal persuasion of an audience of interacting receivers. In a laboratory experiment, I showed that senders benefit from tailoring their communication strategy to the strategic interaction of their audience. In particular, when the audience faces a game of strategic complements, public signals enhance a sender's capability to persuade. In addition, I found that public signals are more persuasive than private signals across environments. This force has not been incorporated in theoretical models so far, yet it is strong enough to offset the potential strategic gains from private signals in games of strategic substitutes.

I ruled out two standard mechanisms that may be driving the superiority of public signals. Neither differences in riskiness nor inequalities can explain why public structures enhance persuasion. Instead, I found evidence for the following two mechanisms. First, receivers struggle with the more complex nature of private signals, as they understand less well what they can learn from them. This increases the noise in behavior. This unpredictability, in turn, reduces how often trusting private signals is a best response. Public signals solve this by relying on common knowledge and common actions, and this symmetry apparently makes them easier to understand and to optimally respond to these signals. Second, receivers exhibit a distaste for differential treatment with private structures if they have experienced unfavorable recommendations early on. Public signals solve this by recommending the same action to all receivers.

This study provides novel evidence on the strength of adapting the communication channel to the strategic environment of the receivers. As even students in a laboratory experiment can capitalize on these gains, it stands to reason that sophisticated players in practice can take advantage of appropriate communication channels to enhance persuasion. However, the senders in the laboratory still underestimate what they can gain from broadly employing public signals.

In practice, senders in these types of setups often use public communication. For example, governments are held accountable with transparent decision-making. Equal treatment is an important cornerstone of democratic governments. The results of this experiment provide an additional, purely strategic, rationale for using public communication. They enhance a government's persuasiveness, particularly strongly in games of strategic complements.

These results can help senders who communicate with strategically interacting audiences in many real-world settings. For example, close to the framing in the experiment, a manager may want to encourage effort on the part of her workers, whose rewards may feature complementarities or substitutabilities. This paper highlights that besides exploiting her knowledge about a project's difficulty, she can maximize effort by (mis)coordinating workers' actions by using private or public signals. In particular, I showed that public signals are a valuable tool for this manager, as they are more persuasive than private signals. Closer to the investment-game framing, a government may want to encourage investments into COVID-19 vaccine-production facilities while holding private information about future waves' severity or planned vaccination campaigns. The interaction of firms may feature strategic substitutes, as stiffer price competition ensues if both firms increase capacity. Alternatively, strategic complements can be introduced by increased public acceptance and subsequent sales of a more widely established vaccine technology, from a better understanding of this new technology with resulting improved production capabilities, or from other network effects on an industry level. This paper provides empirical evidence that the sender should carefully choose the channel in response to the prevailing interaction. Other examples include speculative attacks with strategic complementarities between market participants, which central banks or regulators try to prevent by strategically releasing information publicly.

There is still much to be learned about communication with an audience, with a small empirical and experimental literature. In this paper, I study small audiences, but results for larger audiences are crucial to understand how these strategic forces change with more receivers. In practice, many audiences are large, which increases both the difficulty in reasoning through optimal responses to signals but also the potential gains from optimal persuasion. Similarly, I give the theoretical predictions a good shot by revealing the sender's information structure. Data from an experiment in which this is not revealed, but sender and receivers interact repeatedly to allow them to learn these elements, would move the setup closer to some real-world settings. From a theoretical point of view it would also be interesting to study the benefits of public and private signals in settings without preference misalignment between sender and receivers. For example, Bergemann and

Morris (2016) derive similar insights by incorporating payoff externalities between receivers and a sender maximizing receivers' average payoff.

Appendix

4.A Appendix: Theory

More formally, Bergemann and Morris (2016) consider decision rules σ which for each type t_i and state θ recommend an action to the player. Types t_i in this context capture information about the state revealed to player i. For game G and information structure S, σ is obedient if for all i, t_i and a_i the following inequality holds for all a_i':

$$\begin{aligned}\sum_{a_i,t_i,\theta} \pi\big((a_i,a_{-i})|(t_i,t_{-i})\big)\,\frac{1}{2}(\theta)\sigma\big((a_i,a_{-i})|(t_i,t_{-i}),\theta\big)\,u_i\,\big((a_i,a_{-i}),\theta)\big)\\ \geq \sum_{a_i,t_i,\theta} \pi\big((a_i,a_{-i})|(t_i,t_{-i})\big)\,\frac{1}{2}(\theta)\sigma\big((a_i,a_{-i})|(t_i,t_{-i}),\theta\big)\,u_i\,\big((a_i',a_{-i}),\theta)\big)\end{aligned}$$

That is, the recommended action a_i yields a payoff at least as high as any other action a_i'. Then, a player best responds by implementing the recommendation as long as the other players implement the recommended action. If a decision rule satisfies obedience, it is a Bayes correlated equilibrium (Bergemann and Morris, 2016), and there exists an expansion of the information structure in which following the decision rule constitutes a Bayes Nash equilibrium.

4.A.1 Parameter choice

In Table 4.B.3, I reproduce estimations using only *optimal* information structures, which are just obedient for risk-neutral receivers. These are the information structures for which I chose the parameters to yield the largest treatment differences, e.g., the interaction effect (public (vs. private) × complements (vs. substitutes)) is predicted to be 14.9 percentage points.

In addition, I chose parameters that yield are reasonably large treatment effect, compared to other potential choices. In the notation of Table 4.1, the parameters used in the experiment correspond to $x_{\text{com}} = 0.1$ and $\epsilon_{\text{com}} = 0.3$ for the game of strategic complements, and $x_{sub} = 0.9$ and $\epsilon_{sub} = -0.5$ for the game of strategic substitutes. To obtain the payoffs displayed in Tables 4.3 and 4.4, payoffs are multiplied by 100 and then a constant payoff of 170 is added. This ensures that payoffs are positive round numbers, to minimize loss aversion and mental effort of processing payoffs.

In the parameter choice, I measure the predicted treatment effect for exactly obedient structures. This choice is partially restricted. As they are probabilities, we need that $1 > p_{bad} \geq r_{bad} \geq 0$, as well as to keep signals private. There are two additional considerations. First, I chose parameters such that with private signals, no joint investment arises in the bad state, formally $p_{bad} - r_{bad} < .5$ Second, there are three levels of obedience, where the *high* structures require higher probabilities of investment recommendations than the *optimal* structures I compare here. Taken together, this implies that the highest probability of private signals in the bad state needs to be sufficiently lower than .5, $p_{bad} - r_{bad} << .5$.

For a selection of parameters, I show the predicted treatment effects in Table 4.A.1. Optimally, private structures set $p_{bad} = \epsilon + x$, $r_{bad} = 0$, and public structures set $p_{bad} = r_{bad} = \frac{\epsilon + x}{1-\epsilon}$. The first row is the *optimal* information structure, which is close to the exactly obedient information structure in the experiment, in the second row. Treatment effects are lower when increasing x while holding $p_{bad} - r_{bad}$ constant, see the third and fourth row. When reducing the probabilities to invest, treatment differences again decrease, independent of the x and ϵ chosen, see rows five to eight.

Table 4.A.1: Parameter choices and predicted treatment effects

	Complements		Substitutes		
Parameters $(x_{com}, \epsilon_{com}; x_{sub}, \epsilon_{sub})$	Public r_{bad}	Private $p_{bad} - r_{bad}$	Public r_{bad}	Private $p_{bad} - r_{bad}$	Diff-in-diff TE
(.1, .3; .9, -.5)	.48	.34	.34	.23	.25
(.1, .3; .9, -.5)	.57	.4	.27	.4	.30
(.3, .1; .9, -.5)	.44	.4	.27	.4	.17
(.1, .3; .6, -.2)	.57	.4	.33	.4	.24
(.2, .1; .8, -.5)	.33	.3	.2	.3	.13
(.1, .2; .8, -.5)	.38	.3	.25	.3	.13
(.2, .1; .5, -.2)	.33	.3	.25	.3	.08
(.1, .1; .7, -.5)	.22	.2	.13	.2	.09

Information structure parameters (p_{bad}, r_{bad}) when varying the parameters of the game (x, ϵ). The column Diff-in-diff TE gives the difference-in-differences treatment effect between games and private vs. public structures, which is the difference in probabilities that a recommendation to invest will be sent in the bad state.

4.B Appendix: Additional empirics

4.B.1 Balancing tables

In Tables 4.B.1 and 4.B.2, I show that participant characteristics are balanced across treatments. In the second experiment, Aheadness aversion is not perfectly balanced and significantly different between treatments with a p-value of 0.097. Controlling for this measure, and other characteristics, does not affect results.

Table 4.B.1: Balancing table: First experiment

	Complements		Substitutes		
	Private	Public	Private	Public	p-values
Age	22.7	22.6	22.8	22.6	0.962
% women	54.6	56.5	65.7	53.7	0.398
% Bachelor	69.4	70.4	70.4	63.9	0.815
Risk	3.0	3.4	3.3	3.1	0.174
Numeracy score	2.4	2.4	2.3	2.5	0.770
Behindness aversion	3.5	3.4	3.8	3.6	0.723
Aheadness aversion	5.5	5.0	5.5	5.3	0.513
Quiz attempts	2.7	2.8	2.9	2.1	0.256

Notes: Average characteristic by treatment. In the last column I report p-values of a Kruskal-Wallis test, comparing equality of ranks across all treatments. Risk is the lottery chosen in the Eckel and Grossman (2002)-task, numeracy score the number of correct answers in the Berlin numeracy test (Cokely et al., 2012), behindness and aheadness aversion the switching points in the multiple price list-elicitation of β and α-parameters in Fehr and Schmidt (1999)-preferences by Yang et al. (2016).

Table 4.B.2: Balancing table: Second experiment

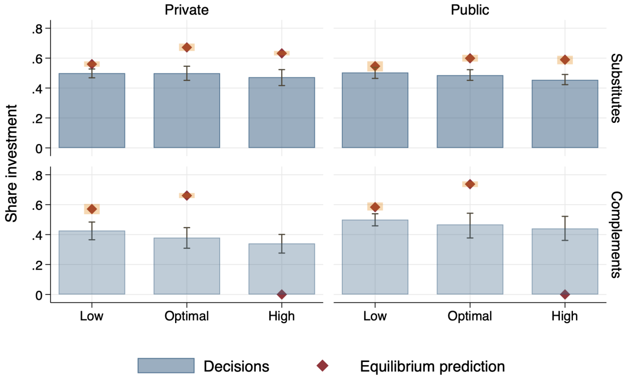
	Complements	Substitutes	p-values
Age	22.4	22.9	0.334
% women	55	57.2	0.672
% Bachelor	70.6	64.4	0.217
Risk	3.2	3.4	0.335
Numeracy score	1.1	1.0	0.548
Behindness aversion	3.7	3.8	0.533
Aheadness aversion	5.5	5.0	0.097
Quiz attempts	4.9	5.2	0.617

Notes: Average characteristic by treatment. In the last column I report p-values of a t-test, comparing equality of means across the two treatments. Risk is the lottery chosen in the Eckel and Grossman (2002)-task, numeracy score the number of correct answers in the Berlin numeracy test using only questions 1 and 3 (Cokely et al., 2012), behindness and aheadness aversion the switching points in the multiple price list-elicitation of β and α-parameters in Fehr and Schmidt (1999)-preferences by Yang et al. (2016).

4.B.2 Investment behavior

In Figure 4.B.1, I show investment rates separately for each level of obedience.

Figure 4.B.1: Investment decisions



Notes: Average frequency of investment by treatment, bars indicate observed choices and red diamonds choices in the Bayes Nash equilibria with the highest investment. Bars and shaded areas indicate 95% bootstrapped confidence intervals.

In Table 4.B.3, I reproduce the treatment effect table from the main text with additional controls, as preregistered. Columns (1), (3), (5) to (7) and (10) present decisions from the experiment, (2), (4), (8), and (11) repeat the regressions when participants use the Nash equilibrium strategy. (9) and (12) interact models (7) and (8) or (10) and (11), respectively.

Columns (1), (3), and (5) show estimates omitted from the table in the main text. Columns (6) and (7) show that results are robust to additional controls. Columns (10) to (12) only use data from *optimal* levels of information structures, which uses one-third of the entire data set. Column (10) shows this level's larger theoretically predicted treatment effects. Column (10) shows that treatment effects in the experiment are robust to only using this level for testing. Columns (10) and (12) show that public structures empirically increase investment compared to the Nash equilibrium prediction, and similar so for both games.

Table 4.B.4 reports logit estimates of the main treatment effects. Results are in line with the OLS results reported in the main text.

Table 4.B.3: Treatment effects with additional controls: Investment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Substitutes		Complements							Diff-in-Diff		
	Obs.	NE	Obs.	NE	Obs.	Obs.	Obs.	NE	NE vs. Obs.	Obs.	NE	NE vs. Obs.
Public	-0.009	-0.043***	0.087**	0.030***	-0.009	-0.001	0.004	-0.043***	-0.009	-0.013	-0.073***	-0.013
	(0.019)	(0.011)	(0.039)	(0.007)	(0.020)	(0.020)	(0.021)	(0.011)	(0.019)	(0.028)	(0.017)	(0.028)
Complements					-0.108***	-0.101***	-0.100***	-0.211***	-0.108***	-0.120***	-0.011	-0.120***
					(0.032)	(0.031)	(0.032)	(0.009)	(0.033)	(0.038)	(0.015)	(0.039)
Public × Complements					0.096**	0.087*	0.084*	0.073***	0.096**	0.101*	0.149***	0.101*
					(0.045)	(0.044)	(0.045)	(0.013)	(0.045)	(0.059)	(0.022)	(0.060)
(1 if level=*optimal*)	-0.009	0.082***	-0.040*	0.122***	-0.025*	-0.024*	-0.025*	0.102***	0.039***			
	(0.018)	(0.014)	(0.021)	(0.016)	(0.014)	(0.014)	(0.014)	(0.011)	(0.010)			
(1 if level=*high*)	-0.038*	0.058***	-0.073***	-0.577***	-0.055***	-0.056***	-0.056***	-0.260***	-0.157***			
	(0.020)	(0.013)	(0.020)	(0.011)	(0.014)	(0.014)	(0.014)	(0.039)	(0.022)			
Period	-0.002**		-0.004***		-0.003***	-0.003***	-0.003***		-0.001**	-0.003**		-0.000
	(0.001)		(0.001)		(0.001)	(0.001)	(0.001)		(0.001)	(0.001)		(0.001)
(1 if part=2)	0.017		-0.032*		-0.007	-0.008	-0.009		0.001	-0.052		-0.026
	(0.021)		(0.018)		(0.014)	(0.014)	(0.014)		(0.019)	(0.034)		(0.020)
(1 if part=3)	0.007		-0.023		-0.008	-0.008	-0.008		0.001	0.008		0.012
	(0.020)		(0.020)		(0.014)	(0.014)	(0.014)		(0.018)	(0.037)		(0.021)
(1 if session in Munich)	0.019		0.106**		0.063***	0.057**	0.053**		0.030**	0.069**		0.037**
	(0.020)		(0.040)		(0.023)	(0.024)	(0.024)		(0.014)	(0.029)		(0.017)
Behindness aversion						0.014**	0.011*					
						(0.006)	(0.006)					
Aheadness aversion						0.004	0.003					
						(0.003)	(0.003)					
Risk						0.012*	0.015**					
						(0.007)	(0.006)					
Numeracy						-0.016**	-0.015**					
						(0.007)	(0.007)					
Age							-0.002					
							(0.003)					
(1 if woman)							0.046***					
							(0.017)					
NE									0.131***			0.173***
									(0.016)			(0.019)
Public × NE									-0.034*			-0.060**
									(0.019)			(0.026)
Complements × NE									-0.103***			0.109***
									(0.031)			(0.038)
Public × Complements × NE									-0.023			0.048
									(0.044)			(0.059)
Constant	0.514***	0.574***	0.445***	0.562***	0.533***	0.461***	0.488***	0.674***	0.532***	0.518***	0.672***	0.496***
	(0.022)	(0.010)	(0.038)	(0.013)	(0.022)	(0.054)	(0.093)	(0.014)	(0.023)	(0.030)	(0.013)	(0.026)
Observations	12960	12960	12948	12948	25908	25860	25800	25908	51816	8628	8628	17256
# clusters	36	36	36	36	72	72	72	72	72	72	72	72
# participants	216	216	216	216	432	432	432	432	432	432	432	432
Levels	All	All	All	All	All	All	All	All	All	Only *optimal*	Only *optimal*	Only *optimal*

Notes: The table reports OLS estimates. The dependent variable is the choice to invest, either observed in the experiment (Obs.) or the investment choice in the equilibrium with the highest investment (NE). (1) and (2) use observed or predicted data from games of strategic substitutes, (3) and (4) from games of strategic complements and (5) to (12) pool all data. (10) to (12) use only data from *optimal* levels, all other models are estimated using all levels. Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. Aheadness aversion and behindness aversion are switching points in the choice lists to elicit α (behindness) and β-parameters (aheadness) of the Fehr and Schmidt (1999)-model, elicited using the task by Yang et al. (2016). Both measures range from 1 to 11, with mean 3.6, standard deviation 1.4 for behindness, and with mean 5.3, standard deviation 2.9 for aheadness aversion. Risk is the lottery chosen in the Eckel and Grossman (2002)-task, ranging from 1 to 6 with mean 3.2, standard deviation 1.5. Numeracy is the number of correct items in the Berlin numeracy test (Cokely et al., 2012), ranging from 0 to 4, mean 2.4, standard deviation 1.2. NE is a dummy variables equal one if this is a predicted choice in the Nash equilibrium with maximal investment, with the omitted category being an observed choice in the experiment. Standard errors in parentheses clustered on matching-group level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.B.4: Logit estimates of the treatment effect: Investment

	(1) Substitutes	(2) Complements	(3) Diff-in-diff
Public	-0.035	0.364**	-0.035
	(0.076)	(0.165)	(0.081)
Complements			-0.445***
			(0.136)
Public × Complements			0.395**
			(0.184)
(1 if level=*optimal*)	-0.036	-0.166*	-0.100*
	(0.073)	(0.089)	(0.058)
(1 if level=*high*)	-0.151*	-0.304***	-0.226***
	(0.081)	(0.083)	(0.058)
Constant	0.054	-0.222	0.138
	(0.087)	(0.160)	(0.090)
Period trend, part and lab FE	Yes	Yes	Yes
Observations	12960	12948	25908
# clusters	36	36	72
# participants	216	216	432

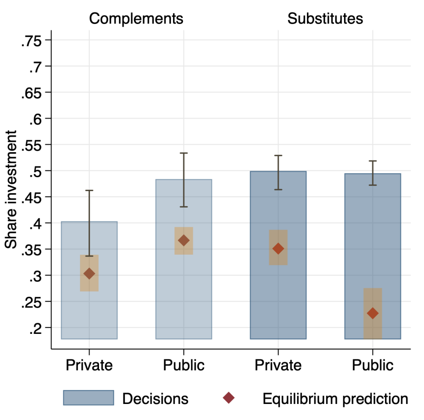
Notes: The table reports logit estimates and includes all data, also *high* structures. The dependent variable is a dummy variable equal to 1 if the participant decided to invest (Data) or was predicted to invest in the Bayes Nash equilibrium with maximal investment (NE). Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. (1 if level=*optimal*) and (1 if level=*high*) are dummy variables equal to 1 if the information structures used *optimal* or *high* probabilities to persuade receivers to invest, relative to the omitted category *low*, respectively. Standard errors in parentheses, clustered at the matching-group level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.B.3 Investment and risk aversion

Observed investment rates in the experiment are, on average, below the predictions. These predictions are based on risk-neutral receivers. Empirically, the two control task measuring risk aversion at the end of the experiment show that an overwhelming majority of participants are risk averse. Furthermore, adding the risk aversion measure introduced by Eckel and Grossman (2002) correlates significantly with investment choice, see Table 4.B.3, patterns are similar using the second control task.

Information from these task can also be used to adjust the equilibrium predictions for the level of risk aversion at the participant level. This is especially relevant for the optimal information structures, at expected profits are slim, while participants face risk. Even only slightly risk averse participants may not be willing to invest at this level. To account for this riskiness, I use the CRRA utility with the coefficients estimated from the lottery choice elicited in the Eckel and Grossman (2002) task, and, as a lower bound, calculate a best response to others' behaving as in equilibrium under risk neutrality. Figure 4.B.2 shows predicted choices, which, if anything, indicate that participants are willing to invest more often than predicted given their measured level of risk aversion.

Figure 4.B.2: Investment decisions



Notes: Average frequency of investment by treatment, bars indicate observed choices and red diamonds choices in the Bayes Nash equilibria with the highest investment when using participants' risk aversion elicited in the Eckel and Grossman (2002) task to calculate their expected utility. The figure only uses data from *low* and *optimal* structures. Bars and shaded areas indicate 95% bootstrapped confidence intervals.

4.B.4 Learning

In Table 4.B.5, I report regressions on learning effects for investment and following. (1) to (4) split data in the first 7 (in (1) and (3)) vs. the last 13 periods (in (2) and

(4)). (5) to (7) repeat the investment regression for each part separately. Results are robust across periods and parts, except the no longer significant estimate on Public × Complements in (6) for part 2.

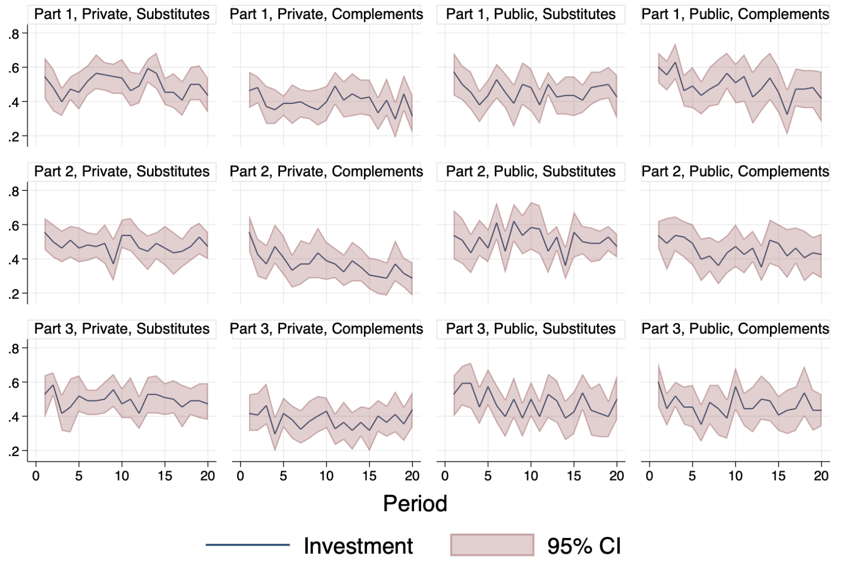
Table 4.B.5: Learning: Investment and following

	Investment		Following		Investment		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Public	0.004	-0.016	0.047**	0.054**	-0.038	0.031	-0.019
	(0.020)	(0.023)	(0.020)	(0.023)	(0.028)	(0.027)	(0.029)
Complements	-0.090***	-0.118***	-0.088***	-0.101***	-0.100***	-0.108***	-0.116***
	(0.032)	(0.034)	(0.028)	(0.032)	(0.033)	(0.040)	(0.036)
Public × Complements	0.085*	0.102**	-0.025	-0.033	0.132**	0.050	0.105*
	(0.043)	(0.049)	(0.039)	(0.046)	(0.050)	(0.054)	(0.056)
(1 if level=*optimal*)	-0.014	-0.030*	-0.109***	-0.134***	0.006	-0.089***	0.010
	(0.017)	(0.016)	(0.015)	(0.016)	(0.030)	(0.031)	(0.033)
(1 if level=*high*)	-0.020	-0.074***	-0.215***	-0.255***	-0.025	-0.079**	-0.062*
	(0.016)	(0.017)	(0.015)	(0.017)	(0.029)	(0.033)	(0.032)
Constant	0.540***	0.551***	0.971***	0.947***	0.522***	0.553***	0.510***
	(0.025)	(0.028)	(0.024)	(0.029)	(0.026)	(0.030)	(0.031)
Period	1-7	13-20	1-7	13-20	1-20	1-20	1-20
Part	1-3	1-3	1-3	1-3	1	2	3
Period trend, part, and lab FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9072	16836	9072	16836	8640	8640	8628
# clusters	72	72	72	72	72	72	72
# participants	432	432	432	432	432	432	432

Notes: The table reports OLS estimates. The dependent variable is a dummy variable equal to 1 if the participant decided to invest in (1), (2), and (5) to (7), or the participant followed the received recommendation in (3) and (4). Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. (1 if level=*optimal*) and (1 if level=*high*) are dummy variables equal to 1 if the information structures used *optimal* or *high* probabilities to persuade receivers to invest, relative to the omitted category *low*, respectively. Standard errors in parentheses, clustered at the matching-group level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In Figure 4.B.3, I plot the average investment rate for the four between-subject treatments, separately for each part. Investment rates are similar over time across both dimensions of learning: between parts and within parts, over periods.

Figure 4.B.3: Learning



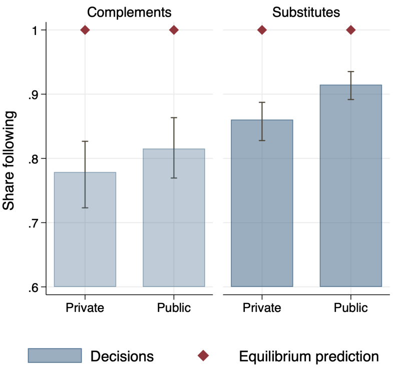
Notes: Average investment per period in the blue line, with 95%, bootstrapped confidence intervals (clustered on matching-group level) shaded in red. Separately by part (part 1, 2, and 3), public vs. private and substitutes vs. complements.

4.B.5 Following behavior

In this section, I present some additional statistics on the following behavior. In Figure 4.B.4, I show the average decision to follow averaged on a between-subject treatment level. In Table 4.B.6, I show regressions of the decisions to follow on treatment dummies with additional controls. In Table 4.B.7, I report estimates when repeating the analysis from the main text, but only using data when participants receive the recommendation to invest, which removes any variation in how often recommendations not to invest are being followed. Results are broadly in line with the analysis from the main text. In addition, I report estimates when regressing the squared distance between observed following decisions and the best response to beliefs in column (5). Empirical behavior is closer to the best response in public structures (estimate on Public, p-value=0.0247), but does not differ significantly in the other between-subject treatment dimensions.

Figure 4.B.5 repeats the best response analysis from Figure 4.2 using the empirical frequencies in the data instead of participants' beliefs. Differences can be attributed to errors in belief updating, either about the state or about others' actions. While broadly similar, especially in games of strategic substitutes receivers underinvest. Participants in games of strategic complements underreact to changes in

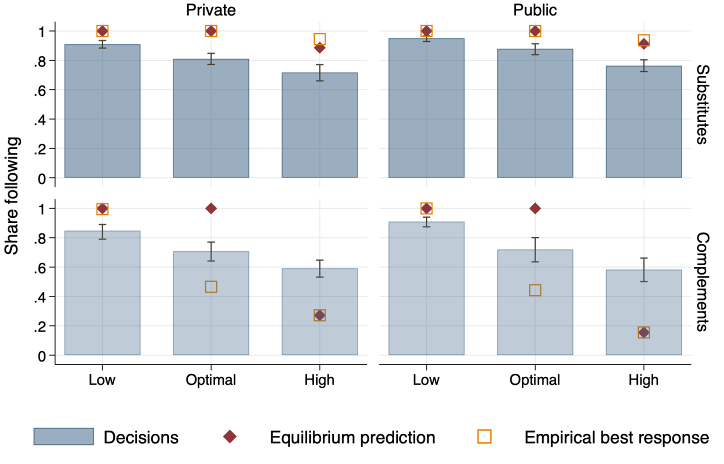
Figure 4.B.4: Following rates



Notes: Average frequency of following a recommendation by treatment, bars indicate observed choices. Error bars indicate 95% bootstrapped confidence intervals.

obedience: they follow not often enough for *low* levels but follow too frequently for *optimal* and *high* levels.

Figure 4.B.5: Following rates



Notes: Average frequency of following a recommendation by treatment and level of the information structure. The variable is a dummy equal to 1 if a recommendation was followed (investment after the recommendation to invest, no investment after the recommendation not to invest). Bars indicate observed choices, diamonds following rate in the equilibrium with the highest following, and squares are empirical best responses based on others' choices in the experiment. Error bars indicate 95% bootstrapped confidence intervals.

Table 4.B.6: Treatment effects with additional controls: Following

	(1)	(2)	(3)
Public	0.052**	0.057***	0.061***
	(0.021)	(0.021)	(0.022)
Complements	-0.096***	-0.092***	-0.090***
	(0.030)	(0.029)	(0.030)
Public × Complements	-0.030	-0.036	-0.041
	(0.043)	(0.042)	(0.043)
(1 if level=*optimal*)	-0.125***	-0.125***	-0.126***
	(0.015)	(0.015)	(0.015)
(1 if level=*high*)	-0.241***	-0.242***	-0.242***
	(0.015)	(0.015)	(0.015)
(1 if part=2)	-0.009	-0.010	-0.010
	(0.014)	(0.014)	(0.014)
(1 if part=3)	-0.005	-0.005	-0.004
	(0.013)	(0.013)	(0.013)
Period	-0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)
(1 if session in Munich)	0.056**	0.051**	0.050**
	(0.022)	(0.023)	(0.023)
Behindness aversion		0.007	0.004
		(0.005)	(0.005)
Aheadness aversion		0.004	0.004
		(0.003)	(0.003)
Risk		0.011*	0.014**
		(0.006)	(0.006)
Numeracy		-0.009	-0.008
		(0.007)	(0.007)
Age			-0.004
			(0.003)
(1 if woman)			0.047***
			(0.017)
Constant	0.956***	0.893***	0.962***
	(0.024)	(0.048)	(0.085)
Observations	25908	25860	25800

Notes: The table reports OLS estimates. The dependent variable is the decision to follow a recommendation (invest after recommended to invest, not invest after not invest). Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. Aheadness aversion and behindness aversion are switching points in the choice lists to elicit α (behindness) and β-parameters (aheadness) of the Fehr and Schmidt (1999)-model, elicited using the task by Yang et al. (2016). Both measures range from 1 to 11, with mean 3.6, standard deviation 1.4 for behindness, and with mean 5.3, standard deviation 2.9 for aheadness aversion. Risk is the lottery chosen in the Eckel and Grossman (2002)-task, ranging from 1 to 6 with mean 3.2, standard deviation 1.5. Numeracy is the number of correct items in the Berlin numeracy test (Cokely et al., 2012), ranging from 0 to 4, mean 2.4, standard deviation 1.2. Standard errors in parentheses clustered on matching-group level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.B.7: Treatment effects: Following the recommendation to invest

	(1) Data	(2) NE	(3) Data	(4) NE	(5) (Data-NE)^2
Public	0.062**	0.000	0.052*	0.006	-0.086**
	(0.029)	(.)	(0.031)	(0.005)	(0.038)
Complements	-0.148***	0.000	-0.158***	-0.318***	-0.044
	(0.048)	(.)	(0.046)	(0.004)	(0.035)
Public × Complements	0.021	0.000	0.023	-0.014**	0.083
	(0.063)	(.)	(0.062)	(0.006)	(0.051)
(1 if level=*optimal*)	-0.164***	0.000	-0.163***	0.004	0.165***
	(0.020)	(.)	(0.020)	(0.003)	(0.022)
(1 if level=*high*)			-0.281***	-0.588***	0.187***
			(0.017)	(0.051)	(0.030)
Constant	0.897***	1.000	0.925***	1.163***	0.338***
	(0.032)	(.)	(0.033)	(0.019)	(0.033)
Level of obedience	*low* & *optimal*		*low*, *optimal* & *high*		
Period trend, part and lab FE	Yes	No	Yes	No	Yes
Observations	10638	10638	17110	17110	17110
# clusters	72	72	72	72	72
# participants	432		432	432	

Notes: The table reports OLS estimates and includes only data where participants received the recommendation to invest. In columns (1) to (4), the dependent variable is a dummy variable equal to 1 if the participant decided to follow a recommendation (invest after recommended to invest, or not invest after recommended not to invest) (Data) or was predicted to follow in the Bayes Nash equilibrium with maximal investment (NE). Columns (1) and (2) use data only from obedient structures, while columns (3) and (4) pool all data. In column (5), the dependent variable is the squared distance between decision to follow the recommendation to invest in the data and the predicted best response to beliefs ((Data-BR)^2). Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. (1 if level=*optimal*) and (1 if level=*high*) are dummy variables equal to 1 if the information structures used *optimal* or *high* probabilities to persuade receivers to invest, relative to the omitted category *low*, respectively. Standard errors in parentheses, clustered at the matching-group level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.B.6 Beliefs

In Table 4.B.8, I report data on all elicited beliefs for all treatments and levels. This now includes beliefs on what participants believed about the state and others' actions after receiving the recommendation not to invest. Beliefs are consistent with three key observations. First, across all treatments, on average, participants understand that the recommendation to invest is good news about the state. In contrast, the recommendation to invest is bad news, as beliefs about the state being good are higher after receiving the recommendation to invest. Second, they understand that others respond reasonably to recommendations, as they are more likely to invest after receiving this recommendation. Third, participants follow the expected pattern across levels, as they are less optimistic about the state and others' investment moving from *low* to *optimal* to *high* levels. Notable is also that participants' beliefs about the state across private structures (comparing complements and substitutes) are virtually identical. These structures were designed to induce identical beliefs, and participants between treatments responded identically. Last, note that beliefs after receiving the recommendation not to invest are likely also surprisingly high because reports were measured for zero or higher; thus, noise in decision-making was only captured for positive errors. E.g., more than 75% of beliefs about the state are 0, as theoretically predicted; only a minority of participants report a positive probability of the state being good even though this is theoretically not possible.

Table 4.B.8: Belief data

		Recommendation to invest		Recommendation not to invest	
Treatment	Level	State	Others invest	State	Others invest
Complements, Public	*Low*	.79	.77	.13	.14
	Optimal	.67	.64	.14	.17
	High	.65	.60	.13	.19
Complements, Private	*Low*	.80	.76	.10	.10
	Optimal	.71	.62	.11	.14
	High	.65	.54	.09	.13
Substitutes, Public	*Low*	.88	.90	.09	.17
	Optimal	.77	.81	.08	.17
	High	.72	.72	.08	.19
Substitutes, Private	*Low*	.79	.80	.11	.20
	Optimal	.69	.70	.13	.17
	High	.65	.67	.14	.21

Average beliefs of the state being good ("State") or others' decision to invest ("Others invest") in response to receiving the recommendation to invest or not to invest. Beliefs are coded as shares, with dummies equal to 1 if the state is good or others invest, respectively.

4.B.7 Risk aversion and following behavior

As an additional measure of risk, I use the separately elicited risk aversion (Eckel and Grossman, 2002). In Table 4.B.9, I regress the decision to follow a recommendation on the risk measure, treatment dummies, and most importantly, their interaction, adding controls from (1) to (3). It does not appear to be the case that the risk measure captures differences in behavior specific to public information structures.

Table 4.B.9: Following and risk aversion

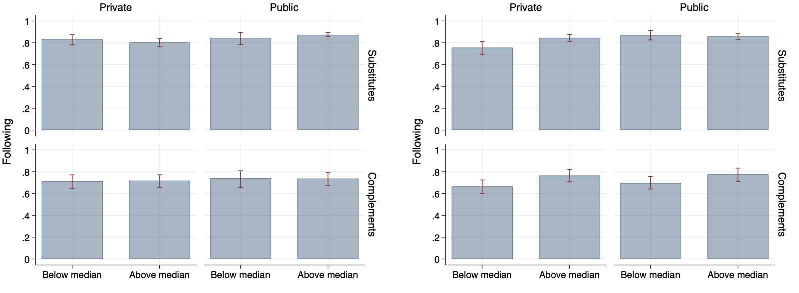
	(1)	(2)	(3)
Public	0.042	0.013	0.024
	(0.047)	(0.056)	(0.056)
Complements	-0.112***	-0.131*	-0.121*
	(0.022)	(0.067)	(0.067)
Public × Complements		0.059	0.042
		(0.095)	(0.096)
Risk	0.009	0.004	0.006
	(0.010)	(0.015)	(0.014)
Risk × Public	-0.002	0.012	0.009
	(0.013)	(0.017)	(0.017)
Risk × Complements		0.011	0.008
		(0.020)	(0.019)
Risk × Public × Complements		-0.028	-0.023
		(0.026)	(0.026)
Constant	0.792***	0.801***	0.900***
	(0.034)	(0.047)	(0.047)
Part, level and lab FE	No	No	Yes
Observations	25860	25860	25860
# clusters	72	72	72
# participants	431	431	431

Notes: The table reports OLS estimates. The dependent variable is the choice to follow a recommendation (investing after being recommended to invest, not investing after being recommended not to invest). Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. Risk is the lottery chosen in the Eckel and Grossman (2002)-task, where higher numbers indicate lower risk aversion. The index ranges from 1 to 6, with mean 2.3 and standard deviation 1.5. Standard errors in parentheses clustered on matching-group level. * p < 0.10, ** p < 0.05, *** p < 0.01.

4.B.8 Inequity aversion and following behavior

Another candidate to explain the superiority of public structures are social preferences. If followed, public structures minimize payoff inequality between participants. In contrast, following a private structure leads to unequal payoffs if the bad state realizes. To test this mechanism, I included an elicitation of the preference parameters of the Fehr and Schmidt (1999)-model, using the task by Yang et al. (2016). In Figure 4.B.6, I show the following rate when performing median splits by the aversion to being ahead in the left panel and by the aversion to being behind in the right panel. No clear pattern may explain higher following rates only in public information structures. Generally, the aversion to being behind appears to lead to more following.

Figure 4.B.6: Following and inequity aversion



Notes: Average following rate. Left panel: Median split by aversion to being ahead. Right panel: Median split by aversion to being behind. Bars indicate observed choices, diamonds the observed target in the data, and error bars indicate 95% bootstrapped confidence intervals.

Result 9. *Inequity aversion cannot explain the higher following in public information structures.*

In Table 4.B.10, I show how the decision to follow recommendations correlates with inequity aversion parameters (Fehr and Schmidt, 1999), especially for public information structures. There is no significant effect of either aversion to being ahead or behind.

Table 4.B.10: Following and inequity aversion

	(1)	(2)
Public	0.088	0.098
	(0.066)	(0.069)
Behindness aversion	0.009	0.008
	(0.007)	(0.007)
Public × Behindness aversion	-0.003	-0.003
	(0.010)	(0.010)
Aheadness aversion	0.009**	0.007
	(0.004)	(0.004)
Public × Aheadness aversion	-0.007	-0.006
	(0.006)	(0.006)
Complements	-0.111***	-0.097***
	(0.022)	(0.030)
Public × Complements		-0.029
		(0.043)
Constant	0.740***	0.851***
	(0.048)	(0.053)
Part, level and lab FE	No	Yes
Observations	25860	25860
Adjusted R^2	0.022	0.083
# clusters	72	72
# participants	431	431

Notes: The table reports OLS estimates. The dependent variable is the choice to follow a recommendation (investing after being recommended to invest, not investing after being recommended not to invest). Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. Aheadness aversion and behindness aversion are switching points in the choice lists to elicit α (behindness) and β-parameters (aheadness) of the Fehr and Schmidt (1999)-model, elicited using the task by Yang et al. (2016). Both measures range from 1 to 11, with mean 3.6, standard deviation 1.4 for behindness, and with mean 5.3, standard deviation 2.9 for aheadness aversion. Standard errors in parentheses clustered on matching-group level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.B.9 Noise in beliefs

Table 4.B.11 documents that beliefs are less noisy in public groups. I regress the standard deviation in beliefs within a matching group, at each level, on treatment dummies. Note that this standard deviation only captures variance within a group: Each participant reported beliefs only once for each level, thus any noise perceived by each participant within a level is not captured.

Table 4.B.11: Noise in beliefs

	(1) SD(beliefs)
Complements	0.163
	(0.238)
Public	-0.415*
	(0.219)
Complements × Public	0.358
	(0.329)
(1 if level=*optimal*)	0.103
	(0.122)
(1 if level=*high*)	0.261**
	(0.124)
Constant	1.910***
	(0.141)
Observations	216
# clusters	72

Notes: The table reports OLS estimates. The dependent variable is the standard deviation in beliefs about others' following a recommendation to invest. This is calculated on the matching group-level level, so one observation is the standard deviation within a matching group for each level (*low*, *optimal* or *high*). Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. (1 if level=*optimal*) and (1 if level=*high*) are dummy variables equal to 1 if the information structures used *optimal* or *high* probabilities to persuade receivers to invest, relative to the omitted category *low*, respectively. Standard errors in parentheses clustered on matching-group level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.B.10 Experiencing bad advice: Robustness

In Section 4.4.4, I show that only in private information structures, experiencing bad advice leads to lower investment and following in future periods. This section presents two robustness checks.

First, I show that the result is robust to different rules to capture who has received bad advice. I repeat the analysis presented in the main text, but count the number of times a participant has received bad advice within each information structure. In addition, I perform a median split of participants who received bad advice more often than the median facing the same information structure, which accounts for the fact that the frequency of receiving bad advice is correlated with the type of structure.

Results in Table 4.B.12 indicate that patterns are similar using the new measures. Columns (1) and (2) report estimates using the number of times bad advice was sent to a participant, columns (3) and (4) report estimates using the median split. Columns (1) and (3) use the decision to invest as dependent variables, (2) and (4) the decision to follow. Note that the bad advice-proxies are not significant in (1) and (3). Yet, across both specifications, public structures lead to higher investment of those participants that initially received bad advice, consistent with the analysis in the main text. Columns (2) and (4) show that those receiving bad advice more often follow less often, but this effect is not present in public structures, in line with the analysis in the main text.

Table 4.B.12: Robustness of bad advice

	(1) Investment	(2) Following	(3) Investment	(4) Following
Public	-0.048*	-0.002	-0.035	0.020
	(0.028)	(0.023)	(0.023)	(0.024)
Complements	-0.062*	-0.059*	-0.083**	-0.080**
	(0.036)	(0.031)	(0.034)	(0.031)
Public × Complements	0.086*	0.041	0.081*	-0.031
	(0.045)	(0.038)	(0.045)	(0.041)
# bad advice	-0.005	-0.020***		
	(0.007)	(0.006)		
Public × # bad advice	0.016*	0.015*		
	(0.009)	(0.009)		
Complements × # bad advice	-0.014*	-0.011		
	(0.008)	(0.008)		
Public × Complements × # bad advice	-0.000	-0.011		
	(0.011)	(0.012)		
Above median bad advice			-0.023	-0.071***
			(0.022)	(0.019)
Public × Above median bad advice			0.070**	0.077**
			(0.030)	(0.031)
Complements × Above median bad advice			-0.057*	-0.035
			(0.032)	(0.030)
Public × Complements × Above median bad advice			0.029	-0.009
			(0.046)	(0.046)
Constant	0.540***	0.982***	0.544***	0.989***
	(0.026)	(0.023)	(0.025)	(0.027)
Period trend, part & lab FE	Yes	Yes	Yes	Yes
Observations	25908	25908	25908	25908

Notes: The table reports OLS estimates and includes all data. The dependent variables are the decision to invest (1) and (3) or the decision to follow a recommendation (2) and (4). # bad advice is the number of times a participant received bad advice when facing an information structure. Above median bad advice is a dummy variable equal one if the participant received bad advice more often than the median times all participants facing that same structure received bad advice. Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. Standard errors in parentheses, clustered at the matching-group level,* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Second, I show that this pattern is driven by participants that receive bad advice. An alternative explanation may be a preference for conformism, or for always receiving the same recommendation. To test this alternative, I rerun the analysis presented in the main text, but instead compare participants that receive the recommendation not to invest to participants who receive the recommendation to invest in the good state in the first period of an information structure, which removes all participants that receive bad advice in the first period. Both remaining groups of participants receive good advice. However, participants that receive the recommendation not to invest with private information structure may experience miscoordinated advice, as their matched participant may receive the recommendation to invest. Instead, participants with public information structures always receive the same recommendation. The alternative explanations would predict that participants respond differently to experience the same or different recommendations. Conformity-driven explanations would imply that participants that experience different recommendation with private structures change their follow-up behavior in patterns similar to those participants who receive bad advice.

The results in Table 4.B.13 indicate that participants that receive the recommendation not to invest in the first period do not invest or follow differently in follow-up periods, irrespective of whether they face public or private information structure, compared to participants that receive the recommendation to invest in the good state. This indicates that the conformity is an unlikely explanation of the data. Instead, the data is consistent with participants disliking experiencing miscoordinated bad advice.

Table 4.B.13: Miscoordinated good advice and future following

	(1) Investment	(2) Following	(3) Following
Public	-0.033	0.019	0.031
	(0.028)	(0.029)	(0.030)
Complements	-0.122***	-0.121***	-0.115***
	(0.041)	(0.040)	(0.040)
Public × Complements	0.128**	0.001	-0.012
	(0.056)	(0.054)	(0.055)
Not invest	-0.013	-0.038	-0.043
	(0.032)	(0.031)	(0.030)
Public × Not invest	0.015	0.038	0.034
	(0.043)	(0.040)	(0.040)
Complements × Not invest	0.029	0.056	0.060
	(0.046)	(0.046)	(0.045)
Public × Complements × Not invest	-0.055	-0.046	-0.045
	(0.072)	(0.069)	(0.068)
Constant	0.548***	0.982***	0.974***
	(0.028)	(0.029)	(0.087)
Period trend, part & lab FE	Yes	Yes	Yes
Additional controls	No	No	Yes
Observations	20615	20615	20558

Notes: The table reports OLS estimates and includes all data after period one in each part. I only use data where participants received good advice, so either the recommendation not to invest in the bad state or the recommendation to invest in the good state. Column (3) uses fewer observations, as some additional controls are not available for all participants. The dependent variables are the decision to invest or the decision to follow a recommendation. Not invest is a dummy variable equal to 1 if a participant received a recommendation not to invest in period 1 of the corresponding information structure. Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. The additional controls are participants' Fehr and Schmidt (1999) preferences, risk aversion, numeracy score, and demographic variables. Standard errors in parentheses, clustered at the matching-group level,* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.B.11 Benchmarking the importance of the two mechanisms

In this section, I provide a rough estimate of the relative contributions of the two mechanisms to advantage of public signals. Table 4.B.14 provides the needed estimates. Model (1) shows that public signals lead to 4 percentage points higher investment across all games. This is an advantage not predicted by theory: model

(2) indicates that in the Bayes Nash equilibrium with maximal investment, no advantage of public structures would be expected.

First, I find that participants who receive bad advice reduce their follow-up investment. Model (3) indicates that in public structures, participants who receive bad advice invest 10 percentage points more than those with private structures. However, receiving such advice is probabilistic: on average, only 16% of participants received bad advice in period 1. This means that the effect on average behavior is only 1.6 percentage points.

Second, I find that in groups with above-median variance, public structures lead to 5 percentage points higher investment. As this effect is only present for half of the groups, those with above-median variance, the total effect is 2.5 percentage points.

Therefore, the total effect of 4 percentage points is can approximately be attributed to a 1.6 percentage point effect of bad advice, and a 2.5 percentage point effect of complexity and high variance. This means that the total effect of complexity is roughly 2.5/(1.6+2.5)=61%.

Table 4.B.14: Decomposing the effect of the two mechanisms

	(1) Investment	(2) NE investment	(3) Investment	(4) Investment
Public	0.039*	-0.007	0.024	0.010
	(0.023)	(0.008)	(0.023)	(0.024)
Complements	-0.060**	-0.174***	-0.061**	-0.056***
	(0.023)	(0.008)	(0.023)	(0.019)
Bad advice			-0.126***	
			(0.026)	
Public × Bad advice			0.102***	
			(0.033)	
High variance				-0.139***
				(0.024)
Public × High variance				0.052
				(0.037)
Constant	0.509***	0.643***	0.511***	0.589***
	(0.024)	(0.025)	(0.024)	(0.024)
Period trend; part, level and lab FE	Yes	Yes	Yes	Yes
Observations	25908	25908	24612	25908

Notes: The table reports OLS estimates. In (1), (3) and (4), the dependent variable is a dummy equal one if the participant chose to invest. In (3), the dependent variable is a dummy equal one if the participant would have been predicted to invest in the Bayes Nash equilibrium with maximal investment. Public and Complements are the treatment indicators. These are dummy variables equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure, or facing a game with strategic complements, with the omitted category being a game of strategic substitutes. Bad advice is a dummy variable equal to 1 if a participant received a recommendation to invest when the state was bad in period 1 of the corresponding information structure. High variance is a dummy variable equal to 1 if the average standard deviation of the matching group (calculated as in (1)) is above the median within each treatment. Standard errors in parentheses, clustered at the matching-group level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.B.12 Additional analysis on the second experiment

Table 4.B.15 reports an analysis of the regressions of the second experiment, Table 4.8, separately for the first third and the last two-thirds, to study learning. Columns

(1), (3) and (5) use data from periods 1 to 7, columns (2), (4) and (6) from periods 8 to 21, as preregistered.

There are clear indications for learning. Comparing columns (1) and (2), we observe that the average use of public signals across both games increases, as the constant increases from 38% to 47%. Columns (4) and (6) also show that receivers persuade less aggressively over time in games of complements. The coefficient on Complements is positive in (3), at the start, but no longer so with experience in (4). Similarly, the coefficient on Complements is not significant at the start in (5), but significant and negative in (6), with experience. For games of substitutes, if anything, receivers become more aggressive over time, as the coefficient on the constant increases in (6), compared to (5), thus senders are more likely to choose *high* instead of *optimal* structures.

Table 4.B.15: Senders: Treatment effects and learning

	(1)	(2)	(3)	(4)	(5)	(6)
	Public		*Optimal* vs. *low*		*High* vs. *optimal*	
Complements	0.244***	0.211***	0.156*	-0.032	-0.123	-0.188*
	(0.064)	(0.072)	(0.086)	(0.118)	(0.085)	(0.110)
Period	0.007	0.001	-0.009	-0.004	-0.014	-0.003
	(0.010)	(0.005)	(0.017)	(0.007)	(0.017)	(0.006)
Constant	0.382***	0.466***	0.133	0.017	0.173*	0.263**
	(0.059)	(0.098)	(0.090)	(0.127)	(0.090)	(0.123)
Period	1-7	8-21	1-7	8-21	1-7	8-21
Lab FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	840	1680	840	1680	840	1680
# clusters	40	40	40	40	40	40
# participants	120	120	120	120	120	120

Notes: The table reports OLS estimates. In (1) and (2), the dependent variable is a dummy equal one if the sender chose a public structure. In (3) and (4), the dependent variable is the difference in level shares, as the share of *optimal* minus the share of *low* structures. In (5) and (6), he dependent variable is the difference in level shares, as the share of *high* minus the share of *optimal* structures. (1), (3) and (5) use data from periods 1 to 7; (2), (4) and (6) from peridos 8 to 21. Complements is the treatment indicator, a dummy variable equal to 1 if the decision was made when receivers face a game with strategic complements, with a game of substitutes as the omitted category. Period is a linear period trend. Standard errors in parentheses, clustered at the matching-group level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4.B.16 complement the analysis on beliefs in the main text. I regress the belief that the state is good after receiving a recommendation to invest on characteristics of the information structure and the game. The estimates show that also beliefs about the state are updated very similarly for receivers in the first and second experiment. Again, senders underestimate the extent to which receivers' update, in response to *optimal* or *high* structures.

Table 4.B.16: Beliefs about the state across all experiments

	(1)	(2)	(3)	(4)
	Belief: Probability state is good			
Public	0.078***	0.081***	0.059**	0.078***
	(0.015)	(0.013)	(0.024)	(0.015)
Complements	0.011	0.024	0.008	0.011
	(0.016)	(0.020)	(0.027)	(0.016)
Public × Complements	-0.099***	-0.105***	-0.033	-0.099***
	(0.023)	(0.017)	(0.029)	(0.023)
(1 if level=*optimal*)	-0.102***	-0.107***	-0.025	-0.102***
	(0.006)	(0.009)	(0.015)	(0.006)
(1 if level=*high*)	-0.148***	-0.151***	-0.042**	-0.148***
	(0.008)	(0.010)	(0.016)	(0.008)
Second exp., receivers				-0.049**
				(0.020)
Second exp., senders				-0.201***
				(0.025)
Public × Second exp., receivers				0.003
				(0.019)
Public × Second exp., senders				-0.019
				(0.028)
Complements × Second exp., receivers				0.014
				(0.025)
Complements × Second exp., senders				-0.004
				(0.031)
Public × Complements × Second exp., receivers				-0.006
				(0.029)
Public × Complements × Second exp., senders				0.066*
				(0.037)
(1 if level=*optimal*) × Second exp., receivers				-0.006
				(0.011)
(1 if level=*optimal*) × Second exp., senders				0.077***
				(0.016)
(1 if level=*high*) × Second exp., receivers				-0.003
				(0.013)
(1 if level=*high*) × Second exp., senders				0.106***
				(0.018)
Constant	0.789***	0.742***	0.584***	0.789***
	(0.011)	(0.019)	(0.025)	(0.011)
Experiment	First	Second	Second	Both
Role	Receivers	Receivers	Senders	Both
Lab FE	Yes	Yes	Yes	Yes
Observations	1293	1440	720	3453
# clusters	72	40	40	112
# participants	431	240	120	791

Notes: The table reports OLS estimates. The dependent variable is the reported belief that the state is good after receiving the recommendation to invest. (1) uses data from the first experiment, with only receivers. (2) and (3) use data from the second experiment. (2) are the receivers, (3) the senders. (4) pools data from both experiments and both roles. Public and Complements are dummy variables equal to 1 if the beliefs was reported for facing a public rather than a private information structure, or facing a game with strategic complements rather than substitutes respectively. (1 if level=*optimal*) and (1 if level=*high*) are dummy variables equal to 1 if the information structures used *optimal* or *high* probabilities to persuade receivers to invest, relative to the omitted category *low*, respectively. Second exp., receivers and Second exp., senders are dummies equal one if the belief is measured in the second experiment, for receivers and senders, respectively. The omitted category are the receivers in the first experiment. Standard errors in parentheses, clustered at the matching-group level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

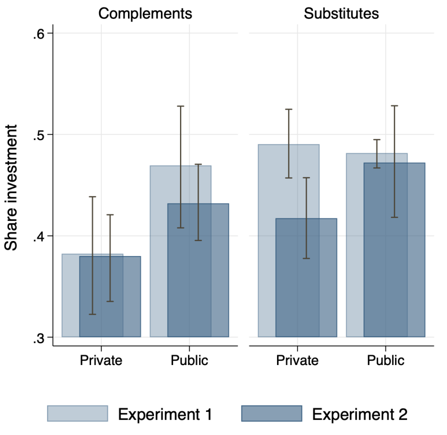
Table 4.B.17: Beliefs about the state across all experiments

	(1)	(2)	(3)
	Belief: Probability state is good		
Public	0.078***	0.081***	0.059**
	(0.015)	(0.013)	(0.024)
Complement	0.011	0.024	0.008
	(0.016)	(0.020)	(0.027)
Public × Complement	-0.099***	-0.105***	-0.033
	(0.023)	(0.017)	(0.029)
(1 if level=*optimal*)	-0.102***	-0.107***	-0.025
	(0.006)	(0.009)	(0.015)
(1 if level=*high*)	-0.148***	-0.151***	-0.042**
	(0.008)	(0.010)	(0.016)
Constant	0.789***	0.742***	0.584***
	(0.011)	(0.019)	(0.025)
Experiment	First	Second	Second
Role	Receivers	Receivers	Senders
Lab FE	Yes	Yes	Yes
Observations	1293	1440	720
# clusters	72	40	40
# participants	432	240	120

Notes: The table reports OLS estimates. The dependent variable is the reported belief that the state is good after receiving the recommendation to invest. (1) uses data from the first experiment, with only receivers. (2) and (3) use data from the second experiment. (2) are the receivers, (3) the senders. Public and Complements are dummy variables equal to 1 if the beliefs was reported for facing a public rather than a private information structure, or facing a game with strategic complements rather than substitutes respectively. (1 if level=*optimal*) and (1 if level=*high*) are dummy variables equal to 1 if the information structures used *optimal* or *high* probabilities to persuade receivers to invest, relative to the omitted category *low*, respectively. Standard errors in parentheses, clustered at the matching-group level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure 4.B.7 and Table 4.B.18 presents data of receiver behavior similar to the first experiment, using data from the second experiment. Note that this is not directly comparable, as senders had chosen the information structure endogenously. This may now reflect that some matching groups responded heterogeneously to specific structures. Senders can anticipate this, so the regressions now compare data under selection, where those groups that respond particularly well, and potentially different from the average group, to a specific structure.

Figure 4.B.7: Investment decisions across the two experiments



Notes: Average frequency of investment by treatment, bars indicate observed choices, bars indicate 95% bootstrapped confidence intervals.

In columns (1), (3) and (5), I regress investment behavior on a treatment dummy for a game of strategic complements, as well as design features of the information structure (public vs. private, level). Columns (2), (4) and (6) repeat this for following decisions.

Table 4.B.18: Receiver behavior in the second experiment

	(1)	(2)	(3)	(4)	(5)	(6)
	Substitutes		Complements		Diff-in-Diff	
	Investment	Following	Investment	Following	Investment	Following
Public	0.048**	0.035*	0.052*	-0.016	0.051**	0.035*
	(0.021)	(0.020)	(0.029)	(0.024)	(0.022)	(0.020)
Complement					-0.039	-0.067***
					(0.029)	(0.022)
Complement × Public					0.006	-0.050*
					(0.037)	(0.029)
(1 if level=*optimal*)	-0.010	-0.121***	-0.036	-0.174***	-0.021	-0.146***
	(0.034)	(0.025)	(0.038)	(0.020)	(0.026)	(0.017)
(1 if level=*high*)	0.017	-0.177***	-0.033	-0.280***	0.001	-0.219***
	(0.027)	(0.025)	(0.047)	(0.029)	(0.027)	(0.020)
Constant	0.389***	0.838***	0.438***	0.896***	0.425***	0.892***
	(0.036)	(0.033)	(0.045)	(0.034)	(0.033)	(0.027)
Lab FE and period trend	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2520	2520	2520	2520	5040	5040
# clusters	20	20	20	20	40	40
# participants	120	120	120	120	240	240

Notes: The table reports OLS estimates. The dependent variable is a dummy variable equal to 1 if the participant decided to invest (Investment) or followed a recommendation (Following) by investing after receiving the recommendation to invest, or not investing after receiving the recommendation not to invest. Complements is a treatment indicator, a dummy variable equal to 1 if the decision was made facing a game with strategic complements, with the omitted category being a game of strategic substitutes. Public is a dummy variable equal to 1 if the decision was made facing a public information structure, with the omitted category being a private structure. (1 if level=*optimal*) and (1 if level=*high*) are dummy variables equal to 1 if the information structures used *optimal* or *high* probabilities to persuade receivers to invest, relative to the omitted category *low*, respectively. Note that the publicness and the level was an endogenous choice by senders in this experiment. Both the level and publicness are now chosen endogenously by senders. Standard errors in parentheses, clustered at the matching-group level, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

4.C Appendix: Instructions and screenshots

This section contains screenshots of the decision screens, receivers' instructions in the first experiment as well as screenshots of the senders' instructions in the second experiment. Receiver instructions in the second experiment were identical, apart from revealing how senders' payoffs depended on their choices.

In the first experiment, instructions were specific to the game (strategic substitutes vs. complements), all information structures one participant received were either public or private. Between parts, the level of the structure was varied.

In the second experiment, instructions were again specific to the game (strategic substitutes vs. complements). In addition, each role assignment (sender vs. receiver) had specific instructions.

4.C.1 Example decision screen

Below are screenshots of the senders' and receivers' decision screens from the second experiment.

Figure 4.C.1: Receivers' decision screen

Figure 4.C.2: Senders' decision screen

Period: 1/21

Please choose a recommendation plan

Both workers always get the same recommendation

Plan 1

Difficult projects	
Worker 1: work, worker 2: work	10%
Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	90%

Easy projects	
Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	0%

A worker decided to work:
The recommendation was followed:
○ **Choose plan 1**

Plan 3

Difficult projects	
Worker 1: work, worker 2: work	23%
Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	77%

Easy projects	
Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	0%

A worker decided to work:
The recommendation was followed:
○ **Choose plan 3**

Plan 5

Difficult projects	
Worker 1: work, worker 2: work	32%
Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	68%

Easy projects	
Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	0%

A worker decided to work:
The recommendation was followed:
○ **Choose plan 5**

Your workers potentially get different recommendations

Plan 2

Difficult projects	
Worker 1: work, worker 2: work	0%
Worker 1: work, worker 2: don't work	14%
Worker 1: don't work, worker 2: work	14%
Worker 1: don't work, worker 2: don't work	72%

Easy projects	
Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	0%

A worker decided to work:
The recommendation was followed:
○ **Choose plan 2**

Plan 4

Difficult projects	
Worker 1: work, worker 2: work	0%
Worker 1: work, worker 2: don't work	34%
Worker 1: don't work, worker 2: work	34%
Worker 1: don't work, worker 2: don't work	32%

Easy projects	
Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	0%

A worker decided to work:
The recommendation was followed:
○ **Choose plan 4**

Plan 6

Difficult projects	
Worker 1: work, worker 2: work	0%
Worker 1: work, worker 2: don't work	48%
Worker 1: don't work, worker 2: work	48%
Worker 1: don't work, worker 2: don't work	4%

Easy projects	
Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	0%

A worker decided to work:
The recommendation was followed:
○ **Choose plan 6**

Payoffs for all possible outcomes

Payoffs for difficult projects

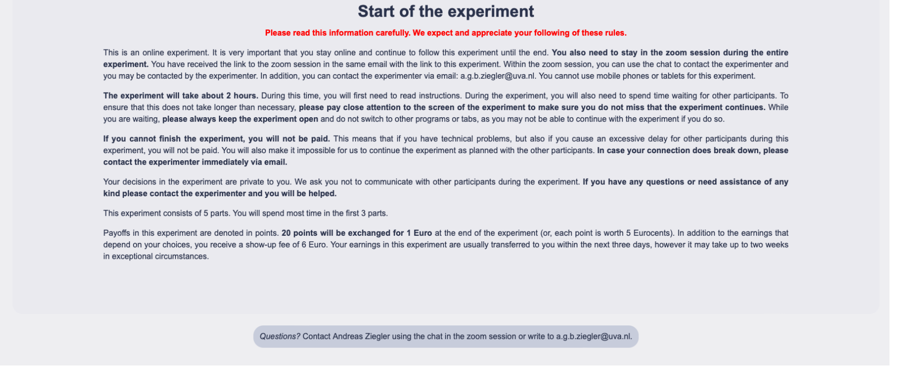
	Worker 2 works	Worker 2 doesn't work
Worker 1 works	20, 20	70, 170
Worker 1 doesn't work	170, 70	170, 170

Payoffs for easy projects

	Worker 2 works	Worker 2 doesn't work
Worker 1 works	210, 210	260, 170
Worker 1 doesn't work	170, 260	170, 170

4.C.2 Receivers' instructions in the first experiment

Figure 4.C.3: Receivers' instructions 1



Start of the experiment

Please read this information carefully. We expect and appreciate your following of these rules.

This is an online experiment. It is very important that you stay online and continue to follow this experiment until the end. **You also need to stay in the zoom session during the entire experiment.** You have received the link to the zoom session in the same email with the link to this experiment. Within the zoom session, you can use the chat to contact the experimenter and you may be contacted by the experimenter. In addition, you can contact the experimenter via email: a.g.b.ziegler@uva.nl. You cannot use mobile phones or tablets for this experiment.

The experiment will take about 2 hours. During this time, you will first need to read instructions. During the experiment, you will also need to spend time waiting for other participants. To ensure that this does not take longer than necessary, **please pay close attention to the screen of the experiment to make sure you do not miss that the experiment continues.** While you are waiting, **please always keep the experiment open** and do not switch to other programs or tabs, as you may not be able to continue with the experiment if you do so.

If you cannot finish the experiment, you will not be paid. This means that if you have technical problems, but also if you cause an excessive delay for other participants during this experiment, you will not be paid. You will also make it impossible for us to continue the experiment as planned with the other participants. **In case your connection does break down, please contact the experimenter immediately via email.**

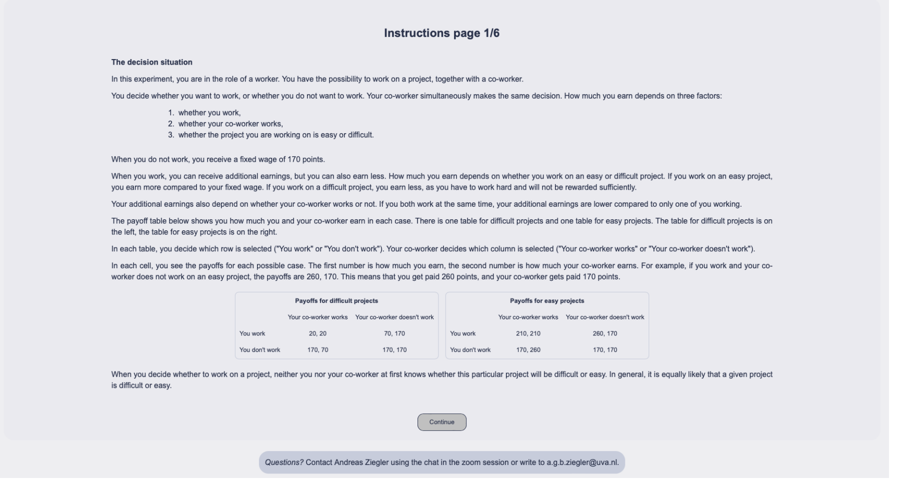
Your decisions in the experiment are private to you. We ask you not to communicate with other participants during the experiment. **If you have any questions or need assistance of any kind please contact the experimenter and you will be helped.**

This experiment consists of 5 parts. You will spend most time in the first 3 parts.

Payoffs in this experiment are denoted in points. **20 points will be exchanged for 1 Euro** at the end of the experiment (or, each point is worth 5 Eurocents). In addition to the earnings that depend on your choices, you receive a show-up fee of 6 Euro. Your earnings in this experiment are usually transferred to you within the next three days, however it may take up to two weeks in exceptional circumstances.

Questions? Contact Andreas Ziegler using the chat in the zoom session or write to a.g.b.ziegler@uva.nl.

Figure 4.C.4: Receivers' instructions 2



Instructions page 1/6

The decision situation

In this experiment, you are in the role of a worker. You have the possibility to work on a project, together with a co-worker.

You decide whether you want to work, or whether you do not want to work. Your co-worker simultaneously makes the same decision. How much you earn depends on three factors:

1. whether you work,
2. whether your co-worker works,
3. whether the project you are working on is easy or difficult.

When you do not work, you receive a fixed wage of 170 points.

When you work, you can receive additional earnings, but you can also earn less. How much you earn depends on whether you work on an easy or difficult project. If you work on an easy project, you earn more compared to your fixed wage. If you work on a difficult project, you earn less, as you have to work hard and will not be rewarded sufficiently.

Your additional earnings also depend on whether your co-worker works or not. If you both work at the same time, your additional earnings are lower compared to only one of you working.

The payoff table below shows you how much you and your co-worker earn in each case. There is one table for difficult projects and one table for easy projects. The table for difficult projects is on the left, the table for easy projects is on the right.

In each table, you decide which row is selected ("You work" or "You don't work"). Your co-worker decides which column is selected ("Your co-worker works" or "Your co-worker doesn't work").

In each cell, you see the payoffs for each possible case. The first number is how much you earn, the second number is how much your co-worker earns. For example, if you work and your co-worker does not work on an easy project, the payoffs are 260, 170. This means that you get paid 260 points, and your co-worker gets paid 170 points.

Payoffs for difficult projects

	Your co-worker works	Your co-worker doesn't work
You work	20, 20	70, 170
You don't work	170, 70	170, 170

Payoffs for easy projects

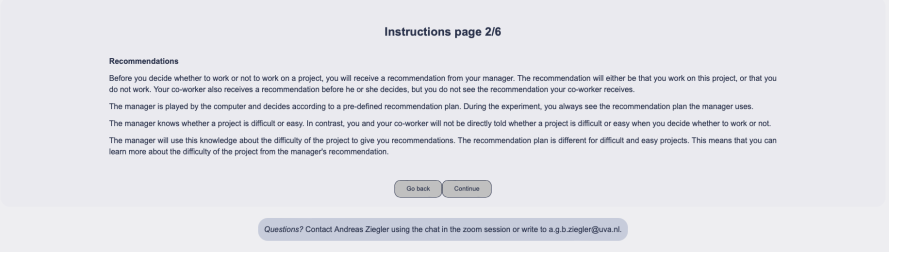
	Your co-worker works	Your co-worker doesn't work
You work	210, 210	260, 170
You don't work	170, 260	170, 170

When you decide whether to work on a project, neither you nor your co-worker at first knows whether this particular project will be difficult or easy. In general, it is equally likely that a given project is difficult or easy.

Continue

Questions? Contact Andreas Ziegler using the chat in the zoom session or write to a.g.b.ziegler@uva.nl.

Figure 4.C.5: Receivers' instructions 3



Instructions page 2/6

Recommendations

Before you decide whether to work or not to work on a project, you will receive a recommendation from your manager. The recommendation will either be that you work on this project, or that you do not work. Your co-worker also receives a recommendation before he or she decides, but you do not see the recommendation your co-worker receives.

The manager is played by the computer and decides according to a pre-defined recommendation plan. During the experiment, you always see the recommendation plan the manager uses.

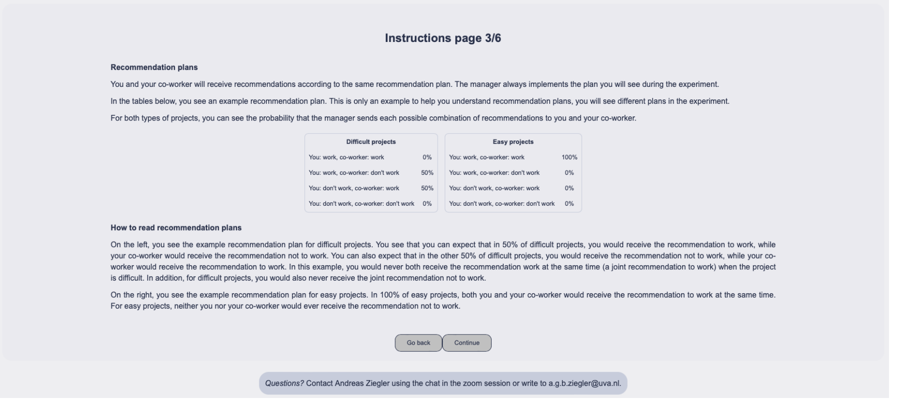
The manager knows whether a project is difficult or easy. In contrast, you and your co-worker will not be directly told whether a project is difficult or easy when you decide whether to work or not.

The manager will use this knowledge about the difficulty of the project to give you recommendations. The recommendation plan is different for difficult and easy projects. This means that you can learn more about the difficulty of the project from the manager's recommendation.

Go back Continue

Questions? Contact Andreas Ziegler using the chat in the zoom session or write to a.g.b.ziegler@uva.nl.

Figure 4.C.6: Receivers' instructions 4



Instructions page 3/6

Recommendation plans

You and your co-worker will receive recommendations according to the same recommendation plan. The manager always implements the plan you will see during the experiment.

In the tables below, you see an example recommendation plan. This is only an example to help you understand recommendation plans, you will see different plans in the experiment.

For both types of projects, you can see the probability that the manager sends each possible combination of recommendations to you and your co-worker.

Difficult projects	
You: work, co-worker: work	0%
You: work, co-worker: don't work	50%
You: don't work, co-worker: work	50%
You: don't work, co-worker: don't work	0%

Easy projects	
You: work, co-worker: work	100%
You: work, co-worker: don't work	0%
You: don't work, co-worker: work	0%
You: don't work, co-worker: don't work	0%

How to read recommendation plans

On the left, you see the example recommendation plan for difficult projects. You see that you can expect that in 50% of difficult projects, you would receive the recommendation to work, while your co-worker would receive the recommendation not to work. You can also expect that in the other 50% of difficult projects, you would receive the recommendation not to work, while your co-worker would receive the recommendation to work. In this example, you would never both receive the recommendation work at the same time (a joint recommendation to work) when the project is difficult. In addition, for difficult projects, you would also never receive the joint recommendation not to work.

On the right, you see the example recommendation plan for easy projects. In 100% of easy projects, both you and your co-worker would receive the recommendation to work at the same time. For easy projects, neither you nor your co-worker would ever receive the recommendation not to work.

Go back Continue

Questions? Contact Andreas Ziegler using the chat in the zoom session or write to a.g.b.ziegler@uva.nl.

Figure 4.C.7: Receivers' instructions 5

Instructions page 4/6

The example recommendation plan

Difficult projects	
You: work, co-worker: work	0%
You: work, co-worker: don't work	50%
You: don't work, co-worker: work	50%
You: don't work, co-worker: don't work	0%

Easy projects	
You: work, co-worker: work	100%
You: work, co-worker: don't work	0%
You: don't work, co-worker: work	0%
You: don't work, co-worker: don't work	0%

What you can learn from recommendation plans about the project

Remember that you will not be directly told whether a project you work on is difficult or easy. At the start, it is equally likely that a project is difficult or easy. Then, you receive your recommendation from your manager. You can combine this recommendation with your manager's recommendation plan to learn more about your project.

Imagine that you receive the recommendation to work. In this example recommendation plan, you cannot tell whether a recommendation to work means for certain that a project is difficult or easy. However, it is more likely that a project is easy whenever you receive the recommendation to work in this example plan. This is the case as whenever your project is easy, you would always receive the recommendation to work. In contrast, you would receive the recommendation to work for only 50% of difficult projects.

What you can learn from recommendation plans about your co-worker's recommendation

Remember also that you will not see the recommendation your co-worker receives. However, you can combine your recommendation with your manager's recommendation plan to learn more about which recommendation your co-worker received.

Imagine that you receive the recommendation to work. In this example recommendation plan, what recommendation your co-worker received depends on whether a particular project is difficult or easy:

- For all difficult projects, your co-worker would receive the recommendation not to work *whenever you receive the recommendation to work*.
- For all easy projects, your co-worker would also receive the recommendation to work.

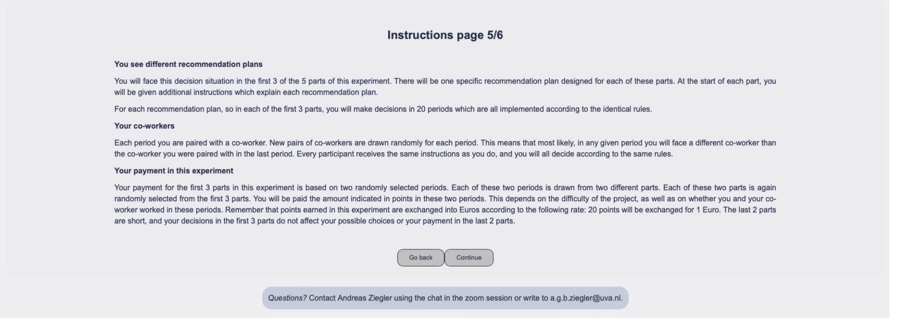
As just explained, a recommendation to work would also mean that it is more likely that the project is easy.

Now, you need to put the two pieces of information together. First, it is more likely that the project would be easy. Second, for easy projects, you would both receive the recommendation to work. Taken together, this means that if you would have received the recommendation to work, it is more likely that you would have both received the recommendation to work in this example plan.

Go back Continue

Questions? Contact Andreas Ziegler using the chat in the zoom session or write to a.g.b.ziegler@uva.nl.

Figure 4.C.8: Receivers' instructions 6



Instructions page 5/6

You see different recommendation plans

You will face this decision situation in the first 3 of the 5 parts of this experiment. There will be one specific recommendation plan designed for each of these parts. At the start of each part, you will be given additional instructions which explain each recommendation plan.

For each recommendation plan, so in each of the first 3 parts, you will make decisions in 20 periods which are all implemented according to the identical rules.

Your co-workers

Each period you are paired with a co-worker. New pairs of co-workers are drawn randomly for each period. This means that most likely, in any given period you will face a different co-worker than the co-worker you were paired with in the last period. Every participant receives the same instructions as you do, and you will all decide according to the same rules.

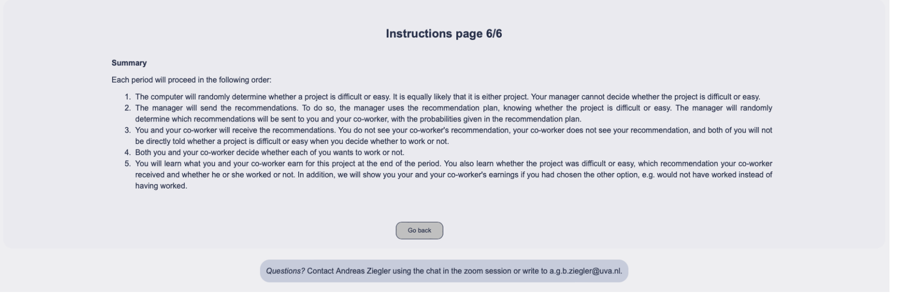
Your payment in this experiment

Your payment for the first 3 parts in this experiment is based on two randomly selected periods. Each of these two periods is drawn from two different parts. Each of these two parts is again randomly selected from the first 3 parts. You will be paid the amount indicated in points in these two periods. This depends on the difficulty of the project, as well as on whether you and your co-worker worked in these periods. Remember that points earned in this experiment are exchanged into Euros according to the following rate: 20 points will be exchanged for 1 Euro. The last 2 parts are short, and your decisions in the first 3 parts do not affect your possible choices or your payment in the last 2 parts.

Go back Continue

Questions? Contact Andreas Ziegler using the chat in the zoom session or write to a.g.b.ziegler@uva.nl.

Figure 4.C.9: Receivers' instructions 7



Instructions page 6/6

Summary

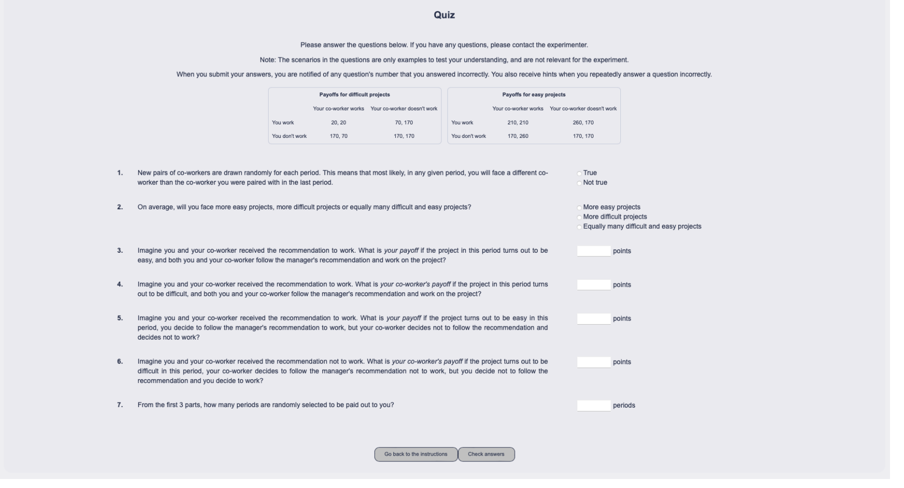
Each period will proceed in the following order:

1. The computer will randomly determine whether a project is difficult or easy. It is equally likely that it is either project. Your manager cannot decide whether the project is difficult or easy.
2. The manager will send the recommendations. To do so, the manager uses the recommendation plan, knowing whether the project is difficult or easy. The manager will randomly determine which recommendations will be sent to you and your co-worker, with the probabilities given in the recommendation plan.
3. You and your co-worker will receive the recommendations. You do not see your co-worker's recommendation, your co-worker does not see your recommendation, and both of you will not be directly told whether a project is difficult or easy when you decide whether to work or not.
4. Both you and your co-worker decide whether each of you wants to work or not.
5. You will learn what you and your co-worker earn for this project at the end of the period. You also learn whether the project was difficult or easy, which recommendation your co-worker received and whether he or she worked or not. In addition, we will show you your and your co-worker's earnings if you had chosen the other option, e.g. would not have worked instead of having worked.

Go back

Questions? Contact Andreas Ziegler using the chat in the zoom session or write to a.g.b.ziegler@uva.nl.

Figure 4.C.10: Receivers' instructions 8



4.C.3 Instructions for new information structures

In the first experiment, the level of the information structure was varied between parts. At the beginning of each part, participants received the following instructions.

Figure 4.C.11: Instructions for new information structure

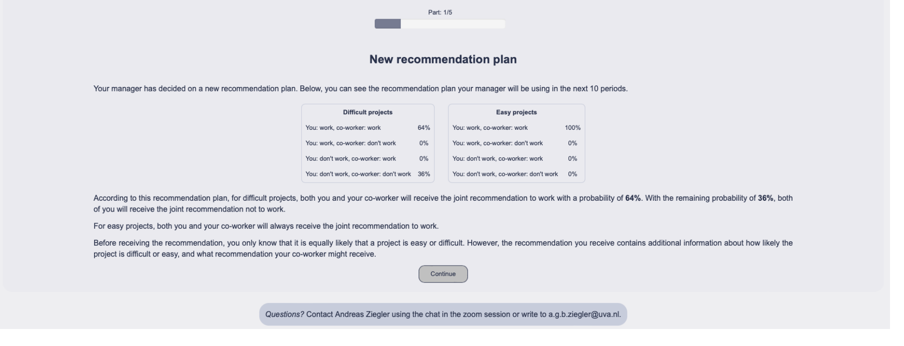
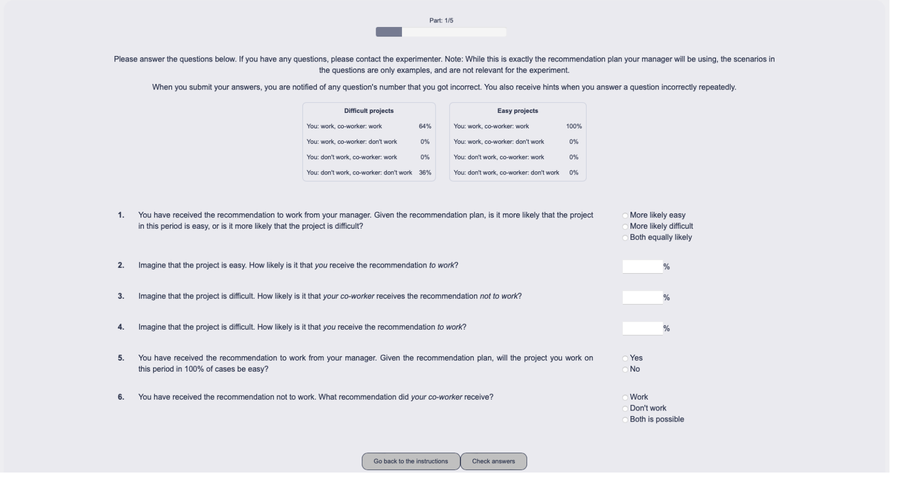


Figure 4.C.12: Quiz for new information structure



Part: 1/5

Please answer the questions below. If you have any questions, please contact the experimenter. Note: While this is exactly the recommendation plan your manager will be using, the scenarios in the questions are only examples, and are not relevant for the experiment.

When you submit your answers, you are notified of any question's number that you got incorrect. You also receive hints when you answer a question incorrectly repeatedly.

Difficult projects	
You: work, co-worker: work	64%
You: work, co-worker: don't work	0%
You: don't work, co-worker: work	0%
You: don't work, co-worker: don't work	36%

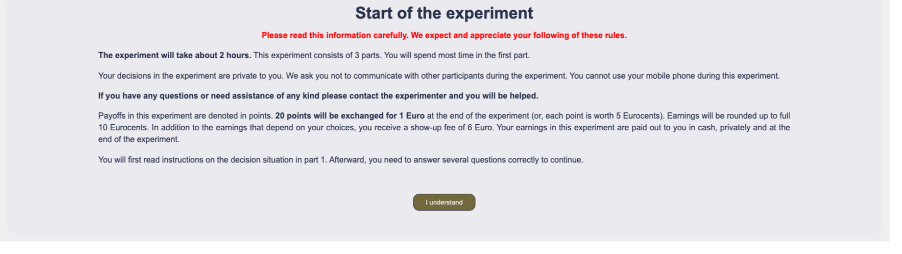
Easy projects	
You: work, co-worker: work	100%
You: work, co-worker: don't work	0%
You: don't work, co-worker: work	0%
You: don't work, co-worker: don't work	0%

1. You have received the recommendation to work from your manager. Given the recommendation plan, is it more likely that the project in this period is easy, or is it more likely that the project is difficult?
   - More likely easy
   - More likely difficult
   - Both equally likely
2. Imagine that the project is easy. How likely is it that *you* receive the recommendation *to work*? %
3. Imagine that the project is difficult. How likely is it that *your co-worker* receives the recommendation *not to work*? %
4. Imagine that the project is difficult. How likely is it that *you* receive the recommendation *to work*? %
5. You have received the recommendation to work from your manager. Given the recommendation plan, will the project you work on this period in 100% of cases be easy?
   - Yes
   - No
6. You have received the recommendation not to work. What recommendation did *your co-worker* receive?
   - Work
   - Don't work
   - Both is possible

Go back to the instructions | Check answers

4.C.4 Senders' instructions in the second experiment

Figure 4.C.13: Senders' instructions 1



Start of the experiment

Please read this information carefully. We expect and appreciate your following of these rules.

The experiment will take about 2 hours. This experiment consists of 3 parts. You will spend most time in the first part.

Your decisions in the experiment are private to you. We ask you not to communicate with other participants during the experiment. You cannot use your mobile phone during this experiment.

If you have any questions or need assistance of any kind please contact the experimenter and you will be helped.

Payoffs in this experiment are denoted in points. **20 points will be exchanged for 1 Euro** at the end of the experiment (or, each point is worth 5 Eurocents). Earnings will be rounded up to full 10 Eurocents. In addition to the earnings that depend on your choices, you receive a show-up fee of 6 Euro. Your earnings in this experiment are paid out to you in cash, privately and at the end of the experiment.

You will first read instructions on the decision situation in part 1. Afterward, you need to answer several questions correctly to continue.

I understand

Figure 4.C.14: Senders' instructions 2

Instructions page 1/8

The decision situation

In this experiment, you are in the role of a manager. You are the manager of two workers. Your task is to recommend to your two workers whether they should work on a project.

You earn 90 points. In addition, your earnings depend on whether you can convince your workers to work. You earn 100 points each time a worker decides to work. So, if no worker works, you earn 90 points. If only one worker works, you earn 190 points. If both workers work, you earn 290 points.

First, you will learn more about the decision that the workers face. The workers receive similar instructions, so you can learn more about how they may reason about this decision situation. Afterward, you will learn more about your own choices.

Continue

Figure 4.C.15: Senders' instructions 3

Instructions page 2/8

The decision situation of the workers

Each worker has to decide whether they want to work or do not want to work. The earnings of each worker depend on three factors:

1. whether they work,
2. whether their co-worker works,
3. whether the project they are working on is easy or difficult.

They receive a fixed wage of 170 points when they do not work.

When they work, they may receive additional earnings, but they may also earn less. How much they earn depends on whether they work on an easy or difficult project. If they work on an easy project, they earn more than their fixed wage. If they work on a difficult project, they earn less.

Their additional earnings also depend on whether their co-worker decides to work or not. If they both work simultaneously, their additional earnings are higher compared to only one of them working.

The payoff tables below show how much the workers earn in each case. The two workers' decision situations are identical; it is entirely random whether each worker is worker 1 or worker 2. There is one table for difficult projects and one table for easy projects. The table for difficult projects is on the left, and the table for easy projects is on the right.

In each table, worker 1 decides which row is selected ("Worker 1 works" or "Worker 1 doesn't work"). Their co-worker, worker 2, decides which column is selected ("Worker 2 works" or "Worker 2 doesn't work").

In each cell, you see the payoffs for each possible case. The first number is how much worker 1 earns; the second number is how much worker 2 earns. For example, if worker 1 works and worker 2 does not work on an easy project, the payoffs are 180, 170 (see the top right cell in the right table). This means that worker 1 gets paid 180 points, and worker 2 gets paid 170 points.

Payoffs for difficult projects

	Worker 2 works	Worker 2 doesn't work
Worker 1 works	100, 100	70, 170
Worker 1 doesn't work	170, 70	170, 170

Payoffs for easy projects

	Worker 2 works	Worker 2 doesn't work
Worker 1 works	210, 210	180, 170
Worker 1 doesn't work	170, 180	170, 170

Your workers know that you earn an additional 100 points each time a worker decides to work.

Recommendations

At first, your workers only know that it is equally likely that they face a difficult or an easy project. Before your workers decide whether to work or not to work on a project, they will receive a recommendation. The recommendation will either be that they work on this project or that they do not work. This can help your workers learn more about the project's difficulty and their decisions' consequences. You will decide what recommendations your workers will receive.

Go back Continue

Figure 4.C.16: Senders' instructions 4

Instructions page 3/8

Sending recommendations

To send recommendations to your workers, you will rely on a recommendation plan. A recommendation plan defines the probabilities that your workers receive a particular combination of recommendations. These probabilities depend on the project's difficulty. In these plans, you choose how often different recommendations will be sent to your two workers. Later, you will learn how to read a recommendation plan.

Timing of recommendations

Decisions are made in the following order:

1. You will choose a recommendation plan.
2. The computer will randomly determine whether the project will be easy or difficult.
3. The computer will send recommendations to your workers in agreement with your chosen recommendation plan.

Before your workers decide whether to work or not, they see which recommendation plan you have chosen and receive their recommendation. Each worker receives only one recommendation from you as their manager, based on your chosen recommendation plan.

The menu of recommendation plans

In the experiment, you can choose from a fixed set of six recommendation plans. You see the menu of all six recommendation plans below. Next, you will learn more about reading recommendation plans.

Plan 1

Difficult projects		Easy projects	
Worker 1: work, worker 2: work	19%	Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%	Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%	Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	81%	Worker 1: don't work, worker 2: don't work	0%

Plan 2

Difficult projects		Easy projects	
Worker 1: work, worker 2: work	0%	Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	14%	Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	14%	Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	72%	Worker 1: don't work, worker 2: don't work	0%

Plan 3

Difficult projects		Easy projects	
Worker 1: work, worker 2: work	48%	Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%	Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%	Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	52%	Worker 1: don't work, worker 2: don't work	0%

Plan 4

Difficult projects		Easy projects	
Worker 1: work, worker 2: work	0%	Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	34%	Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	34%	Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	32%	Worker 1: don't work, worker 2: don't work	0%

Plan 5

Difficult projects		Easy projects	
Worker 1: work, worker 2: work	71%	Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%	Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%	Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	29%	Worker 1: don't work, worker 2: don't work	0%

Plan 6

Difficult projects		Easy projects	
Worker 1: work, worker 2: work	0%	Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	48%	Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	48%	Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	4%	Worker 1: don't work, worker 2: don't work	0%

Go back Continue

Figure 4.C.17: Senders' instructions 5

Instructions page 4/8

Reading recommendation plans

You will now learn how to read recommendation plans, as will your workers. For this, we will describe two recommendation plans from the six recommendation plans available to you, Plan 1 and Plan 6. The other plans work similarly; we randomly chose these two plans to describe them in more detail.

Below, you see Plan 1.

Plan 1

Difficult projects		Easy projects	
Worker 1: work, worker 2: work	19%	Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%	Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%	Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	81%	Worker 1: don't work, worker 2: don't work	0%

You can see the probability that your workers will receive each possible combination of recommendations for both difficult and easy projects. Again, worker 1 and worker 2 are in identical roles. For example, the first row in the left table tells you the probability that both your workers receive the recommendation to work for difficult projects.

On the left, you see how likely each recommendation will be sent whenever a project is difficult. You see that you can expect that in 19% of difficult projects, both worker 1 and worker 2 would receive the joint recommendation to work (left table, top row). You can also expect that in the other 81% of difficult projects, both worker 1 and worker 2 would receive the joint recommendation not to work (left table, bottom row). In this plan, a worker would never receive the recommendation to work for a difficult project at a point where the other worker received the recommendation not to work (left table, middle two rows).

On the right, you see how likely each recommendation will be sent whenever a project is easy. In 100% of easy projects, both workers would receive the recommendation to work simultaneously (right table, top row). Neither worker would ever receive the recommendation not to work for easy projects.

Below, you see Plan 6.

Plan 6

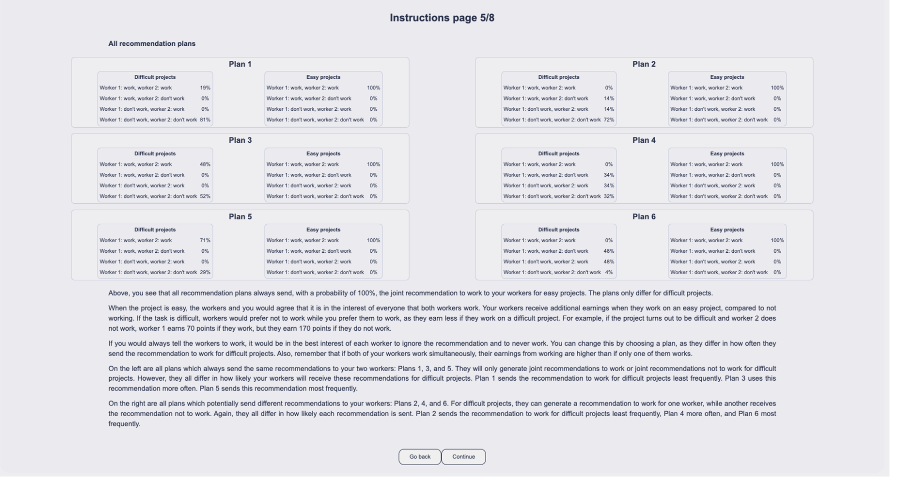
Difficult projects		Easy projects	
Worker 1: work, worker 2: work	0%	Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	48%	Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	48%	Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	4%	Worker 1: don't work, worker 2: don't work	0%

On the left, you see that you can expect that in 48% of difficult projects, worker 1 receives the recommendation to work, while worker 2 receives the recommendation not to work. You can also expect that in other 48% of difficult projects, worker 2 receives the recommendation to work, while worker 1 receives the recommendation not to work. Lastly, you can expect that in the remaining 4% of difficult projects, both workers receive the joint recommendation not to work.

On the right, you see that in 100% of easy projects, both workers would receive the recommendation to work simultaneously, as in Plan 1.

Go back Continue

Figure 4.C.18: Senders' instructions 6



Instructions page 5/8

All recommendation plans

Plan 1

Difficult projects		Easy projects	
Worker 1: work, worker 2: work	19%	Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%	Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%	Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	81%	Worker 1: don't work, worker 2: don't work	0%

Plan 2

Difficult projects		Easy projects	
Worker 1: work, worker 2: work	0%	Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	14%	Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	14%	Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	72%	Worker 1: don't work, worker 2: don't work	0%

Plan 3

Difficult projects		Easy projects	
Worker 1: work, worker 2: work	48%	Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%	Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%	Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	52%	Worker 1: don't work, worker 2: don't work	0%

Plan 4

Difficult projects		Easy projects	
Worker 1: work, worker 2: work	0%	Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	34%	Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	34%	Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	32%	Worker 1: don't work, worker 2: don't work	0%

Plan 5

Difficult projects		Easy projects	
Worker 1: work, worker 2: work	71%	Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%	Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%	Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	29%	Worker 1: don't work, worker 2: don't work	0%

Plan 6

Difficult projects		Easy projects	
Worker 1: work, worker 2: work	0%	Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	48%	Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	48%	Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	4%	Worker 1: don't work, worker 2: don't work	0%

Above, you see that all recommendation plans always send, with a probability of 100%, the joint recommendation to work to your workers for easy projects. The plans only differ for difficult projects.

When the project is easy, the workers and you would agree that it is in the interest of everyone that both workers work. Your workers receive additional earnings when they work on an easy project, compared to not working. If the task is difficult, workers would prefer not to work while you prefer them to work, as they earn less if they work on a difficult project. For example, if the project turns out to be difficult and worker 2 does not work, worker 1 earns 70 points if they work, but they earn 170 points if they do not work.

If you would always tell the workers to work, it would be in the best interest of each worker to ignore the recommendation and to never work. You can change this by choosing a plan, as they differ in how often they send the recommendation to work for difficult projects. Also, remember that if both of your workers work simultaneously, their earnings from working are higher than if only one of them works.

On the left are all plans which always send the same recommendations to your two workers: Plans 1, 3, and 5. They will only generate joint recommendations to work or joint recommendations not to work for difficult projects. However, they all differ in how likely your workers will receive these recommendations for difficult projects. Plan 1 sends the recommendation to work for difficult projects least frequently. Plan 3 uses this recommendation more often. Plan 5 sends this recommendation most frequently.

On the right are all plans which potentially send different recommendations to your workers: Plans 2, 4, and 6. For difficult projects, they can generate a recommendation to work for one worker, while another receives the recommendation not to work. Again, they all differ in how likely each recommendation is sent. Plan 2 sends the recommendation to work for difficult projects least frequently, Plan 4 more often, and Plan 6 most frequently.

Go back Continue

Figure 4.C.19: Senders' instructions 7

Instructions page 6/8

What your workers can learn from recommendation plans

Your workers will see the recommendation plan you have chosen. They can use the plan to decide how informative the recommendation is and whether they want to implement the recommendation they receive.

Your workers can combine the recommendation they receive with the recommendation plan you chose to learn more about their project's difficulty. Your workers will not be directly told whether a project they work on is difficult or easy. At the start, it is equally likely that a project is difficult or easy. Imagine that a worker receives the recommendation to work. In all plans, they cannot tell whether a recommendation to work means for sure that a project is difficult or easy because they may receive this recommendation for both types of projects. However, it is more likely that a project is easy whenever they receive the recommendation to work in these plans. This is the case as whenever their project is easy they would always receive the recommendation to work. In contrast, they would only sometimes receive the recommendation to work for difficult projects.

Your workers can also combine their recommendation and your recommendation plan to learn more about their co-worker's recommendation. Each worker only sees the recommendation they receive themselves, but not their co-worker's recommendation. However, they can use that the recommendation plan tells them more about their co-workers recommendation for the particular recommendation they have received. For example, for all plans on the left (Plans 1, 3 and 5), they know that their co-worker will always receive the same recommendation.

The decision screen of the workers

This is a screenshot of the decision screen of the workers.

You receive the following recommendation from your manager: "Please **work**".

Now, please decide whether you want to work in this period.

Your decision: ○ Work ○ Don't work

Submit decision

Payoffs and recommendation plans

Below, you can see the payoffs for all possible decisions and projects as well as your manager's recommendation plan for this period.

Payoffs for all possible outcomes

Payoffs for difficult projects

	Your co-worker works	Your co-worker doesn't work
You work	100, 100	70, 170
You don't work	170, 70	170, 170

Payoffs for easy projects

	Your co-worker works	Your co-worker doesn't work
You work	210, 210	180, 170
You don't work	170, 180	170, 170

Figure 4.C.20: Senders' instructions 8

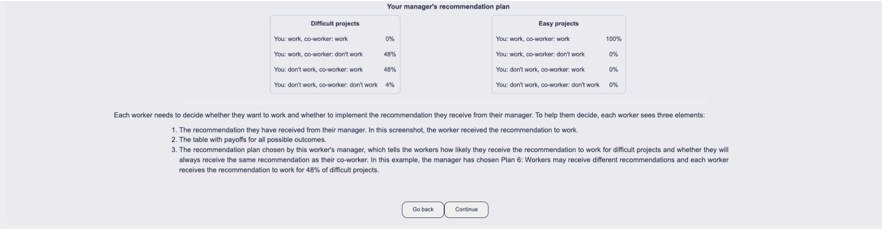


Figure 4.C.21: Senders' instructions 9

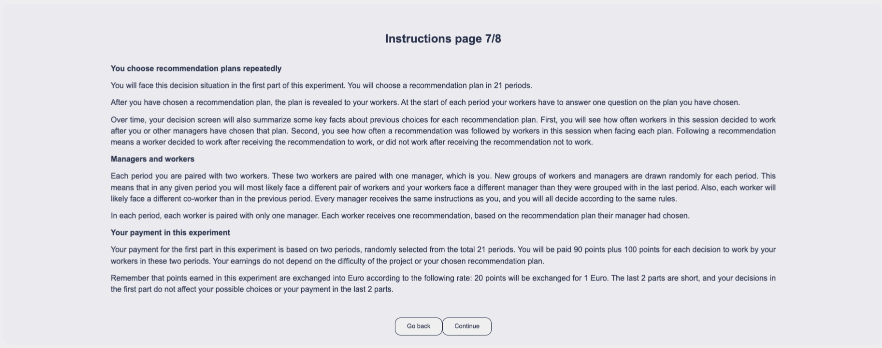


Figure 4.C.22: Senders' instructions 10

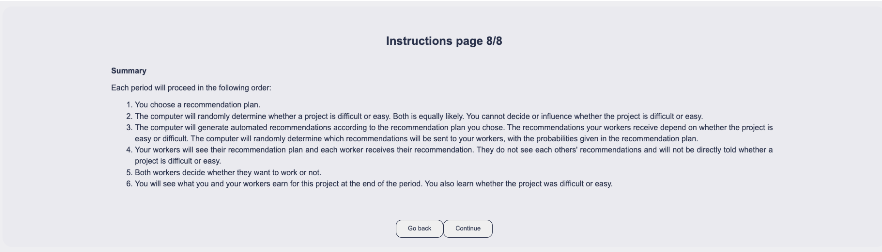
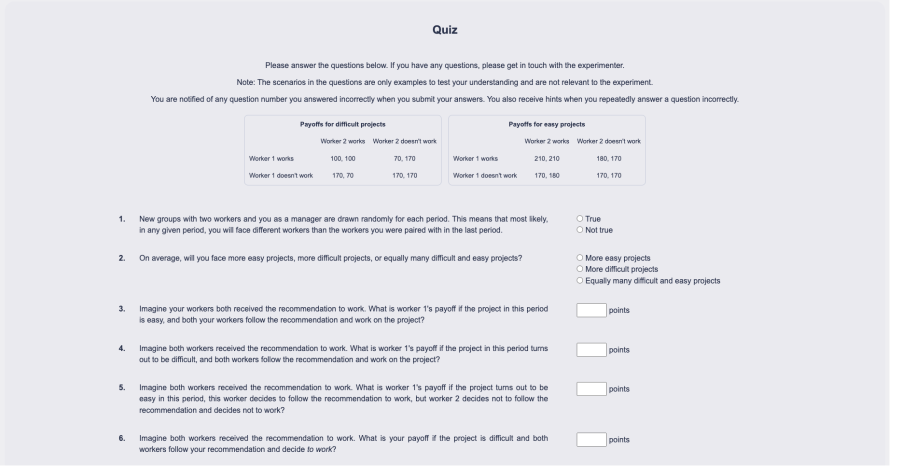


Figure 4.C.23: Senders’ instructions 11



Quiz

Please answer the questions below. If you have any questions, please get in touch with the experimenter.

Note: The scenarios in the questions are only examples to test your understanding and are not relevant to the experiment.

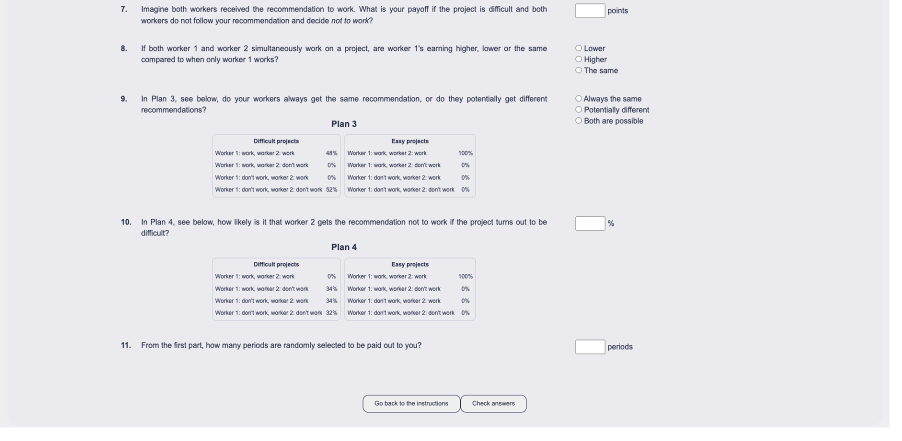
You are notified of any question number you answered incorrectly when you submit your answers. You also receive hints when you repeatedly answer a question incorrectly.

Payoffs for difficult projects	Worker 2 works	Worker 2 doesn't work
Worker 1 works	100, 100	70, 170
Worker 1 doesn't work	170, 70	170, 170

Payoffs for easy projects	Worker 2 works	Worker 2 doesn't work
Worker 1 works	210, 210	180, 170
Worker 1 doesn't work	170, 180	170, 170

1. New groups with two workers and you as a manager are drawn randomly for each period. This means that most likely, in any given period, you will face different workers than the workers you were paired with in the last period.
   - True
   - Not true
2. On average, will you face more easy projects, more difficult projects, or equally many difficult and easy projects?
   - More easy projects
   - More difficult projects
   - Equally many difficult and easy projects
3. Imagine your workers both received the recommendation to work. What is worker 1's payoff if the project in this period is easy, and both your workers follow the recommendation and work on the project? ___ points
4. Imagine both workers received the recommendation to work. What is worker 1's payoff if the project in this period turns out to be difficult, and both workers follow the recommendation and work on the project? ___ points
5. Imagine both workers received the recommendation to work. What is worker 1's payoff if the project turns out to be easy in this period, this worker decides to follow the recommendation to work, but worker 2 decides not to follow the recommendation and decides not to work? ___ points
6. Imagine both workers received the recommendation to work. What is your payoff if the project is difficult and both workers follow your recommendation and decide *to work*? ___ points

Figure 4.C.24: Senders’ instructions 12



7. Imagine both workers received the recommendation to work. What is your payoff if the project is difficult and both workers do not follow your recommendation and decide *not to work*? ___ points
8. If both worker 1 and worker 2 simultaneously work on a project, are worker 1's earning higher, lower or the same compared to when only worker 1 works?
   - Lower
   - Higher
   - The same
9. In Plan 3, see below, do your workers always get the same recommendation, or do they potentially get different recommendations?
   - Always the same
   - Potentially different
   - Both are possible

Plan 3

Difficult projects	
Worker 1: work, worker 2: work	48%
Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	52%

Easy projects	
Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	0%

10. In Plan 4, see below, how likely is it that worker 2 gets the recommendation not to work if the project turns out to be difficult? ___ %

Plan 4

Difficult projects	
Worker 1: work, worker 2: work	0%
Worker 1: work, worker 2: don't work	34%
Worker 1: don't work, worker 2: work	34%
Worker 1: don't work, worker 2: don't work	32%

Easy projects	
Worker 1: work, worker 2: work	100%
Worker 1: work, worker 2: don't work	0%
Worker 1: don't work, worker 2: work	0%
Worker 1: don't work, worker 2: don't work	0%

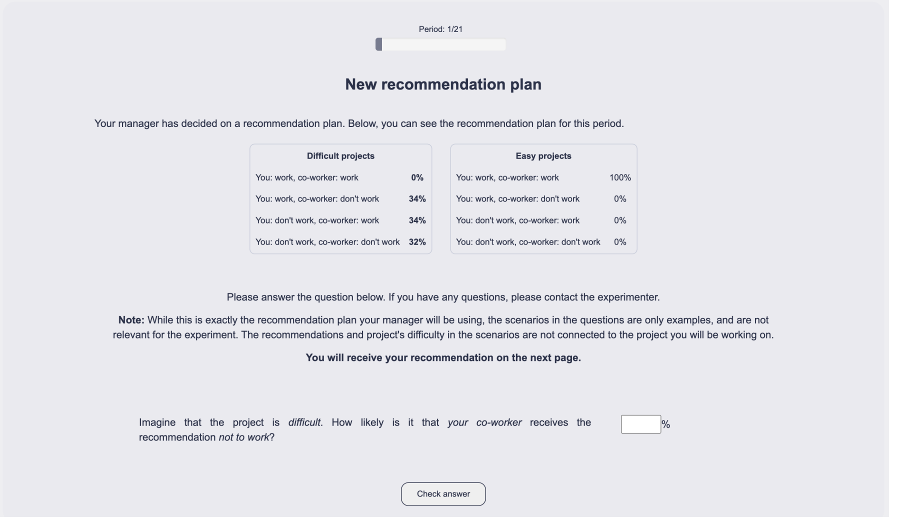
11. From the first part, how many periods are randomly selected to be paid out to you? ___ periods

Go back to the instructions | Check answers

4.C.5 Instructions for new information structures

In the second experiment, the level of the information structure was varied between period. At the beginning of each period, participants received a quiz question. The questions were randomized out of a set of questions similar to the questions in the first experiment.

Figure 4.C.25: Instructions for new information structure



Period: 1/21

New recommendation plan

Your manager has decided on a recommendation plan. Below, you can see the recommendation plan for this period.

Difficult projects	
You: work, co-worker: work	**0%**
You: work, co-worker: don't work	**34%**
You: don't work, co-worker: work	**34%**
You: don't work, co-worker: don't work	**32%**

Easy projects	
You: work, co-worker: work	100%
You: work, co-worker: don't work	0%
You: don't work, co-worker: work	0%
You: don't work, co-worker: don't work	0%

Please answer the question below. If you have any questions, please contact the experimenter.

Note: While this is exactly the recommendation plan your manager will be using, the scenarios in the questions are only examples, and are not relevant for the experiment. The recommendations and project's difficulty in the scenarios are not connected to the project you will be working on.

You will receive your recommendation on the next page.

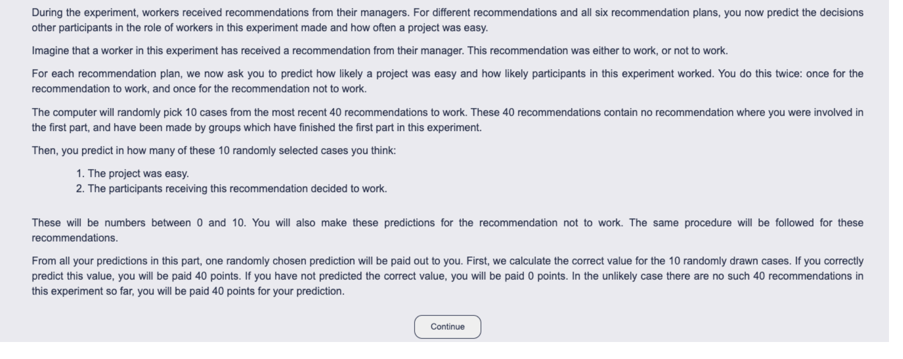
Imagine that the project is *difficult*. How likely is it that *your co-worker* receives the recommendation *not to work*? %

Check answer

4.C.6 Instructions for tasks at the end of the experiment

After the game, I elicited participants' beliefs, for all structures they faced in the experiment.

Figure 4.C.26: Belief instructions



During the experiment, workers received recommendations from their managers. For different recommendations and all six recommendation plans, you now predict the decisions other participants in the role of workers in this experiment made and how often a project was easy.

Imagine that a worker in this experiment has received a recommendation from their manager. This recommendation was either to work, or not to work.

For each recommendation plan, we now ask you to predict how likely a project was easy and how likely participants in this experiment worked. You do this twice: once for the recommendation to work, and once for the recommendation not to work.

The computer will randomly pick 10 cases from the most recent 40 recommendations to work. These 40 recommendations contain no recommendation where you were involved in the first part, and have been made by groups which have finished the first part in this experiment.

Then, you predict in how many of these 10 randomly selected cases you think:

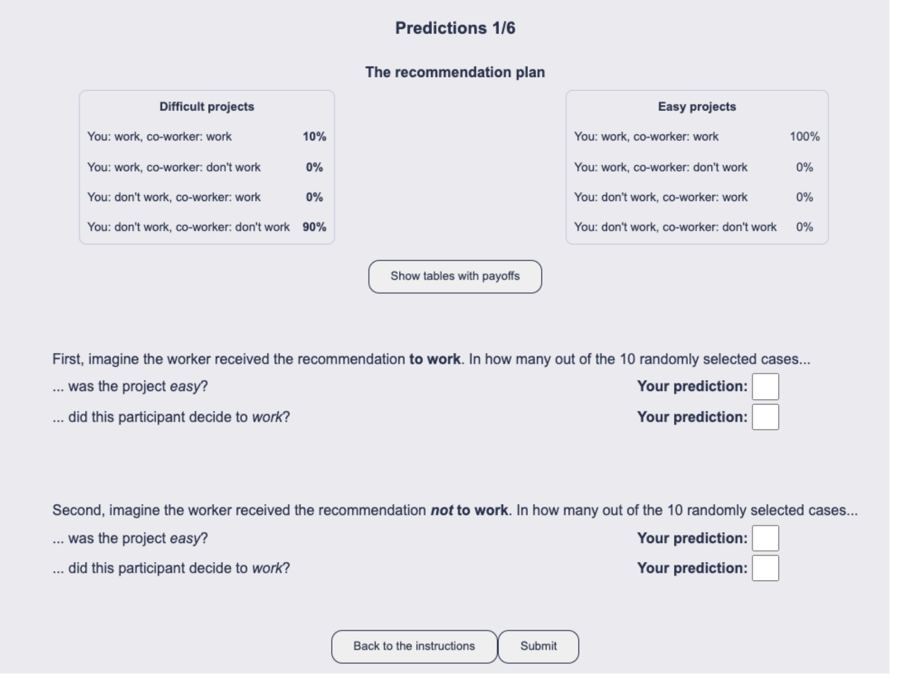
1. The project was easy.
2. The participants receiving this recommendation decided to work.

These will be numbers between 0 and 10. You will also make these predictions for the recommendation not to work. The same procedure will be followed for these recommendations.

From all your predictions in this part, one randomly chosen prediction will be paid out to you. First, we calculate the correct value for the 10 randomly drawn cases. If you correctly predict this value, you will be paid 40 points. If you have not predicted the correct value, you will be paid 0 points. In the unlikely case there are no such 40 recommendations in this experiment so far, you will be paid 40 points for your prediction.

Continue

Figure 4.C.27: Example belief decision screen



At the end, participants faced two different risk elicitations. In the first experiment, they saw only the lotteries associated with their treatment. In the second experiment, they saw lotteries for both public and private structures (as below).

Figure 4.C.28: Risk 1

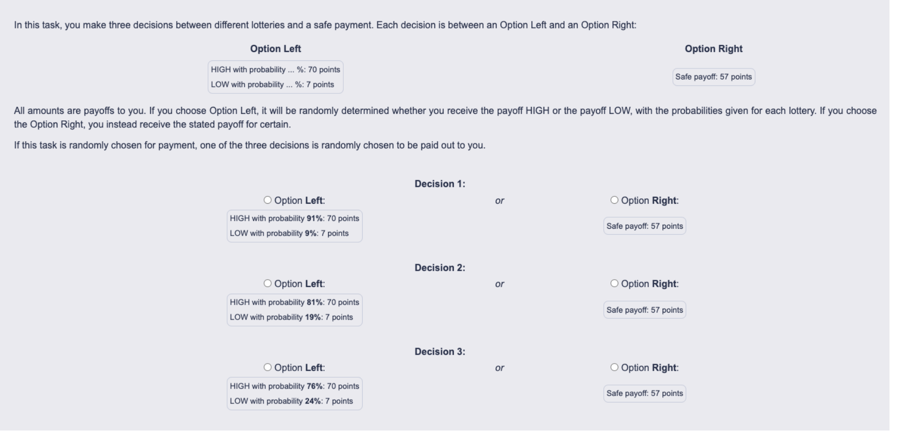


Figure 4.C.29: Risk 2

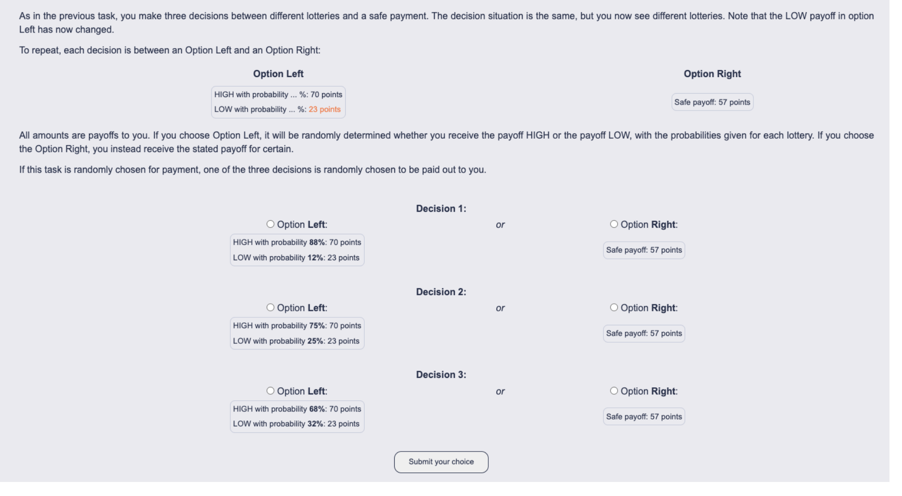


Figure 4.C.30: Risk 3

For this task, you choose one gamble you would like to play from six different gambles. The six different gambles are listed below. You must select one and only one of these gambles.

Each gamble has two possible outcomes (Roll Low or Roll High). For every gamble, each Roll has a 50% probability of occurring. At the end of the experiment, it will be randomly determined which event will occur.

For example, if you select Gamble 4 and Roll High occurs, you will be paid 130 points. If Roll Low occurs, you will be paid 40 points.

Your decision:

Gamble	Choice	Roll	Payoff	Probabilities
1	○	High	70 points	50%
		Low	70 points	50%
2	○	High	90 points	50%
		Low	60 points	50%
3		High	110 points	50%
	○	Low	50 points	50%
4	○	High	130 points	50%
		Low	40 points	50%
5	○	High	150 points	50%
		Low	30 points	50%
6	○	High	170 points	50%
		Low	0 points	50%

Submit your choice

Then, participants' social preferences were elicited.

Figure 4.C.31: Social preferences 1

In this task, you make 20 decisions, across two tables. Each decision involves a choice between an Option 1 and an Option 2:

Option 1	**Option 2**
Your payoff	Your payoff ....
Other's payoff	Other's payoff ....

The options refer to payments in points to you and one of the other participants in this experiment. For each option, two amounts will be displayed: one amount that you will receive yourself, and one amount that the other participant will receive.

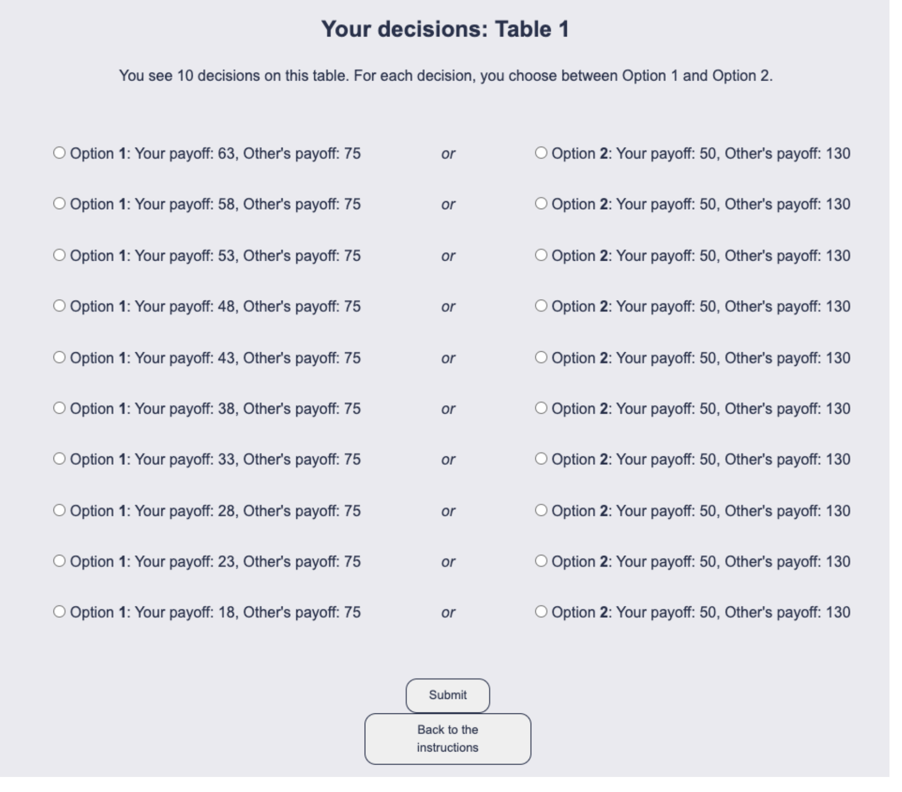
At the end of the experiment, all participants will be randomly matched into pairs. In each pair, one participant will be randomly chosen to be **Player A**, and the other will be **Player B**. If you are chosen to be **Player A**, one out of your 20 decisions made in this task will be randomly selected. For this randomly selected decision and the option you have chosen, you will receive *Your payoff* and your paired **Player B** will receive the *Other's payoff*. Otherwise, if you are chosen to be **Player B**, you will receive the *Other's payoff* as decided by your paired **Player A**. In this case that you are chosen to be **Player B**, your choices in this task do not affect anyone's payment.

If the number of the participants in this task is odd, we cannot combine all of them in pairs at the end of the experiment. In this case, one participant will receive a fixed payment of 130 points as his or her payoff in this task.

Within each table, only Option 1 changes between decisions. To simplify your choice, the computer will pre-fill choices as soon as you click on one decision. If you choose Option 1, all decision on the table above that choice will be pre-filled with Option 1; while all decision below a choice where you choose Option 2 will be pre-filled with Option 2. You can always change your decision until you click on "Submit".

Continue

Figure 4.C.32: Social preferences 2



Your decisions: Table 1

You see 10 decisions on this table. For each decision, you choose between Option 1 and Option 2.

Option **1**: Your payoff: 63, Other's payoff: 75	*or*	Option **2**: Your payoff: 50, Other's payoff: 130
Option **1**: Your payoff: 58, Other's payoff: 75	*or*	Option **2**: Your payoff: 50, Other's payoff: 130
Option **1**: Your payoff: 53, Other's payoff: 75	*or*	Option **2**: Your payoff: 50, Other's payoff: 130
Option **1**: Your payoff: 48, Other's payoff: 75	*or*	Option **2**: Your payoff: 50, Other's payoff: 130
Option **1**: Your payoff: 43, Other's payoff: 75	*or*	Option **2**: Your payoff: 50, Other's payoff: 130
Option **1**: Your payoff: 38, Other's payoff: 75	*or*	Option **2**: Your payoff: 50, Other's payoff: 130
Option **1**: Your payoff: 33, Other's payoff: 75	*or*	Option **2**: Your payoff: 50, Other's payoff: 130
Option **1**: Your payoff: 28, Other's payoff: 75	*or*	Option **2**: Your payoff: 50, Other's payoff: 130
Option **1**: Your payoff: 23, Other's payoff: 75	*or*	Option **2**: Your payoff: 50, Other's payoff: 130
Option **1**: Your payoff: 18, Other's payoff: 75	*or*	Option **2**: Your payoff: 50, Other's payoff: 130

Submit

Back to the instructions

Figure 4.C.33: Social preferences 3

Your decisions: Table 2

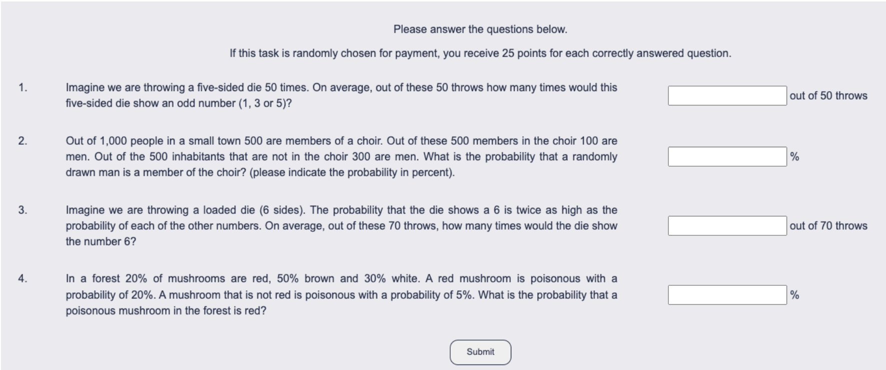
You see 10 decisions on this table. For each decision, you choose between Option 1 and Option 2.

○ Option **1**: Your payoff: 93, Other's payoff: 45	*or*	○ Option **2**: Your payoff: 85, Other's payoff: 25
○ Option **1**: Your payoff: 88, Other's payoff: 45	*or*	○ Option **2**: Your payoff: 85, Other's payoff: 25
○ Option **1**: Your payoff: 83, Other's payoff: 45	*or*	○ Option **2**: Your payoff: 85, Other's payoff: 25
○ Option **1**: Your payoff: 78, Other's payoff: 45	*or*	○ Option **2**: Your payoff: 85, Other's payoff: 25
○ Option **1**: Your payoff: 73, Other's payoff: 45	*or*	○ Option **2**: Your payoff: 85, Other's payoff: 25
○ Option **1**: Your payoff: 68, Other's payoff: 45	*or*	○ Option **2**: Your payoff: 85, Other's payoff: 25
○ Option **1**: Your payoff: 63, Other's payoff: 45	*or*	○ Option **2**: Your payoff: 85, Other's payoff: 25
○ Option **1**: Your payoff: 58, Other's payoff: 45	*or*	○ Option **2**: Your payoff: 85, Other's payoff: 25
○ Option **1**: Your payoff: 53, Other's payoff: 45	*or*	○ Option **2**: Your payoff: 85, Other's payoff: 25
○ Option **1**: Your payoff: 48, Other's payoff: 45	*or*	○ Option **2**: Your payoff: 85, Other's payoff: 25

Submit

Last, participants faced the Berlin numeracy test. In the second experiment, I used only two out of the four questions.

Figure 4.C.34: Numeracy task in the first experiment



Please answer the questions below.

If this task is randomly chosen for payment, you receive 25 points for each correctly answered question.

1. Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3 or 5)? ____ out of 50 throws

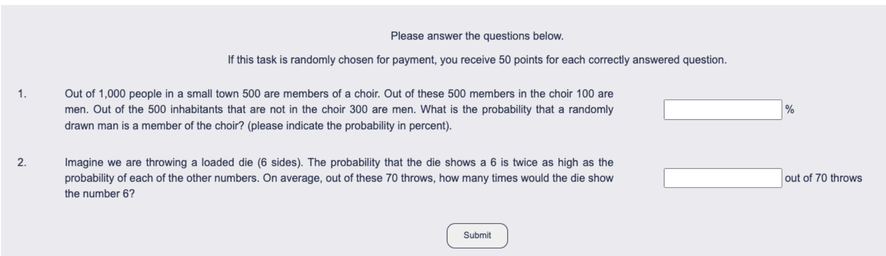
2. Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in the choir 100 are men. Out of the 500 inhabitants that are not in the choir 300 are men. What is the probability that a randomly drawn man is a member of the choir? (please indicate the probability in percent). ____ %

3. Imagine we are throwing a loaded die (6 sides). The probability that the die shows a 6 is twice as high as the probability of each of the other numbers. On average, out of these 70 throws, how many times would the die show the number 6? ____ out of 70 throws

4. In a forest 20% of mushrooms are red, 50% brown and 30% white. A red mushroom is poisonous with a probability of 20%. A mushroom that is not red is poisonous with a probability of 5%. What is the probability that a poisonous mushroom in the forest is red? ____ %

Submit

Figure 4.C.35: Numeracy task in the second experiment



Please answer the questions below.

If this task is randomly chosen for payment, you receive 50 points for each correctly answered question.

1. Out of 1,000 people in a small town 500 are members of a choir. Out of these 500 members in the choir 100 are men. Out of the 500 inhabitants that are not in the choir 300 are men. What is the probability that a randomly drawn man is a member of the choir? (please indicate the probability in percent). ____ %

2. Imagine we are throwing a loaded die (6 sides). The probability that the die shows a 6 is twice as high as the probability of each of the other numbers. On average, out of these 70 throws, how many times would the die show the number 6? ____ out of 70 throws

Submit

CHAPTER 5

Summary:
The Strategic Role of Information in Markets and Games: Essays in Behavioral Economics

This thesis investigates the behavioral role of information in strategic interactions through laboratory experiments, focusing on three different environments where information plays a crucial role.

In Chapter 2, we explore the popularity of open ascending auctions. These auctions, commonly seen on platforms like eBay and in auction houses, reveal bid information to participants. Standard economic theory suggests that open auctions facilitate information aggregation through openly observable bidding. However, another reason for their frequent use may be the activation of revenue-enhancing biases. We conduct a laboratory experiment comparing three auctions with varying levels of information revelation and behavioral bias activation: (i) the ascending Vickrey auction, a closed format; (ii) the Japanese-English auction; and (iii) the Oral Outcry auction, an open format. Contrary to theoretical predictions, we find that the Oral Outcry auction generates higher revenue due to unprofitable jump bids and the triggering of a quasi-endowment effect. In both open formats, information fails to be aggregated, as prices are less predictive of the value of the item for sale, compared to the sealed-bid Vickrey auction.

Chapter 3 examines whether market environments erode the morality of participants. We study the impact of individual traders' market power on morality, norms, and norm compliance. Previous research focused on single-unit markets, which limited the impact of the replacement logic and market selection forces. Our laboratory experiment compares single-unit markets to more common multi-unit markets, where these forces may be active. We find that all markets lead to partial norm erosion. Crucially, unlike single-unit markets, multi-unit markets result in full erosion of morals and norm compliance. The replacement logic drives this finding, as treatment differences in the market design affect beliefs about others' activity. In addition, participants with higher beliefs about others' activity trade more frequently.

In Chapter 4, I investigate how to effectively persuade audiences comprising multiple receivers. Governments, central banks, and private organizations often face the challenge of convincing their audience to take a specific action. They can choose between public messages that coordinate the audience's actions and private messages that may miscoordinate actions. I conduct a laboratory experiment to determine the persuasiveness of public and private messages and to study how this depends on the audience's strategic environment. I find that public signals are more persuasive than predicted by economic theory and are particularly effective in environments featuring strategic complements. Surprisingly, public signals are equally persuasive as private signals under strategic substitutes, contrary to predictions of economic theory. Senders adjust their communication strategies accordingly, engaging more frequently in public communication when the receivers' environment features strategic complements.

CHAPTER 6

Samenvatting
De Strategische Rol van Informatie in Markten en Spellen: Essays in Gedragseconomie

Dit proefschrift onderzoekt de gedragsrol van informatie in strategische interacties door middel van laboratoriumexperimenten en richt zich op drie verschillende omgevingen waarin informatie een cruciale rol speelt.

In Hoofdstuk 2 onderzoeken we de populariteit van open opbodveilingen. Deze veilingen, die veel voorkomen op platforms zoals eBay en in veilinghuizen, onthullen biedinformatie aan de deelnemers. De standaard economische theorie suggereert dat open veilingen informatie-aggregatie vergemakkelijken door openlijk waarneembaar biedgedrag. Er kan echter een andere reden zijn voor hun veelvuldig gebruik, namelijk het activeren van opbrengstverhogende biases. We voeren een laboratoriumexperiment uit waarin we drie veilingen vergelijken met verschillende niveaus van informatie-onthulling en activering van biases: (i) de ascending Vickrey-veiling, een gesloten formaat; (ii) de Japanese-English veiling; en (iii) een openbare opbodveiling, de Oral Outcry veiling. In tegenstelling tot theoretische voorspellingen blijkt dat de openbare opbodveiling een hogere opbrengst genereert als gevolg van onrendabele sprongbiedingen en het activeren van een quasi-endowment-effect. In beide openbare formats wordt informatie niet geaggregeerd, aangezien de prijzen minder voorspellend zijn voor de waarde van het te koop aangeboden item dan in de gesloten-bod Vickrey-veiling.

Hoofdstuk 3 onderzoekt of marktomgevingen de moraliteit van deelnemers aantasten. We bestuderen de impact van de marktmacht van individuele handelaren op moraliteit, normen en de naleving van normen. Eerder onderzoek richtte zich op markten waarin handelaren hooguit één eenheid konden kopen of verkopen, wat de impact van de replacement logic en market selection beperkte. Ons laboratoriumexperiment vergelijkt markten met één eenheid met meer gangbare markten met meerdere eenheden, waarin deze krachten actief kunnen zijn. We ontdekken dat alle markten leiden tot gedeeltelijke aantasting van normen. Cruciaal is dat, in tegenstelling tot markten met één eenheid, markten met meerdere eenheden leiden tot volledige erosie van moraliteit en naleving van normen. De replacement logic drijft dit resultaat, omdat verschillen in marktontwerp van invloed zijn op het handelen in de markt. Bovendien handelen deelnemers die verwachten dat meer andere deelnemers actief zijn, zelfs nog vaker.

In Hoofdstuk 4 onderzoek ik hoe je effectief publieksgroepen kunt overtuigen die uit meerdere ontvangers bestaan. Overheden, centrale banken en particuliere organisaties worden vaak geconfronteerd met de uitdaging om hun publiek ervan te overtuigen een specifieke actie te ondernemen. Ze kunnen kiezen tussen openbare boodschappen die de acties van het publiek coördineren en privéboodschappen die acties mogelijk anti-coördineren. Ik voer een laboratoriumexperiment uit om de overtuigingskracht van openbare en privéboodschappen te bepalen en te bestuderen hoe die afhangt van de strategische omgeving. Ik ontdek dat openbare signalen overtuigender zijn dan voorspeld door de economische theorie en met name effectief zijn in omgevingen met strategische complementen. Verrassend genoeg zijn openbare signalen even overtuigend als privésignalen bij strategische substituten, in tegenstelling tot voorspellingen van de economische theorie. Verzenders passen hun communicatiestrategieën dienovereenkomstig aan en maken vaker gebruik van openbare communicatie wanneer de omgeving van de ontvangers strategische complementen bevat.

Bibliography

Abeler, J., D. Nosenzo, and C. Raymond (2019): “Preferences for truth-telling,” *Econometrica*, 87, 1115–1153.

Agranov, M., and A. Schotter (2013): “Language and government coordination: An experimental study of communication in the announcement game,” *Journal of Public Economics*, 104, 26–39.

Alpert, A., W. N. Evans, E. M. Lieber, and D. Powell (2022): “Origins of the opioid crisis and its enduring impacts,” *The Quarterly Journal of Economics*, 137, 1139–1179.

Anbarci, N., N. Feltovich, and M. Y. Gürdal (2018): “Payoff inequity reduces the effectiveness of correlated-equilibrium recommendations,” *European Economic Review*, 108, 172–190.

Anderson, B. L. R., and C. A. Holt (1997): “Information Cascades in the Laboratory,” *American Economic Review*, 87, 847–862.

Andreoni, J., Y. K. Che, and J. Kim (2007): “Asymmetric information about rivals’ types in standard auctions: An experiment,” *Games and Economic Behavior*, 59, 240–259.

Angeletos, G.-M., and A. Pavan (2007): “Efficient use of information and social value of information,” *Econometrica*, 75, 1103–1142.

Arieli, I., and Y. Babichenko (2019): “Private bayesian persuasion,” *Journal of Economic Theory*, 182, 185–217.

Ariely, D., A. Ockenfels, and A. Roth (2005): “An experimental analysis of ending rules in internet auctions,” *RAND Journal of Economics*, 36, 890–907.

Aristidou, A., G. Coricelli, and A. Vostroknutov (2019): “Incentives or Persuasion? An Experimental Investigation,” *Working paper*.

Au, P. H., and K. K. Li (2018): “Bayesian persuasion and reciprocity: Theory and experiment,” *Working paper*.

Aumann, R. J. (1987): “Correlated equilibrium as an expression of Bayesian rationality,” *Econometrica*, 55, 1–18.

Ausubel, L. M. (2004): “An efficient ascending-bid auction for multiple objects,” *American Economic Review*, 94, 1452–1475.

Avery, C. (1998): “Strategic jump bidding in English auctions,” *The Review of Economic Studies*, 65, 185–210.

Avoyan, A. (2022): “Communication in global games: Theory and experiment,” *Working paper*.

Barr, S., G. Shaw, and T. Coles (2011): “Times for (Un)sustainability? Challenges and opportunities for developing behaviour change policy. A case-study of consumers at home and away,” *Global Environmental Change*, 21, 1234–1244.

Bartling, B., E. Fehr, and Y. Özdemir (2023): “Does market interaction erode moral values?” *Review of Economics and Statistics*, 105, 226–235.

Bartling, B., T. Gesche, and N. Netzer (2016): “Does the Absence of Human Sellers Bias Bidding Behavior in Auction Experiments?,” *Journal of the Economic Sciences Association*, 44–61.

Bartling, B., and N. Netzer (2016): “An externality-robust auction: Theory and experimental evidence,” *Games and Economic Behavior*, 97, 186–204.

Bartling, B., and Y. Özdemir (2023): “The limits to moral erosion in markets: Social norms and the replacement excuse,” *Games and Economic Behavior*, 138, 143–160.

Bartling, B., V. Valero, and R. Weber (2019): “On the scope of externalities in experimental markets,” *Experimental Economics*, 22, 610–624.

Bartling, B., R. A. Weber, and L. Yao (2015): “Do markets erode social responsibility?” *The Quarterly Journal of Economics*, 130, 219–266.

Battaglini, M., and U. Makarov (2014): “Cheap talk with multiple audiences: An experimental analysis,” *Games and Economic Behavior*, 83, 147–164.

Bénabou, R., A. Falk, L. Henkel, and J. Tirole (2020): “Eliciting moral preferences: Theory and Experiment,” *Working paper*.

Benjamin, D. J. (2019): “Errors in probabilistic reasoning and judgment biases,” in *Handbook of Behavioral Economics - Foundations and Applications 2* ed. by Bernheim, B. D., DellaVigna, S., and Laibson, D.: North-Holland, 69–186.

Bergemann, D., B. Brooks, and S. Morris (2019): “Revenue guarantee equivalence,” *American Economic Review*, 109, 1911–29.

Bergemann, D., and S. Morris (2016): “Bayes correlated equilibrium and the comparison of information structures in games,” *Theoretical Economics*, 11, 487–522.

——— (2019): “Information design: A Unified Perspective,” *Journal of Economic Literature*, 57, 44–95.

Betz, R., B. Greiner, S. Schweitzer, and S. Seifert (2017): “Auction Format and Auction Sequence in Multi-item Multi-unit Auctions: An Experimental Study,” *The Economic Journal*, 127, F351–F371.

Bholat, D., N. Broughton, A. Parker, J. Ter Meer, and E. Walczak (2018): “Enhancing central bank communications with behavioural insights,”Technical report, Bank of England.

Bikhchandani, S., P. A. Haile, and J. G. Riley (2002): “Symmetric separating equilibria in English auctions,” *Games and Economic Behavior*, 38, 19–27.

Bikhchandani, S., and J. G. Riley (1991): “Equilibria in open common value auctions,” *Journal of Economic Theory*, 53, 101–130.

Binder, C. (2017): “Fed speak on main street: Central bank communication and household expectations,” *Journal of Macroeconomics*, 52, 238–251.

Blume, A. (2018): “Failure of common knowledge of language in common-interest communication games,” *Games and Economic Behavior*, 109, 132–155.

Blume, A., and O. Board (2013): “Language barriers,” *Econometrica*, 81, 781–812.

Blume, A., E. K. Lai, and W. Lim (2020): “Strategic information transmission: A survey of experiments and theoretical foundations,” in *Handbook of Experimental Game Theory*: Edward Elgar Publishing.

——— (2023): “Mediated talk: An experiment,” *Journal of Economic Theory*, 208, 105593.

Bone, J., M. Drouvelis, and I. Ray (2013): “Co-ordination in 2 x 2 Games by Following Recommendations from Correlated Equilibria,” *Working paper*.

Boone, J., R. Chen, J. K. Goeree, and A. Polydoro (2009): “Risky procurement with an insider bidder,” *Experimental Economics*, 12, 417.

Brandts, J., and C. A. Holt (1992): “An Experimental Test of Equilibrium Dominance in Signaling Games,” *American Economic Review*, 82, 1350–1365.

Brooks, B., and S. Du (2021): "Optimal auction design with common values: An informationally robust approach," *Econometrica*, 89, 1313–1360.

Brütt, K., A. Schram, and J. Sonnemans (2020): "Endogenous group formation and responsibility diffusion: An experimental study," *Games and Economic Behavior*, 121, 1–31.

Burlando, R. M., and F. Guala (2005): "Heterogeneous agents in public goods experiments," *Experimental Economics*, 8, 35–54.

Buser, T., M. Niederle, and H. Oosterbeek (2020): "Can competitiveness predict education and labor market outcomes? Evidence from incentivized choices and validated survey measures," *Tinbergen Institute Discussion Paper*, 20-048/I.

Byambadalai, U., C.-t. A. Ma, and D. Wiesen (2019): "Changing Preferences: An Experiment and Estimation of Market-Incentive Effects on Altruism," *Working paper*.

Cabrales, A., R. Nagel, and R. Armenter (2007): "Equilibrium selection through incomplete information in coordination games: An experimental study," *Experimental Economics*, 10, 221–234.

Camerer, C., S. Nunnari, and T. R. Palfrey (2016): "Quantal response and nonequilibrium beliefs explain overbidding in maximum-value auctions," *Games and Economic Behavior*, 98, 243–263.

Carlsson, H., and E. van Damme (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61, 989–1018.

Casella, A., E. Friedman, and M. P. Archila (2020): "Mediating Conflict in the Lab," *National Bureau of Economic Research*.

Cason, T. N., and T. Sharma (2007): "Recommended play and correlated equilibria: An experimental study," *Economic Theory*, 33, 11–27.

Charness, G., and D. Levin (2009): "The Origin of the Winner's Curse: A Laboratory Study," *American Economic Journal: Microeconomics*, 1, 207–236.

Cho, S.-J., H. J. Paarsch, and J. Rust (2014): "Is the 'Linkage Principle' Valid? Evidence from the Field," *The Journal of Industrial Economics*, 62, 346–375.

Choi, S., J.-A. Guerra, and J. Kim (2019): "Interdependent value auctions with insider information: Theory and experiment," *Games and Economic Behavior*, 117, 218–237.

Cokely, E. T., M. Galesic, E. Schulz, S. Ghazal, and R. Garcia-Retamero (2012): "Measuring Risk Literacy: The Berlin Numeracy Test," *Judgment and Decision Making*, 7, 25–47.

Cooper, D. J., J. R. Hamman, and R. A. Weber (2020): "Fool me once: An experiment on credibility and leadership," *The Economic Journal*, 130, 2105–2133.

Cornand, C., and F. Heinemann (2008): "Optimal degree of public information dissemination," *The Economic Journal*, 118, 718–742.

——— (2014): "Measuring agents' reaction to private and public information in games with strategic complementarities," *Experimental Economics*, 17, 61–77.

Cramton, P. (1998): "Ascending auctions," *European Economic Review*, 42, 745–756.

Crawford, V. P., and N. Iriberri (2007): "Level-K Auctions: Can a Nonequilibrium Model of Strategic Thinking Explain the Winner's Curse in Private-Value Auctions?," *Econometrica*, 75, 1721–1770.

Crawford, V. P., and J. Sobel (1982): "Strategic information transmission," *Econometrica*, 50, 1431–1451.

Dana, J., R. A. Weber, and J. X. Kuang (2007): "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness," *Economic Theory*, 33, 67–80.

Das, S., E. Kamenica, and R. Mirka (2017): "Reducing congestion through information design," in *2017 55th annual allerton conference on communication, control, and computing (allerton)*, 1279–1284, IEEE.

DeGroot, M. H. (2005): *Optimal statistical decisions* Volume 82: John Wiley & Sons.

Drugov, M., R. Hernán-González, P. Kujal, and M. Troya-Martinez (2021): "Cheap Talk with Two Audiences: An Experiment," *Working paper*.

Duffy, J., and N. Feltovich (2010): "Correlated equilibria, good and bad: An experimental study," *International Economic Review*, 51, 701–721.

Duffy, J., and E. O. Fisher (2005): "Sunspots in the laboratory," *American Economic Review*, 95, 510–529.

Dufwenberg, M., and U. Gneezy (2002): "Information disclosure in auctions: An experiment," *Journal of Economic Behavior and Organization*, 48, 431–444.

Dyer, D., J. H. Kagel, and D. Levin (1989): "A Comparison of Naive and Experienced Bidders in Common Value Offer Auctions: A Laboratory Analysis," *The Economic Journal*, 99, 108–115.

Eckel, C. C., and P. J. Grossman (2002): "Sex differences and statistical stereotyping in attitudes toward financial risk," *Evolution and human behavior*, 23, 281–295.

Ehrhart, K. M., M. Ott, and S. Abele (2015): “Auction fever: Rising revenue in second-price auction formats,” *Games and Economic Behavior*, 92, 206–227.

Ely, J. C. (2017): “Beeps,” *American Economic Review*, 107, 31–53.

Embrey, M., F. Mengel, and R. Peeters (2019): “Strategy revision opportunities and collusion,” *Experimental Economics*, 22, 834–856.

Engelmann, D., J. Friedrichsen, and D. Kübler (2018): “Fairness in markets and market experiments,” *WZB Discussion Paper*.

Enke, B., T. Graeber, and R. Oprea (2022): “Confidence, Self-Selection and Bias in the Aggregate,” *Working paper*.

Enke, B., and F. Zimmermann (2019): “Correlation Neglect in Belief Formation,” *The Review of Economic Studies*, 86, 313–332.

Esponda, I., and E. Vespa (2014): “Hypothetical Thinking and Information Extraction: Pivotal Voting in the Laboratory,” *American Economic Journal: Microeconomics*, 6, 180–202.

Ettinger, D., and F. Michelucci (2016): “Hiding Information in Open Auctions with Jump Bids,” *The Economic Journal*, 126, 1484–1502.

Exley, C. L. (2016): “Excusing selfishness in charitable giving: The role of risk,” *The Review of Economic Studies*, 83, 587–628.

Eyster, E., and M. Rabin (2005): “Cursed Equilibrium,” *Econometrica*, 73, 1623–1672.

Eyster, E., M. Rabin, and G. Weizsacker (2018): “An Experiment on Social Mislearning,” *Working paper*.

Falk, A., T. Neuber, and N. Szech (2020): “Diffusion of Being Pivotal and Immoral Outcomes,” *The Review of Economic Studies*, 87, 2205–2229.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018): “Global evidence on economic preferences,” *The Quarterly Journal of Economics*, 133, 1645–1692.

Falk, A., and N. Szech (2013): “Morals and markets,” *Science*, 340, 707–711.

——— (2014): “Diffusion of Being Pivotal and Immoral Outcomes,” *Working paper*.

Farrell, J., and R. Gibbons (1989): “Cheap talk with two audiences,” *American Economic Review*, 79, 1214–1223.

Fehr, D., F. Heinemann, and A. Llorente-Saguer (2019): “The power of sunspots: An experimental analysis,” *Journal of Monetary Economics*, 103, 123–136.

Fehr, E., and K. M. Schmidt (1999): "A theory of fairness, competition, and cooperation," *The Quarterly Journal of Economics*, 114, 817–868.

Fehr, E., and J.-R. Tyran (2008): "Limited rationality and strategic interaction: the impact of the strategic environment on nominal inertia," *Econometrica*, 76, 353–394.

Fischbacher, U., C. M. Fong, and E. Fehr (2009): "Fairness, errors and the power of competition," *Journal of Economic Behavior & Organization*, 72, 527–545.

Fischbacher, U., S. Gächter, and E. Fehr (2001): "Are people conditionally cooperative? Evidence from a public goods experiment," *Economics Letters*, 71, 397–404.

Forges, F. (1993): "Five legitimate definitions of correlated equilibrium in games with incomplete information," *Theory and decision*, 35, 277–310.

Forsythe, R., and R. Lundholm (1990): "Information Aggregation in an Experimental Market," *Econometrica*, 58, 309–347.

Frechette, G. R., A. Lizzeri, and J. Perego (2022): "Rules and Commitment in Communication: An Experimental Analysis," *Econometrica*, 90, 2283–2318.

Friedman, D., J. P. Rabanal, O. A. Rud, and S. Zhao (2022): "On the empirical relevance of correlated equilibrium," *Journal of Economic Theory*, 205, 105531.

van Gils, F., W. Müller, and J. Prüfer (2022): "Social Media and Democracy: Experimental Results," *Working paper*.

Giovannoni, F., and S. Xiong (2019): "Communication under language barriers," *Journal of Economic Theory*, 180, 274–303.

Goeree, J. K., and T. Offerman (2002): "Efficiency in Auctions with Private and Common Values: An Experimental Study," *American Economic Review*, 92, 625–643.

——— (2003a): "Competitive bidding in auctions with private and common values," *The Economic Journal*, 113, 598–613.

——— (2003b): "Winner's curse without overbidding," *European Economic Review*, 47, 625–644.

Goeree, J. K., and L. Yariv (2015): "Conformity in the lab," *Journal of the Economic Science Association*, 1, 15–28.

Gonçalves, R., and J. Hey (2011): "Experimental evidence on English auctions: Oral outcry versus clock," *Bulletin of Economic Research*, 63, 313–352.

Gössling, S., and A. Humpe (2020): "The global scale, distribution and growth of aviation: Implications for climate change," *Global Environmental Change*, 65, 102194.

Greiner, B. (2015): “Subject pool recruitment procedures: organizing experiments with ORSEE,” *Journal of the Economic Science Association*, 1, 114–125.

Grossman, Z. (2014): “Strategic ignorance and the robustness of social preferences,” *Management Science*, 60, 2659–2665.

Grossman, Z., and J. J. Van Der Weele (2017): “Self-image and willful ignorance in social decisions,” *Journal of the European Economic Association*, 15, 173–217.

Haile, P. A., and E. Tamer (2003): “Inference with an incomplete model of English auctions,” *Journal of Political Economy*, 111, 1–51.

Haldane, A., A. Macaulay, and M. McMahon (2021): “The three E’s of central-bank communication with the public,” *Working paper*.

Haldane, A., and M. McMahon (2018): “Central Bank Communications and the General Public,” *AEA Papers and Proceedings*, 108, 578–83.

Hanson, R., R. Oprea, and D. Porter (2006): “Information aggregation and manipulation in an experimental market,” *Journal of Economic Behavior and Organization*, 60, 449–459.

Heinemann, F., R. Nagel, and P. Ockenfels (2004): “The theory of global games on test: Experimental analysis of coordination games with public and private information,” *Econometrica*, 72, 1583–1599.

——— (2009): “Measuring strategic uncertainty in coordination games,” *The Review of Economic Studies*, 76, 181–221.

Herweg, F., and K. M. Schmidt (2022): “How to Regulate Carbon Emissions with Climate-Conscious Consumers,” *The Economic Journal*, 132, 2992–3019.

Heyman, J. E., Y. Orhun, and D. Ariely (2004): “Auction fever: The effect of opponents and quasi-endowment on product valuations,” *Journal of Interactive Marketing*, 18, 7–21.

Hoelzl, E., and A. Rustichini (2005): “Do You Put Your Money on It?,” *The Economic Journal*, 115, 305–318.

Hossain, T., and R. Okui (2018): “Belief Formation Under Signal Correlation,” *Working Paper*.

Huck, S., and K. A. Konrad (2005): “Moral cost, commitment, and committee size,” *Journal of Institutional and Theoretical Economics*, 575–588.

Inostroza, N., and A. Pavan (2021): “Persuasion in global games with application to stress testing,” *Working paper*.

Irlenbusch, B., and D. J. Saxler (2019): "The role of social information, market framing, and diffusion of responsibility as determinants of socially responsible behavior," *Journal of Behavioral and Experimental Economics*, 80, 141–161.

Ivanov, A., D. Levin, and M. Niederle (2010): "Can Relaxation of Beliefs Rationalize the Winners Curse?: An Experimental Study," *Econometrica*, 78, 1435–1452, 10.3982/ECTA8112.

Kagel, J. H., and D. Levin (1986): "The Winner's Curse and Public Information in Common Value Auctions," *American Economic Review*, 76, 894–920.

——— (2014): "Auctions: A survey of experimental research," 1–135.

Kagel, J. H., D. Levin, and R. M. Harstad (1995): "Comparative Static Effects of Number of Bidders and Public Information on Behavior in Second-Price Common Value Auctions," *International Journal of Game Theory*, 24, 293–319.

Kamenica, E., and M. Gentzkow (2011): "Bayesian persuasion," *American Economic Review*, 101, 2590–2615.

Kapoor, S., and A. Magesan (2014): "Paging inspector Sands: The costs of public information," *American Economic Journal: Economic Policy*, 6, 92–113.

Ketcham, J., V. L. Smith, and A. W. Williams (1984): "A comparison of posted-offer and double-auction pricing institutions," *The Review of Economic Studies*, 51, 595–614.

Kirchkamp, O., and B. Moldovanu (2004): "An experimental analysis of auctions with interdependent valuations," *Games and Economic Behavior*, 48, 54–85.

Kirchler, M., J. Huber, M. Stefan, and M. Sutter (2016): "Market design and moral behavior," *Management Science*, 62, 2615–2625.

Klemperer, P., and P. Temin (2001): "An Early Example of the "Winner's Curse" in an Auction," *Journal of Political Economy*, 109, Back Cover.

Koch, C., and S. P. Penczynski (2018): "The winner's curse: Conditional reasoning and belief formation," *Journal of Economic Theory*, 174, 57–102.

Kremer, I. (2002): "Information Aggregation in Common Value Auctions," *Econometrica*, 70, 1675–1682.

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

Kurz, V., A. Orland, and K. Posadzy (2018): "Fairness versus efficiency: How procedural fairness concerns affect coordination," *Experimental Economics*, 21, 601–626.

van Leeuwen, B., T. Offerman, and A. Schram (2020): "Competition for status creates superstars: An experiment on public good provision and network formation," *Journal of the European Economic Association*, 18, 666–707.

Levin, D., J. H. Kagel, and J.-F. Richard (1996): "Revenue Effects and Information Processing in English Common Value Auctions," *American Economic Review*, 86, 442–460.

Levin, D., J. Peck, and A. Ivanov (2016): "Separating Bayesian updating from non-probabilistic reasoning: An experimental investigation," *American Economic Journal: Microeconomics*, 8, 39–60.

Levin, J., and A. Skrzypacz (2016): "Properties of the Combinatorial Clock Auction," *American Economic Review*, 106, 2528–2551.

Linde, J., and J. Sonnemans (2012): "Social comparison and risky choices," *Journal of Risk and Uncertainty*, 44, 45–72.

List, J. A. (2003): "Does market experience eliminate market anomalies?" *The Quarterly Journal of Economics*, 118, 41–71.

Magnani, J., and R. Oprea (2017): "Why Do People Violate No-Trade Theorems? A Diagnostic Test," *Working paper*.

Martínez-Marquina, A., M. Niederle, and E. Vespa (2019): "Failures in contingent reasoning: The role of uncertainty," *American Economic Review*, 109, 3437–74.

Mathevet, L., J. Perego, and I. Taneva (2020): "On information design in games," *Journal of Political Economy*, 128, 1370–1404.

McMillan, J. (1994): "Selling Spectrum Rights," *Journal of Economic Perspectives*, 8, 145–162.

Mermer, A. G., W. Müller, and S. Suetens (2021): "Cooperation in infinitely repeated games of strategic complements and substitutes," *Journal of Economic Behavior & Organization*, 188, 1191–1205.

Milgrom, P. R. (1989): "Auctions and Bidding: A Primer," *Journal of Economic Perspectives*, 3, 3–22.

Milgrom, P. R., and R. J. Weber (1982): "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50, 1089–1122.

——— (2004): *Putting auction theory to work*: Cambridge University Press.

Moreno, D., and J. Wooders (1998): "An experimental study of communication and coordination in noncooperative games," *Games and Economic Behavior*, 24, 47–76.

Morris, S., and H. S. Shin (1998): "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks," *American Economic Review*, 88, 587–597.

——— (2002): "Social value of public information," *American Economic Review*, 92, 1521–1534.

NERA (2017): "Effective Spectrum Pricing in Europe: Policies to support better quality and more affordable mobile services."

Ngangoué, M. K., and G. Weizsäcker (2021): "Learning from unrealized versus realized prices," *American Economic Journal: Microeconomics*, 13, 174–201.

Ockenfels, A., P. Werner, and O. Edenhofer (2020): "Pricing externalities and moral behaviour," *Nature Sustainability*, 3, 872–877.

Offerman, T., G. Romagnoli, and A. Ziegler (2019a): "Endogenous information revelation in common value auctions," *AEA RCT Registry*, https://doi.org/10.1257/rct.2965-3.0.

——— (2019b): "Morals in multi-unit markets," *AEA RCT Registry*, https://doi.org/10.1257/rct.4603-1.0.

——— (2021): "Morals in multi-unit markets: Study II," *AEA RCT Registry*, https://doi.org/10.1257/rct.8306-1.0.

——— (2022): "Why are open ascending auctions popular? The role of information aggregation and behavioral biases," *Quantitative Economics*, 13, 787–823.

Offerman, T., J. Sonnemans, and A. Schram (1996): "Value orientations, expectations and voluntary contributions in public goods," *The Economic Journal*, 106, 817–845.

Oprea, R. (2020): "What makes a rule complex?" *American Economic Review*, 110, 3913–51.

Perry, M., and P. J. Reny (1999): "On the failure of the linkage principle in multi-unit auctions," *Econometrica*, 67, 895–900.

Pesendorfer, W., and J. M. Swinkels (2000): "Efficiency and Information Aggregation in Auctions," *American Economic Review*, 90, 499–525.

Plott, C. R. (1983): "Externalities and corrective policies in experimental markets," *The Economic Journal*, 93, 106–127.

Plott, C., and S. Sunder (1988): "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets," *Econometrica*, 56, 1085–1118.

Potters, J., and S. Suetens (2009): "Cooperation in experimental games of strategic complements and substitutes," *The Review of Economic Studies*, 76, 1125–1147.

PowerAuctions (2015): "Auction design considerations for the public sector spectrum release," Prepared for Hutchinson 3G.

Prasnikar, V., and A. E. Roth (1992): "Considerations of fairness and strategy: Experimental data from sequential games," *The Quarterly Journal of Economics*, 107, 865–888.

Riehm, T., N. Fugger, P. Gillen, V. Gretschko, and P. Werner (2022): "Social norms, sanctions, and conditional entry in markets with externalities: Evidence from an artefactual field experiment," *Journal of Public Economics*, 212, 104701.

Roth, A. E., V. Prasnikar, M. Okuno-Fujiwara, and S. Zamir (1991): "Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study," *American Economic Review*, 81, 1068–1095.

Roth, A., and A. Ockenfels (2002): "Last-Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet," *American Economic Review*, 92, 1093–1103.

Rothenhäusler, D., N. Schweizer, and N. Szech (2018): "Guilt in voting and public good games," *European Economic Review*, 101, 664–681.

Samuelson, P. A., and W. D. Nordhaus (2005): *Economics*: Columbus, OH: McGraw-Hill Irwin, 18th edition.

Schneider, F., F. Brun, and R. A. Weber (2020): "Sorting and wage premiums in immoral work," *University of Zurich, Department of Economics, Working Paper*, 353.

Scholl, L., P. Seth, M. Kariisa, N. Wilson, and G. Baldwin (2019): "Drug and opioid-involved overdose deaths—United States, 2013–2017," *Morbidity and Mortality Weekly Report*, 67, 1419.

Serra-Garcia, M., and N. Szech (2022): "The (in) elasticity of moral ignorance," *Management Science*, 68, 4815–4834.

Smith, V. L. (1962): "An Experimental Study of Competitive Market Behavior," *Journal of Political Economy*, 70, 111–137.

Sobel, J. (2007): "Do Markets Make People Look Selfish?" *Working paper*.

Sonnemans, J., F. van Dijk, and F. van Winden (2006): "On the dynamics of social ties structures in groups," *Journal of Economic Psychology*, 27, 187–204.

Sutter, M., J. Huber, M. Kirchler, M. Stefan, and M. Walzl (2020): "Where to Look for the Morals in Markets?" *Experimental Economics*, 23, 30–52.

Szkup, M., and I. Trevino (2021): “Information acquisition and self-selection in coordination games,” *Working paper*.

Taneva, I. (2019): “Information design,” *American Economic Journal: Microeconomics*, 11, 151–85.

Taneva, I., and T. Wiseman (2022): “Strategic Ignorance and Information Design,” *Working paper*.

Tappin, B. M., C. Wittenberg, L. B. Hewitt, A. J. Berinsky, and D. G. Rand (2022): “Quantifying the Persuasive Returns to Political Microtargeting,” *Working paper*.

Thaler, R. (1980): “Toward a positive theory of consumer choice,” *Journal of Economic Behavior & Organization*, 1, 39–60.

Trevino, I. (2020): “Informational channels of financial contagion,” *Econometrica*, 88, 297–335.

Van Huyck, J. B., A. B. Gillette, and R. C. Battalio (1992): “Credible assignments in coordination games,” *Games and Economic Behavior*, 4, 606–626.

Vuillemey, G. (2020): “Evading Corporate Responsibilities: Evidence from the Shipping Industry,” *Working paper*.

Wilson, R. (1977): “A Bidding Model of Perfect Competition,” *The Review of Economic Studies*, 44, 511–518.

Yang, Y., S. Onderstal, and A. Schram (2016): “Inequity aversion revisited,” *Journal of Economic Psychology*, 54, 1–16.

Ziegler, A. (2021): “Optimal (mis)coordination under uncertainty: Testing information design in the laboratory,” *AEA RCT Registry*, https://doi.org/10.1257/rct.7060.

——— (2022): “Optimal (mis-)coordination under uncertainty: Testing information design in the laboratory; study 2,” *AEA RCT Registry*, https://doi.org/10.1257/rct.9780.

——— (2023): “Persuading an audience: Testing information design in the laboratory,” *Working paper*.

Ziegler, A., G. Romagnoli, and T. Offerman (2022): “Morals in multi-unit markets,” *Tinbergen Institute Discussion Paper*.

List of co-authors and contributions

Chapter 4 of this thesis is single-authored, Chapter 2 and Chapter 3 are based on co-authored work. All references of co-authored work are provided in the chapters. The contributions of the individual authors in the co-authored chapters are outlined below.

Chapter 2: "Why are open ascending auctions popular? The role of information aggregation and behavioral biases"
Co-authors: Theo Offerman and Giorgia Romagnoli.

The project is based on the MPhil thesis of Andreas Ziegler, on an initial experimental idea by Theo Offerman. All authors jointly developed the experimental design. Andreas Ziegler programmed the experiment, conducted the experimental sessions, analyzed the data, and wrote the first draft of the working paper. All authors contributed to revisions of the working paper.

Chapter 3: "Morals in multi-unit markets."
Co-authors: Giorgia Romagnoli and Theo Offerman.

All authors contributed to the experimental idea and design. Andreas Ziegler programmed the experiment, conducted the experimental sessions, analyzed the data, and wrote the first draft of the working paper. All authors contributed to revisions of the working paper.

The Tinbergen Institute is the Institute for Economic Research, which was founded in 1987 by the Faculties of Economics and Econometrics of the Erasmus University Rotterdam, University of Amsterdam and Vrije Universiteit Amsterdam. The Institute is named after the late Professor Jan Tinbergen, Dutch Nobel Prize laureate in economics in 1969. The Tinbergen Institute is located in Amsterdam and Rotterdam. For a full list of PhD theses that appeared in the series we refer to List of PhD Theses – Tinbergen.nl. The following books recently appeared in the Tinbergen Institute Research Series:

768. Y. Xu, *Eliciting Preferences and Private Information: Tell Me What You Like and What You Think*
769. S. Rellstab, *Balancing Paid Work and Unpaid Care over the Life-Cycle*
770. Z. Deng, *Empirical Studies in Health and Development Economics*
771. L. Kong, *Identification Robust Testing in Linear Factor Models*
772. I. Neamţu, *Unintended Consequences of Post-Crisis Banking Reforms*
773. B. Klein Teeselink, *From Mice to Men: Field Studies in Behavioral Economics*
774. B. Tereick, *Making Crowds Wiser: The Role of Incentives, Individual Biases, and Improved Aggregation*
775. A. Castelein, *Models for Individual Responses*
776. D. Kolesnyk, *Consumer Disclosures on Social Media Platforms: A Global Investigation*
777. M.A. Rola-Janicka, *Essays on Financial Instability and Political Economy of Regulation*
778. J.J. Klingen, *Natural Experiments in Environmental and Transport Economics*
779. E.M. Aermann, *Educational Choices and Family Outcomes*
780. F.J. Ostermeijer, *Economic Analyses of Cars in the City*
781. T. Özden, *Adaptive Learning and Monetary Policy in DSGE Models*
782. D. Wang, *Empirical Studies in Financial Stability and Natural Capital*
783. L.S. Stephan, *Estimating Diffusion and Adoption Parameters in Networks New Estimation Approaches for the Latent-Diffusion-Observed-Adoption Model*
784. S.R. Mayer, *Essays in Financial Economics*
785. A.R.S. Woerner, *Behavioral and Financial Change – Essays in Market Design*
786. M. Wiegand, *Essays in Development Economics*
787. L.M. Treuren, *Essays in Industrial Economics - Labor market imperfections, cartel stability, and public interest cartels*
788. D.K. Brands, *Economic Policies and Mobility Behaviour*
789. H.T.T. Nguyen, *Words Matter? Gender Disparities in Speeches, Evaluation and Competitive Performance*
790. C.A.P Burik, *The Genetic Lottery. Essays on Genetics, Income, and Inequality*
791. S.W.J. Olijslagers, *The Economics of Climate Change: on the Role of Risk and*

Preferences

792. C.W.A. van der Kraats, *On Inequalities in Well-Being and Human Capital Formation*
793. Y. Yue, *Essays on Risk and Econometrics*
794. E.F. Janssens, *Estimation and Identification in Macroeconomic Models with Incomplete Markets*
795. P.B. Kastelein, *Essays in Household Finance: Pension Funding, Housing and Consumption over the Life Cycle*
796. J.O. Oorschot, *Extremes in Statistics and Econometrics*
797. S.D.T. Hoey, *Economics on Ice: Research on Peer Effects, Rehiring Decisions and Worker Absenteeism*
798. J. Vidiella-Martin, *Levelling the Playing Field: Inequalities in early life conditions and policy responses*
799. Y. Xiao, *Fertility, parental investments and intergenerational mobility*
800. X. Yu, *Decision Making under Different Circumstances: Uncertainty, Urgency, and Health Threat*
801. G. Gianluca, *Productivity and Strategies of Multiproduct Firms*
802. H. Kweon, *Biological Perspective of Socioeconomic Inequality*
803. D.K. Dimitrov, *Three Essays on the Optimal Allocation of Risk with Illiquidity, Intergenerational Sharing and Systemic Institutions*
804. J.B. Bloomfield, *Essays on Early Childhood Interventions*
805. S. Yu, *Trading and Clearing in Fast-Paced Markets*
806. M.G. Gregori, *Advanced Measurement and Sampling for Marketing Research*
807. O.C. Soons, *The Past, Present, and Future of the Euro Area*
808. D. Garces Urzainqui, *The Distribution of Development. Essays on Economic Mobility, Inequality and Social Change*
809. A.C. Peker, *Guess What I Think: Essays on the Wisdom in Meta-predictions*
810. A. Akdeniz, *On the Origins of Human Sociality*
811. K. Brütt, *Strategic Interaction and Social Information: Essays in Behavioural Economics*
812. P.N. Kusumawardhani, *Learning Trends and Supply-side Education Policies in Indonesia*
813. F. Capozza, *Essays on the Behavioral Economics of Social Inequalities*
814. D.A. Muslimova, *Complementarities in Human Capital Production: The Role of Gene-Environment Interactions*
815. J.A. De Jong, *Coordination in Market and Bank Run Experiments*
816. Y. Kim, *Micro studies of macroprudential policies using loan-level data*
817. S.R. Ter Meulen, *Grade retention, ability tracking, and selection in education*