An investigation of prosocial behavior in different economic contexts

kumulative Dissertation

Universität Hamburg,
Fakultät Wirtschafts- und Sozialwissenschaften

Dissertation
Zur Erlangung der Würde der Doktorin
der Wirtschafts-und Sozialwissenschaften
(gemäß der PromO vom 24. August 2010)

vorgelegt von
Jana Freundt
aus Bochum, Deutschland

Hamburg, den 15. Dezember 2017

Datum der Disputation:	24.01.2018
Vorsitz:	Prof. Dr. Gerd Mühlheusser
Erstgutachter:	Prof. Dr. Andreas Lange
Zweitgutachterin:	Prof. Dr. Dr. Lydia Mechtenberg

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.0.1 Social Preferences in Behavioral Economics

In the 1980s behavioral economists increasingly began to study social preferences.[1] An early example are Kahneman et al. (1986a) and Kahneman et al. (1986b) who conducted surveys to learn people's fairness views on various microeconomic behaviors. The increasing interest in pro-social behavior was to a large extend motivated by the first evidence from experiments that elicited pro-social motivations. In the ultimatum game Gueth et al. (1982), players usually reject low offers and they rarely offer zero. In dictator games (Forsythe et al., 1994), participants also make positive offers but at the same time one observes more zero giving than in the ultimatum game, indicating that part of the positive offers are driven by beliefs about acceptance rates among responders. Trust games confirm the above intuitions (Berg et al., 1995). Similarly, experiments on social dilemma such as the prisonner's dilemma (Cooper et al., 1996) and public good games (Dawes and Thaler, 1988) show that a substantial fraction of participants cooperate and contribute to common goods, thereby contradicting the Nash equilibrium predictions of economic models.

This early experimental evidence and a large body of later studies including extensions as well as replications of the early findings have led to several theoretical models that attempt to describe the observed behavioral regularities. They represent cases of *social* or *other-regarding* preferences, meaning that an individual's utility varies with own and others' payoff. They can broadly be classified as distributional preferences, or outcome-based preferences, and belief-based preferences, which includes especially (but not only) models of intention-based preferences. Fehr and Schmidt (1999), Bolton and Ockenfels (2000) and Bolton (1991) propose models of inequality aversion, stating that people dislike inequalities between players and especially so if they are to their own disadvantage. The model of social welfare preferences by Charness and Rabin (2002) describes a combination

[1]However, already Becker (2010) introduced "taste-based discrimination" describing a desire to reduce the payoffs of outgroup members and in 1981, Becker introduced the concept of altruism within the family.

of preferences over social welfare and so-called maxi-min preferences, denoting a special concern for the payout of the least well-of person (or, in other words, inequality aversion w.r.t. the least well of player). In principle, most of these models explaining pro-social behavior in social interactions assume some form of reference-dependence, but they differ in what constitutes the reference group or the reference payoff. Parallel, another stream of literature modelled pro-social concerns as intention-based preferences, e.g., Rabin (1993), Falk and Fischbacher (2006), Dufwenberg and Kirchsteiger (2004), Charness and Rabin (2002), Battigalli and Dufwenberg (2009), incorporating the idea that an individual's kindness depends on her perceived kindness of the other person.

More recently, these early seminal contributions have been extended in several directions, one of them being choices under risk and uncertainty. Introducing risk in two people's payoffs gives rise to social comparisons based on either ex-ante chances to win or on final outcomes after uncertainty has been resolved. Thus, distribution choices under uncertainty can be influenced by different fairness ideals and people have been found to be heterogeneous with respect to the valuation they assign to fairness in ex-ante chances or in the ex-post distribution of outcomes Bolton and Ockenfels (2000), Krawczyk and LeLec (2010), Brock et al. (2013), Cappelen et al. (2013). However, how to extend models of social preferences to risky environments greatly depends on the social decision environment. Especially, social comparisons as above are unlikely to play a role for, for example, individual donations to a charity. Models that describe motivations to give to charity (under certainty) often assume that individuals derive utility from their contribution to the charitable cause (i.e. altruistic motives) or from the act of giving, as expressed by the concept of impure altruism (Andreoni, 1989, 1990).[2] How behavior in individual donation decisions and in public good contributions changes as a response to the introduction of risk in the decision environment has been less well investigated in the literature. For the case of public good games, (Gangadharan and Nemes, 2009, e.g.) compare contributions to public goods with certain or risky returns to the contributors and find that the latter diminishes giving. My research in chapters 3 and 4 extends the experimental literature in this field for public good contributions and charitable donations in uncertain environments.

Another stream of literature explored how social preferences play a role on markets. Already Fehr and Schmidt (1999) apply their model of inequity aversion to games with

[2]To what extend altruism or a feeling of "warm-glow" from the act of giving play a role not only in donations but also public good games is subject to debate in the literature (e.g. Palfrey and Prisbre (1997), (Bardsley and Moffatt, 2007), Anderson et al. (2011), Chaudhuri (2011)). Surely, it is not the only driver of positive contributions, but motives such as preferences for conditional cooperation play an important role (Bardsley and Moffatt, 2007) aswell.

proposer and responder competition. Recently, the popular experiment by Falk and Szech (2013) stimulated a body of experimental research on the importance of positive or negative externalities of uninvolved third parties in market interactions (e.g. Bartling et al. (2015), Bartling et al. (2015)). The finding by Falk and Szech (2013) that competitive market interactions erode moral values could not be replicated by subsequent studies, thus the evidence on how competition on markets affects pro-social behavior remains an open question. In chapter 5, we extend this literature to investors' behavior on financial markets.

The notion of *procedural fairness* has been adopted by behavioral economists to describe a concern for equality of ex-ante expected values in pro-social decisions (Fudenberg and Levine, 2012, Saito, 2013, Brock et al., 2013, Freundt and Lange, 2017), however, it has also been used in a more general way to describe a concern for fair procedures or mechanisms. In this field of literature, a ”fair” procedure usually denotes an unbiased and inclusive procedure to determine rules ad outcomes. Experimental evidence for the idea that inclusive democratic procedure might foster cooperative behavior in a society has been provided by experiments comparing behavior under exogeneously versus endogenously implemented institutions (e.g. Dal Bó et al., 2010, Tyran and Feld, 2006, Ertan et al., 2009, Sutter et al., 2010). Related, a couple of experiments find evidence in favor of the idea that people directly value decision rights, i.e. that people value the possibility to participate in a decision procedure (e.g., Bonin et al., 1993; Bardhan, 2000 and Bartling et al., 2014). In chapter 2, we establish that voluntary compliance to elected rules is significantly higher if the rules have been implemented by a fair, i.e. unbiased an inclusive, procedure.

Lastly, the same chapter directly builds up on previous experimental literature that established evidence on heterogeneity in fairness ideals. In particular, Cappelen et al. (2007) and Almås et al. (2017) (among others) have shown that people systematically differ in their judgments regarding whether income received through luck should be redistributed. In the experiment in chapter 2, we study behavior under rules that promote egalitarian values and rules that promote libertarian values.

1.0.2 Summaries of Chapters 2 to 5 and their Contributions to the Literature

1.0.2.0.1 Chapter 2: Rule Compliance. Ideally, democratic elections not only aggregate individual preferences but they can also confer legitimacy on elected outcomes. This can imply important efficiency gains for democratic societies as they would have to engage less in punishing people who do not follow rules like, for example, tax laws. In chapter 2, we investigate how undermining democratic voting procedures affects people's compliance with rules prescribing whether to redistribute. More precisely, we use an online experiment to establish a causal effect of electoral malpractice in a referendum on compliance with rules that were implemented by this referendum. The rules are codes of conduct that ask people to share or not to share their income with unlucky agents in modified dictator games. Treatments vary between subjects whether a rule has been selected by a democratic procedure to aggregate votes (i.e., by majority vote), or by a manipulating voting procedure where the majority vote is undermined either by the possibility to accept bribes, by introducing a fee for voting, or by excluding poor voters. The results show that a democratically selected rule shifts behavior towards the action prescribed by the rule because a significant fraction of people follow it even if they preferred to opposite action in an individual decision. We find that the decision to follow a pro-social rule (prescribing to redistribute one's income) is not driven by a desire to follow others. Rather, electoral malpractice has a strong and significant adverse effect on individuals' intrinsic motivation to comply with a pro-social rule. In particular, a subject is less likely to follow the rule when voter manipulation prevents her from casting her vote or when it leads her to believe that the outcome of the referendum will be biased compared to a fair majority vote.

Taken together, our study complements existing research in economics on the positive behavioral effects of democratic institutions by showing that such effects may be sensitive to attempts of electoral manipulation. We contribute evidence on whether procedural changes in how an election is conducted affect the *intrinsic* motivation of subjects to follow rules.

1.0.2.0.2 Chapters 3 and 4: Impure Public Goods and Giving under Risk. In the first studies of my PhD research (chapters 3 and 4) I investigate people's willingness to privately provide public goods with uncertain payoffs. The experiments aim at informing social investments, such as crowdinvestments or microlending, that have become increasingly popular means to provide (impure) public goods, as for example clean energy, in recent years. Social investments are characterized by a frequent simultaneous presence of risk in private and public returns from investments.

The study in chapter 3 investigates to what extend the risk inherent in social investments influences their attractiveness for investors. We analyze how risk in the provision of the public benefit generated by an investment and in the financial return to the investor each affect investment decisions and how, in addition, their correlation influences investments when both risks are simultaneously present. By identifying treatment effects for subgroups of participants who are more (less) risk-averse and who show greater (lower) pro-social concern compared to their peers in the experimental sample, we establish important behavioral heterogeneities in investors' behavior.

In the study in chapter 4, we investigate the impact of risk in a strategic setting and use a public good game experiment to isolate the impact of risk in returns from public good contributions. We separate the return of a subject's contribution to herself vs. the return to other group members that her contribution generates and introduce risk one-by-one in each dimension. Individuals' contributions particularly respond to the downside risk of investments, suggesting that participants do not take the whole distribution of risk into account and that the correlation of the coexistent risks matters to the extent that it affects the overall downside risk. This finding indicates the importance of attenuating the risk of complete failure for attracting investments in environmental protection projects like abatement of emissions.

With this, our findings in chapters 3 and 4 support—in the context of impure public goods—models that assume utility being driven by the *success* in giving, rather than the act of giving alone, extending evidence gathered by, e.g., Anderson et al. (2011), Goeree et al. (2002) and Palfrey and Prisbre (1997) to risky situations.

1.0.2.0.3 Chapter 5: Green Assets on Financial Markets. The previous two studies analyze people's motivations to invest in bundled goods that may generate risky public and private benefits, which can be applied, among others, to explaining the demand for socially responsible investments. In the project of chapter 5, we investigate the question how socially responsible investments perform in a competitive market with conventional investment opportunities. We thus set up experimental asset markets for 'green', i.e. socially responsible, and non-green stocks to identify the impact of introducing green investment opportunities on market developments, such as prices and trading volumes. On the aggregate, we cannot identify a price premium for the green asset. In contrast, our results show that introducing an asset with a costly positive externality increases the market value for conventional assets.

The results add to mixed previous findings in empirical and experimental studies on the financial performance of SRI. Our study furthermore complements previous work on pro-social behavior on markets by testing to to what extend pro-social behavior on a financial market, where an individual action does not impact the final provision of the public good, correlates with generosity in individual donations.

Chapter 2

Manipulated Votes and Rule Compliance[1]

Abstract

Allegations of voter fraud accompany many real-world elections. How does electoral malpractice affect the acceptance of elected institutions? Using an online experiment in which people distribute income according to majority-elected rules, we show that those who experience vote buying or voter disenfranchisement during the election are subsequently less likely to comply with a rule. On average, the detrimental impact of electoral malpractice on compliance is of the same magnitude as removing the election altogether and imposing a rule exogenously. Our experiment shows how corrupting democratic processes can impact economic behavior and sheds light on the psychological mechanisms underlying "rule legitimacy".

JEL Codes: D02, D72, D91, C92

Keywords: rule compliance, endogenous institutions, corruption, procedural fairness, legitimacy

[1]This chapter is co-authored by Arno Apffelstaedt (University of Cologne).

2.1 Introduction

People follow rules for different reasons. One reason is the existence of incentive and deterrence mechanisms such as implicit or explicit rewards for compliance, or punishment for non-compliance. Another reason, stressed by legal scholars and political scientists, comes from people accepting the procedure by which the rule came into force as *legitimate*: When the rule setting procedure is seen as being fair, people may change their behavior and follow the rule "*voluntarily out of obligation rather than out of the fear of punishment or anticipation of reward*" (Tyler, 2006, p.375). This paper is about such latter type of rule compliance.

An important source of rule legitimacy are thought to be democratic voting procedures. Consider, for instance, the introduction of a CO_2 tax or a policy that changes the rules of organ donation from an 'opt-in' to an 'opt-out' organ donation system. Intuition suggests that such policies will see higher acceptance and will be voluntarily complied with to a larger extent if people perceive the rule setting mechanism to be participatory and inclusive. Indeed, all but a handful of countries in today's world hold elections or referenda of some kind—often in an attempt to confer legitimacy on a policy addressing some critical political issue on which the electorate is divided (LeDuc et al., 2014).[2] The extent to how well democratic procedures are implemented, however, varies widely. In many countries, promises of a "free and fair" vote are openly undermined by practices ranging from systematic vote buying to the outright exclusion of social groups, often minorities or poor voters. In other instances, unintentional disenfranchisement or alleged manipulation of parts of the electorate leads people to question the integrity of elections and referenda.[3]

When many people perceive a voting procedure to be "corrupt" or "flawed", legitimacy of the elected outcome may suffer, possibly leading citizens to show lower compliance with elected rules and policies. Suggestive evidence for this claim can be found in survey data, see Figure 2.1: In countries with higher perceived levels of electoral malpractice (X-axis), the average citizen is significantly more likely to say that it is justifiable to break social rules (Y-axis), ranging from wrongfully claiming government benefits to not paying the

[2]Recent examples include the 2016 Brexit referendum, the 2017 constitutional referendum in Turkey, and the 2019 referendum in Romania about whether to prohibit amnesties and pardons for corruption offenses.

[3]Brusco et al. (2004) and Gonzalez-Ocantos et al. (2012) document vote buying schemes in Argentina (2002) and Nicaragua (2008), respectively. Enikolopov et al. (2013) presents data on the extent of electoral fraud during the Russian parliamentary elections of 2011. In the UK and the US, allegations of voter fraud have recently been extensively discussed in the popular press (Cottrell et al., 2018; UK Electoral Commision, 2018). Both, actual instances of electoral malpractice as well as allegations thereof—even if entirely unfounded—can lead voters to question the integrity of elections (Norris, 2014).

fare on public transport.

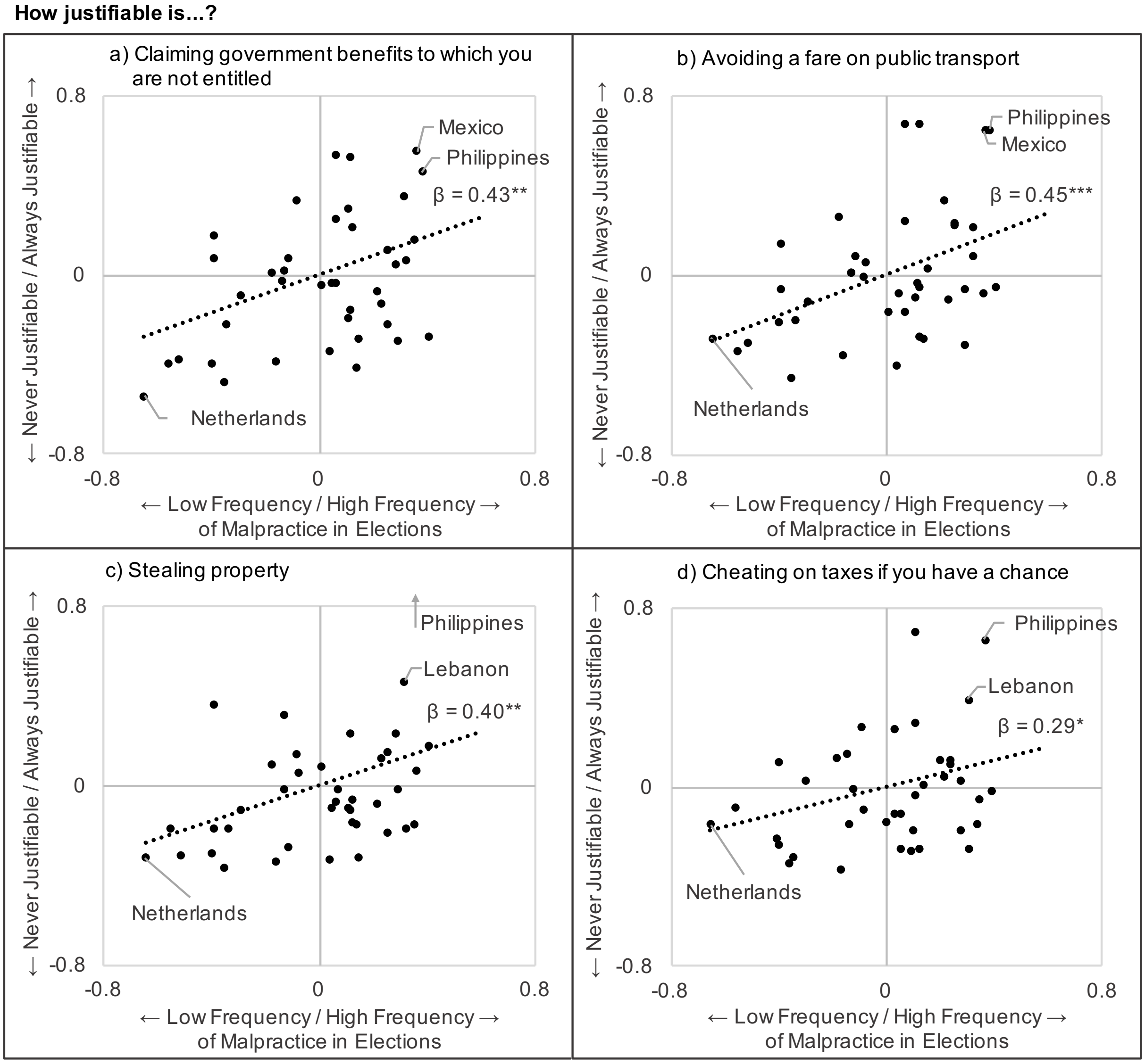


Figure 2.1: Country-level correlations between citizens' perceived frequency of malpractice in elections and their statements about the justifiability of violating rules and laws (Country averages calculated from the WVS (2014)). Y-axis: Average answers in a country to questions V198-V201 ("*How justifiable is...?*"). X-axis: Index of perceived malpractice in elections, calculated from average answers in a country to questions V228 B,C,D,G, and H ("*How often do the following things occur in your country? B: Opposition candidates are prevented from running, C: TV news favor the governing party, D: Voters are bribed, G: Rich people buy elections, H: Voters are threatened with violence at the polls.*"). Data is normalized to show relative deviations from the average across all countries. Univariate OLS regressions without intercept: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Causal evidence on whether voting procedures directly affect behavior can be gathered through experiments. However, no causal evidence exists on whether the power of democracy to change behavior is sensitive to electoral malpractice such as vote buying and voter disenfranchisement. In this paper, we investigate this question using a novel experiment. Modelling a typical referendum situation in which the electorate is split between two competing policies, our experiment allows us to systematically investigate whether and due to which psychological mechanisms the willingness to accept and comply with democratically elected rules may suffer when people perceive the voting procedure to be “corrupt”.

The key take away of our study is that the power of an election to increase the acceptance of a rule can be significantly reduced when democratic voting procedures are tempered with. In fact, we find that electoral malpractice can wipe out the entire democracy premium, meaning that rule compliance after a “corrupt” election is equivalent to imposing the rule exogenously. Voluntary compliance decreases especially among people who are personally excluded from the ballot as well as among people who believe the voting outcome to be biased. This does not mean, however, that malpractice always reduces rule compliance: In our experiment, if people follow a rule for reasons other than its perceived legitimacy (i.e., if there does not exist a democracy premium in the first place), electoral malpractice leaves behavior unaffected.

The setting of our experiment is as follows. In each session, 100 subjects have to decide, each individually, whether to share one’s experimental income with another subject who is less well of. Before subjects make that decision, they vote on whether to introduce a policy that asks everyone in the session to voluntarily share (*Rule:Give*) or to introduce a policy that asks everyone to not share (*Rule:Don’t*). We measure the strength of the elected rule by its power to convince people to change their behavior relative to a setting without a rule. In the baseline treatment, a majority vote among all 100 subjects selects the rule. With three further treatments, we measure the causal effect of electoral malpractice on rule compliance: In one treatment, we demand that subjects pay for their vote, excluding everyone from the ballot who does not pay. In another, we manipulate votes by paying subjects for reversing their initial vote. In a third, we exclude subjects with a low household income from the ballot.

Our main result is that electoral malpractice drastically reduces the power of democracy to convince people to follow *Rule:Give*, but does not affect the power of *Rule:Don’t*. In the baseline treatment, the election of *Rule:Give* has the power to decrease non-giving rates by more than 60%. Malpractice reduces this power by half ($p < .01$). With the help of an additional (fifth) treatment, we show that this “malpractice effect” on compliance with *Rule:Give* is equivalent to the effect of removing the election altogether and imposing the rule exogenously. In other words, electoral malpractice wipes out the democracy premium entirely. For *Rule:Don’t*, on the other hand, we find that there is neither a malpractice effect nor a democracy premium: Across all treatments, the power of *Rule:Don’t* is strong yet constant (its implementation decreases giving rates by roughly 50%).

To shed light on the psychological reasons why individual behavior may respond to how a rule has been selected, we study two possible mechanisms. We first explore the role of beliefs about how other subjects behave under the same rule. The rationale is as follows. Notice that an election in which parts of the electorate have been excluded (or when their votes have been manipulated) leads to a noisier signal of the modal policy preferences in the population than an unbiased majority vote. Hence, if people care to align their behavior with what others do or value, then a "corrupt" voting procedure may lead to a weaker response to the election result, and thus, to lower individual rule compliance.[4] We explore this idea in our experimental framework by analyzing elicited beliefs about the behavior of others across treatments. Using an exogenous shock to these beliefs, we measure their causal effect on the willingness to comply. While we do not find evidence for this idea when studying behavior under *Rule:Give*, compliance with *Rule:Don't*—the rule for which we do *not* find an effect of malpractice–is to a large extent driven by preferences for following the behavior of others.

The second mechanism we investigate are "intrinsic" concerns about the fairness of the voting procedure. In particular, we study whether the effect of electoral malpractice on compliance is associated with (1) subjects who have been personally excluded from taking part in the election and (2) subjects who believe that the voting outcome is biased.[5] Indeed, we find that roughly 80% of the treatment variance under *Rule:Give* is captured with these two variables, suggesting that people intrinsically care about personal participation as well as about the overall unbiasedness of the procedure.

Our experiment is conducted online with subjects from different countries and demographic backgrounds. Using a post-experimental questionnaire, this variance allows us to investigate how treatment effects relate to standpoints on various political issues such as redistribution, corruption, democratic values, and trust in institutions. We find that treatment effects are more significant and of larger magnitude among subjects who live in (relatively) democratic countries and among those who self-report to have stronger concerns for democratic values. This finding indicates that the effect of malpractice we identify in our experimental game relates to psychological domains that are also relevant in corresponding real-world decision making. Moreover, it corroborates our analysis of mechanisms in showing that it is indeed people with a preference for democratic elections

[4]While there are no monetary coordination incentives in our experiment, it is reasonable to assume that some subjects may nonetheless care about aligning their behavior with what others do or value (see, for instance, Bernheim, 1994; Bénabou and Tirole, 2012; Krupka and Weber, 2013).

[5]Measured as the belief about the difference between the share of votes for a given rule with and without malpractice.

who show negative reactions to electoral manipulation.

2.1.0.0.1 Related Literature. To our knowledge, this paper is the first to provide causal evidence for the negative effects of electoral malpractice on the acceptance of elected institutions. We complement earlier research in public and political economics that has provided evidence for the positive effects of democratic compared to exogenously imposed institutions. For instance, Frey (1997) shows that tax compliance is higher in Swiss cantons that see more democratic participation. Subsequent experiments, for example by Tyran and Feld (2006), Ertan et al. (2009), Sutter et al. (2010), Grossman and Baldassarri (2012), and Dal Bó et al. (2010), have shown that in social dilemma situations, punishments and rewards work better when endogenously elected rather than exogenously imposed.[6] Note that these experiments compare cooperation rates under an endogenously elected versus an exogenously selected institution instead of directly measuring individual rule compliance as we do. Overall, the existing literature suggests that giving citizens decision rights through majority votes can bring important efficiency gains to societies. We show that for such efficiency gains to materialize it matters *how* these institutions are introduced. More specifically, we provide evidence that the positive dividend of democracy is sensitive to interventions in the voting procedure that disenfranchise or manipulate voters. Because our design allows us to isolate and study the effect of endogenous institutions on the *intrinsic* component of preferences better than earlier studies, we also generate new insights into the psychological mechanisms driving democracy effects.

Probably closest to the aim of our study, Dickson et al. (2015) experimentally show that people are more willing to actively help (and less willing to actively hinder) the punishment authority in a public good game if this authority has been elected by a majority vote rather than exogenously imposed. They interpret their finding as showing differences in the perceived legitimacy of the authority.[7] We study the (indirect) behavioral consequences of legitimate procedures that can affect the efficiency of the working of institutions rather than direct expressions of support for an authority.

With this, we add to a different stream of research in psychology and behavioral economics suggesting that procedural aspects of decision making can affect behavior. In

[6]This list of studies is not meant to be exhaustive. See, e.g., Dal Bó (2014) for further studies. A related literature in organizational economics studies the value of "democratic" compared to "autocratic" decision-making mechanisms within firms and organizations. For example, Bonin et al. (1993), Black and Lynch (2001) and Zwick (2004) provide empirical support that higher levels of employee participation are associated with increased worker productivity, leading to potentially large efficiency gains. Similarly, Fehr et al. (2013) show that giving away decision-rights leads to an under-provision of working effort.

[7]A similar approach is followed by Berman et al. (2014). Here, the authors measure the effect of an election fraud intervention in the field on multiple survey measures of attitudes toward government, including the willingness to report insurgent behaviors to security forces.

particular, studies have shown that people seem to care about the "fairness" of decision-making processes in a more general sense (see, e.g., Tyler, 1990; Frey et al., 2004; Cappelen et al., 2013) as well as about personally partaking in them (see, e.g., Bonin et al., 1993; Bardhan, 2000; Bartling et al., 2014). The idea that procedural concerns may lower the normative appeal of elected rules and thus directly affect the willingness of people to comply is also related to theories of "legitimate authority" (Weber, 1978; Tyler, 2006; Akerlof, 2017). Supporting this view, Besley et al. (2015) find that a change in property taxes in the UK—which was perceived as highly unfair by the public—led to an increase in tax evasion. The authors suggest to attribute this increase to a shock in intrinsic motivation; they cannot, however, pin down the exact motives.

The remainder of the paper is structured as follows. Section 2.2 explains the experimental design in detail. Sections 2.3 presents our results: We first estimate the effects of malpractice on rule compliance (section 2.3.2) and then study the behavioral mechanisms that drive these effects (section 2.3.3). Our findings are discussed in section 2.4, before we conclude in section 3.4. Screenshots of the experimental instructions and the questionnaire can be found in the appendix.

2.2 Experimental Design

The main prediction guiding the design of our experiment and our analysis is as follows:

Malpractice Effect: Electoral malpractice lowers voluntary compliance with the elected policy:

$$E(Compliance \,|\, Malpractice = 1) < E(Compliance \,|\, Malpractice = 0).$$

Our goal is to design an experiment which can (1) identify a causal effect of malpractice on compliance and that (2) can shed light on the psychological mechanisms driving this effect. Satisfying this goal comes with different requirements for our design.

First, we want to make sure that the effect we measure is a general malpractice effect and not a feature of a specific malpractice intervention. Our experiment for that reason implements three malpractice treatments in order to mimic the variation in corruptive practices in the real world. With this, we can robustly test the hypothesis that—independent of the particular practice—compliance with elected policies will decrease if democratic principles are violated.

Second, in order to identify a causal effect of malpractice, we need to control for possibly unbalanced treatment groups as different people may have different inclinations *ex-ante* to prefer and therefore follow a given policy (see Dal Bó et al., 2010; Dal Bó et al., 2019). By eliciting individual giving choices and votes *before* the introduction of each treatment, we are able to control for different distributions of types across treatment groups.

Third, we aim to set up an environment in which people disagree about what is the "right" thing to do and therefore vote for different policies. We achieve this by letting subjects vote on policies in the environment of a (binary) dictator game. Numerous studies show that people differ in their judgements regarding whether income received through luck should be redistributed or not (see, e.g., Cappelen et al., 2007; Almås et al., 2017). Our design allows us to measure the power of elected rules to change behavior away from what people ex-ante preferred as an action or as a policy.

Finally, there are a few design elements that we require in order to study the psychological mechanism driving behavior. To make sure that we measure *voluntary* compliance with elected policies, we do not implement any form of punishment or reward for certain behavior. That is, subjects are free to choose to follow (or not follow) the elected rule without having to fear any monetary consequences. Because there are no classical coordination incentives in a one-shot dictator game (and no reputation effects),[8] voting mechanisms can then only work by their normative appeal.[9] As outlined in the introduction, the normative appeal of a democratic election may be due to (a) people having intrinsic preferences for a fair and unbiased election procedure or due to (b) the election producing a good signal about what other people do and value, making it easier to "do what others do". To be able to shed light on these two mechanisms, we elicit subjects' beliefs about what other subjects do and introduce an exogenous shock to these beliefs in order to estimate a causal effect of these beliefs on behavior.

[8] From the perspective of standard game theory, treatment effects cannot be driven by people adjusting their behavior to "equilibrium effects" (as, for example, in Dal Bó et al., 2017).

[9] The possibility to construct a well-defined *behavioral* measure of voluntary compliance is a major advantage of using an experiment. With surveys, researchers have to rely on the self-reported willingness to comply (see, e.g., Berman et al., 2014), while in field studies, deterrence mechanisms often interfere with clean measures of intrinsic motives (see, e.g., Fjeldstad et al., 2018). Being able to measure voluntary compliance is not merely a technically desirable feature. In many cases where deterrence mechanisms are in place in the real world, expected punishments are usually not high enough to explain the high levels of compliance observed (e.g., in the case of tax compliance, see Feld and Tyran, 2002, p.88). In some instances deterrence might not be feasible to implement, as for the case of littering. In situations with deterrence mechanisms, it is important to understand the intrinsic component of rule compliance in order to properly isolate and understand the effect of punishments on compliance (see, e.g., Dwenger et al., 2016).

2.2.1 The Experiment in Detail

For each session, 100 individual subjects are recruited on the online platform Prolific.ac with a small, fixed base payment and the prospect of a lottery that has one of them wining GBP 100.[10] The lottery is used to naturally form voting groups and to construct an experimental currency for the redistribution choices.

2.2.1.0.1 Redistribution Choices. The following situation provides the background for our referendum: There are 100 subjects, each of them aware of the presence of the 99 others. Subjects are informed that one of them will be drawn to receive a cash prize of GBP 100. However, lottery tickets will be distributed unequally: While 50 randomly chosen participants (called "*receivers*") will get 10 lottery tickets each, the remaining 50 participants ("*nonreceivers*") will get no tickets, and thus have no chance to win the prize. Subjects are then told that they will not learn until after the experiment whether they have been chosen to be a receiver or nonreceiver—however, they can now decide whether they would like to conditionally redistribute. Specifically, we use the unequal distribution of lottery tickets to construct a binary dictator game with role uncertainty: Each subject is asked to decide whether—in case of being a receiver—she wants to $Give_i \in \{0, 1\}$ three out of her ten lottery tickets to a randomly selected non-receiver. The subject is informed that in case of being a receiver (50% probability), her decision is automatically implemented and determines the number of lottery tickets for herself and for one random other person. She is also informed that in case of being a nonreceiver (50% probability), her decision does not play a role for the distribution of lottery tickets. The question of whether or not one *should* redistribute forms the basis for our referendum.

In each experimental session, subjects face the redistribution choice twice. Participants are informed that there will be two rounds but learn about the details of round 2 only after having completed round 1. After the session, one round is randomly drawn to determine the final distribution of lottery tickets among subjects. One lottery ticket is then drawn uniform randomly and the holder of this ticket awarded the cash prize of GBP 100. All decisions are taken anonymously and in private. The timeline of a session is summarized in Figure 2.2.

2.2.1.0.2 Round 1. Round 1 consists of two stages: A choice stage and and information stage.

Choice stage. Round 1 implements the redistribution choice without a rule for behavior

[10] For details on recruitment see paragraph Implementation below. For demographics of the Prolific.ac subject pool, see `https://www.prolific.ac/demographics`.

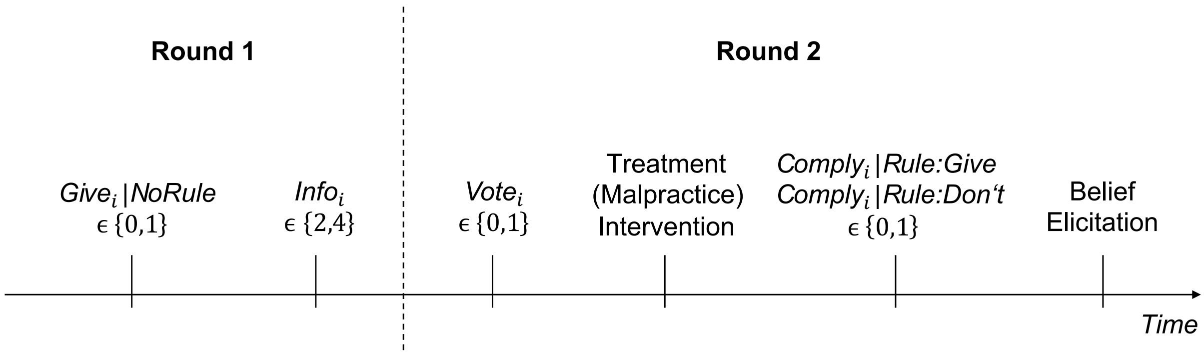


Figure 2.2: Timeline of experimental session

being in place. Each subject decides individually whether to give, $(Give_i|NoRule) = 1$, or not give, $(Give_i|NoRule) = 0$. To ease notation, we introduce the following definition:

Definition 1 (Givers and Non-Givers)**.** *If* $(Give_i|NoRule) = 1$, *we call individual* i *a* Giver*. If* $(Give_i|NoRule) = 0$, *we call individual* i *a* Non-Giver*.*

Information stage. After a subject has made her choice in round 1, she is presented with a screen that shows her information on how five other people in "*an earlier study*" (participants in our pilot sessions) decided in the exact same situation. With probability one half we show the subject a pre-selected sample that features two Givers and three Non-Givers ($info_i = 2$), and with probability one half a sample that features four Givers and one Non-Giver ($info_i = 4$). We introduce $info_i \in \{2, 4\}$ in order to generate exogenous variance to the beliefs of a subject about how other subjects will behave in round 2 of the dictator game.

2.2.1.0.3 Round 2. Round 2 consists of four stages: A voting stage, a treatment stage, a rule compliance stage, and a belief elicitation stage.

Voting Stage. At the beginning of round 2, subjects are informed that they will shortly have to make the redistribution choice $Give_i \in \{0, 1\}$ again. They are also informed that in this round, a "*code of conduct*" will be implemented for all participants. Each subject is then asked to vote for the code that she "*prefers to have implemented as the code of conduct for all participants.*" The subject can cast her vote either for *Rule:Give* ("*everybody should choose Give*") or for *Rule:Don't* ("*everybody should choose Don't Give*"). All participants of a lottery decide in one large voting group of 100 subjects on the rule they prefer to have implemented for everyone. With this, subjects are very unlikely to cast a pivotal vote (which would potentially lead to strategic voting considerations) and our results are thus scalable to larger societies. The decision of the subject in the voting stage is coded $Vote_i \in \{Rule{:}Give, Rule{:}Don't\}$. Subjects are not informed about how other participants voted until after the experiment.

Treatment Stage. Treatments are introduced after the voting stage. We employ a between-subject design making it difficult for subjects to infer our research question. There are four treatments, see Table 2.1.

Treatment	Malpractice?	Description	n
T_Baseline	No	Standard majority vote	100
T_Pay4Vote	Yes	Subjects have to pay GBP 0.20 to make vote count	100
T_MoneyOffer	Yes	Subjects are offered GBP 0.20 to reverse their vote	100
T_ExcludePoor	Yes	Only the votes of subjects with annual household income > GBP 40K are counted in the referendum	100

Table 2.1: Overview of Treatments

In the baseline treatment (*T_Baseline*), the rule is selected by simple majority vote among all 100 participants. After a subject has submitted her vote, she is informed that "*the rule that receives more votes in total will be implemented as the code of conduct.*" In treatment *T_Pay4Vote*, subjects learn that "*only the votes of participants who pay GBP 0.20 will be counted.*" Each subject can decide whether or not to pay. If a subject decides to pay, her vote is counted toward the majority vote; otherwise, her vote is not counted. In *T_MoneyOffer*, subjects learn that "*all participants are offered an extra payment of GBP 0.20 to vote for the rule that is opposite to what they originally wanted to vote for.*" Each subject can decide whether or not to accept the offer. If a subject decides to accept, her vote is reversed and counts for the opposite rule. Otherwise, her original vote is counted. In *T_ExcludePoor*, subjects are informed that "*only the votes of participants with a household income above GBP 40,000 are counted.*" Each subject learns whether her individual vote has not been counted toward the majority vote.[11] In all treatments, participants know that everyone in their session is subject to the same voting mechanism. They are not informed, however, about the number of participants who decide to pay the fee in *T_Pay4Vote*, about the number of participants who accept the bonus payment in *T_MoneyOffer*, or about the number of participants whose votes are excluded due to their household income in *T_ExcludePoor*.

Rule Compliance Stage. After the treatment stage, subjects make the redistribution choice a second time. Each subject decides whether she wants to $(Give_i|\textit{Rule:Give}) \in \{0,1\}$ conditional on *Rule:Give* being elected and whether she wants to $(Give_i|\textit{Rule:Don't}) \in \{0,1\}$ conditional on *Rule:Don't* being elected. Thus, all subjects make the decision

[11] To identify a subject as having a household income above or (weakly) below GBP 40,000, we use self-declared information provided by Prolific.ac.

whether or not to follow *each* rule conditional on it being elected.[12] These two choices form our measure of rule compliance:

Definition 2 (Rule Compliance)**.** *We say that a subject complies with Rule:Give,* $(Comply_i|Rule{:}Give) = 1$*, if and only if* $(Give_i|Rule{:}Give) = 1$*. We say that a subject complies with Rule:Don't,* $(Comply_i|Rule{:}Don't) = 1$*, if and only if* $(Give_i|Rule{:}Don't) = 0$*.*

Belief Elicitation Stage. At the end of round 2, we ask participants to state their beliefs about how many of the other 99 participants in their treatment (a) voted for *Rule:Give*, (b) decided to comply with *Rule:Give*, and (c) decided to comply with *Rule:Don't*. Subjects give their answer by indicating a bracket in the set [(0-9), (10-19),..., (90-99)], following Schlag and Tremewan (2016). In order to incentivize agents to state their true empirical expectations, a GBP 0.50 bonus payment is awarded for each correct answer.[13] In *T_Pay4Vote*, *T_MoneyOffer* and *T_ExcludePoor*, we additionally elicit beliefs about the impact of the intervention on final voting outcomes. In *T_Pay4Vote* we ask participants to guess (d) what share of *Rule:Give*-voters in their session were willing to pay for their vote, and (e) what share of *Rule:Don't*-voters in their session were willing to pay. We do the same regarding the share of *Rule:Give*-voters (*Rule:Don't*-voters) who accept the monetary offer in In *T_MoneyOffer*. Finally, in *T_ExcludePoor*, we ask subjects to guess the share of votes for *Rule:Give* separately among high income (income > GBP 40,000) and low income participants (income $\leq$ GBP 40,000).

2.2.1.0.4 Post-Experimental Questionnaire. In a post-experimental questionnaire, we ask participants about their experience with and attitudes toward, e.g., redistribution, corruption and democratic institutions. Most of the questions in this part are either directly taken or adapted from questions featuring in the 6th wave of the World Value Survey (WVS, 2014). We also collect data on personality characteristics such as risk preferences (self-reported and hypothetical lottery choice), trust, and the Big Five personality traits (using the question format in Gosling et al. (2003)). The questionnaire was posted on Prolific.ac as an unrelated survey using a different visual design and researcher profile no

[12]Eliciting such state-dependent compliance choices has major advantages for us: There is no selection into *Rule:Give* or *Rule:Don't* and the decision whether to give under each rule is made without yet knowing the voting outcome. The latter is important for eliciting beliefs at the end of the experiment. Importantly, having large voter groups of 100 subjects—which we prioritize for external validity and scalability of the results—makes a real-time matching of all votes a practical problem that we avoid with this design choice.

[13]Simply put, the subject is asked to guess (up to a certain precision) an empirical frequency that is observed by the experimenter. A prize is then awarded if and only if her guess coincides with the realized frequency. Schlag and Tremewan (2016) show that this method is not only easy to implement, but also particularly robust: Inference does not require postulating any assumptions on the utility function beyond assuming that the subject strictly prefers the prize.

earlier than two weeks after a subject had participated in the experiment. These measures are meant to minimize the risk of spillovers from decisions in the experiment and especially from exposure to the different treatments to questionnaire answers. Only subjects who participated in our experiment were able to enter the survey. The follow-up-rate is close to 100 percent.[14] The full list of questions can be found in the appendix.

2.2.1.0.5 Implementation. The experiment was implemented in February and March 2017 online using a subject pool of international participants on the platform Prolific.ac based in Oxford, UK. Our population sample differs in several respects from the typical subject pool at Western university labs: The mean age is 31, almost two thirds of the participants are not students (64%), and about one third have a non-Western nationality (32%). We programmed the experiment using the software *LimeSurvey* (Schmitz et al., 2012). Detailed instructions and screenshots can be found in the appendix. To ensure understanding and common knowledge thereof, control questions at the end of each screen had to be answered correctly in order to proceed with the experiment. Registered participants on Prolific.ac have a unique ID that is used to identify subjects, to prevent repeated participation and to process payments. In addition, subjects' unique Prolific-ID allows us to access an extensive set of self-reported socio-demographic data, including gender, nationality and income. Everyone who filled out information on at least gender, nationality and country of birth was eligible to participate.[15] When selecting into the experiment, *all* subjects see that they will take part in a lottery that pays GBP 100 to one out of 100 participants and that they will receive a fixed base payment of GBP 1.30 for completing the study which takes roughly 15 minutes to complete.[16] Additional payments are announced during the course of the experiment. Subjects receive all payments and an e-mail with a summary of all outcomes through the online survey platform Prolific.ac within two days after the experiment. For completing the 10 minute post-experimental questionnaire, subjects receive a compensation of GBP 1. Subjects had to give informed consent before they were able to enter the experiment and questionnaire, respectively.

[14]Of 400 subjects, 387 filled out the questionnaire, i.e. 96.75 percent.

[15]In treatment *T_ExludePoor* we additionally required that participants had filled out information on household income.

[16]In the case of *T_Pay4Vote*, we increase the base payment by GBP 0.20 to counter adverse wealth effects when subjects pay to make their vote count. This is only announced after they selected into the study; the base payment announced on the prolific website is the same across all treatments.

2.3 Results

Figure 2.3 nicely summarizes the key insight of our experiment. We find that the power of rules to change behavior can be strongly and significantly reduced by the presence of electoral malpractice. When implemented by a fair majority vote (*Baseline*), *Rule:Give* (left-hand side) has the power to decrease non-giving rates by more than 60% relative to the situation without a rule. Malpractice reduces this power by nearly half to roughly 30%. Malpractice has no effect on the power of *Rule:Don't* (right-hand side).

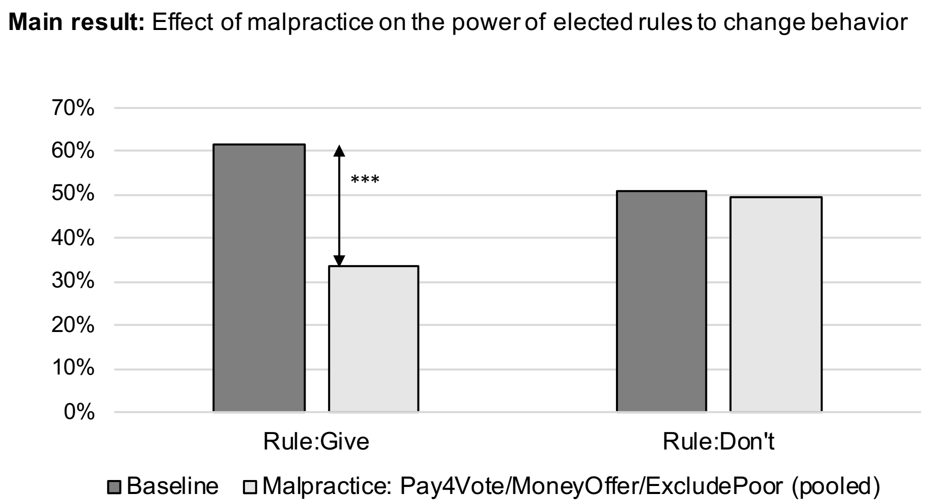


Figure 2.3: Effect of interventions *Pay4Vote*, *MoneyOffer* and *ExcludePoor* (pooled) on the power of elected rules to change behavior. Bars show decrease (in %) in the share of subjects choosing to not give (give) after the election of *Rule:Give* (*Rule:Don't*) relative to the share of subjects choosing to not give (give) in the absence of a rule. The graph is based on type-weighted averages, stars denote significance of population average treatment effect of *Malpractice* (Pooled) on compliance with *Rule:Give*, *** $p < .01$, see Table 2.3.

We thus confirm our prediction that malpractice in an election can substantially impact compliance decisions, however, only for one type of rule. Why do we observe such an asymmetry and who are the people whose compliance decisions are sensitive to the procedure that implements *Rule:Give*? In the remainder of this section we will provide the results that lead to the above general finding. First, sections 2.3.1 and 2.3.2.1 set the stage by summarizing giving and voting behavior as well as baseline rule compliance. In 2.3.2.2, we explain in detail how we compute treatment effects to then continue analysing the behavioral mechanisms driving these effects in 2.3.3.

2.3.1 Setting the Stage

We begin by providing summary statistics of how subjects behave in round 1, how they vote in round 2, and how interventions *Pay4Vote*, *MoneyOffer*, and *ExludePoor* affect the voting process. This information is summarized in Table 2.2.

	Base-line	*Pay 4Vote*	*Money Offer*	*Exclude Poor*
Round 1				
Share of subjects choosing...				
$(Give_i \mid NoRule) = 1$	.57	.57	.71	.60
Round 2				
Share of...				
initial votes cast = *Rule:Give*	.64	.75	.81	.71
if $(Give_i \mid NoRule) = 0$	.35	.47	.45	.38
if $(Give_i \mid NoRule) = 1$	.86	.97	.96	.93
subjects paying for vote		.65		
if $Vote_i$ = *Rule:Give*		.69		
if $Vote_i$ = *Rule:Don't*		.52		
subjects accepting money offer			.39	
if $Vote_i$ = *Rule:Give*			.31	
if $Vote_i$ = *Rule:Don't*			.74	
subjects excluded by income $\leq 40K$				.50
if $Vote_i$ = *Rule:Give*				.52
if $Vote_i$ = *Rule:Don't*				.45
final votes counted = *Rule:Give*	.64	.80	.70	.68
Measures of Election Bias				
Outcome_Bias[a]	0	.05	.11	.03
Lost_Votes[b]	0	.35	.39	.50
Observations	100	100	100	100

[a] |(Share of initial votes cast = *Rule:Give*) − (Share of final votes counted = *Rule:Give*)|
[b] $Lost\_Vote_i = 1$ if i does not pay for vote, accepts money offer, or has income $\leq 40K$

Table 2.2: Summary Statistics. Giving in round 1, voting behavior, and measures of election bias by treatment.

In the absence of a rule, subjects are roughly split between giving and non-giving: On average, 61% of subjects (245/400) choose to give in round 1 (row 1 of Table 2.2).[17] Voting behavior in round 2 (summarized in the second to fourth rows) strongly correlates with giving behavior in round 1: Among *Givers* ($(Give_i \mid NoRule) = 1$), an overwhelming

[17] While the specific set-up of our dictator game is atypical (role uncertainty, binary decisions, risky prospects with a small probability to win a high price, online participant pool), observed behavior in round 1 of our experiment does not deviate much from typical findings on dictator game behavior in the literature. For instance, in a meta-study of 129 dictator game studies covering 41,433 observations, Engel (2011, p.6) finds a share of 63.89% of subjects giving non-zero amounts.

majority (93% on average) vote for *Rule:Give*. Among *Non-Givers* ($(Give_i|NoRule) = 0$), *Rule:Don't* always receives more than half of the votes (59% on average). Overall, between 64% and 81% of the 100 subjects in a treatment group cast their vote for *Rule:Give*. As a result of the treatment interventions, a considerable share of votes are either not counted or reversed: 35% of participants in *T_Pay4Vote* refuse to pay a fee to make their vote count, 39% of participants in *T_MoneyOffer* are willing to reverse their vote in exchange for the small bonus payment, and, by design, 50% of voters are excluded due to a low household income in *T_ExcludePoor*, see the second to last row of Table 2.2. We introduce the variable $Lost\_Vote_i \in \{0, 1\}$ to identify a subject whose vote is either uncounted (*T_Pay4Vote* and *T_ExcludePoor*) or reversed (*T_MoneyOffer*) due to the intervention as one of our measures of election bias. Intuitively, excluding a substantial fraction of voters can affect the voting outcome. We measure *Outcome_Bias* as the (absolute) difference between the share of votes for *Rule:Give* before and after the intervention. While a large share of participants lose their vote, the effects on voting outcomes are relatively minor: *Outcome_Bias* ranges between three and eleven percentage points, see the third to last row of Table 2.2.

2.3.2 Rule Compliance

Because compliance with either rule likely depends on whether the individual is a *Giver* ($(Give_i|NoRule) = 1$) or *Non-Giver* ($(Give_i|NoRule) = 0$), as well as on whether the individual voted for *Rule:Give* or *Rule:Don't* we take a type-weighted approach to studying rule compliance.[18] We first assess, for each $Type_i = (Give_i|NoRule) \times Vote_i$, the level of rule compliance in the baseline treatment and the effect of interventions *Pay4Vote*, *MoneyOffer* and *ExcludePoor* against this benchmark. We then weight types according to the relative frequency with which they appear in our sample. This approach, which closely follows Dal Bó et al. (2010), prevents a misestimation of compliance that can result from an unbalanced distribution of types across our four treatments, which can hide or exaggerate actual changes in behavior.

In this section, we present estimates for rule compliance on the population level as well as for subgroups defined by giving behavior in round 1 ($Give_i|NoRule$) and voting behavior in the referendum ($Vote_i$). Type-level estimates of all treatment effects can be found in table A.1 in the appendix.

[18] In the appendix we provide a theoretical framework supporting the claim that rule compliance and voting behavior likely depends on the intrinsic giving preferences (i.e., $(Give_i|NoRule)$) of the individual.

2.3.2.1 Baseline Rule Compliance

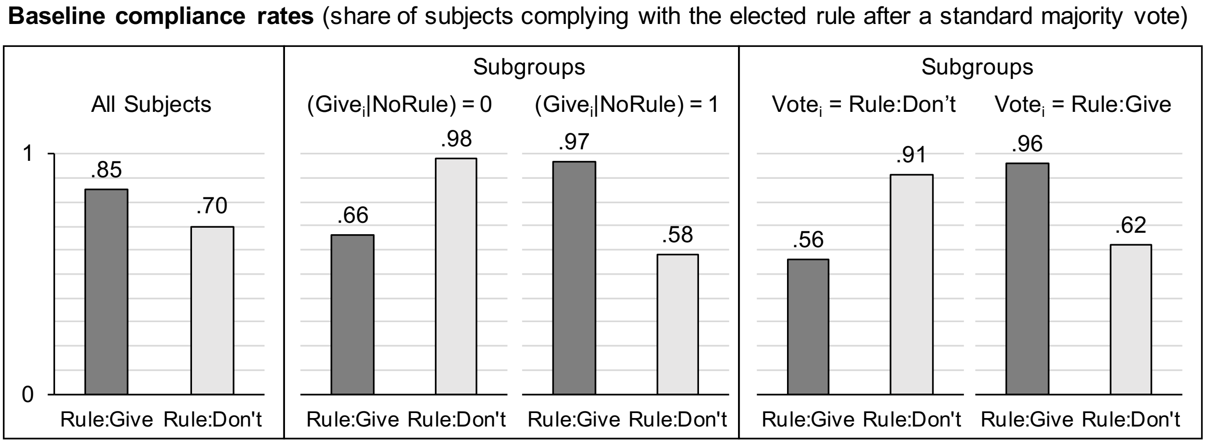


Figure 2.4: Share of subjects complying with majority-elected rules. Graphs show type-weighted averages. For details see Table A.1 in the appendix.

We observe high compliance with both *Rule:Give* and *Rule:Don't* when rules are selected by a standard majority vote, see Figure 2.4. As expected, a subject is more likely to follow *Rule:Give* if she is a *Giver* and if she voted for *Rule:Give*. A symmetric observation holds for *Rule:Don't*. The probability with which subjects comply with rules that are opposite to their original choice is striking: 66% of *Non-Givers* (56% of *Rule:Don't*-voters) voluntarily follow *Rule:Give* when it is elected by the majority of participants. Similarly, 58% of *Givers* (62% of *Rule:Give*-voters) comply with *Rule:Don't*. Taking the weighted average across all types, we find that the unconditional probability of compliance is .85 for *Rule:Give* and .70 for *Rule:Don't*. This compares to a probability of giving (non-giving) in the absence of a rule of only .61 (.39).

The average difference between an individual's choice in round 2 ($Give_i|Rule{:}Give$ and $Give_i|Rule{:}Don't$, respectively) and the same individual's choice in round 1 ($Give_i|NoRule$) is used as an estimator of the power of the majority-elected rule to change individual behavior. Analyzing $\Delta Give_i|Rule := (Give_i|Rule) - (Give_i|NoRule)$ in *T_Baseline* we find:

Result 1 (Rules selected by majority vote shift behavior)**.** *When selected by a standard majority vote, the share of subjects complying with Rule:Give (Rule:Don't) is substantially larger than the share of subjects choosing to give (to not give) in the absence of a rule.*

2.3.2.1.1 Support. Within *T_Baseline*, the average of $\Delta Give_i|Rule{:}Give := (Give_i|Rule{:}Give) - (Give_i|NoRule)$ is $+.24$, which implies a large (24 percentage points) and highly significant ($p < 0.001$, one-sample t-test, two-tailed) increase in giving rates under *Rule:Give*. Similarly, the average of $\Delta Give_i|Rule{:}Don't := (Give_i|Rule{:}Don't) -$

($Give_i|NoRule$) is $-.29$, which implies a large (29 percentage points) and highly significant ($p < 0.001$, one-sample t-test, two-tailed) decrease in giving rates under *Rule:Don't*.[19] Confirming these results, non-parametric McNemar tests of the null hypotheses that subjects are equally likely to choose to give in round 1 and round 2 are rejected for both rules ($p < 0.001$).

2.3.2.2 Treatment Effects

How does malpractice affect compliance with elected rules? Table 2.3 and Figure 2.5 report the estimated difference between the share of subjects complying with *Rule:Give* (*Rule:Don't*) after intervention *Pay4Vote*/*MoneyOffer*/*ExcludePoor* and the share of subjects complying with *Rule:Give* (*Rule:Don't*) in the baseline.

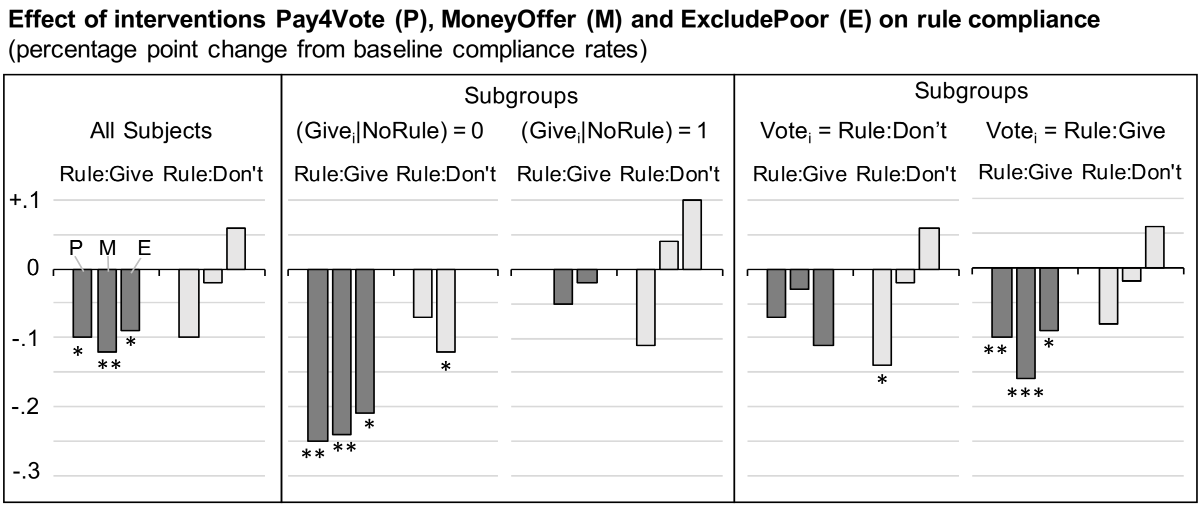


Figure 2.5: Effect of interventions *Pay4Vote* (P), *MoneyOffer* (M) and *ExcludePoor* (E) on rule compliance. Graphs show type-weighted averages, see Table 2.3. Stars denote statistically significant differences to the baseline compliance rate: * $p < .1$, ** $p < .05$, *** $p < .01$.

We see strong, systematic, and statistically significant effects on compliance with *Rule:Give*. When subjects are asked to pay for their vote (*T_Pay4Vote*), when they are offered money to reverse their vote (*T_MoneyOffer*), or when a large share of them is excluded from the ballot due to household income (*T_ExcludePoor*), compliance with the prosocial rule decreases between 9 and 12 percentage points in the overall population (see column 1 in Table 2.3 as well as the first panel of Figure 2.5). The second column in Table 2.3 (the second panel in Figure 2.5, respectively) shows that this effect is largely driven by *Non-Givers*: Only roughly 40% of *Non-Givers* follow *Rule:Give* after an election that saw one of the three interventions, compared to roughly 65% in the baseline. This

[19]These estimates control for correlation in error terms that are due to unobserved individual fixed effects when comparing the behavior of the same group of individuals in round 1 and round 2.

Share of n complying with...	All Subjects	Subgroups $Give_i \mid NoRule$ $= 0$	$= 1$	$Vote_i = Rule{:}$ *Don't*	*Give*
...Rule:Give					
Pay4Vote	-.10	-.25	.00	-.07	-.10
	(.05)	(.11)	(.04)	(.14)	(.04)
MoneyOffer	-.12	-.24	-.05	-.03	-.16
	(.05)	(.12)	(.04)	(.14)	(.05)
ExcludePoor	-.09	-.21	-.02	-.11	-.09
	(.05)	(.11)	(.04)	(.13)	(.04)
Malpractice (Pooled)	-.11	-.23	-.03	-.09	-.11
	(.04)	(.09)	(.03)	(.10)	(.04)
Constant (*T_Baseline*)	.85	.66	.97	.56	.96
	(.03)	(.08)	(.03)	(.09)	(.03)
...Rule:Don't					
Pay4Vote	-.10	-.07	-.11	-.14	-.08
	(.06)	(.05)	(.09)	(.08)	(.08)
MoneyOffer	-.02	-.12	.04	-.02	-.02
	(.06)	(.06)	(.09)	(.08)	(.08)
ExcludePoor	.06	-.00	.10	.05	.06
	(.06)	(.05)	(.09)	(.07)	(.08)
Malpractice (Pooled)	-.01	-.06	-.02	-.01	-.01
	(.05)	(.04)	(.08)	(.06)	(.07)
Constant (*T_Baseline*)	.70	.98	.58	.91	.62
	(.04)	(.04)	(.07)	(.05)	(.06)
Observations	400	155	245	109	291

Standard errors in parentheses.

Table 2.3: Effect of interventions *Pay4Vote*, *MoneyOffer*, and *ExcludePoor* on compliance rates. Average treatment effects (ATE) calculated as the weighted average of treatment effects by $Type_i = (Give_i \mid NoRule) \times Vote_i$ assuming normally distributed standard errors. See Table A.1 in the appendix for treatment effects on type-level.

is intuitive: First and foremost, malpractice should be affecting those subjects who need to be convinced to follow the behavior promoted by the rule.[20] The strongest effect is found for *Non-Givers* who voted for *Rule:Give*.While other types show smaller effects, the negative impact on compliance with *Rule:Give* is systematic across the entire sample. Although the nature of the interventions is quite different, their effect on compliance with *Rule:Give* is strikingly similar.

Regarding subjects' compliance with *Rule:Don't*, Table 2.3 and Figure 2.5 show

[20]See Appendix for a theoretical framework which formalizes this claim.

smaller, inconsistent, and mostly insignificant treatment effects. Given the systematic changes we observe for the opposite rule, this might be surprising.

We conclude:

Result 2 (**Main Result**) (Electoral malpractice decreases compliance with *Rule:Give* but not with *Rule:Don't*)**.** *Subjects display strong, systematic, and statistically significant reductions in compliance with Rule:Give when the rule is elected in the presence of interventions Pay4Vote, MoneyOffer, and ExludePoor. We observe smaller, inconsistent, and insignificant effects of the same interventions on compliance with Rule:Don't.*

2.3.2.2.1 Support. Using a type-weighting approach (see also Dal Bó et al., 2019), we find that the population average treatment effect (ATE) of interventions *Pay4Vote*, *MoneyOffer*, and *ExludePoor* on compliance with *Rule:Give* is $-.10$ ($p = 0.053$), $-.12$ ($p = 0.013$), and $-.09$ ($p = 0.059$), respectively (see Table 2.3, column 1).[21] When pooling interventions, the ATE on compliance with *Rule:Give* is $-.11$ ($p = 0.008$). While *Non-Givers* show the strongest decline, a weakly negative effect is found for all subgroups (see columns 2 to 5). Treatment effects on *Rule:Don't*, on the other hand, are sometimes positive and sometimes negative, mostly insignificant and generally smaller. On average, the interventions are estimated to have little to no effect on compliance with *Rule:Don't*: The pooled ATE is $-.01$ ($p = 0.823$).[22]

2.3.3 Understanding Rule Compliance and Treatment Effects

In order to shed light on potential psychological mechanisms underlying the treatment effects we find, we now analyze elicited beliefs about the rule compliance of other participants. With this, we can say more about the potential role of "peer effects" in compliance decisions. In particular, it might be that subjects change their behavior as a reaction to our interventions because the intervention changed their beliefs about what others will do. In section 2.3.3.2, we explore two explanations that are directly related to procedural preferences subjects may have about rule-setting mechanisms. Are people less willing to comply with rules if they did not personally participate in selecting them? And, does compliance vary with beliefs about a potential bias in the voting outcome?

[21]For treatment effects on type-level see Table A.1 in the appendix.

[22]Identical effects as those reported in Table 2.3 (usually with higher levels of significance) are found with other methods that account for type-dependent treatment effects, for example, inverse probability weighting or regression adjustment. Note that the type-weighted approach we follow is identical to a matching estimator with exact matching on (discrete) type covariates.

2.3.3.1 Beliefs About the Rule Compliance of Other Subjects

Do subjects follow rules because they want to follow others? While in the dictator game payoffs are not interdependent, subjects may still be inclined to condition their compliance choices on the expected behavior of the 99 other participants in their group, for example due to preferences for conditional cooperation or conformity. Following this conjecture, we study to what extent beliefs about the voting and compliance behavior of other subjects can explain rule compliance in general and treatment differences in particular. Figure 2.6 displays the frequencies of beliefs (pooled across all treatments) by answer bracket.

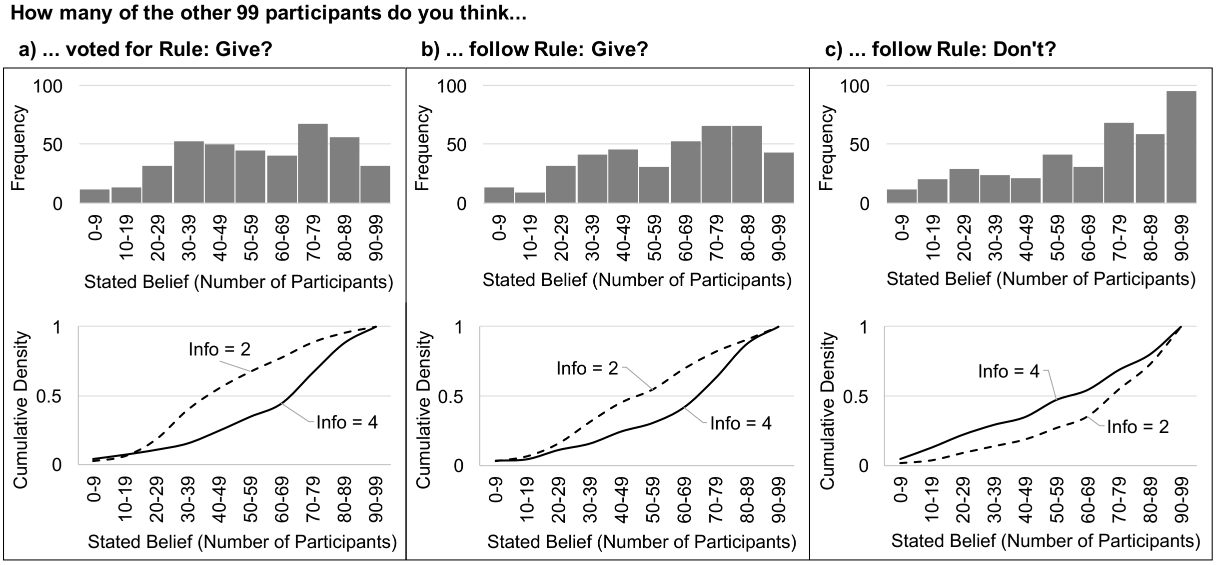


Figure 2.6: Beliefs about the choices of other participants (data from all treatments pooled, N=400). Top: Frequency of beliefs by answer bracket. Bottom: Cumulative density of answers among subjects having received *info*= 2 and *info*= 4, respectively.

Comparing the distributions of individual beliefs about the behavior of other participants in treatment *T_Baseline* with *T_Pay4Vote*, *T_MoneyOffer* and *T_ExcludePoor*, we do not observe systematic differences.[23] This makes beliefs about others an unlikely candidate to explain the treatment differences we find. Nonetheless, they may be an important determinant of rule compliance in general: Understanding the causal effect of beliefs about others on the decision to comply with *Rule:Give* and *Rule:Don't*, respectively, may help us to better understand the overall pattern of choices observed in the experiment.

In a regression of beliefs on behavior, beliefs are very likely to be endogenous, i.e., correlated with the error term. In the case of rule compliance, for example, attitudes about how one "ought" to behave (injunctive social norms) will most likely affect both how an

[23]Beliefs in each treatment follow very much the same distribution as the pooled data shown in Figure 2.6.

individual behaves herself and what the individual believes about how others will behave (see also the discussion in Costa-Gomes et al., 2014). Likewise, other unobserved individual characteristics can lead to an omitted variable bias. To overcome the endogeneity issue and to estimate a causal effect of beliefs on behavior, we use variable $info_i \in \{2, 4\}$ as an instrument for beliefs. Variable $info_i$ records whether, at the end of round 1, individual i was i.i.d. randomly shown a sample in which four out of five subjects chose to give in the dictator game ($info_i = 4$) or, alternatively, a sample in which two out of five subjects chose to give ($info_i = 2$). As this is the only information that participants receive about the behavior of others throughout the entire experiment, $info_i$ is very likely to have a strong effect on subjects' beliefs about the distribution of pro-social types in the population. Figure 2.6 (bottom panel) confirms this intuition: Subjects who randomly received $info_i = 4$ have consistently higher beliefs about the number of other subjects (a) voting for or (b) complying with *Rule:Give*, as well as consistently lower beliefs about (c) the number of other subjects complying with *Rule:Don't*.

Table 2.4 presents the results of an instrumental variable approach to estimating the role of beliefs about others' behavior in guiding a subject's own choices under *Rule:Give* (panel a) and *Rule:Don't* (panel b). The main covariate of interest in this analysis is $E_i(Comply_{-i})$, which is the share of the 99 other participants whom individual i believes to comply with *Rule:Give* or *Rule:Don't*, respectively.[24] Columns (1) in Table 2.4 present the results of OLS regressions on $E_i(Comply_{-i})$, using $info_i$, a binary variable *Malpractice* (equal to one if individual i is a subject in treatment *T_Pay4Vote*, *T_MoneyOffer* or *T_ExcludePoor*, zero otherwise), and type controls $(Give_i|NoRule) \times Vote_i$ as covariates. The large and highly significant coefficients on $info_i$ confirm the observation from Figure 2.6 that variable $info_i$ is a powerful instrument to assess the causal effect of beliefs on behavior under both rules.

Columns (2) report results of OLS regressions of $E_i(Comply_{-i})$ on compliance with *Rule:Give* (panel a) and with *Rule:Don't* (panel b), respectively. The strong and highly significant coefficients on $E_i(Comply_{-i})$ show that beliefs about the behavior of others and individual compliance decisions are highly correlated. To identify the causal effect of beliefs on behavior, we use an IV (2SLS) estimator with $info_i$ instrumenting for $E_i(Comply_{-i})$ in columns (3). Columns (4) and (5) present variations on the same scheme: Columns (4) show the result of an OLS regression using $info_i$ directly as an explanatory variable instead of using it as an instrument for $E_i(Comply_{-i})$. This way, we control for *any* systematic

[24] We ask subjects to state their belief about the *number* of compliant others in their treatment. The response of individual i identifies a bracket, $E_i(\#Compliers_{-i}) \in \{0\text{-}9, 10\text{-}19, ..., 90\text{-}99\}$. $E_i(Comply_{-i})$ is the median of this bracket divided by 99. For example, if $E_i(\#Compliers_{-i}) =$ 40-49, then the median is 44.5 and $E_i(Comply_{-i}) = 44.5/99 \approx 0.45$.

	(a) ***Rule:Give***					(b) ***Rule:Don't***				
	$E_i(Comply_{-i})$	$Comply_i = 1$				$E_i(Comply_{-i})$	$Comply_i = 1$			
	(1) OLS	(2) OLS	(3) 2SLS	(4) OLS	(5) OLS	(1) OLS	(2) OLS	(3) 2SLS	(4) OLS	(5) OLS
$info_i$	.13			-.04	-.04	-.13			-.11	-.09
	(.02)			(.04)	(.04)	(.03)			(.04)	(.04)
$E_i(Comply_{-i})$		.46	-.32				.51	.87		
		(.07)	(.30)				(.07)	(.33)		
Malpractice	-.02	-.10	-.11	-.11	-.10	.01	-.02	-.02	-.01	-.05
	(.03)	(.04)	(.05)	(.04)	(.04)	(.03)	(.05)	(.05)	(.05)	(.05)
Constant	.51	.31	.74	.58	.30	.79	.57	.30	.99	.59
	(.03)	(.06)	(.18)	(.05)	(.16)	(.04)	(.08)	(.25)	(.06)	(.19)
Control for $Type_i$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Add. Controls					Yes					Yes
Observations	400	400	400	400	375	400	400	400	400	375

Standard errors in parentheses.

Table 2.4: The role of others in guiding behavior. $E_i(Comply_{-i})$ is individual i's belief about the share of other participants complying with the rule. *Malpractice* = 1 if individual i is in treatment *T_Pay4Vote*, *T_MoneyOffer* or *T_ExcludePoor*. IV regressions are 2SLS with $E_i(Comply_{-i})$ being instrumented by $1.[info_i = 4]$. Control for $Type_i$ includes $Give_i|NoRule$, $Vote_i$, and $(Give_i|NoRule) \times Vote_i$. Additional controls in (5) are: $Female_i$, $Risk\_Seeking_i$, $Betrayal\_Aversion_i$, $Western_i$, $Student_i$, $UGrad_i$, number of mistakes in control questions, factor variables measuring political and social values in questionnaire, as well as *Big Five* personality test measures. $Female_i$ and $Risk\_Seeking_i$ (answer on 11-point Likert-scale to "Are you a person who is generally willing to take risks (10) or do you try to avoid taking risks (0)?") are weakly significant for compliance with *Rule:Give* (.08 and .02, respectively, $p < .10$). $Betrayal\_Aversion_i$ (answer on 11-point Likert-scale to "Do you think that most people would try to take advantage of you if they got the chance (10), or would they try to be fair (0)?") is highly significant for compliance with *Rule:Don't* (.04, $p < .01$). All other demographic and questionnaire controls are insignificant.

dependency between individual behavior and beliefs about the share of pro-social agents in the population that are shifted by $info_i$. Columns (5) include individual characteristics and questionnaire answers as controls. The following result summarizes our findings:

Result 3 (Beliefs about others only affect compliance with *Rule:Don't*)**.** *Variance in subjects' beliefs about the rule compliance of others cannot explain the negative effect of interventions Pay4Vote, MoneyOffer, and ExcludePoor on compliance with Rule:Give. Moreover, there is no evidence that beliefs about others' compliance causally affect baseline compliance with Rule:Give. A subject's compliance with Rule:Don't, on the other hand, is strongly and positively affected by beliefs about the rule following of others.*

2.3.3.1.1 Support. Two-sample Kolmogorov-Smirnov tests cannot reject the equality of belief distributions across treatments regarding the number of other subjects who vote for *Rule:Give* (smallest p-value is $p = .468$), comply with *Rule:Give* (smallest p-value is $p = .813$), or comply with *Rule:Don't* (smallest p-value is $p = .699$). In line with these results, variable *Malpractice* is insignificant in an OLS regression on $E_i(Comply_{-i})$, both for *Rule:Give* and for *Rule:Don't*, see Table 2.4, columns (1). Also, variance in $E_i(Comply_{-i})$ cannot explain the negative effect of interventions *Pay4Vote*, *MoneyOffer*, and *Exclude-Poor* on compliance with *Rule:Give*: Irrespective of whether one includes beliefs directly as a control (Table 2.4, column (2)) or via instrument $info_i$ (column (3)), *Malpractice* is identified to have virtually the same average treatment effect (ATE) on rule compliance as in Table 2.3. That is, it reduces compliance with *Rule:Give* by approximately 10 percentage points.

Regarding rule compliance in general, Table 2.4 column (3) shows that beliefs about the rule compliance of others causally impact compliance with *Rule:Don't* but do not affect compliance with *Rule:Give*. Specifically, using $info_i$ as an instrument for $E_i(Comply_{-i})$, a 1 percentage point increase in $E_i(Comply_{-i})$ is estimated to increase the probability of individual i to comply with *Rule:Don't* by 0.87 percentage points ($p < 0.01$). Accounting for this effect, no other explanatory variable is significant at the 5 percent level. For compliance with *Rule:Give*, on the other hand, the effect of $E_i(Comply_{-i})$ (when instrumented with $info_i$) is insignificant. Our results are robust to using $info_i$ directly as an explanatory variable (columns 4 of Table 2.4) and to including a battery of individual characteristics and questionnaire answers as controls (columns 5).

2.3.3.2 Lost Votes and Beliefs about Outcome Bias

While treatments *T_Pay4Vote*, *T_MoneyOffer*, and *T_ExcludePoor* differ in the particular form of electoral malpractice, they have in common that due to the intervention many votes are not counted or not counted for the rule the individual originally preferred. In the beginning of this section, we observed that a substantial fraction of participants are excluded from having their vote count due to the intervention in each treatment (35%, 39% and 50%, see binary variable $Lost\_Vote_i$ in Table 2.2 and Figure 2.7 panel (a)). If between-treatment differences in rule compliance vary with $Lost\_Vote_i$, this can be an indication that part of the malpractice effect we see can be explained by subjects disregarding rules that were elected without their *personal* vote being accounted for.

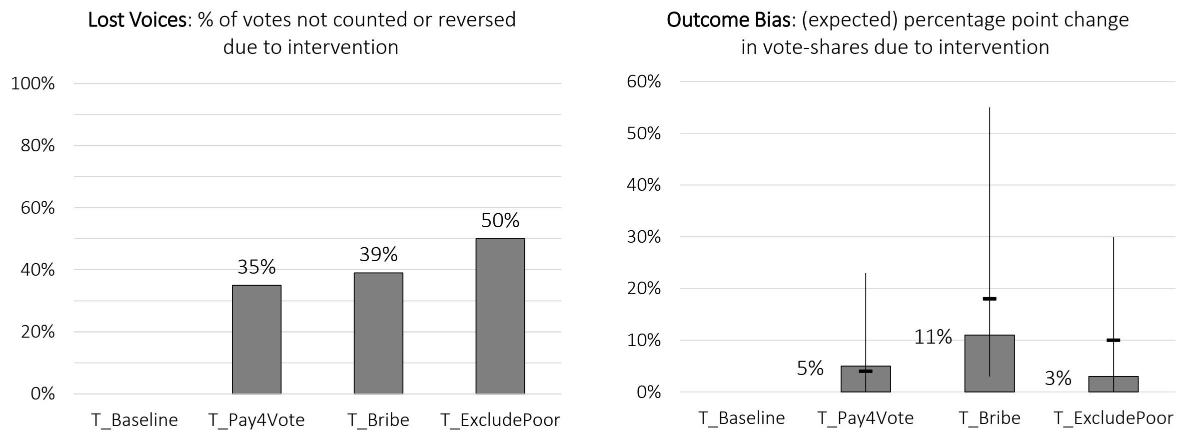


Figure 2.7: Measures of Election Bias. Panel a): Share of votes not counted or reversed due to interventions *Pay4Vote*, *MoneyOffer*, *ExcludePoor*. Panel b): Outcome bias (percentage point change in vote shares due to intervention, bar plot) and subjects' beliefs about outcome bias (10th to 90th percentile with median, whisker plot).

Intuitively, the exclusion or manipulation of votes can lead to vote shares being shifted relative to a standard majority vote without interventions. The absolute shift in vote shares in our treatments, which we call *Outcome_Bias*, is minor (5 (*T_Pay4Vote*), 11 (*T_MoneyOffer*) and 3 percentage points (*T_ExcludePoor*), respectively, see figure 2.7 panel (b)) and is never critical in shifting the voting outcome to the other rule. Because subjects are not informed about how many votes were lost due to the intervention, however, individuals' *beliefs* about the outcome bias may vary. Figure 2.7 panel (b) plots the median and the 10th to 90th percentile of beliefs about this bias for each of our treatments.[25] A relatively large proportion of subjects expresses beliefs implying that they expect vote shares to shift by more than 10 percentage points (26%, 70%, and 53%, re-

[25]Note that to avoid responses that are influenced by social desirability, we do not ask subjects to directly report their beliefs about a potential outcome bias. Instead, we compute $E_i[Outcome\_Bias]$ from elicited beliefs regarding the share of subjects accepting to pay for their vote (*T_Pay4Vote*), the share of subjects accepting the monetary offer (*T_MoneyOffer*), or the voting behavior among "poor" and "rich" subjects (*T_ExcludePoor*). In particular, we calculate individual i's belief about the outcome

spectively). We can exploit the variance in $E_i[Outcome\_Bias]$ to explore in how far beliefs about the referendum's overall representativeness may explain the shift in rule compliance observed across our treatments.

$Lost\_Vote_i$ and $E_i[Outcome\_Bias]$ thus form our two measures of (perceived) election bias. Can the variance in these two measures explain the variance in compliance with *Rule:Give* between treatments?

Table 2.5 presents results from OLS regressions of binary treatment variables and controls on $Comply_i|Rule{:}Give$, to which we successively add $Lost\_Vote_i$ (column (2)) and $E_i[Outcome\_Bias]$ (column (3)) as additional explanatory variables; column (4) includes both. We also run analyses of variance (ANOVA) to learn more about the share of variance in treatment effects that is captured by variance in $Lost\_Vote_i$ and $E_i[Outcome\_Bias]$.

We find:

Result 4 (Explanatory power of lost votes and beliefs about outcome bias.)**.** *Subjects whose (original) vote is not counted and subjects who hold the belief that the referendum is not representative drive the decline in compliance with Rule:Give in treatments T_Pay4Vote, T_MoneyOffer, and T_ExcludePoor.*

2.3.3.2.1 Support. Table 2.5 shows that the addition of $Lost\_Vote_i$ (column (2)), $E_i[Outcome\_Bias]$ (column (3)), or both (column (4)) as explanatory variables for compliance with *Rule:Give* considerably lowers the explanatory power of binary treatment variables for treatments *T_Pay4Vote*, *T_MoneyOffer*, and *T_ExcludePoor*: Column (1) reproduces our main finding that all three forms of malpractice significantly reduce compliance with *Rule:Give* by roughly 10 percentage points. Including just one of the two variables in the regression (columns (2) and (3)) lowers the estimated coefficients on treatment variables to roughly one third to two thirds of their original effect. Including both variables simultaneously (column (4)) leads to the average residual effects of the treatment variables being further reduced to an estimated residual effect of -.05 ($p = 0.36$) for *T_Pay4Vote* and effects close to zero for the other two treatments. When running the same regression with the pooled treatment indicator *Malpractice* instead of including each

bias as $E_i[Outcome\_Bias]$

$$:= \begin{cases} 0 & \text{if } i \text{ is in } T\_Baseline, \\ \left| \dfrac{E_i[Accept\_Pay_j | Vote_j = 1] E_i[Vote_j]}{E_i[Accept\_Pay_j]} \right| & \text{if } i \text{ is in } T\_Pay4Vote, \\ |E_i[Accept\_MoneyOffer_j | Vote_j = 1] E_i[Vote_j] & \\ +E_i[Accept\_MoneyOffer_j | Vote_j = 0](1 - E_i[Vote_j])| & \text{if } i \text{ is in } T\_MoneyOffer, \\ |E_i[Vote_j | Income_j > 40\text{K}] - E_i[Vote_j]| & \text{if } i \text{ is in } T\_ExcludePoor. \end{cases}$$

	$Comply_i$ | ***Rule:Give***			
	(1) OLS	(2) OLS	(3) OLS	(4) OLS
Lost_Vote$_i$		-.11		-.10
		(.04)		(.04)
E_i[*Outcome_Bias*]			-.34	-.33
			(.12)	(.12)
T_Pay4Vote	-.11	-.07	-.08	-.05
	(.05)	(.05)	(.05)	(.05)
T_MoneyOffer	-.12	-.08	-.04	.00
	(.05)	(.05)	(.06)	(.06)
T_ExcludePoor	-.09	-.04	-.06	-.01
	(.05)	(.05)	(.05)	(.05)
Constant	.56	.57	.56	.57
	(.05)	(.05)	(.05)	(.05)
Add. Controls	Yes	Yes	Yes	Yes
Observations	400	400	400	400

Standard errors in parentheses.

Table 2.5: Explaining treatment variance in compliance with *Rule:Give* with variance in *Lost_Vote*$_i$ $\in \{0,1\}$ and with variance in subjects' beliefs about outcome bias E_i[*Outcome_Bias*] $\in [0,1]$. Controls are: *Give*$_i$|*NoRule*, *Vote*$_i$, (*Give*$_i$|*NoRule*) $\times$ *Vote*$_i$, and $info_i$.

treatment separately, the average residual effect amounts to -.03 (p = .57). Analysis-of-variance (ANOVA) models suggest that including *Lost_Vote*$_i$ and E_i[*Outcome_Bias*] as explanatory variables for rule compliance decreases the variance in behavior explained by binary treatment variables by roughly 80%. In a general sense, the effect sizes close to zero of the treatment variables in column (4) imply that participants who are *not* excluded and who do *not* hold the belief that the voting outcome looses its representativeness show the same compliance behavior as the average participant in *T_Baseline*.

Table 2.5 thus confirms our expectation that both *Lost_Vote*$_i$ and E_i[*Outcome_Bias*] are associated with significantly lower rates of rule compliance.[26] Interestingly though, our analysis shows that it is *not* only the subjects losing their vote who show negative responses to interventions *Pay4Vote*, *MoneyOffer*, and *ExcludePoor*: In column (2),

[26]Note that the exact coefficients on *Lost_Vote*$_i$ should be interpreted with caution: While the decrease in treatment effect size implies that part of the effect *must* be causal (because treatment exposure is random on the individual), the variable is very likely to also capture selection effects in treatments *T_Pay4Vote* and *T_MoneyOffer*. In these two treatments, whether a subject's vote is counted in the ballot is endogenous to her decision of whether to pay the fee or to accept the bribe, respectively. We included *T_ExcludePoor* in our experiment in order to have one treatment with an exogenous exclusion criterion where subjects do not select into "being treated".

residual treatment effects are smaller but remain consistently negative. This suggests that the experience of malpractice alone—even without one's personal vote being directly affected—can negatively affect compliance rates. Indeed, while $E_i[Outcome\_Bias]$ is not independent from treatment exposure, the results show that reductions in compliance are associated with holding the belief that the voting outcome is not representative of voting preferences in the population (see columns (3) and (4)).

2.4 Discussion

2.4.0.0.1 Relation of malpractice and democracy effects. Our paper shows that experimentally induced "malpractice" during the election of a rule governing voluntary social behavior can lead to lower compliance with the elected rule. One way to interpret the result is that malpractice erodes the positive "democracy effect" that earlier studies have found in experimental games in which subjects can vote for similar institutions. Dal Bó et al. (2010), for example, study the effect on cooperation when subjects endogenously—i.e., through voting—choose to convert a prisoners' dilemma game into a coordination game compared to the effect of changing the game exogenously (by random choice of the computer). They find an endogeneity premium in cooperation of roughly 14 percentage points.

How does the "malpractice effect" we find compare to a potential "democracy premium" in the same game? To answer this question, we discuss the results of an additional treatment, *T_Exo*.[27] In this treatment, everything is equal to our baseline treatment except that the rule (*Rule:Give* or *Rule:Don't*) is now exogenously implemented. Before playing the second round of the dictator game, participants are informed that "*(t)he code of conduct will be randomly selected by the computer*" using a "*coin flip*" with equal probabilities. We find that in *T_Exo*, 75% of subjects comply with *Rule:Give* and 70% with *Rule:Don't*. Compared to our baseline treatment, this amounts to a decline in compliance of $-.10$ ($p = .037$) and $\pm.00$ (i.e., no significant reduction, $p = .96$), respectively. In other words, measured against the implementation of an exogenous rule, we find a democracy premium of +10 percentage points for *Rule:Give* when the rule is selected by a standard majority vote, but no such premium for *Rule:Don't*.

Strikingly, the positive democracy premium for *Rule:Give* that we establish against *T_Exo* is virtually identical to the negative malpractice effect we find in treatments

[27]The treatment was run with 100 new participants in summer 2018 on *Prolific.ac*. Instructions and implementation were identical to the main treatments except for the description of the vote aggregation procedure as described here. The mean age of participants is 29 years, 53% are female, and 37% are students.

T_Pay4Vote, *T_MoneyOffer*, and *T_ExcludePoor* (−10, −12 and −9 percentage points). At the same time, for *Rule:Don't*, where malpractice on average does not affect compliance rates, *T_Exo* can also not establish a democracy effect. This finding suggests that, indeed, the mechanism by which malpractice erodes compliance is by undermining the democracy premium on domains in which such a premium exists.

2.4.0.0.2 Do treatment effects relate to how people perceive violations of democratic principles in the real world? Our experiment establishes how personal disenfranchisement and voters' beliefs about biases in the voting outcome affect subsequent compliance with elected rules of behavior in a neutrally framed experimental setup. With this, we aim to establish a finding that relates to the behavioral consequences of electoral malpractice in real world elections.

One way to find suggestive evidence for this relation to behavior in real world institutions is to study whether treatment effects are more likely to be found among participants who place a high value on democratic institutions and who are sensitive to mechanisms that may corrupt these institutions (such as bribing and lobbying). If this is the case, then the reactions of these participants to instances of real world malpractice can be thought to be governed by similar concerns as their reactions in our experiment. In Table 2.6, we perform this exercise by exploiting the variation in demographic characteristics in our online subject pool as well as in participants' answers in the post-experimental questionnaire to empirically identify types with a relatively lower or higher value for—or expectation of—democratic procedures.

Table 2.6 demonstrates that interventions *Pay4Vote*, *MoneyOffer*, and *ExcludePoor* tend to produce treatment effects of larger magnitude and higher statistical significance among participants who have more experience with democratic institutions (1,2), among participants who self-identify as placing high value on democratic decision-making processes (3,4), and, finally, among subjects who believe that it is never justifiable to offer or take a bribe, or to lobby politicians (5,6,7).[28] Column (5) provides maybe the strongest support for our claim: Those who indicate a very high sensitivity to bribery in the real world also react very sensitively to electoral malpractice in our experiment, the strongest negative effect being found in treatment *T_MoneyOffer*. Overall, the observations in Table 2.6 suggest that, indeed, our findings in the (context-free) online experiment relate to

[28]Recall that the questionnaire is sent to subjects using a different researcher profile and visual design more than two weeks after they have taken part in the experiment, making spillovers from our treatments to the questionnaire answers highly unlikely. Indeed, we find that the probability for a subject to be identified as "High" or "Low" in Table 2.6 does not significantly depend on the treatment to which the subject was assigned. There is only one exception: In column (3), a subject is more likely to be identified as "High Dem_Importance=1" if she participated in treatment *T_Pay4Vote*.

	$Comply_i$|***Rule:Give***													
	(1)		(2)		(3)		(4)		(5)		(6)		(7)	
			High		High		High		Low		Low		Low	
	Western		Dem_Country		Dem_Importance		Dem_Voice		Just_Bribes		Just_Lobby		Just_Money	
	1	0	1	0	1	0	1	0	1	0	1	0	1	0
T_Pay4Vote	-.14	-.08	-.15	-.09	-.14	-.06	-.12	-.15	-.16	-.05	-.18	-.09	-.14	-.03
	(.06)	(.08)	(.09)	(.08)	(.08)	(.09)	(.08)	(.08)	(.07)	(.08)	(.08)	(.07)	(.07)	(.08)
T_MoneyOffer	-.12	-.12	-.21	-.10	-.19	-.00	-.19	-.11	-.22	-.01	-.19	-.07	-.19	-.07
	(.06)	(.09)	(.09)	(.08)	(.08)	(.08)	(.08)	(.08)	(.07)	(.08)	(.08)	(.07)	(.07)	(.09)
T_ExcludePoor	-.11	-.01	-.18	-.11	-.09	-.09	-.18	.-04	-.13	-.05	-.12	-.07	-.11	-.07
	(.06)	(.09)	(.09)	(.08)	(.08)	(.08)	(.08)	(.08)	(.07)	(.07)	(.08)	(.07)	(.07)	(.09)
Constant	.57	.63	.58	.52	.52	.51	.52	.55	.59	.49	.60	.50	.60	.53
	(.06)	(.09)	(.09)	(.07)	(.07)	(.07)	(.08)	(.07)	(.06)	(.07)	(.07)	(.07)	(.07)	(.08)
Add. Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	272	128	139	156	183	129	140	184	214	170	168	179	189	142

Standard errors in parentheses.

Table 2.6: Treatment effects on compliance with *Rule:Give* (OLS estimates) by nationality and by questionnaire responses to the following questions: (2): "*How democratic do you think your country is overall?*" (median = 7/10), (3): "*How important is it for you to live in a country that is governed democratically*?" (median = 9/10); (4): "*How important is it for you to personally express your voice when it comes to political decision making?*" (median = 8/10); (5), (6), and (7): "*Please indicate to what extent you think the following actions can be justified:*" (5) "*Accepting a bribe in the course of one's duties.*" (median = 0/10), (6) "*Lobbying politicians to influence legislation.*" (median = 3/10), (7) "*Influencing the actions of people by giving them money.*" (median = 2/10). Except for column (5), "High" and "Low" identifies subjects with answers strictly above or strictly below the median (subjects with answers on the median are not included in regressions). In column (5), "Low" identifies subjects with answer = 0 (median) and "High" subjects with answers > 0. Controls are: $Give_i|NoRule$, $Vote_i$, $(Give_i|NoRule) \times Vote_i$ and $info_i$.

psychological domains that are also relevant in corresponding real-world decision making.

2.4.0.0.3 Discussion of behavioral mechanisms. The findings in table 2.6 also support our interpretation of the results in section 2.3.3. Together, they suggest that procedural concerns about the inclusiveness and unbiasedness of the election procedure might drive the decline in compliance observed for *Rule:Give*. This resonates with theories of "legitimate authority" (e.g., Weber, 1978; Tyler, 2006; Dickson et al., 2015; Akerlof, 2017) and with empirical findings suggesting that people care about the " fairness" of decision making processes (see, e.g., Tyler, 1990; Frey et al., 2004; Cappelen et al., 2013). In line with our findings, the previously established "democracy effect" in Dal Bó et al. (2010) (see, in particular, p.2222f) also does not seem to work via differences in informational content (of the election) and strategic motives, but rather by the appeal of the endogenous institution itself.

The additional treatment *T_Exo* sheds a new light on our surprising finding that malpractice seems to have an asymmetric effect: we find a strong and systematic malpractice effect for *Rule:Give* but not for *Rule:Don't*. Interestingly, the same asymmetric pattern can be found for the existence of a democracy effect. In other words, in our setting a malpractice effect can *always* be found in cases where a democracy effect exists.

Compliance with *Rule:Don't* is strongly driven by beliefs about what others do and since beliefs about others' behavior are not affected by the corrupted voting procedures, no differences in average compliance can be found. We can thus speculate that rules that are being complied with due to peer effects are one type of rule where procedural aspects do not play a role for compliance. In contrast, in the case of *Rule:Give*, compliance seems to rather occur due to a preference for following the rule and we find no evidence for beliefs about the rule compliance of others playing a role for own decisions.[29] This type of intrinsically motivated rule compliance seems to be sensitive to procedural aspects. Indeed, our experiment establishes a democracy effect for a fair majority vote as in Dal Bó et al. (2010) and at the same time shows how the same sensitivity to the procedure leads to a complete erosion of this effect if the majority vote has been corrupted. We thus speculate that democracy effects as well as malpractice effects might not be effective in all domains. Whether this speculation holds true in a more general sense and outside of our experimental setup will need to be uncovered by future research.

[29] For detecting a significant effect under *Rule:Give* with a power of 80% (which means that the effect will be significant 80% of the time with $\alpha = 0.05$), the minimum detectable effect size of the coefficient of $info_i$ (Table 2.4, column (4)) is $2.8 * 0.04 = 0.11$, where 0.04 is the standard error of the estimated coefficient. This is an effect we are able to find for *Rule:Don't* where the standard error is very similar. While the effect of beliefs on behavior is not zero under *Rule:Give*, the relationship is both, statistically insignificant and smaller in magnitude than for the case of *Rule:Don't*.

2.5 Conclusion

In this paper, we demonstrated how introducing a voting fee, offering subjects money to reverse their vote, or excluding low-income voters from the ballot during a referendum causally impact subsequent compliance with elected rules of behavior. We find a strong and systematic reduction in voluntary compliance with *Rule:Give* but not with *Rule:Don't*. We demonstrate that the effects we observe under *Rule:Give* correspond to a complete erosion of a democracy effect on the same rule. Compliance with *Rule:Don't*, however, is driven by peer-effects and is not sensitive to procedural aspects. A sensitivity of rule compliance to the implementation procedure is mainly found among subjects who are themselves excluded from the ballot and those who believe the voting outcome to no longer be representative due to the corruption of the vote.

Overall, the experimental results presented in this paper imply that the positive behavioral effects of democratic procedures that earlier studies have established (for example, Frey, 1997; Tyran and Feld, 2006; Ertan et al., 2009; Sutter et al., 2010; Dal Bó et al., 2010) are sensitive to the manipulation of votes. We see this study as a first step towards understanding the effects of electoral malpractice on behavior for democratically elected institutions; more research is needed to draw general conclusions. We chose to study rule compliance in the domain of redistribution for its important economic and social role. Extending the analysis to other domains such as cheating and tax evasion, as well as to other forms of centralized and de-centralized manipulation (such as ballot box stuffing and subject-to-subject bribes), will allow to establish results about compliance with social rules in general.

Appendix

Appendix A: Theoretical Framework

We provide a simple theoretical framework to guide the analysis of giving behavior and compliance rates across treatments. Consider first the decision to give in the absence of a code of conduct. Let $u_i(Give_i)$, $Give_i \in \{0, 1\}$ denote individual i's utility when deciding to give or not give, respectively. Define $\Delta u_i = u_i(Give_i = 1) - u_i(Give_i = 0)$. It follows that

$$(Give_i | NoRule) = 1 \Leftrightarrow \Delta u_i \geq 0.$$

A positive Δu_i may reflect social preferences of individual i such as inequality aversion or "warm glow" utility.[30] Let Δu_i be distributed in the population with cumulative density function $F[\cdot]$. The share of *Givers* in the population is then given by $1 - F[0]$ as illustrated in Figure A.1, panel a), below.

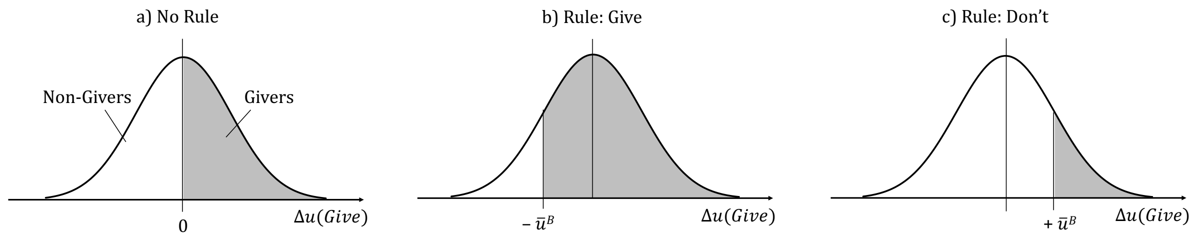


Figure A.1: Theory: Illustration of population shares choosing to give ($Give_i = 1$) and not to give ($Give_i = 0$) when there exists no code of conduct (panel a) and when there exists a code of conduct that came into force with a standard majority vote (panels b and c).

Consider next the situation with a code of conduct, either *Rule:Give* or *Rule:Don't*. If the code has come into force with a standard majority vote (*T_Baseline*) we assume that it adds fixed utility $D \geq 0$ to the action that is prescribed by the code. This constant can

[30] Typical examples in standard settings are Fehr and Schmidt (1999), Bolton and Ockenfels (2000) and Andreoni (1989, 1990). Inequity aversion over chances to win a prize has been modeled by, for example, Saito (2013). Experimental evidence showing how prosocial behavior extends to choices over risky payoffs can be found in Brock et al. (2013) and Freundt and Lange (2017), among others.

be interpreted as an emotional utility some people derive from following a rule elected by the majority. It follows that

$$\begin{aligned} \text{If } \textit{Malpractice} = 0, \quad & (\textit{Comply}_i | \textit{Rule:Give}) = 1 \Leftrightarrow \Delta u_i \geq -D, \\ \text{and} \quad & (\textit{Comply}_i | \textit{Rule:Don't}) = 1 \Leftrightarrow \Delta u_i < +D. \end{aligned}$$

Compared to the case without a code, the share of subjects choosing to give increases or decreases, see Figure A.1, panel (b) and (c), respectively. Note, importantly, that rules only affect the behavior of those individuals who in the absence of a code would have chosen the opposite action. While *Rule:Give* may convince a *Non-Giver* to give, it will leave the behavior of a *Giver* ($\Delta u_i \geq 0$) unaffected. Similarly, *Rule:Don't* may induce some *Givers* to stop giving, but will not affect the choice of *Non-Givers* ($\Delta u_i < 0$). We assume that electoral malpractice (in our experiment, *Pay4Vote*, *MoneyOffer*, *ExcludePoor*) alters the value some people derive from obeying the elected code. Instead of generating utility D, rule compliance is now associated with a lower utility $D - M$. Constant $M \geq 0$ measures the loss in utility induced by malpractice. As a result, individual i's propensity to comply with the elected rule is reduced. In particular,

$$\begin{aligned} \text{If } \textit{Malpractice} = 1, \quad & (\textit{Comply}_i | \textit{Rule:Give}) = 1 \Leftrightarrow \Delta u_i \geq -(D - M), \\ \text{and} \quad & (\textit{Comply}_i | \textit{Rule:Don't}) = 1 \Leftrightarrow \Delta u_i < +(D - M). \end{aligned}$$

First and foremost, we thus expect that malpractice leads people to revert back to their individually preferred behavior: As M increases, a lower share of *Non-Givers* will follow *Rule:Give*, see Figure A.2, panel b). Similarly, a lower share of *Givers* will be willing to follow *Rule:Don't* (Figure A.2, panel c)). As M becomes sufficiently large such that $D - M$ turns negative, people may even turn against rules that match their individual giving preferences. For example, it is theoretically possible that giving under *Rule:Give* will deteriorate below rates observed in the absence of a code, although such a strong reaction might be unlikely to be observed in the experiment.

2.A.0.0.1 Voting Behavior. We can extend above theory to yield predictions about voting behavior. Note that in all treatments, subjects vote before interventions take place that may undermine the democratic election. Voting decisions are therefore unbiased by the exposure to a particular treatment. We assume that each subject votes *sincerely* in the sense that she chooses to vote for the outcome that yields her a higher expected utility. Let $U_i[\textit{Rule}]$ denote i's expected utility given $\textit{Rule} \in \{\textit{Rule:Give}, \textit{Rule:Don't}\}$. When voting, individual i takes into account how her own giving behavior will be affected by the rule as

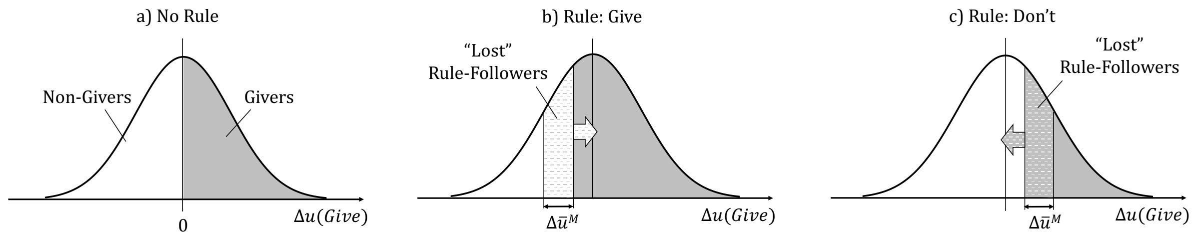


Figure A.2: Theory: Illustration of population shares choosing to give ($Give_i = 1$) and not to give ($Give_i = 0$) when there exists no code of conduct (panel a) and when there exists a code of conduct that came into force with malpractice during the election (panels b and c).

well as how the behavior of *other* subjects will be affected. Conditional on i not receiving tickets from the computer (which happens with probability 0.5), let $\Delta u(Receive) > 0$ denote the difference in utility between receiving three tickets from another subject and not receiving any tickets. Because the average subject in the population is more likely to give under *Rule:Give* than under *Rule:Don't*, the conditional probability that i will receive three tickets from another subject increases by

$$\Delta F[D] = F[+D] - F[-D]$$

when going from *Rule:Don't* to *Rule:Give*. In our setup, voting behavior depends on the individual's giving preferences $\Delta u_i(Give)$ as follows:

1. *Unconditional Givers:* If $\Delta u_i \geq +D$, individual i will choose $Give_i = 1$ irrespective of the rule. Individual i will then always vote for *Rule:Give*:

$$U_i[\mathit{Rule{:}Give}\ |(Give_i|Rule) = 1)] \geq U_i[\mathit{Rule{:}Don't}\ |(Give_i|Rule) = 1]$$

$$0.5 \cdot [u_i(Give_i = 1) + D] + 0.5 \cdot \Delta F[D] \cdot \Delta u_i(Receive) \geq 0.5 \cdot u_i(Give_i = 1)$$

$$\Leftrightarrow \underbrace{\Delta F(D)}_{>0} \geq -\underbrace{\frac{D}{\Delta u(Receive)}}_{<0}.$$

2. *Unconditional Non-Givers:* If $\Delta u_i < -D$, individual i will choose $Give_i = 0$ irrespective of the rule. Individual i will then vote for *Rule:Give* if

$$U_i[\mathit{Rule{:}Give}\ |(Give_i|Rule) = 0] \geq U_i[\mathit{Rule{:}Don't}\ |(Give_i|Rule) = 0]$$

$$0.5 \cdot u_i(Give_i = 0) + 0.5 \cdot \Delta F[D] \cdot \Delta u_i(Receive) \geq 0.5 \cdot [u_i(Give_i = 0) + D]$$

$$\Leftrightarrow -D \geq -\Delta F(D) \cdot \Delta u(Receive)$$

$$\Leftrightarrow \Delta F(D) \geq \frac{D}{\Delta u(Receive)}$$

and otherwise will vote for *Rule:Don't*.

3. *Rule-Followers:* If $-D \leq \Delta u_i < +D$, individual i will choose $Give_i = 1$ under *Rule:Give* and $Give_i = 0$ under *Rule:Don't*. Individual i will then vote for *Rule:Give* if

$$U_i[Rule{:}Give\ |(Give_i|Rule) = 1] \geq U_i[Rule{:}Don't\ |(Give_i|Rule) = 0]$$

$$0.5 \cdot [u_i(Give_i = 1) + D] + 0.5 \cdot \Delta F[D] \cdot \Delta u_i(Receive) \geq 0.5 \cdot [u_i(Give_i = 0) + D]$$

$$\Leftrightarrow\ \Delta u_i \geq -\Delta F(D) \cdot \Delta u(Receive)$$

$$\Leftrightarrow\ \Delta F(D) \geq -\frac{\Delta u_i}{\Delta u(Receive)},$$

and otherwise will vote for *Rule:Don't*. Note that this implies that *Givers* ($\Delta u_i \geq 0$) always vote for *Rule:Give*, while Non-Givers ($\Delta u_i < 0$) do the same if and only if $\Delta F(D)$ is sufficiently large.

We can see that there is a monotonic relation between $\Delta u_i(Give)$ and the tendency to vote for *Rule:Give*. *Givers* always vote for *Rule:Give*. This is true for both, unconditional givers and rule-followers. *Non-Givers*, on the other hand, only vote for *Rule:Give* if they expect that rules have sufficiently large effect on the giving behavior of others. Otherwise, they vote for *Rule:Don't*. If $\Delta F[D]$ is close to zero, all *Non-Givers* vote for *Rule:Don't*. This case is illustrated in Figure A.3, panel a). Increasing $\Delta F[D]$ shifts voting preferences of non-givers in favor of *Rule:Give*. This first affects rule-following *Non-Givers* who indeed would choose to give under the pro-social rule, i.e., those individuals who satisfy $-D \leq \Delta u_i(Give) < 0$, see Figure A.3, panel (b). Only once $\Delta F(D) \geq \frac{D}{\Delta u(Receive)}$, also unconditional non-givers (and thus, all individuals) vote for *Rule:Give*, see Figure A.3, panel c).

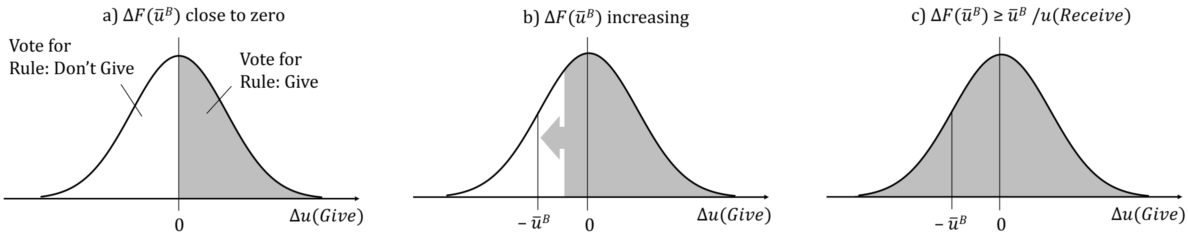


Figure A.3: Theory: Share of Population voting for Rule: Give

Appendix B: Type-level analysis

(a) All treatments: n by $Type_i$

	$Give_i \| NoRule$		
$Vote_i$	0	1	Σ
Rule:Don't	92	17	109
Rule:Give	63	228	291
Σ	155	245	**400**

(b) *T_Baseline*: Share of n complying with...

	...**Rule:Give**			...**Rule:Don't**		
	$Give_i \| NoRule$			$Give_i \| NoRule$		
$Vote_i$	0	1	w.avg.	0	1	w.avg.
Rule:Don't	.57	.50	.56	.96	.63	.91
Rule:Give	.80	1	.96	1	.51	.62
w.avg.	.66	.97	**.85**	.98	.58	**.70**

(c) Treatment Effects (vs. *T_Baseline*):

		...**Rule:Give**			...**Rule:Don't**		
		$Give_i \| NoRule$			$Give_i \| NoRule$		
	$Vote_i$	0	1	w.avg.	0	1	w.avg.
T_Pay4Vote	*Rule:Don't*	-.18	.50	-.07	-.05	-.63	-.14
		(.14)	(.42)	(.14)	(.07)	(.30)	(.08)
	Rule:Give	-.35	-.04	-.10	-.10	-.07	-.08
		(.16)	(.03)	(.04)	(.08)	(.10)	(.08)
	w.avg.	-.25	.00	**-.10**	-.07	-.11	**-.10**
		(.11)	(.04)	(.05)	(.05)	(.09)	(.06)
T_MoneyOffer	*Rule:Don't*	-.01	-.17	.-03	-.09	.38	-.02
		(.16)	(.36)	(.14)	(.08)	(.26)	(.08)
	Rule:Give	-.57	-.04	-.16	-.15	.02	-.02
		(.18)	(.03)	(.05)	(.09)	(.09)	(.08)
	w.avg.	-.24	-.05	**-.12**	-.12	.04	**-.02**
		(.12)	(.04)	(.05)	(.06)	(.09)	(.06)
T_ExcludePoor	*Rule:Don't*	-.13	.00	-.11	.00	.38	.06
		(.14)	(.33)	(.13)	(.07)	(.23)	(.07)
	Rule:Give	-.33	-.02	-.09	.00	.08	.06
		(.17)	(.03)	(.04)	(.09)	(.10)	(.08)
	w.avg.	-.21	-.02	**-.09**	.00	.10	**.06**
		(.11)	(.04)	(.05)	(.05)	(.09)	(.06)
Pooled	*Rule:Don't*	-.12	.06	-.09	-.04	.15	-.01
		(.11)	(.26)	(.10)	(.06)	(.23)	(.06)
	Rule:Give	-.40	-.03	-.11	-.08	.01	-.01
		(.14)	(.03)	(.04)	(.07)	(.08)	(.07)
	w.avg.	-.23	-.03	**-.11**	-.06	.02	**-.01**
		(.09)	(.03)	(.04)	(.04)	(.08)	(.05)

Standard errors in parentheses.

Table A.1: Number of subjects (a), baseline compliance rates (b) and treatment effects by $Type_i = (Give_i|NoRule) \times Vote_i$. Gray cells in (b) and (c) show weighted averages. Weights follow the type-distribution in panel (a). Weighted standard errors calculated assuming normally distributed standard errors (Delta method).

Appendix C: Questionnaire

Questionnaire: Politics

Overall, there are 15 questions. The first 10 questions relate to your views on politics.

1. In political matters, people talk of "the left" and "the right". On a scale from 0 to 10, where would you place your views, generally speaking?
(Scale: 0 = Left, 10 = Right)

2. On a scale from 0 to 10, how important is it for you to live in a country that is governed democratically?
(Scale: 0 = not at all important, 10 = extremely important)

3. How democratic do you think your country is overall?
(Scale: 0 = not at all democratic, 10 = completely democratic)

4. How important is it for you to personally express your voice when it comes to political decision making?
(Scale: 0 = not at all important, 10 = extremely important)

5. It is important that you pay attention to this study. Please tick number 7 to show that you pay attention. The scale below does not play a role.
(Scale: 0 = not at all important, 10 = very important)

6. On a scale from 0 to 10, where 0 means "no trust at all" and 10 means "very much trust", how much do you personally trust...
...politicians?
...large corporations?
...the results of elections?

7. Please indicate for each of the following actions to what extent you think that action can be justified:
(Scale: 0= can never be justified, 10= can always be justified)

- Violating the instructions of one's superiors (for example at work or school).

- Accepting a bribe in the course of one's duties.
- Cheating on taxes if one has the chance.
- Influencing the actions of people by giving them money.
- Lobbying politicians to influence legislation.

8. Below you find two opposing statements on redistribution. How would you place your personal standpoint between the two statements *(0 means that you agree completely with the statement on the left, 10 means that you agree completely with the statement on the right)*

0:	10:
"The rich have an obligation to subsidize the poor. If necessary, they have to be forced to do so."	"Everybody is responsible for himself. Forcefully taking from the rich to subsidize the poor is theft."

9. Below you find two opposing statements on inequality. How would you place your personal standpoint between the two statements *(0 means that you agree completely with the statement on the left, 10 means that you agree completely with the statement on the right)*

0:	10:
"For a society to be fair, the incomes of all people should be equal."	"There is nothing unfair in having more money than somebody else, no matter how large the difference."

10. When elections take place, do you vote always, usually, or never?

Never Rarely Usually Almost always Always

Questionnaire: General questions

These are the final 5 questions of our study. They concern your views in general and your personality.

1. How do you see yourself: Are you a person who is generally willing to take risks, or do you try to avoid taking risks?

(Scale: 0 = Completely unwilling to take risks, 10 = Very willing to take risks)

2. How much do you agree with the following statement: "Money brings out the worst in people."?
(Scale: 0 = Do not agree at all, 10 = Agree completely)

3. Do you think that most people would try to take advantage of you if they got the chance, or would they try to be fair?
(Scale: 0 = All people would try to be fair, 10 = All people would try to take advantage of you)

4. Assume that you had the opportunity to take part in the following gamble: There are 100 balls in an urn. Of these balls, 99 are black and 1 is red. One ball is randomly drawn from the urn. If it is red you win 1000 GBP. If it is black you win 0 GBP. What would be the maximal amount of money you would be willing to pay in order to take part?
Would be willing to pay at most... (dropdown menu with answer choices from 0 GBP to 20 GBP in steps of 1)

5. Here are a number of personality traits that may or may not apply to you. Please indicate to what extent you agree or disagree that these personality traits apply to you. Note: You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other.
I see myself as...

- Extraverted, enthusiastic (NOT reserved or shy)
- Agreeable, kind (NOT quarrelsome or critical)
- Dependable, self-disciplined (NOT careless or disorganized)
- Emotionally stable, calm (NOT anxious or easily upset/stressed)
- Open to new experiences, creative (NOT conventional)

(Scale: 1 = Disagree strongly, 2 = Disagree moderately, 3 = Disagree a little, 4 = Neither agree nor disagree, 5 = agree a little, 6 = agree moderately, 7 = agree strongly)

Appendix D: Instructions and Screenshots

Welcome

This study is hosted by:

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG [https://www.uni-hamburg.de/en.html]

Thank you for participating in our study! Your participation is very important to our research. The study takes about 15 minutes to complete and we ask you to please finish the study in one sitting.

Please read the following consent form before continuing:

I consent to participate in this research study. I am free to withdraw at any time without giving a reason (knowing that any payments only become effective if I complete the study).

I understand that all data will be kept confidential by the researchers. All choices are made in private and anonymously. Individual names and other personally identifiable information are not available to the researchers and will not be asked at any time. No personally identifiable information will be stored with or linked to data from the study.

I consent to the publication of study results as long as the information is anonymous so that no identification of participants can be made.

The study has received approval from the Dean's Office of the University of Hamburg, Germany.

If you have any questions about this research, please feel free to contact us at experiments@wiso.uni-hamburg.de.

To proceed, please give your consent by ticking the box below:

☐ I have read and understand the explanations and I voluntarily consent to participate in this study.

Figure A.4: Screenshot: Welcome and Consent Form

General Instructions

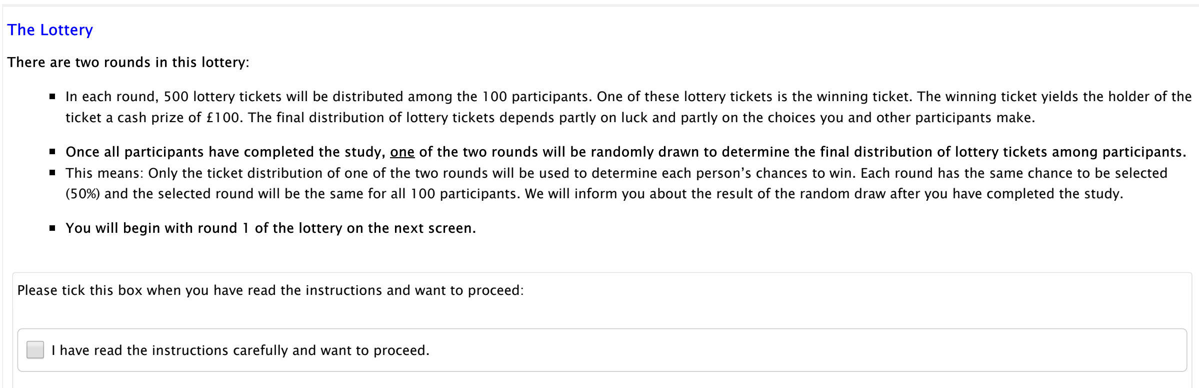
Please read the following instructions *very* carefully before proceeding with the study.

- This study has 100 participants. You are one of them.
- Each participant receives a base payment of £1.50 for completing the study. During the study, you may choose to invest £0.20 of this money. The minimum payment any participant receives is £1.30 (as announced on prolific.ac).
- One participant will receive an extra cash prize of £100. The winner of this cash prize is determined by a lottery. The chance of a participant to win the lottery depends on how many lottery tickets he/she holds at the end of the study.
- The number of lottery tickets you receive depends partly on luck and partly on yours and other participants' choices during this study. The final number of lottery tickets a participant holds ranges from 0 to 10. Each lottery ticket has the same chance to be the winning ticket.
- The winner of the £100 cash prize will be drawn once all 100 participants have completed the study and will be notified one week from now at the latest. You receive all payments through your Prolific.ac account.
- Completion of the study at normal pace should not take more than 15 minutes.

Please tick this box when you are done reading the information and want to proceed.

☐ I have read the information and want to proceed.

Figure A.5: Screenshot: General Instructions ($T\_Pay4Vote$)



The Lottery

There are two rounds in this lottery:

- In each round, 500 lottery tickets will be distributed among the 100 participants. One of these lottery tickets is the winning ticket. The winning ticket yields the holder of the ticket a cash prize of £100. The final distribution of lottery tickets depends partly on luck and partly on the choices you and other participants make.
- **Once all participants have completed the study, one of the two rounds will be randomly drawn to determine the final distribution of lottery tickets among participants.**
- This means: Only the ticket distribution of one of the two rounds will be used to determine each person's chances to win. Each round has the same chance to be selected (50%) and the selected round will be the same for all 100 participants. We will inform you about the result of the random draw after you have completed the study.
- **You will begin with round 1 of the lottery on the next screen.**

Please tick this box when you have read the instructions and want to proceed:

I have read the instructions carefully and want to proceed.

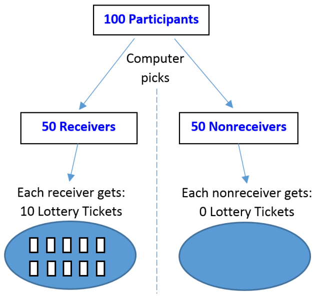
Figure A.6: Screenshot: Instructions about the Lottery

Distribution of lottery tickets

In both rounds 1 and 2, the lottery tickets are distributed in two steps.

Step 1: The computer picks 50 receivers and 50 nonreceivers:

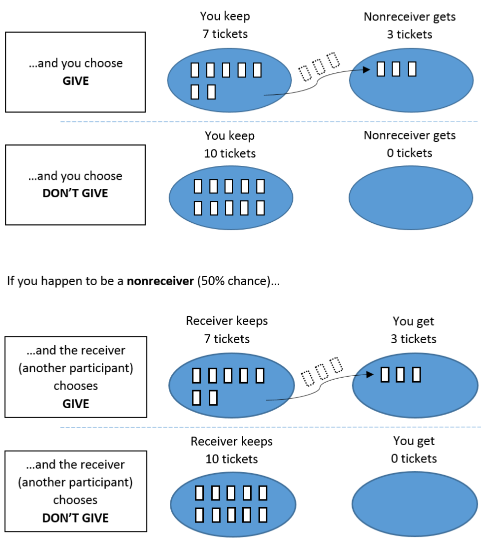
- The computer randomly selects 50 out of 100 participants to be "Receivers". Each receiver gets 10 lottery tickets from the computer.
- The other 50 participants are "Nonreceivers". Nonreceivers get *no* tickets from the computer.
- No participant learns whether he/she has been chosen to be a receiver or a nonreceiver until the end of the study.



Step 2: Participants decide whether they want to share tickets with nonreceivers:

- All participants decide—for the case they happen to be a receiver—whether they want to give **3 lottery tickets** to a nonreceiver.
- This decision (**GIVE** or **DON'T GIVE**) has the following consequences:

If you happen to be **receiver** (50% chance)...



When taking the decision whether to GIVE or DON'T GIVE, you will *not* know whether you have been selected to be a receiver or a nonreceiver. Nor will anybody else. You will receive a message with this information after all participants have finished the study.

If you happen to be a receiver (50% chance), your choice whether to GIVE or DON'T GIVE determines the final number of lottery tickets for you *and* for one other participant.

If you happen to be a nonreceiver (50 % chance), your choice whether to GIVE or DON'T GIVE does *not* play a role. In this case, the choice of *another* participant (who happens to be a receiver) determines the number of lottery tickets that you will receive.

You will take the decision whether to GIVE or DON'T GIVE in both rounds 1 and 2.

Please make sure that you have understood the instructions given above. Once you are sure to have understood the instructions, please tick here to proceed.

☐ I have read and understood the instructions and would like to proceed.

Figure A.7: Screenshot: Instructions about the Distribution of Lottery Tickets

Round 1

Your Choice: Give or Don't Give

If you happen to be a receiver in round 1, **do you want to GIVE or DON'T GIVE** 3 of your 10 lottery tickets to a randomly selected participant who has received no tickets?

- We ask all participants to make this choice.
- If you happen to be a receiver, your choice will be automatically implemented.
- If you happen to be a nonreceiver, your choice does not play a role.
- Your choice remains private and anonymous to other participants.

Click here to be reminded of how lottery tickets are distributed to all participants of this study.

☑ Remind me of the way lottery tickets are distributed.

Lottery tickets are distributed in two steps:

Step 1: The computer randomly selects 50 receivers and 50 nonreceivers. Each receiver gets 10 lottery tickets. Nonreceivers get no lottery tickets. No participant will learn whether he/she has been selected to be a receiver or a nonreceiver until the end of the study.

Step 2: Each participant decides privately whether he/she wants to GIVE or DON'T GIVE 3 lottery tickets to a nonreceiver for the case that he/she happens to be a receiver.

Please choose now:

◉ GIVE 3 lottery tickets to a nonreceiver.

○ DON'T GIVE 3 lottery tickets to a nonreceiver.

Once you have made your decision, please tick below:

☐ This is my final answer. Please proceed.

Figure A.8: Screenshot: Choice $Give_i \in \{0, 1\}$ (Round 1)

End of Round 1

- Your choice in round 1 has been saved.
- You will be informed about the outcome of this round (whether you have been chosen to be a receiver or nonreceiver and how many lottery tickets you hold) via a private prolific.ac-message within one week of the end of this study.

Information about the choices of other people:

- To give you some information on how other people choose in the same situation, below you can see the choices of 5 participants *from an earlier study:*

Participant 1	Participant 2	Participant 3	Participant 4	Participant 5
Don't Give	Give	Give	Don't Give	Don't Give

- Of these participants, **2 (out of 5) chose GIVE** and **3 (out of 5) chose DON'T GIVE**.

Please tick this box when you are done reading the information and want to proceed to round 2:

☐ I have read the information and want to proceed to round 2.

Figure A.9: Screenshot: Information $info_i \in \{2, 4\}$ (following Round 1)

Round 2

A code of conduct

In this round, lottery tickets will be distributed in the same way as in round 1.

Click here to be reminded of how lottery tickets are distributed to all participants of this study.

☑ Remind me of the way lottery tickets are distributed.

Lottery tickets are distributed in two steps:

Step 1: The computer randomly selects 50 receivers and 50 nonreceivers. Each receiver gets 10 lottery tickets. Nonreceivers get no lottery tickets. No participant will learn whether he/she has been selected to be a receiver or a nonreceiver until the end of the study.

Step 2: Each participant decides privately whether he/she wants to GIVE or DON'T GIVE 3 lottery tickets to a nonreceiver for the case that he/she happens to be a receiver.

However, before anyone decides anew whether to choose GIVE or DON'T GIVE, a code of conduct will be set.

- The code of conduct says whether everyone should choose GIVE (⇒**RULE: GIVE**) or whether everyone should choose DON'T GIVE (⇒**RULE: DON'T GIVE**). Only one of the two rules will be implemented for this study.
- Once a rule has been set, all participants decide privately and anonymously whether they want to follow the rule or not.

Your vote: We ask each participant to vote for the rule (RULE: GIVE or RULE: DON'T GIVE) he/she prefers to have implemented as the code of conduct for <u>all</u> participants. Please select a rule below.

○ Vote for RULE: GIVE

◉ Vote for RULE: DON'T GIVE

Once you have made your decision, please tick below:

☐ This is my final answer. Please proceed.

Figure A.10: Screenshot: $Vote_i \in \{Rule{:}Give, Rule{:}Don't\}$ (Round 2)

Round 2

Pay £0.20 to make your vote count

- You just selected RULE: DON'T GIVE as the rule you want to vote for.
- You have to pay £0.20 to make your vote count.

The code of conduct will be determined as follows:

- The rule that receives more votes in total will be implemented as the code of conduct.*
- The votes of participants who pay £0.20 will be counted. Other votes will not be counted.

*Tie Breaker: In case there are exactly the same number of votes counted for RULE: GIVE as for RULE: DON'T GIVE, a coin-flip decides which of the two rules will be implemented.

- If you pay £0.20, your vote for RULE: DON'T GIVE will be counted. If you don't pay, your vote will not be counted.
- This payment is independent of which rule you have selected (and whether or not the rule you have selected will be implemented).
- If you choose to pay, £0.20 will be substracted from your base payment. All other payments are unaffected.
- We ask all 100 participants to make this choice. This means: Only the votes of those participants who pay £0.20 will be counted.

Please choose now:

○ Don't pay £0.20. Your vote will NOT be counted.

◉ Pay £0.20. Your vote will be counted.

Once you have made your decision, please tick below:

☐ This is my final answer. Please proceed.

Figure A.11: Screenshot: $Accept\_Pay4Vote \in \{0, 1\}$ (Round 2, $T\_Pay4Vote$)

Round 2

Receive £0.20 for changing your vote

You just selected RULE: DON'T GIVE as the rule you want to vote for.

- The rule that receives more votes in total will be implemented as the code of conduct.*

*Tie Breaker: In case there are exactly the same number of votes counted for RULE: GIVE as for RULE: DON'T GIVE, a coin-flip decides which of the two rules will be implemented.

For an extra payment of £0.20: Are you willing to vote for the opposite rule instead?

- If you vote for the rule that is opposite to what you wanted to vote for (RULE: GIVE instead of RULE: DON'T GIVE), you will receive an extra payment of £0.20 on top of your base payment.
- This will be your final vote. Only the vote that you cast on this page will be counted.
- We ask all 100 participants to make the same choice. This means: All participants are offered an extra payment of £0.20 to vote for the rule that is *opposite to* what they originally wanted to vote for. Only the final vote of each participant will be counted.

Please choose now:

(●) Accept extra payment of £0.20 and change my vote to RULE: GIVE.

( ) Reject extra payment of £0.20 and keep my vote for RULE: DON'T GIVE.

Once you have made your decision, please tick below:

☐ This is my final answer. Please proceed.

Figure A.12: Screenshot: $Accept\_MoneyOffer \in \{0, 1\}$ (Round 2, $T\_MoneyOffer$)

Round 2

Your choice: Follow the rule or not

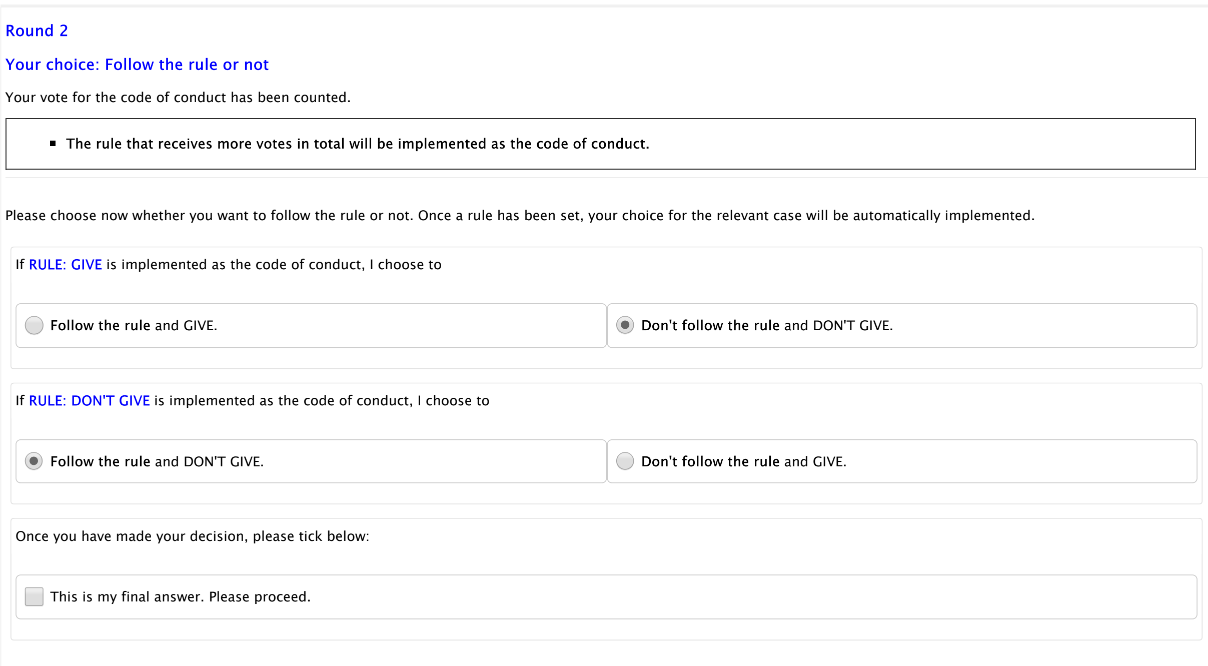
- The rule that receives more votes in total will be implemented as the code of conduct.
- **Only the votes of participants with household income above £40,000 are counted.* The votes of other participants are not counted.**

*according to the household income a participant indicated on Prolific.ac.

According to your prolific.ac profile, your household income is below £40,000:

- **Your vote for the code of conduct has NOT been counted.**

Figure A.13: Screenshot: Information about intervention *Exclude_Poor* (Round 2)



Round 2

Your choice: Follow the rule or not

Your vote for the code of conduct has been counted.

- The rule that receives more votes in total will be implemented as the code of conduct.

Please choose now whether you want to follow the rule or not. Once a rule has been set, your choice for the relevant case will be automatically implemented.

If RULE: GIVE is implemented as the code of conduct, I choose to

Follow the rule and GIVE.

Don't follow the rule and DON'T GIVE.

If RULE: DON'T GIVE is implemented as the code of conduct, I choose to

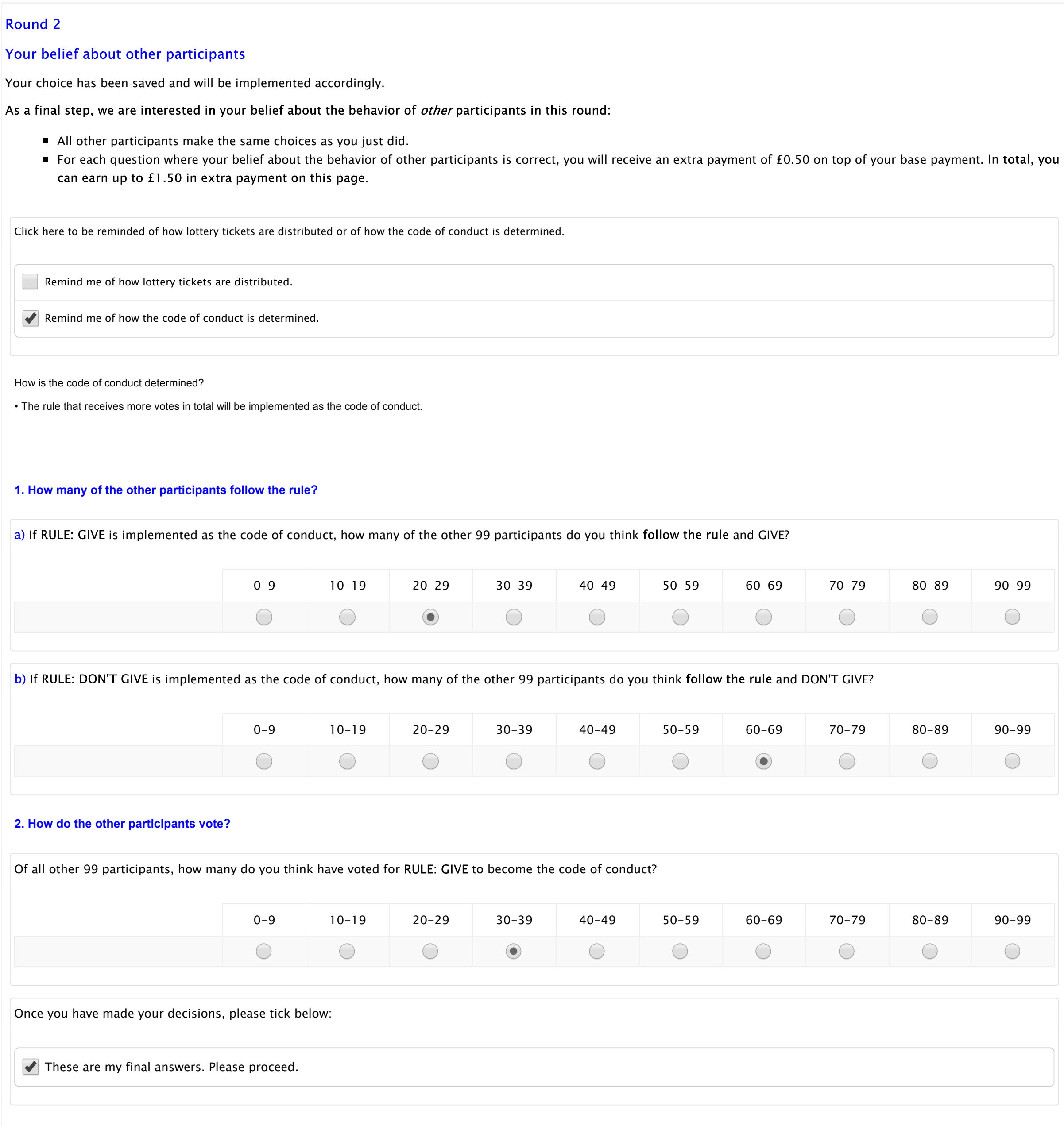
Follow the rule and DON'T GIVE.

Don't follow the rule and GIVE.

Once you have made your decision, please tick below:

This is my final answer. Please proceed.

Figure A.14: Screenshot: $Give_i|Rule \in \{0, 1\}$ (Round 2, $T\_Baseline$)



Round 2

Your belief about other participants

Your choice has been saved and will be implemented accordingly.

As a final step, we are interested in your belief about the behavior of *other* participants in this round:

- All other participants make the same choices as you just did.
- For each question where your belief about the behavior of other participants is correct, you will receive an extra payment of £0.50 on top of your base payment. **In total, you can earn up to £1.50 in extra payment on this page.**

Click here to be reminded of how lottery tickets are distributed or of how the code of conduct is determined.

- [ ] Remind me of how lottery tickets are distributed.
- [x] Remind me of how the code of conduct is determined.

How is the code of conduct determined?

• The rule that receives more votes in total will be implemented as the code of conduct.

1. How many of the other participants follow the rule?

a) If **RULE: GIVE** is implemented as the code of conduct, how many of the other 99 participants do you think **follow the rule** and GIVE?

0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	80–89	90–99
○	○	●	○	○	○	○	○	○	○

b) If **RULE: DON'T GIVE** is implemented as the code of conduct, how many of the other 99 participants do you think **follow the rule** and DON'T GIVE?

0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	80–89	90–99
○	○	○	○	○	○	●	○	○	○

2. How do the other participants vote?

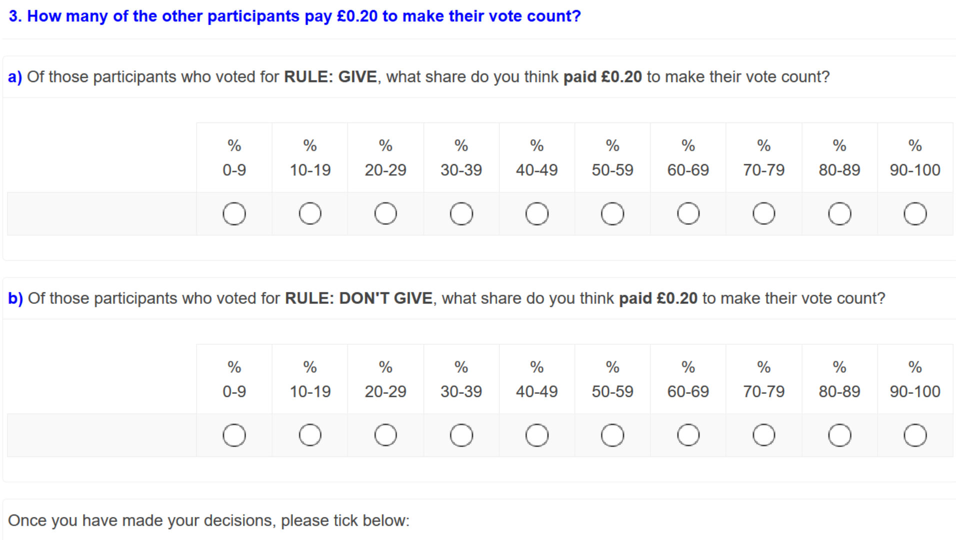
Of all other 99 participants, how many do you think have voted for **RULE: GIVE** to become the code of conduct?

0–9	10–19	20–29	30–39	40–49	50–59	60–69	70–79	80–89	90–99
○	○	○	●	○	○	○	○	○	○

Once you have made your decisions, please tick below:

- [x] These are my final answers. Please proceed.

Figure A.15: Screenshot: Beliefs about Others (Round 2, *T_Baseline*)



3. How many of the other participants pay £0.20 to make their vote count?

a) Of those participants who voted for **RULE: GIVE**, what share do you think **paid £0.20** to make their vote count?

	% 0-9	% 10-19	% 20-29	% 30-39	% 40-49	% 50-59	% 60-69	% 70-79	% 80-89	% 90-100
	○	○	○	○	○	○	○	○	○	○

b) Of those participants who voted for **RULE: DON'T GIVE**, what share do you think **paid £0.20** to make their vote count?

	% 0-9	% 10-19	% 20-29	% 30-39	% 40-49	% 50-59	% 60-69	% 70-79	% 80-89	% 90-100
	○	○	○	○	○	○	○	○	○	○

Once you have made your decisions, please tick below:

Figure A.16: Screenshot: Beliefs about Intervention (Round 2, *T_Pay4Vote*)

Chapter 3

Investments in Impure Public Goods

Abstract

This paper experimentally investigates to what extend the type of risk inherent in social investments influences their attractiveness for investors. In particular, we analyze how risk in the provision of the public benefit and in the financial return to the investor each affect investment decisions separately and how, in addition, their correlation influences investments when both risks are simultaneously present. The results show that the reaction to risk in the private return to the investor and in the public benefit (that is paid to a charity) crucially depends on (1) the correlation of the co-existent risks and (2) the type of the investor. Identifying heterogeneous treatment effects shows a particularly strong reaction of pro-social and risk averse participants to co-existent risks when random draws are independent. It furthermore suggests that less inherently pro-social and less risk averse participants can be attracted to invest in risky impure public goods. The findings not only inform social investments such as crowdinvestments and microlending but also theories of giving under risk: the data rather support models that assume donors (also) care about the impact of their donation on the public good.

JEL Codes: C91, D01, D63, D64, D81, D90, H41

Keywords: impure public goods, giving under risk, crowdfunding, donations, bundled goods, correlated risks, experiment, altruism

3.1 Introduction

We create an experimental design that models investments in bundled investment goods, that may generate a private return to the investor as well as a public benefit. To illustrate, consider an investment in a green technology that will yield the investor a monetary return and with which she also contributes to the reduction of CO_2 emissions and to the mitigation of climate change. In sectors where such bundled goods, or "impure public goods", are being produced, for example the energy and the environmental sector, so-called "crowdinvestments", meaning small-scale investments by small businesses, private investors or regular citizens have recently been gaining importance as a financing alternative for social ventures and environmental projects (Lehner, 2013). According to the annual market report of crowdinvestments in Germany, the overall volume of crowdinvestments in Germany in 2015 was 48.9 million € and the market grew by 169% compared to 2014. As a comparison, in 2011, the total market for crowdinvestments only reached a volume of 1.8 million € . Crowdinvestments in green energy projects are one very popular branch attracting investment of about 6.9 million € in 2015 (thereby growing by 167% compared to 2014). The biggest share was financing of energy efficiency projects, followed by solar and wind energy (Harms, 2016).

These investments are often characterized by the simultaneous presence of risk in the financial investment of the investor and in the provision of the public good that the investment is supposed to generate. In the experiment, we investigate participants' responses to co-existent private and public risk in an investment situation. We focus on two aspects: first, we separate the two domains and introduce risk in the public and in the private domain one-by-one. Second, we vary the relationship between the risks in the two components of the bundled investment. By identifying crowdinvestors' willingness to invest dependent on the type of risk inherent in the investment, this study indicates under which conditions microlending or crowdinvesting might be able to best attract investors.

Recently, the riskiness of crowdinvestments has been debated in the context of new laws that have been proposed to regulate this new industry and to protect private investors from taking too high risks.[1] However, these regulations focus on the private part of the investment only, without taking into account that the investment decision in s bundled good might be determined by both payoffs it generates and that the perceived riskiness

[1] In Germany, the federal government implemented the "Kleinanlegerschutzgesetz" in 2015, which was proceeded by many controversies between the industry and the regulation authorities. One element of this laws is, for example, that an investor who invests more than 1000€ has to prove that he can afford to do so by issuing a self-disclosure (http://www.bundesfinanzministerium.de/Content/DE/Monatsberichte/2015/08/Inhalte/Kapitel-3-Analysen/3-3-kleinanlegerschutzgesetz.html, November 2017).

of the overall investment might depend on the relationship between the co-existent risks.

By exogenously manipulating the riskiness of the private or the public component of the investment, we establish a causal effect of different allocations of risk on average investments as well as identify heterogeneous treatment effects. The controlled decision environment in the laboratory allows to track the choices of different types of investors that differ in their concern for the public good as well as in their attitude towards risk.

We create a series of modified investment games where an investment is linked to giving to a charity. In individual decisions, subjects can allocate tokens from a safe account to one with risky payouts. Between treatments, we vary whether the risky account only generates a return from investment to the investor or whether it additionally generates a public dividend that is paid to a charity outside the lab. The variable of interest is the allocation of risk across these two returns, the private and the public one: the design varies whether one or both returns are risky and whether and how the risks are correlated. We measure how subjects change their investment behavior in response to those risks. In order to understand heterogeneity in treatment differences, we elicit subjects risk and social preferences and conduct a type-based analysis. In an additional part, we elicit participants' *social* risk preferences by identifying their preferences over risky prospects for themselves and for a charity.

The results show no differences in average investments when risk is exogeneously introduced in one component of the bundled investment good compared to a situation with no risk. When risk in the public component is introduced in addition to risk in the private return from a bundled investment good, mean investments significantly decrease if the risks are independent—but not if they are positively or negatively correlated. This decline in average investments is mainly driven by the subgroup of risk averse and pro-social participants. Furthermore, the data suggests that the not inherently pro-social and not risk averse participants can be attracted to invest in risky impure public good.

The implications of risk either in purely private or in purely altruistic decisions has been investigated by several studies on risk preferences and on giving under risk. The treatments in this experiment are based on the so-called investment game as proposed by Gneezy and Potters (1997) which has since been used in many experimental studies as a simple task to elicit risk attitudes of participants (Charness and Viceisza, 2012, Charness and Gneezy, 2012). This game is combined with a dictator game with donations as first experimentally examined by Eckel and Grossman (1996). Dictator games with student recipients have been used to investigate how risk for another person affects giving decisions by Krawczyk and LeLec (2010), Brock et al. (2013) and Freundt and Lange (2017). The authors find that prosociality generally decreases when risk for the receiver is being introduced or does not change significantly. Similar findings have been obtained in public good game experiments (Gangadharan and Nemes, 2009). In an auction experiment by Gueth et al. (2008) where participants state their willingness to accept to forego the payoffs of a prospect that pays money to the decision-maker and another passive participant, participants are other-regarding under certainty but not as much when risk is involved.

Many studies of pro-social behavior under risk are conducted in games where a small number of people interact. In these situations, changes in pro-social behavior compared to environments with certain payoffs can be driven by different preferences as for example fairness preferences over outcomes and over procedures or social comparisons that will not play a role in individual decisions. As I am interested in the interrelation of altruistic concerns and risk attitudes I choose an individual decision framework with giving to a charity – which can be regarded as an approximation of investment decisions in a market setting with many actors where individuals are price takers (an argument also mentioned in the discussion of Falk and Szech (2013) in (Kirchler et al., 2015, 9)).

Exley (2014) investigates risk in donations and finds that the deterring effect of risk is much stronger when giving involves a cost for the decision maker. When making decisions over lotteries for a charity that do not involve own payoff consequences the risk attitudes do not change significantly between charity risk and own risk. Exley (2014) concludes that the reaction to risk in giving is stronger when the risk can serve as an excuse not to give.

In a field experiment among US households, Landry et al. (2006) find that using lotteries significantly increases contributions for a charitable cause compared to simply voluntary contributions which is mainly explained by an increase in participation rates rather than in the magnitude of contributions. However, in their theoretical framework higher contributions with lotteries can be explained by externalities rather than based on (risk) preferences, which is outside the scope of this article. The theoretical model and experimental evidence in Lange et al. (2007) demonstrate the importance of having information about contributors' risk preferences –and heterogeneity of those preferences in the population– for choosing an optimal charity fundraising mechanism.

The question whether people have some concern for the actual impact of their donation on the public good rather than caring about the cost they have to incur when donating relates to the discussions in the literature on charitable giving about rebates and matches and about overhead costs. Charitable giving has been shown to be significantly influenced by the price of giving as shown by the introduction of matches and rebates in experimental studies. For example Eckel and Grossman (2006), Karlan and List (2006), and Scharf and Smith (2010) show that donors increase their giving when rebates are introduced and they increase it even more with matches. The argument that people might care about the impact of their donation (instead of simply deriving utility from the act of donating) relates to a discussion on so-called 'overhead aversion' that donors seem to exhibit in charitable giving. Gneezy et al. (2014) show that large overhead costs lead to lower donations but only if the donors pay for the overhead costs themselves. This can be seen as an example of donors caring about the actual impact of the donation.

The present study relates to this field by informing about whether the cost of donating or the impact of a donation motivate pro-social behavior. Furthermore, it extends the environments being studied in the context of charitable giving to bundled goods and to risk. How charitable giving responds to risks is essential to understand in order for charities to decide whether disclosing to donors how their donation will be used. A donor who cares about the *impact* of their own donation on the public good might want to know whether her donation has been used to cover overhead costs or not and, in case it is used directly for the public benefit, whether it has really been provided as expected. A charity might furthermore want to know how the demand for charitable contributions changes when bundled with private goods and in which circumstances this is beneficial for her.

An impure public good denotes a bundled good that yields a private and a public payoff and consumers derive utility from both its private and its public component. My research project extends existing models (Cornes and Sandler 1994, Kotchen 2005, Chan and Kotchen 2014) and experimental studies on demand for impure public goods. to risky environments by drawing on evidence from the literature on rebates and matches in giving decisions and on giving under risk. In lab and in field experiments, previous studies have found a willingness to pay a price premium for a public benefit bundled with a private consumption good, like organic cotton (Casadesus-Masanell et al., 2009b), certified toilet paper (**?**), charity-linked products (lab: (Frackenpohl and Pönitzsch, 2013), field: Elfenbein and McManus (2010)) or electricity from renewable energies (Kotchen and Moore, 2007). In the latter study, the authors show moreover that altruistic attitudes influence the demand for such goods. Lange et al. (2017) empirically assess the question whether impure public goods might be a substitute for direct donations in the context of climate change mitigation. The survey data does not confirm this hypothesis but rather suggests an overall complementary relationship. Lai et al. (2017) provide a theoretical model analyzing in which situations it is financially profitable for firms and for charities to bundle private with public goods. The concept of bundled goods allows to jointly examine individuals' reactions to risk in the private and the public good, thereby informing about how to model social preferences under risk.

The paper is organized as follows. Section 3.2 outlines the experimental design together with the behavioral predictions. In section 3.3 we present the summary statistics and the regression analyzes on the aggregate and the individual level and the main results. Section 3.4 concludes.

3.2 Experimental Design

The experiment consists of three parts. The main part consists of a series of investment games linked to donations. The remaining two parts are meant to obtain more detailed information on individual preferences. Specifically, part 2 aims at decomposing the bundle such that subjects can freely allocate risks and returns. The design and the results from part 3 are provided in Appendix B.

3.2.1 The bundled investment game

An impure public good is represented in the lab in the following way: Participants get an endowment of 100 units of the experimental currency (ECU) in each treatment. They can choose to transfer an amount $0 \leq x_i \leq 100$ from a private *Account A* to an *Account B*. The nature of the payout from *Account B* differs between treatments. The basic structure builds up on the investment game as originally proposed in Gneezy and Potters (1997). Here, the invested amount x_i is either multiplied by a constant return r^H (high return) or it is lost, $r^L = 0$ (low return). Both outcomes can happen with a probability of 50% (see also Charness and Gneezy, 2010). To create an impure public good in the lab, an investment game is combined with a donation game as originally used in Eckel and Grossman (1996). We modify the investment game such that the return generated in *Account B* is split between the investor and a charity at a fixed proportion. In other words, in the 'impure public good'-treatments, the investment of an individual i may generate a private payoff to herself as well as a public payoff to a charity in case the investment is "successful" (probability 1/2).

Across treatments, the state-dependent private payoff of an individual i is:

$$\pi_s(s_s) = m - x_i + r_s(s_s)x_i \tag{3.1}$$

Accordingly, the public payoff to the charity is:

$$\pi_{ch}(s_{ch}) = r_{ch}(s_{ch})x_i \tag{3.2}$$

where m denotes the endowment and x_i the amount invested by i, r_s and r_{ch} are the state-dependent returns (to self and to charity):

$$r(s_s) = \begin{cases} r^H \; if \; s_s = 1 \\ r^L \; if \; s_s = 0 \end{cases}$$

$$h(s_{ch}) = \begin{cases} h^H \; if \; s_{ch} = 1 \\ h^L \; if \; s_{ch} = 0 \end{cases}$$

In the experiment, we denote the state-dependent returns from investments to the investor and to the charity by r_s^H, r_{ch}^H for high return and r_s^L, r_{ch}^L for low return, which is always equal to zero.

3.2.2 Treatments

The experiment uses a within-subject design with one-shot decisions. The treatments are played in random order so that we can control for order effects in the regression analysis.One decision is randomly chosen for payment after the experiment has been completed. Feedback about the outcomes of the random draws is not given until the end of the experiment. Treatment *DG* is a standard dictator game in which participants can donate a desired amount to a charity instead of to a randomly selected other participant, see Eckel and Grossman (2006). *DG_RCharity* modifies this donation game by making the donation risky: with a likelihood of 50% the amount donated is multiplied by r^H, otherwise the transfer to the charity is zero. In the experiment, the realized low return is $r^L = 0$, the high return is $r^H = 2.6$ and $p = 1/2$. The parameters are chosen such that the returns in the treatments without risk satisfy $\bar{r} = pr^H + (1-p)r^L$. In *DG*, the donation is multiplied by 1.3 so that the expected value of *Account B* is kept constant.[2] Treatment *IG* is a standard investment game as described in subsection 3.2.1 (Gneezy and Potters, 1997, parameters based on Charness and Gneezy, 2010). We first replicate standard settings to ensure that preferences in this sample do not differ significantly from previous experiments so that the results of this study can be related to previous findings in the literature. Importantly, these games elicit participants' risk preferences and social preferences which will serve for classifying subjects' types in the statistical analysis of the 'impure public good'—treatments.

These basic games are extended to investigate the impact of (co-existent) risk in bun-

[2]This reduced price of giving corresponds to a match by the experimenter as analyzed and described in detail for example in Eckel and Grossman (2003, 2006, 2008b).

	Account B		
Treatment	r_s	r_{ch}	EV
DG	-	1.3	1.3
DG_RCharity	-	2.6-0	1.3
IG	2.6-0	-	1.3
IPG_NoRisk	0.65	0.65	1.3
IPG_RCharity	0.65	1.3-0	1.3
IPG_RSelf	1.3-0	0.65	1.3
IPG_RBoth	1.3-0	1.3-0	1.3

Table 3.1: Returns from Account B in all treatments of Part 1, EV=expected value

dled goods. The payoffs of all treatments are summarized in Table 3.1. In all impure public good treatments the high return r^H is split between the investor and the charity at a *fixed* proportion. For simplicity (and to make the donation non-negligible) we chose an equal split in all treatments. In *IPG_NoRisk*, *Account B* generates a bundled return without risk. Thereby, it resembles treatment *DG* with the difference that the return of 1.3 is split equally between the investor and the charity, $\bar{r} = \bar{r}_s + \bar{r}_{ch} = 1.3$ and $\bar{r}_s = \bar{r}_{ch} = 0.65$. This setup can be interpreted as introducing a rebate in the donation game. At the same time, this treatment serves as the benchmark for the remaining *IPG*—treatments with risky returns. *IPG_RBoth* divides the high return of the investment game *IG* equally between the investor and the charity. Thus, the investment x_i may generate a bundled return $r^H = r^H_s + r^H_{ch} = 2.6$, $r^H_s = r^H_{ch} = 1.3$, with $p = 1/2$, zero otherwise ($r^L_s = r^L_{ch} = 0$). With this, it holds that $\mathbb{E}[r(s_r)] < 1$, $\mathbb{E}[r(s_r)] + \mathbb{E}[h(s_r)] > 1$, meaning that it is only worthwhile to invest for people who care about the public component. In order to assess the impact of introducing risk in each dimension separately, we add treatments *IPG_RCharity* and *IPG_RSelf*. In the former, only the return from *Account B* to the charity is risky and the investor receives a risk-free payoff of $\bar{r}_s * x_i$, $\bar{r}_s = 0.65$. As in *IPG_RBoth*, the risky payoff to the charity is $p_{ch} * r^H_{ch} * x_i$. Correspondingly, in *IPG_RSelf* the charity receives a safe payoff and the payoff to the investor from *Account B* is risky. Importantly, the parameters are chosen such that the risk-less returns satisfy $\bar{r}_s = p_s r^H_s + (1 - p_s) r^L_s$ and $\bar{r}_{ch} = p_{ch} r^H_{ch} + (1 - p_{ch}) r^L_{ch}$ and expected returns therefore stay the same across treatments, see Table 3.1.

In the case of risk in both components of the bundled return from investment, the two risks can be either independent random draws or positively correlated or negatively correlated. We integrate all three cases because the interrelation of the two risks determines the overall riskiness of the bundled investment good and is therefore expected to impact individuals' investment decisions. We distinguish *IPG_RBoth_Ind* with independent private and public risk, *IPG_RBoth_Neg* with (perfectly) negatively correlated private and public risk and *IPG_RBoth_Pos* with (perfectly) positively correlated private and public risk. To illustrate, the bundled investment good in *IPG_RBoth_Neg* generates a high return for the investor when the provision of the public good fails and the other way round. *IPG_RBoth_Pos* generates either a high return for the investor and a public benefit or neither, meaning that in case of a financial failure the public good is also not provided. In *IPG_RBoth_Ind*, the successful provision of the public good and the high return for the investor are independent of each other. Taking up the example of investments in a CO_2-reducing technology, *IPG_RBoth_Pos* represents a case in which the project can completely fail such that the private return from investment and the reduction in CO_2 emissions would equal zero. However, one could imagine that the project can be financially successful but fail to provide the public good, or the other way round (*IPG_RBoth_Ind*). This is the case if the two components are driven by different underlying processes. The financial success might be influenced by the financial skills of the manager or the economic situation whereas the environmental success might be determined by biological or technological factors.

3.2.3 Decomposing the Bundle

When investing in or buying bundled goods, their composition, the share of the public component and the riskiness of the bundle, is determined beforehand due to the nature of the product or according to the objectives of the producer. Therefore, it is modeled as being exogenously fixed in part 1 of the experiment. In part 2, however, we elicit the features of an individual's preferred bundle to see how it compares to the fixed bundle in two dimensions, the allocation of risks and the allocation of the return. We will use the information about how the fixed bundle relates to the subjects' *preferred* bundle for analyzing the demand for the bundles offered to subjects in part 1. Participants make three choices in part 2: one on the distribution of the high return, r^H, from investment and two choices on the distribution of risks within a bundle. The additional information about individuals' preferences will be used in the statistical analysis of part 1.

	Outcome A	**Outcome B**
Investor	65+$Transfer_s$	65-$Transfer_s$
Charity	65+$Transfer_{ch}$	65-$Transfer_{ch}$

Table 3.2: Distribution of Risks in *T_Distr_Risk* and *T_Distr_Risk_Ind* (Part 2)

In treatment *T_Distr_Return* ("distribution of return") the whole endowment of 100 ECU of a subject is allocated to one account that pays a high return of $r^H = 2.6$ and a low return of $r^L = 0$, both with a probability of $p = 1/2$ (as in *DG_RCharity* and in *IG*). However, now participants decide *ex-ante* how to split the high return r^H between themselves and the charity in case they win the lottery. Thus, the expected payoff of the investor is $p * (r^H - r^H_{Transfer})100 + (1-p)0$ and the expected payoff to the charity is $p*(r^H_{Transfer})100+(1-p)0$, where $r^H_{Transfer}$ denotes the share of the high return the subject allocates to the charity. This game is a version of a dictator game under risk, tailored to the decision environment of the investment games in part 1. Because the giving choice in this game is state dependent, i.e. the person gives *conditional on winning the lottery*, we expect decisions to not deviate from giving in *DG*.[3] Importantly, the division of the return creates a direct measure of an individual's distance of the fixed bundle to her preferred split of the return.

In treatment *T_Distr_Risk_Ind* (distribution of risk), a subject is presented two lotteries, one for herself and one for the charity. Each lottery has two outcomes of 65 ECU each, outcome A and outcome B, that can be obtained with a probability of 1/2 at the beginning of the task. Each participant decides whether to transfer an amount $0 \leq Transfer^{IND}_s \leq 65$ to the own lottery that will be added to outcome A and at the same time subtracted from outcome B. This means that a higher transfer increases the variance of the lottery while the expected value stays the same. In the same way, the decision-maker can decide to transfer $0 \leq Transfer^{IND}_{ch} \leq 65$ to the lottery of the charity. Comparing the two independent transfer decisions allows to directly elicit a subject's risk attitudes in both domains. Note that subjects play two versions of this task (in random order), as displayed in Table 3.2. In the modified version *T_Distr_Risk* (see T6 in Brock et al., 2013) subjects make the same decision over the own lottery as described above but with one additional constraint: The sum of the transfers has to equal exactly 65 ($Transfer_s + Transfer_{ch} = 65$). Thereby, the amount a subject does not transfer to her own lottery is automatically put on the lottery for the charity. In other words, reducing the variance of a subject's private lottery automatically increases the variance of the lottery for the charity and the other way round. Assuming that an individual has the same risk preferences over both domains (which we test in the experiment), this constraint introduces a tradeoff between allocating the 65 tokens to one of the two lotteries. Securing a safe payoff for herself ($Transfer_s = 0$) implies a risky lottery with outcomes 130 and 0 for the charity (and the other way round).[4]

[3] To make the decisions directly comparable, we could rewrite an individual's donation decision in *DG* in terms of the fraction of the endowment donated.

[4] We implement negatively correlated random draws for the two lotteries (and subjects are informed about this) such that the decision-maker cannot bring the final outcomes closer to each other by "giving risk" as the differences in outcomes between her and the charity stay the same. For example, with $Transfer_s = 0$, the difference in final outcomes is 65; with $Transfer_s = 65$, the difference in final

The difference in transfers $Transfer_s^{IND} - Transfer_s$ with and without this constraint additionally elicits a subject's *social risk preference*, defined as the amount of risk a risk-averse subject is willing to incur to reduce the variance of the charity's lottery, or, as the amount of risk a risk-loving subject is willing to forgo in favor of the lottery for the charity.

The task relates to experiments on risk sharing in group decisions (e.g. Bone et al., 2004). They find that teams fail to allocate prospects in an ex-ante efficient way, taking into account individual risk preferences, i.e. their potentially different individual certainty equivalents for a given prospect. In our simple task, the allocation of risks according to one's risk attitudes over own and the charity's payoff is stripped off the allocation choice and thus simple and in the focus of the decision-maker.[5] Furthermore, due to the constant expected values and the negative correlation of the random draws any concerns about fairness and comparison of *payoffs* should be excluded. Thus, we are confident to obtain a measure of a subject's isolated *social risk preference*, i.e. a measure of her willingness to take on the cost of increased (decreased) risk in order to benefit the charity if she is risk-averse (risk-seeking).[6] One drawback is obviously that this measure can only be used after having established that an individual has the same risk preferences in both domains.

Intuitively, risk preferences can be an important determinant of investment in impure public goods and the treatments in part 2 are designed to shed light on *how* exactly they impact investment decisions. This includes measuring to what extend risk preferences in the private and public domain differ in order to then establish which risk preference dominates the investments and treatment differences in willingness to invest. In addition, the social risk preference provides a measure of pro-sociality with respect to the allocation of risks (instead of (expected) payoff allocations as measured in dictator games). Previous experiments suggested that risk-aversion and generosity might be correlated characteristics in individuals and risk-aversion tends to dominate the latter in giving decisions where the own payoff is risky (Freundt and Lange, 2017). As a consequence, for risk-averse subjects giving under certainty might not predict giving under (private) risk very well and the measure of social risk preference might better capture pro-social motives in the cases where transferring tokens to *Account B* makes own payoff risky.

outcomes is 65; with $Transfer_s = 30$, the difference in final outcomes is $100 - 35 = 95 - 30 = 65$.

[5]The results in Bone et al. (2004) suggest that team members are diverted from the agreement over the choice of prospects and thus fail to pay attention to ex-ante efficiency in the allocation of the chosen prospect. This explains choices in their experiment better than a desire to split prospects equally.

[6]The term has been used in a different way by Gueth et al. (2008) to denote a participants' choice over risky prospects similar to *DG* versus *DG_RCharity* and *IPG_NoRisk* versus *IPG_RCharity*. Subjects in their experiment can give expected payoffs to another participant but do not face trade-offs between the *allocations of risk*–with expected payoffs remaining constant under each allocation choice.

3.2.4 Predictions

We will begin by assuming an additively separable individual utility function that allows for individuals being heterogeneous with respect to their degree of pro-social concern and with respect to their risk preference. It also allows for heterogeneity in the difference between risk preferences over the two components of subjects' utility, the private and the public payout. Such a utility framework is presented in equation 3.3.

$$U_i(\pi_s, \pi_{ch}) = u_i(\pi_s) + \alpha_i v_i(\pi_{ch}), 0 \leq \alpha_i \leq 1 \tag{3.3}$$

The concavity of u_i and v_i describes individual i's risk aversion over each component. α_i describes i's concern for the public benefit generated by her donation to the charity. Π_s and π_{ch} are as defined in equations 3.1 and 3.2. We will consider a decision-maker who exhibits some degree of pro-social concern, i.e. who has positive utility from donating, $\alpha_i > 0$. As noted in section 3.2.1, the parameters are chosen such that a participant who only cares about her own payoff should not invest in the impure public goods. Among the pro-social participants, we expect investments to be influenced by their risk preference and the difference between their risk attitudes over private and public lotteries. Note however, that risk aversion is not sufficient to make comparative statics predictions over the sizes of investments across treatments for the average individual in our sample. A risk-averse player's expected utility from a risky gamble is always lower than her utility from the expected value. However, her optimal investment depends on the marginal utility function and it's shape is determined by the third derivative, her prudence. From a measure of risk aversion alone we cannot determine whether the third derivative is positive or not and thus we can not make predictions on subjects' investments without assuming a specific functional form on participants' preferences. It remains an empirical question how average investments change as a response to the introduction of risks in each treatment, which we will address in this study.

However, for the parameters used in the experiment, $p_s = p_c = p = 0.5$, we can show that an expected utility function with additively separable utility does not predict differences in the investor's utility between the correlation treatments *IPG_Both_Ind*, *Neg* and *Pos*. To see that the expected utility representations for independent risk, perfect positive correlation and perfect negative correlation are equivalent, compare the general case in the first row of equation 3.4 with the rewritten equation for perfect positive (second row) and for perfect negative correlation (third row) for $p_s = p_c = p = 0.5$.

$$\begin{aligned} &p_s u_i(\pi_s^H) + (1-p_s)u_i(\pi_s^L) + p_c\alpha_i v_i(\pi_c^H) + (1-p_c)\alpha_i v_i(\pi_c^L) \\ &= p\left[u_i(\pi_s^H) + \alpha_i v_i(\pi_c^H)\right] + (1-p)\left[u_i(\pi_s^L) + \alpha_i v_i(\pi_c^L)\right] \\ &= p\left[u_i(\pi_s^H) + \alpha_i v_i(\pi_c^L)\right] + (1-p)\left[u_i(\pi_s^L) + \alpha_i v_i(\pi_c^H)\right] \end{aligned} \tag{3.4}$$

It clearly shows that, for the parameters we use in the experiment, the correlation of the random draws of the co-existent risks does not influence an individual's expected utility and we cannot predict differences in investments between the three treatments *IPG_RBoth*. Thus we expect to observe *IPG_RBoth_Ind=IPG_RBoth_Pos=IPG_RBoth_Neg*.

A different reasoning directly builds up on previous findings on the nature of charitable giving showing that donors might derive utility from the act of giving (i.e. warm-glow preferences, Andreoni, 1989) as well as from the level of the public good that is provided by the donations (i.e. altruism, Andreoni, 1989). Andreoni (1989) provides a general formulation of so-called "impure altruism" that allows for both motivations to play a role in donation decisions. However, the introduction of a lottery over the payoffs for the charity drives a wedge between the individual donation and the *impact* on the public good. Therefore, we can expect a donor who is motivated by altruism, to exhibit an adverse reaction to the introduction of risk to the charity's payoff, while warm-glow types of donors should not show a reaction because their utility depends on the act of giving. Thus, an agent with warm-glow type of preferences is expected to behave *as if* risk-neutral over the charity's payoff. Only a donor or investor who cares about her impact can be affected by risk in giving.

3.2.5 Implementation

The experiment has been conducted at the experimental laboratory of the School of Economics and Social Sciences, University of Hamburg in 2015. Participants are students from all departments of the University of Hamburg. The experiment is programmed in ztree (Fischbacher, 2007) and recruitment was administered via hroot (Bock et al., 2014). In total, we conducted 6 sessions with 151 participants in total. The payoffs consist of a 5€ show-up fee plus the payoff from one randomly chosen treatment. The average payoff of a participant was 12.21€ and the average donation was 2.24€ . As part 1 is the part of primary interest in this study, it was always the first part of the experiment, while the order of part 2 and part 3 (Appendix B) was randomized at the sessions level.

Experimental instructions can be found in section 3.4. Donations were made via online transfers to a project from the donation platform BetterPlace.org. Before beginning with the experiment, each participant could choose her preferred project out of a list of three projects on Betterplace.org to which her donations will be transferred if applicable. The list contained one local project for children in Hamburg, one animal protection project and one environmental project in a developing country so that participants could chose their preferred cause to donate for. Choosing the one or other project is assumed to not influence the treatment differences of interest. In order to make it credible that we indeed donate the amounts indicated and to foster trust, three precautionary measures were taken: First, we handed out a leaflet with information about the charity and a link to the webpage betterplace.org that participants were allowed to take home in order to be able to verify the information. Second, a webpage with a documentation of all donations was provided via e-mail to all participants upon the completion of the whole experiment to prove that the transfers have been made (this was announced during the experiment). Third, the experimenter made the individual online transfers together with the cash payments at the end of the experiments such that participants could watch her transferring the donation.

3.3 Results

We will begin with reviewing the main summary statistics of the impure public good treatments to then briefly discuss the observed choices over the decomposed bundles in part 2. A regression analysis in subsection 3.3.3 sheds light on the determinants of individual choices.

3.3.1 Summary Statistics of Investments in Impure Public Goods

To assess subjects' reaction to the introduction of risk in the payout to the charity, we first compare mean transfers in the two donation games, *DG* and *DG_RCharity*. Risk in the payoff to the charity significantly reduces average giving from 24.72 to 21.24 token (out of an endowment of 100 token, $p \leq 0.01$, Wilcoxon rank sum test of equality of distributions (WRS in the following)).[7] The decline in donations when giving is risky replicates previous

[7] In dictator games with student receivers, average giving is close to this result with averages of 25% of the endowment reported in Camerer (2003, 57) and 28.4% of the endowment reported in a meta study by Engel (2011, 6). Grossman and Eckel (1996, 187) find average donations of 11% for a student receiver and 31% for a charity receiver in dictator games.

findings in the literature, such as in a donation game by Exley (2014). In dictator games with a student receiver, Krawczyk and LeLec (2010) and Brock et al. (2013), among others, find significantly lower giving with risk for the receiver, whereas this effect is not significant in Freundt and Lange (2017).[8] Mean investments in *IG* are 44.85 token. 10.9% of the participants chose to invest 0 in the risky asset while 15.8% chose to invest 100. This is very close to previous findings in investment games, see for example Charness and Gneezy (2012). Overall, the distribution of risk preferences and social preferences in the experimental population does not seem to differ much from student samples in previous experiments.

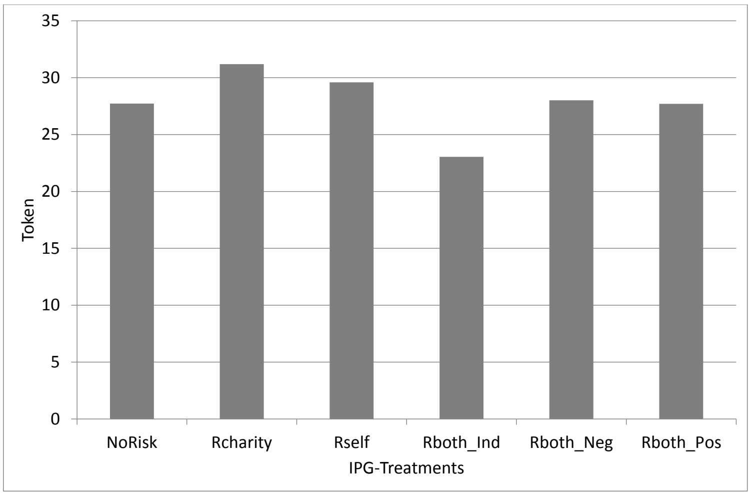


Figure 3.1: Mean transfers in token in each 'Impure Public Good'—treatment (N=151, Endowment=100 token)

Surprisingly, for investments in the bundled goods the decline in mean transfers as a reaction to the introduction of risk in the public return cannot be confirmed. Mean giving under *IPG_NoRisk* and *IPG_RCharity* does not differ significantly, mean transfers in *IPG_RCharity* are even slightly higher, see Figure 3.1. A closer look at the data suggests that this result might not be due to behavior in *IPG_RCharity* but rather due to an underinvestment in *IPG_NoRisk*: As the private payoff component in this treatment functions as a rebate, giving should be higher compared to the standard dictator game. More precisely, whereas a subject pays 1 token for a donation of 1.3 token in *DG*, she can donate the same amount in *IPG_NoRisk* for 0.7 token. The only slightly higher

[8]Note that in dictator games with student receivers different behavioral aspects might play a role that are not relevant in a donation game (such as social comparisons and procedural fairness concerns) such that treatment differences observed in dictator games are not directly comparable to those in donation games with risk in giving.

mean transfer compared to *DG* indicates that, on average, subjects do not appropriately account for this. This argument presupposes that subjects care about the *impact* of their donation rather than only about the amount they transfer (which would be in line with a model of warm glow, Andreoni, 1989). The difference in mean transfers to *IPG_RSelf*, where subjects have to transfer part of their endowment in a risky asset in order to donate, is also small and insignificant.

How do average investments react to introducing risk in the second component of each bundle? Whether or not the introduction of additional risk in the charity's payoff in the presence of own risky returns from the bundle reduces transfers depends on whether the risks are correlated: Mean investments significantly drop from 29.6 in *IPG_RSelf* to 23 in *IPG_RBoth_Ind* ($p \leq 0.001$, WRS), whereas no such decline can be observed for *IPG_RBoth_Neg* and *IPG_RBoth_Pos* (28 and 27.7 token, respectively). Also the differences between *IPG_RBoth_Ind* and the two treatments with correlated risks are statistically significant ($p \leq 0.1$, WRS, for *IPG_RBoth_Neg* and $p \leq 0.05$, WRS, compared to *IPG_RBoth_Pos*). The observation that mean investment with perfectly positively correlated risks are significantly higher than with independent risks and very close to the case with negatively correlated risks is very surprising and will be investigated in more detail in section 3.3.3.

Median decisions (that are less sensitive to outliers) across treatments mostly confirm the above behavioral pattern, however, some treatment differences appear more extreme: Giving declines from 17 token in *DG* to 10 token in *DG_RCharity*. Underinvestments in *IPG_NoRisk* seem even stronger with a mean of 10 token, compared to 20 in *IPG_RCharity* and 25 in *IPG_RSelf*. The results for the treatments with co-existent private and public risk is reproduced: As with a comparison of means, independent private and public risks lead to lower median investments than in the two cases when risks are correlated (15 versus 20 token in *IPG_RBoth_Neg* and *IPG_RBoth_Pos*)—however, in all three cases median investments are lower than in the treatment with only risky private returns and a sure payoff to the charity (25 in *IPG_RSelf*).

Result 5 (Risk in one Component)**.** *We observe no differences in average investments when risk is exogenously introduced in one component of the bundled investment good compared to a situation with no risk.*

The finding that giving is reduced from *DG* to *DG_RCharity*, but not from *IPG_NoRisk* to *IPG_RCharity* is actually surprising. It implies that the finding that risk in giving reduces donations—that has been previously established by Exley (2014) and is in line with behavior in dictator games with student receivers—might not be replicated with bundled goods. As the payoff from a bundled good is divided between a private

Figure 3.2: Frequency of choices in each treatment from top left to bottom right: *DG*, *DG_RCharity*, *IG*, *IPG_NoRisk*, *IPG_RCharity*, *IPG_RSelf*, *IPG_RBoth_Ind*, *IPG_RBoth_Neg*, *IPG_RBoth_Pos*, Number of choices on y-axis and transfer in token on x-axis in brackets of 0-19, 10-19,...,90-100

payoff to the investor and a public payoff to the charity, one might argue that payoff differences become too small to care about and especially the impact of risk becomes negligible. Thus, let us compare the impact of risk in *DG_RCharity* (where donations significantly declined compared to *DG*) and *IPG_RCharity* (where investments were more or less equal to *IPG_NoRisk*). In *DG*, donating 1.3 token imposes a cost of 1 token on the decision-maker. In the risky version *DG_RCharity*, the charity gets 2.6 or 0 token from this cost of 1 token to the investor. In *IPG_NoRisk* and *IPG_RCharity*, the cost of donating 0.65 token is 0.35 (due to the "rebate"). Thus for a cost of 1.05 token the investor donates 1.95 token to the charity or, accordingly, in *IPG_RCharity* she would give a lottery with the outcomes 3.9 and 0. This demonstrates that the impact of risk in giving does not become negligible in *IPG_RCharity* compared to *DG_RCharity*.

To get a more detailed understanding of individual choices behind the averages reported in Figure 3.1, we report the whole distributions of the outcome variable in each treatment (Figure 3.2). In particular due to the high number of zero transfers, the distributions of transfers in the impure public goods look more similar to the distribution of choices in the dictator games than in the investment game (upper right panel). This result has been expected because the parameters haven been chosen such that a purely self-regarding individual should not be willing to invest in the impure public goods. When looking at the share of participants transferring $x_i = 0$ in each treatment, we observe similar treatment differences as in the above comparison of mean and median transfers, with a few interesting differences: First, while *IPG_NoRisk* and *IPG_RCharity* have similar participation rates of about 63% (defined as the share of participants transferring $x_i > 0$), it increases to 75% in *IPG_RSelf*, see Figure 3.2. This might indicate a "'crowding-in"' of players that are attracted by the gamble in *IPG_RSelf*–an interpretation resonating with the findings by Lange et al. (2007) where public good contributions increased when coupled with participation in a lottery. This interpretation would suggest that different types of participants react differently to the treatment variations. As the experiment is designed to elicit information about subjects' risk and social preferences and to follow individual changes in behavior across games we are able to explore such possible explanations in a type-based analysis of impure public good investments below.

3.3.2 Summary Statistics from Decomposing the Bundle

In *T_Distr_Return*, the average subject keeps 1.89 out of the overall return of 2.6, thus giving 27.31% to the charity conditional on winning the lottery. This share is remarkably close to average giving in the standard dictator game *DG* (24.72), thus confirming our hypothesis that giving between the two tasks should not differ.[9]

In the two independent choices in *T_Distr_Risk_Ind*, the average subject transfers $Transfer_s^{IND} = 22.54$ out of 65 token to the her own lottery, leading to an "average" gamble between an *Outcome A* of 87.54 and an *Outcome B* of 42.46 instead of the sure bet. With $Transfer_{ch}^{IND} = 27.13$ token, the average transfer to the lottery of the charity is only slightly higher, leading to a gamble between 92.13 and 37.87 (weakly significant with $p < 0.1$, WSR, median choices are 18 and 25). Participants' preferred risk allocations are thus not significantly different (in line with Exley, 2014), even though risk allocations on behalf of the charity exhibit slightly less risk aversion on average. When imposing the additional constraint of $Transfer_s + Transfer_{ch} = 65$ in *T_Distr_Risk*, mean transfers are $Transfer_s = 29.04$ and $Transfer_{ch} = 35.96$ ($p < 0.05$, WSR). Given that subjects on average prefer a smaller variance in their own lottery, they are on average willing to increase their own risk in order to not increase the variance in the gamble for the charity too much. Subjects did on average take significantly more risk upon themselves ($p < 0.01$, WSR of equality of distributions of $Transfer_s^{IND}$ and $Transfer_s$). This can be interpreted as

[9]Median choices are 2 for the own return and 0.6 for the charity and 30.46% of the subjects kept the whole return of 2.6 for themselves (which is exactly the same fraction of subjects as those who gave zero in the standard dictator game). The only difference in the distribution of choices compared to DG is that the equal split of 1.3 and 1.3 is chosen more often here, which might have been induced by anchoring from part 1.

the attempt of an on average moderately risk averse subject to 'share the burden". This interpretation is supported by the observation that in *T_Distr_Risk_Ind*, 34.44% of the participants chose a transfer of zero to the own lottery, whereas only 14.57%, i.e. about half as many, do so in the case of a trade-off with the charity's lottery. Furthermore, in the latter, about 1/3 (35.75%) of participants chose a transfer to the own lottery between 35 and 30 token, which corresponds to a (slightly biased) equal risk allocation.

Note however, that the above interpretation holds for risk averse subjects but that *T_Distr_Risk* imposes a different trade-off on people with different risk preferences. Thus, we look at subgroups of players in order to better understand individuals' motivations behind the aggregate choice pattern. Overall, we observe that the observations for the whole experimental population hold also for those classified as (moderately) risk-averse in this experiment. However, the subgroup of players that can be labeled "'risk-neutral or risk-seeking"' does not seem be affected by a possible trade-off between the lottery for the charity or their own lottery. Subjects who invest their whole endowment in the investment game in *IG* ($x_i = 100$, 21 / 151 participants) chose on average almost exactly the same transfers, whether this is independent of the charity's lottery or not (39.4 vs 40.8 token). Thus, they largely implement their preferred allocation regardless of the trade-off. Furthermore, the fraction of people transferring the maximum amount to their own lottery remains almost unchanged (17 and 18 subjects out of 151) across the two allocation tasks, suggesting that those participants who prefer the maximum risk in their own payoff do not alter their decision by a trade-off with a lottery for a charity.[10] [11]

The data thus indicate that the not risk averse subgroup of players does not exhibit significant social risk preferences. While there is no a-prioiri reason why risk attitudes and pro-social behavior should be correlated, we do find that risk-aversion and pro-social behavior are negatively correlated among the participants in this experimental study (according to both, choices in part 1, *DG* and *IG*, Pearson corr. coeff.= 0.27, p<0.01, and in part 2, distribution of risks and of returns, corr. coeff.=0.19, p<0.05). On the other hand, we can establish that the average (risk averse) subject exhibits pro-social concerns concerning the distribution of risks and is willing to take on some additional risk on herself in order to not make the charity's payoff less risky, thus exhibiting what we will label *social risk preferences*. The two observations together imply that there might be

[10] As a comparison, those who invested $x_i < 100$ in *IG* choose $Transfer_s = 27.1$ token to the own lottery when mutually exclusive and $Transfer_s^{IND} = 19.8$ when they are independent, thus adjusting the size of the transfer in order to reduce $Transfer_{ch}$.

[11] Note again, the behavior of subjects who hold opposite risk preferences over own payoff and payoff to the charity cannot be meaningfully interpreted in this task. However, we established that the hypothesis of same risk attitudes over payoffs for oneself and for the charity cannot be rejected for the experimental population in *T_Distr_Risk*.

important differences between a person's social preferences and her *social risk preferences.*

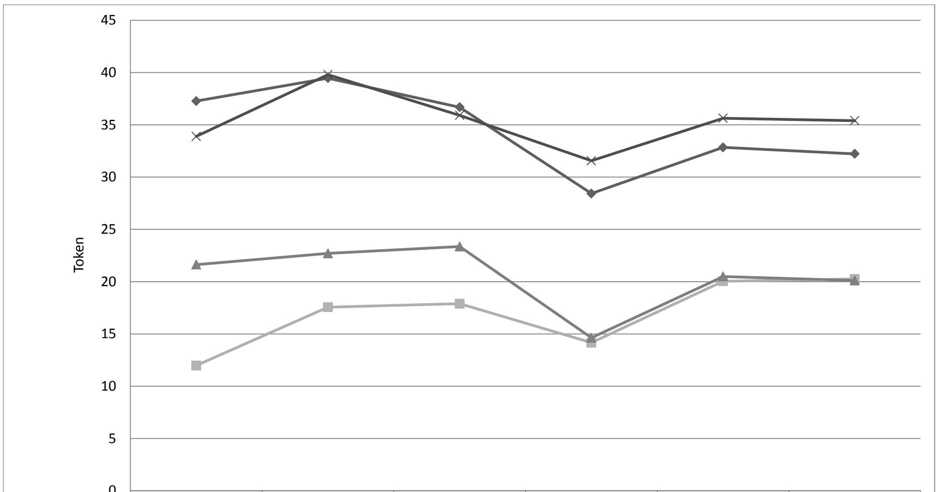
Result 6 (Risk Preferences)**.** *Average choices over lotteries for oneself and on behalf of a charity recipient do not differ significantly. Additionally, participants exhibit significant social risk preferences by being willing to increase own risk to lower the riskiness in the payout to the charity if they are risk averse. In this sample, risk aversion is negatively correlated with pro-social concern.*

3.3.3 Individual Level Analysis of Investments

Based on the control treatments, we define different types of players in order to analyze investments in the bundled goods for the following individual types. Standard dictator games have been shown to be a reliable predictor also for giving under risk (Brock et al., 2013, Freundt and Lange, 2017), but as the *IPG*—treatments are composed from dictator and investment games, we will not use the choices in those games as control variables for defining subgroups for the regression analysis. rather, we will draw on the choices in part 2. In particular, the preferred bundle in *T_Distr_Return* and the $Transfer_s$ in *T_Distr_Risk_Ind* will be used to classify relatively (not) pro-social and relatively (not) risk-averse subjects within this sample.[12] We define a binary variable $Prosocial \in 0, 1$ that is equal to 1 if a person assigned more than the median share of 0.6 (mean=0.71) of the high return of 2.6 to the charity, zero otherwise. In the same way, risk averse types are defined relative to the distribution of choices in the experimental sample by $RA \in 0, 1$, equal to 1 if the person transferred less than or equal the median amount of 18 token, zero otherwise. These measures provide us with similarly large subgroups for analyzing behavior in the impure public good treatments. 94 subjects are classified as $Prosocial = 1$, 57 as $Prosocial = 0$, 76 subjects are labeled as $RA = 1$, and 75 are of the type $RA = 0$.

The differences in magnitudes of transfers by $Prosocial = 1$ types and by $Prosocial = 0$ types in all *IPG*—treatments imply that the investments in impure public goods are regarded as pro-social decisions by participants. The pattern across treatments furthermore suggests that $Prosocial = 0$ types seem to rather increase their transfers as a response to risk in the bundle, i.e. they transfer least in *NoRisk* and almost twice as much in the treatment with the highest risk in *Account B*, *RBoth_Pos*. This observa-

[12] We are mainly interested in how behavior of those who are more risk averse or more pro-social compares to the behavior of the respective other participants rather than making claims about subjects characteristics. Thus, our only claim is that the measure we use to elicit subjects' types allows us to draw conclusions about how subjects rank on this characteristic *relative to the other participants* in the sample instead of claiming that it provides an absolute measure of a person's risk preference or her absolute degree of pro-social concern.



	NoRisk	Rcharity	Rself	Rboth_Ind	Rboth_Neg	Rboth_Pos
Prosocial=1	37.27	39.46	36.68	28.43	32.84	32.22
Prosocial=0	11.98	17.56	17.89	14.16	20.04	20.25
RA=1	21.63	22.7	23.36	14.63	20.49	20.11
RA=0	33.89	39.8	35.91	31.56	35.63	35.4

Figure 3.3: Mean Investments in each IPG-treatment by Type *Prosocial* $\in [0, 1]$ and *RA* $\in [0, 1]$

tion hints at the interpretation that not inherently altruistic people can be motivated to invest in bundled investment goods that provide a public good because they are attracted by the gamble.[13] Figure 3.3 also reveals that the drop in average investments in *IPG_RBoth_Ind*, observed in Figure 3.1, can be found -to some degree- for *every* type of participant. Also, the pattern of average investments across the three bundles in *IPG_RBoth* can be observed for *all* types defined in Figure 3.3. Note that with respect to the *IPG_RBoth*—treatments, participants do not react to the introduction of risk in a fashion compatible with *any* rational risk preferences. No matter if people are on average risk seeking or risk averse, average transfers in *IPG_RBoth_Ind* should lie in the middle between *IPG_RBoth_Neg* and *IPG_RBoth_Pos*. Only a small minority exhibits a behavioral pattern that can be rationalized by an individual's risk preference, i.e.: 10 subjects transferred *IPG_RBoth_Ind>IPG_RBoth_Pos>IPG_RBoth_Neg*, 7 subjects transferred *IPG_RBoth_Neg>IPG_RBoth_Pos>IPG_RBoth_Ind* and 8 subjects transferred *IPG_RBoth_Ind=IPG_RBoth_Neg=IPG_RBoth_Pos*. The robustness of the pattern between the treatments with co-existent risks is very surprising and calls for further investigation in future studies.

[13] Looking only at the "‘selfish"’ participants who gave zero token in DG (46 out of 151 subjects), makes the above pattern indicated by the above subgroup comparison even more prominent.

	(1) Diff	(2) Diff	(3) Diff	(4) Diff	(5) Diff
Prosocial	6.681	4.201	5.285*	6.505*	7.374
	(4.926)	(5.220)	(3.187)	(3.771)	(4.955)
RA	1.230	5.432	5.121	3.504	3.781
	(4.722)	(4.992)	(3.434)	(3.541)	(4.693)
_cons	-6.646	-8.819	0.682	-4.230	-4.606
	(5.342)	(5.821)	(2.439)	(3.118)	(5.556)
N	151	151	151	151	151

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.3: OLS regressions of Type=Prosocial∈[0,1]*RA∈[0,1] on treatment differences between amounts invested: (1) IPG_NoRisk-IPG_RSelf, (2) IPG_NoRisk-IPG_RCharity, (3) IPG_RSelf-IPG_RBoth_Ind, (4) IPG_RSelf-IPG_RBoth_Neg, (5) IPG_RSelf-IPG_RBoth_Pos, robust standard errors were used after conducting Breusch-Pagan tests of homoskedasticity

Comparing investments by $RA \in [0, 1]$ types, we observe that the $RA = 0$ types invest much higher amounts in all *IPG*—treatments. This is in line with what we expected to observe for the risky bundles. However, it is surprising that the same difference can be observed in *NoRisk*.[14] This finding is in line with the negative correlation between risk aversion an pro-social behavior among the experimental sample that we discussed in section 3.3.2.

A parametric test of the treatment effects and the importance of individual characteristics is provided by a OLS regression of the differences between treatments in Table 3.3 and by a random effects regression of treatments on individuals' investment decisions, coded as binary variables equal to 1 if the treatment applies, zero otherwise. The estimation results in Table 3.3 indicate that risk preferences and pro-social attitudes (measured by the subject's preferred bundle) have very little power to explain *differences* in investment decisions across treatments. The random effect regressions are performed for each $Type = Prosocial \in [0, 1] * RA \in [0, 1]$ separately. In all estimations in Table 3.4, treatment *IPG_NoRisk* is the baseline. Confirming the observations from the summary statistics, most treatment differences are insignificant compared to the baseline condition without risk for all subgroups. As before, the only significant decline in investments is observed under treatment *IPG_RBoth_Ind*, however, the estimations in Table 3.4 reveal that this effect mainly occurs among the subgroup of risk-averse *and* pro-social

[14] Taking only the subgroups of individuals into account who invested the full endowment of $x_i = 100$ token in *IG* (21 out of 151), we also do not observe a systematic increase in transfers to *IPGs* with increasing riskiness of the bundles.

	(1)	(2)	(3)	(4)
	Inv	Inv	Inv	Inv
IPG_RCharity	1.214	0.882	2.981	12.52
	(3.790)	(3.932)	(4.445)	(8.541)
IPG_RSelf	1.810	1.618	-2.519	12.26
	(3.600)	(4.103)	(4.288)	(7.610)
IPG_RBoth_Ind	-9.381**	-4.059	-8.404	11.39
	(3.795)	(3.095)	(5.646)	(7.491)
IPG_RBoth_Neg	-4.500	3.000	-4.365	15.52*
	(4.191)	(4.798)	(4.871)	(8.111)
IPG_RBoth_Pos	-3.929	1.441	-5.942	18.35***
	(2.792)	(4.211)	(5.916)	(7.115)
_cons	30.93***	10.15***	42.38***	14.70**
	(4.796)	(3.159)	(5.547)	(6.302)
N	252	204	312	138

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.4: Random effects regressions of investments on treatments, coded as binary variables equal to 1 if the treatment applies, zero otherwise, on investments for different subgroups: (1) $Prosocial = 1$ & $RA = 1$, (2) $Prosocial = 0$ & $RA = 1$, (3) $Prosocial = 1$ & $RA = 0$,(4) $Prosocial = 0$ & $RA = 0$. Baseline is *IPG_NoRisk*.

types, see column (1). Only among $Prosocial = 0$ & $RA = 0$ players, the coefficient is even positive (but insignificant). Likewise, the treatment effect of *IPG_RBoth_Pos* and *IPG_RBoth_Neg* is statistically significant and positive among the $Prosocial = 0$ & $RA = 0$ subgroup. For interpreting the treatment differences in the subgroups it is important to note that giving in the baseline differs notably between the $Prosocial = 1$ (columns(1, 3)and the $Prosocial = 0$ types (columns(2, 4)).[15] Comparing the magnitudes of the coefficients in columns (2) and (4), i.e. treatment effects for the RA=1 and RA=0 types, show that—among those not classified as being pro-social—the not risk-averse participants start investing much higher amounts when risk is being introduced, thus confirming our intuition from the summary statistics.

Note that by testing the effect of treatment on our dependent variable ($investment_i$) for different subgroups we test multiple hypotheses at the same time (see also List et al., 2016). The chance of observing at least one positive significance test due to chance (i.e. making a type I error) increases with the number of dependent tests made simultaneously and we will have to adjust the p-values to take this into account. Therefore, we additionally report the multiplicity-adjusted p-values calculated according to the ”step-down” pro-

[15]Mean investment of each subgroup in IPG_NoRisk: (1):30.93, (2): 10.15, (3): 42.38, (4): 14.7 token.

	(1) Inv	(2) Inv	(3) Inv	(4) Inv
IPG_RBoth_Ind	-11.19***	-5.676	-5.885	-0.870
	(3.481)	(3.528)	(3.648)	(1.417)
IPG_RBoth_Neg	-6.310	1.382	-1.846	3.261
	(3.853)	(4.896)	(2.445)	(3.109)
IPG_RBoth_Pos	-5.738*	-0.176	-3.423	6.087
	(2.946)	(4.014)	(4.470)	(8.181)
_cons	32.74***	11.76***	39.87***	26.96***
	(3.673)	(3.593)	(4.262)	(6.399)
N	168	136	208	92

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: Random effects regressions of treatments, coded as binary variables equal to 1 if the treatment applies, zero otherwise, on investments for different subgroups: (1) $Prosocial = 1 \& RA = 1$, (2) $Prosocial = 0 \& RA = 1$, (3) $Prosocial = 1 \& RA = 0$,(4) $Prosocial = 0 \& RA = 0$. Baseline is *IPG_RSelf*, treatments *IPG_NoRisk* and *IPG_RCharity* are excluded from the sample.

cedure suggested in Romano and Wolf (2005). When applying these corrections, we find that, while the strong significances of the negative treatment effects of *IPG_RBoth_Ind* among the pro-social and risk averse types (column (1)) and of *IPG_RBoth_Pos* in column (4) remain (albeit at a slightly lower significance level), the weak significance for the effect of treatment *IPG_RBoth_Neg* in column (4) vanishes, with the multiplicity adjusted p-value being slightly greater than 0.1.

Result 7 (Crowding-In)**.** *The data suggest that the not pro-social and not risk averse participants can be attracted to invest in risky impure public good.*

In Table 3.5 we repeat the estimations from Table 3.4, but here we take *IPG_RSelf* as the baseline (and exclude treatments *IPG_NoRisk* and *IPG_RCharity* from the sample) in order to estimate the effect of introducing risk in the public component of the bundle *in addition to* already existing private risk. Again, we observe that *IPG_RBoth_Ind* significantly and negatively affects individuals' investments for the $Prosocial = 1$ & $RA = 1$ types, as does *IPG_RBoth_Pos*. Multiplicity-adjusted p-values do not change the above conclusion considerable, the only difference is that the negative treatment effect of *IPG_RBoth_Pos* (column (1)) is no longer significant. The adjustments in both cases clearly show that it is important to consider the dependency of the tests on the four subgroups and adjust the p-values accordingly in order to correctly identify the subgroup-specific effects of treatment. Taking all results together, we mainly observe an effect of introducing risk in the returns from the bundle in the case of co-existent private and public risks with *independent* random draws. This effects is driven by the relatively more risk-averse *and* pro-socially minded types.

Result 8 (Correlated Risks)**.** *When risk in the public component is introduced in addition to risk in the private return from a bundled investment good, mean investments significantly decrease if the risks are independent—but not if they are positively or negatively correlated. This decline in average investments is mainly driven by the subgroup of risk averse and pro-social participants.*

In dictator games, self-regarding players are predicted to give zero, leading to excess zeros in the data (which is also visible in the IPG investments, see Figure 3.2). Using a tobit model interprets these choices as being censored at zero, i.e. it assumes an underlying latent variable that can take negative numbers.[16] This would mean that some of the zero givers would actually *take* something from the charity if they were allowed to, while—from a behavioral point of view—one would prefer to view those types as "selfish" types who act according to the prediction for rational agents. Hurdle models (e.g. Engel and Moffatt, 2014) allow for both interpretations and in addition they allow for analyzing how the probability to be a "zero" type depends on individual characteristics. This models investment decisions as a two-step process. For example, one could imagine that selfish people who just never give cannot possibly be affected by treatments or by their risk aversion but those who are "natural" donors might well be affected by them in their decision *how much* to give.

We are interested to test to what extent treatment differences between the IPG—treatments with different risks are driven by risk aversion—and to what extent introducing risk in the two payoff components affects the decision how much to give. As before, we compare *IPG_RCharity* and *IPG_RSelf* both to the baseline treatment without risks, *IPG_NoRisk* (column (1) in Table 3.6), and we compare the three cases with co-existent risks (*IPG_RBoth_Ind, _Pos, _Neg*) to the case of only risk for the investor (column (2)) in order to observe how introducing additional risk in the provision of the public good affects investment decisions. Due to construction of the panel hurdle model, treatment variables can by definition not predict the likelihood to pass the first hurdle as it captures those participants who always transfer zero across all treatment. The first hurdle test if the participants who never invest can be predicted by their social or risk preferences

[16]The latent variable can be seen as a subject's intended or preferred contribution that is a linear function of the covariates (plus a normally distributed error term). Due to the bound at zero, she can only implement her preferred contribution if it is positive, otherwise she has to give zero, leading to the data pattern that we observe in the experiment.

	(1)	(2)
	Inv	Inv
hurdle		
Tr_Return_Ch	1.562***	25.43
	(0.567)	(267.139)
Tr_Risk_Self	0.00285	0.0148
	(0.008)	(0.011)
_cons	0.601**	0.0751
	(0.266)	(0.269)
above		
IPG_RCharity	4.220	
	(3.385)	
IPG_RSelf	4.825	
	(3.352)	
IPG_RBoth_Ind	-5.939*	-10.35***
	(3.409)	(2.725)
IPG_RBoth_Neg	1.926	-2.749
	(3.369)	(2.687)
IPG_RBoth_Pos	2.113	-2.674
	(3.362)	(2.681)
Tr_Return_Ch	9.921**	10.47**
	(4.589)	(4.267)
Tr_Risk_Self	0.404***	0.245**
	(0.138)	(0.104)
_cons	10.89*	15.76***
	(6.017)	(6.028)
sigma_u		
_cons	26.36***	24.78***
	(3.034)	(1.765)
sigma_e		
_cons	26.30***	21.20***
	(0.857)	(0.859)
transformed_rho		
_cons	-1.038**	0.215
	(0.528)	(0.527)
N	906	604

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.6: Panel double hurdle estimation (based on probit and tobit regressions) of investment decisions on treatments and types, treatments coded as binary variables equal to 1 if the treatment applies, zero otherwise. Baseline is *IPG_NoRisk* in column (1). In column (2), the baseline is *IPG_RSelf*, treatments *IPG_NoRisk* and *IPG_RCharity* are excluded from the sample.

as elicited in part 2 of the experiment.[17] Those players who might revise their giving decision as a response to the treatment modification are captured in the second part of the estimation.

Table 3.6 confirms the negative and significant treatment effect of *IPG_RBoth_Ind* on the magnitude of investments, conditional on investing a positive amount. Preferring a bundled investment that gives a higher share of the return to the charity in the case of a successful investment is significantly and positively correlated with the likelihood to invest in the impure public goods and with the magnitude of the investments. As observed already in Figure 3.3, risk loving participants invest higher amounts.

3.4 Conclusion

This paper investigated to what extend the risk inherent in social investments influences their attractiveness for investors. In particular, we analyzed how risk in the provision of the public benefit and in the financial return to the investor each affect investment decisions separately and how, in addition, their correlation influences investments when both risks are simultaneously present. We identified heterogeneous treatment effects for participants who are more (less) risk-averse and who show greater (lower) pro-social concern compared to their peers in the experimental sample.

The results show no differences in average investments when risk is exogenously introduced in one component of the bundled investment good compared to a situation with no risk. When risk in the public component is introduced in addition to risk in the private return from a bundled investment good, mean investments significantly decrease if the risks are independent—but not if they are positively or negatively correlated. This decline in average investments is mainly driven by the subgroup of risk averse and pro-social participants. Furthermore, the data suggest that the not inherently pro-social and not risk averse participants can be attracted to invest in risky impure public good.

With respect to modeling preferences for giving, our results suggest that—in the context of impure public goods—models of individual giving decisions should take the *impact* of a participant's donation into account. The decisions on the allocation of risk within the decomposed bundles in part 2 as well as the importance of distinguishing between investors with different risk preferences and the strong decline in investments in *IPG_RBoth_Ind* all suggest that at least a fraction of subjects cares about the variance in the donation in addition to their own risk. Thus, models of giving that are based on a feeling of warm-glow

[17]Note that we refrain from the binary classification in order to estimate the probit model of the first hurdle.

from the act of giving (that can be interpreted for example as social-image or identity concerns) fall short of describing the overall observed investment pattern for impure public goods. The findings close an important gap in the experimental evidence on individual's pro-social behavior under risk when more than one risk co-exist, which can inform models of the demand for impure public goods under risk (Lai et al., 2017). By identifying determinants of crowdinvestors' willingness to invest in impure public goods and by outlining important heterogeneities in people's reaction to the risks in our treatments, this study helps to inform under which conditions microlending or crowdinvesting might be able to best attract investors. The observation that relatively less pro-social participants seem to be attracted to participate in risky investments in impure public goods if they are not risk averse seems interesting to further explore with respect to its consequences for charities and social entrepreneurship. In summary, the results show some interesting new aspects but also some puzzles that remain to be solved by further experimental investigations.

Appendix

Appendix A: Additional Analysis

Treatment	Mean	Median	x_i=0
DG	24.722 (27.644)	17	30.46
DG_RCharity	21.238 (26.39)	10	36.42
IG	41.192 (31.14)	40	12.58
IPG_NoRisk	27.722 (34.2)	10	38.41
IPG_RCharity	31.192 (35.26)	20	37.09
IPG_RSelf	29.589 (28.44)	25	25.83
IPG_RBoth_Ind	23.04 (27.27)	15	37.09
IPG_RBoth_Neg	28.007 (29.49)	20	29.80
IPG_RBoth_Pos	27.702 (28.72)	20	27.15

Table 3.7: Summary Statistics of transfers in part 1, mean with standard deviations in brackets (column 1), median (column 2), share of participants transferring zero token ($x_i = 0$) in percent (column 3), N=151

Appendix B: Experimental Instructions

Welcome to the experimental laboratory,

And thank you for participating in this experiment.

Please switch off your phones during the entire duration of the experiment. It is not allowed to communicate with other participants and not following this rule might lead to and exclusion from the experiment as well as from all payments. If you have a question during the course of the experiment, please raise your hand, we will come to your cabin.

Experimental Session

The experiment consists of three independent tasks. The instructions for the second and third part will be distributed and read loud after the previous part has been finished. The decisions you make during one part have no relevance for the decisions and payoffs from the respective other two parts. After the experiment is over, we will ask you to fill out a brief questionnaire.

Procedure

Part 1 consists of 9 decision-making situations, part 2 of 3 decision-making situations and part 3 of 2 decision-making situations. Out of these 14 situations, one single situation will be selected for payment by a random draw and each situation has the same likelihood to be drawn. The payoff from this one situation then determines your final payoff. You will be informed about the result of this random drawn at the end of the experiment.
All decisions will be made anonymously, i.e. neither another participant nor the experimenter can match them with personally identifiable characteristics.

Payouts

Your payout will be partly determined by your decisions, partly by chance. Because your decisions determine what payment you finally receive, it is important that you read the instructions carefully before making a decision. In case something is unclear to you, please do not hesitate to ask!
Your income in the experiment will be calculated in Taler. These will be converted into Euro at the end of the experiment with an exchange rate of

$$100 \text{ Taler}{=}8\text{Euro}.$$

In addition to the payoff from your decision in the experiment, you will receive a show-up fee of 5 Euro. The show up fee will not be used in the experiment. Payment will be made in cash after the experiment is over. The other participants will not be able to see how much money you earned. In some decision-making situations, your choice will affect not only your own payout but also a payout to the non-profit organization BetterPlace.

Description of the organization BetterPlace:
BetterPlace is a web-based donation platform and the largest online donation platform in Germany. Via their website, non-profit and non-governmental organization can collect

money for charitable causes. 100% of the money collected is transferred to the respective project. At the beginning of the experiment, you will be able to choose between three projects that are supported by BetterPlace to which you want to transfer a possible payout. Your decision is binding for the whole experiment.

BP-Project 1: upbringing of orphaned elephants in Kenia
'Aktionsgemeinschaft Artenschutz e.V.' (Action Group Species Protection) looks after young elephants, whose parent often died due to poaching. The young animals get veterinary care, they are brought up and later released into the wild. Thereby, the project fosters the protection of elephants in Kenia, who are threatened by illegal poaching and ivory trade, and contribute to preserving biodiversity.

BP-Project 2: Mentorship for children in Hamburg
'Zeit für die Zukunft - Mentoren für Kinder e.V.' (Time for the Future - Mentorship for children) is a volunteer mentorship programme for the individual support of children and youth ages 6-16 years-old in Hamburg and surroundings. The children are accompanied by a chosen mentor for at least one year, who is available as a caregiver. Through this children from underpriviledged families and children in challenging life situations receive individual support and improved education opportunities.

BP-Project 3: Open-Source small hydropower plant
'Ingenieure Ohne Grenzen e.V.' (Engineers without borders) build 250W small hydropower plants for households in African developing countries. There the state owned electricity grids are poorly developed, so that especially households in rural areas do not have access to electricity. Engineers without borders has developed a micro water turbine, which enables an efficient and environmentally friendly power generation with low fix costs. This ensures an independent power supply for households. The construction manual is available in accordance with the open source principle, to promote on the ground local expertise in the field of environmentally friendly technology.

In case of a bank transfer to BetterPlace after the end of the experiment the money will be transferred for every participant to the chosen project. This will take place simultaneously with the cash payment in your presence via online bank transfer. On an extra sheet, which you'll be able to take home, you can find once more all information in regard to the projects. A few days after the experiment you will receive an E-Mail from the WiSo research lab with a link, with which you can review the original receipts of today's bank transfers to BetterPlace.

PART 1

During the first part of the experiment you will make nine independent decisions. Your decision in one situation does not have any impact on the decision or the payment of another situation. The participants do not run through the decision-making situations in the same order. It is therefore possible, that participants make different decisions at the same time, yet every participant overall runs through the same nine situations.

Decision-making situations

For every decision that you will take, you will receive 100 Taler to your private account, account A, to your disposal. Of these 100 Taler you can transfer an amount chosen by yourself to account B. The leftover Taler will remain in your private account A; these will be paid out to you. The payment from account B will be put together differently in every situation. It can include a private payment to you and/or a payment to the chosen BetterPlace project.
In a few of these decision-making situations the level of the actual payment from account B is dependent on a lottery drawing. Should this be the case, the mechanism of the lottery will be described on your screen. The lottery drawing as described there will be carried out by the computer. Your final payment will be disclosed to you by the end of the experiment.
Attention: As your decisions in every situation have different impacts on your payments, previous to every decision a detailed description of the respective decision-making situation will be displayed on your screen. It is important, that you read carefully through the changing descriptions on your screen, to know the consequences of your respective decision. If questions should arise, contact us, we will get in touch with you!

Do you have any questions regarding the instructions? If not, the experiment begins now. On the first screen a few questions will be posed to you to ensure, that you have understood the process of the experiment. As soon as you have answered these, you will get to the actual experiment.

PART 2

In part 2 of the experiment you will make three completely independent decisions. This means, every decision has no impact on the other decisions and the respective payments. The decision-making situation will be described here and show up in a random order on your screen.

Decision-making situations:

Situation A

In situation A 100 Taler are in your account B (and 0 Taler in your account A). Account B generates two payments, one to you and one to BetterPlace. Both payments are determined by a lottery: with a probability of 50% the payment will be 0 Taler and with a probability of 50% the 100 Taler are multiplied with a rate of return. The overall rate of return in account B is 2.6 and is the sum of Return1 to you and Return2 to BetterPlace. You can now decide the allocation of the overall rate of return by determining Return1 and Return2.

Your Payments	Payments to BetterPlace
Return1*100 or 0	Return2*100 or 0

Your payment with a probability of 50:50 will be Return1*100 Taler or 0 Taler. The payment to BetterPlace with a probability of 50:50 will be Return2*100 Taler or 0 Taler. As the overall rate of return is 2.6, the following must apply:

$$\text{Return2} = 2.6 \quad \text{Return1}.$$

Lottery:

The payments to you and to BetterPlace are determined by one single lottery drawing. So either both you and BetterPlace receive the high rate of return or both of you receive 0 Taler with a probability of 50%.

Situation B

For the decision in situation B 65 Taler will be placed at your disposal. You can use these Taler, to alter the potential payment of two random lottery drawings, one for yourself and one for BetterPlace.

Both, a payment to yourself as well as a payment to BetterPlace are determined by a lottery, in which respectively two possible payments with a probability of 50% can occur. Both of these possible payments are named in the displayed tables 'left payment' and 'right payment'. In the starting situation you would receive 65 Taler if the left payment is drawn, and a payment of 65 Taler, if the right payment is drawn. BetterPlace would in the starting situation receive the same possible left and right payment.

Your possible payments:	
left	right
65+Transfer 1	65-Transfer 1

Possible payments to BetterPlace:	
left	right
65+Transfer 2	65-Transfer 2

By choosing a transfer, Transfer 1, you can change both of the possible payments in your lottery. The chosen value, Transfer 1, will be added to your left payment and consequently automatically deducted from your right payment. The same applies for the possible payments of the lottery for BetterPlace: The Transfer 2 chosen by you will be added to the left payment and consequently automatically deducted from the right payment. Both payments can occur with the same probability of 50%.
For the transfers you should use the above mentioned 65 Taler, which are available to you. Please note the following restriction in the choice of transfer 1 and 2: You have to divide up the full 65 Taler. This means:

$$\text{Transfer 1} + \text{Transfer 2} = 65 \text{ Taler.}$$

Lottery:
The actual payment to you and to BetterPlace are determined by one single lottery drawing. So either you or BetterPlace receive the left payment with a probability of 50%. Should you receive the left payment, BetterPlace will receive the right payment and vice versa.

Situation C
In situation C you will make two separate independent decisions, for which you will each be provided 65 Taler. In a decision you can use 65 Taler to change the possible payments in a lottery for yourself. In a second decision you can use the 65 Taler to change the possible payments in a lottery for BetterPlace.

Your possible payments:	
left	right
65+Transfer S	65-Transfer S

Your payment will be determined by a lottery drawing, in which two possible payments with a probability of 50% can occur. Both of these payments are named in the displayed table 'left payment' and 'right payment'. In the starting situation you would receive a payment of 65 Taler, if the left payment is drawn, and a payment of 65 Taler, if the right payment is drawn. You can now choose a transfer S, which will be added to your left payment and consequently automatically deducted from your right payment. The following applies: Transfer S≤65 Taler.

In a second decision you will make an analogue decision for both possible payments in a lottery drawing for BetterPlace. The chosen transfer B will be added to the left payment and consequently automatically deducted from the right payment. Here applies as well: Transfer B≤65 Taler.

Possible payments to BetterPlace:	
left	right
65+Transfer B	65-Transfer B

Your decision on transfer S has no impact on a payment to BetterPlace and your decision on transfer B has no impact on your payment.

Lottery:
The payment to you and the payment to BetterPlace are determined by one single lottery drawing. So either you or BetterPlace receive the left payment with a probability of 50%. Should you receive the left payment, BetterPlace will receive the right payment and vice versa.

The lottery will be carried out by the computer at the end of the experiment.
If you should not have any questions about the instructions in advance, the second part of the experiment will start now. Before the actual decisions you will receive a few brief questions to ensure, that you have understood the process.

PART 3
You will receive 100 Taler to be at your disposal, which you can transfer to a project 1 and to a project 2. You will divide the full 100 Taler here to both these projects, so both your transfers must add up to 100 Taler all together! There are nine decision-making situations, which will be displayed in a list on your screen. Every row in the list represents one new decision-making situation. In every situation 1 to 9 you will each receive 100 Taler, which you can divide up between project1 and project2 in a row. At the end there

will be a lottery drawing one of the rows, while each row can be drawn with the same probability. The payments from project1 and project2 in that row will be added to your payment.

Situation	R1 Chance von 50%	R2 Chance von 50%	Projekt 1 mit Rendite an BetterPlace	Projekt 2
1	2,6 0	2,6 0		
2	2,4 0	2,6 0		
3	2,2 0	2,6 0		
4	2,0 0	2,6 0		
5	1,8 0	2,6 0		
6	1,6 0	2,6 0		
7	1,4 0	2,6 0		
8	1,2 0	2,6 0		
9	1,0 0	2,6 0		

Figure 3.4: Screenshot Part 3

In the screenshot you can see such a list, as it will appear on your screen. In project 2 your transfer with a probability of 50% will be multiplied with a rate of return R2 of 2.6, which means you will receive a payment of 2.6 times the amount of Taler in project 2. With a probability of also 50% you will receive a payment of 0 Taler.
The transfer in project 1 generates two payments, one to you and one to BetterPlace. For your private payment from project 1 you will receive a transfer with a probability of 50% multiplied with a rate of return R1, which means you will receive R1 times the amount of Taler from project 1. With a probability of also 50% you will receive a payment of 0 Taler. The rate of return R1 changes from row to row. The lowest payment, which can be drawn with a probability of 50%, is always 0 Taler.
There will be 2 different lists, which will appear after one another on your screen. The payment to BetterPlace from project 1 is in these two lists composed differently.
In both lists you will divide up 100 Taler between a project 1 and a project 2. As previously described in the chapter 'procedure', one of these two lists or a situation from part 1 or 2 can be relevant for your final payment.

Do you have any questions in regard to the instructions? If not, the third part of the experiment will start now. In advance to the actual decisions you will again receive a few brief questions to ensure, that you have understood the process.

Appendix C: Additional Treatment (Part 3: Competition)

Part 3 of the experiment extends the previous investigations of bundled investment goods by introducing a simple form of competition. In a price list participants can allocate 100 token to either an impure public good or to a purely private investment good (projects 1 and 2 in Table 3.8). The private return of the latter diminishes in each row of the list such that participants have to give up more of the private return in order to generate the public benefit. As subjects go down the price list they have to sacrifice more of the private component of the bundled good for the public benefit, see Table 3.8.

Number	**R1** **Chance of 50%**	**R2** **Chance of 50%**	**Project 1** (IPG)	**Project 2** (PrivG)
1	2.6 0	2.6 0		
2	2.4 0	2.6 0		
3	2.2 0	2.6 0		
4	2.0 0	2.6 0		
5	1.8 0	2.6 0		
6	1.6 0	2.6 0		
7	1.4 0	2.6 0		
8	1.2 0	2.6 0		
9	1.0 0	2.6 0		

Table 3.8: List Choice (Part 3)

In the price list in Table 3.8 each row is a new decision situation and each row can be drawn for payment with equal probability. The return of the private good, R2, equals the return in *IG* and probabilities are always 1/2. The private return of the impure public good, R1, equals this return in row 1 and subsequently the high return diminishes in steps of 0.2 down to 1.0. Thus, behaving in a prosocial manner becomes more and more costly. Note that in row 1, the public component is complementary, i.e. a public benefit is added "for free" to the private good. Two lists are presented to participants in random order: In one the public benefit of the impure public good is riskless (return of 0.65) and in the other a risky public component is added (return of 1.3 or 0).

The private returns of the two options are exactly the same in row 1 such that individuals who do not care about the public component should balance their portfolio and allocate 50 token to each project as random draws are independent. In the subsequent rows they should allocate less and less to project 1 with the diminishing return depending on the degree of their risk aversion. Prosocial subjects who are concerned about the public benefit should start at a higher level of transfers to project 1 in row one and then diminish their transfers accordingly.

The results can give insights on the question how firms can benefit from choosing crowdinvesting as a financing mean for impure public goods compared to investors who do not take the altruistic benefit into account. It indicates how much lower they can possibly set the private interest rate in the presence of a public benefit when potential crowdinvestors have also a purely private investment option.

We find that mean allocations to project 1, the impure public good, are declining with the declining private return in each row in both lists. Allocations in the two lists with a safe or a risky public benefit do not differ systematically.

20 (20)[18] participants started by allocating 100 token to Project 1. 17 (20) participants still allocate more to project 1 with the public benefit in the last row than to the purely private project 2. 114 (114) participants never allocated zero tokens to project 1, which shows a strong concern for the benefit of the charity. 82 (90) subjects allocated 50 -50 in row 1. Out of those, 3 (3) subjects followed the prediction for risk-neutral and selfish agents and chose 50-50 in row 1 and then transferred zero to project 1 from row 2 on in both lists. Thus, most of the other subjects might have had additional motives for their allocation choice than pure portfolio balancing. 38 (37) subjects deviated from the 50-50 allocation to favor the charity and allocated more than 50 token to project 1 in row 1.

As the random draws are independent, in the first row all selfish subjects should evenly allocate their budget to the two projects to balance their portfolio. In the remaining rows, risk-neutral and selfish subjects should switch to project 2 from the second row on. Risk-averse subjects are assumed to slowly shift their allocations more towards project 2 until they allocate zero to project 1. Subjects who have a preference for donating to the charity should start at a higher level of allocations to project 1 in row 1, i.e. they deviate from the optimal 50-50 allocation in favor of the charity. Then they also allocate less to project 1 with the diminishing return R1, but might still allocate a nonzero amount in the last row, depending on the strength of their social preference. Thus, whereas the social preference affects the level of allocations to project 1, the individual risk preference leads them to shift

[18] I present the number from the price list with a risky public benefit in brackets.

allocations more or less slowly. In order to test this prediction I estimate a random effects model with transfer to project 1 as the dependent variable. Social preference (giving in *DG*), return R1 and risk preference (investments in *IG*) are included as covariates plus an interaction term of R1 and risk reference, see Table 3.9. A significant interaction effect indicates that the slope of the continuous variable is different for different levels of risk aversion. The results show that the coefficient of R1 is positive and statistically significant (p≤0.05). The coefficients of risk attitude and of pro-social concern are highly significant and, due to how they are defined, of opposite signs: The more pro-social and the more risk averse, the more they allocate to project 1. Interestingly, the coefficient of the interaction term is positive and significant (p≤0.05), which supports the above hypothesis.

	(1) Project 1
Giv_DG	0.299***
	(0.056)
R1	5.167**
	(2.335)
Inv_IG	-0.487***
	(0.110)
R1*Inv_IG	0.181***
	(0.061)
_cons	29.74***
	(5.332)
N	1359

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.9: Allocations to project 1 in LC (RE)

Chapter 4

On the voluntary provision of public goods under risk: the role of risk in private or public returns.[1]

Abstract

This paper studies how risk can impact the successful provision of public goods. In a laboratory experiment, we identify the impact of risk in the private and public dimensions of individual investments: Creating variants of a public good game, we separate the return a subject's contribution generates for herself vs. the return her contribution generates to others. Risk is then introduced in either one or both dimensions of the return. We find a detrimental effect of risk on public good provision when returns in both dimensions are risky and positively correlated or independent. A negative correlation, however, provides an insurance effect and leads to more stable investments. Disentangling the impact of risk in each dimension, we find that investments particularly respond to risk in the public return dimension.

JEL Codes: C91, D64, D81, H41

Keywords: public goods, giving under risk, correlated risks, social investments

[1] This chapter is co-authored by Andreas Lange (University of Hamburg).

4.1 Introduction

Most investment decisions are inherently subject to uncertain returns. This applies to both, private and public benefits from investments. Addressing these risk, several papers have experimentally investigated how voluntary contributions to public goods are impacted by uncertainty (e.g., Dickinson, 1998; Levati and Morone, 2009; Théroude and Zylbersztejn, 2020). Most of this literature focuses on situations where the uncertain benefits accrue to all. In this paper, we argue that it is important to investigate settings that separate the impact of a subject's decision on her *own* payoff from the return that *others* receive from this investment. Based on a laboratory experiment, we explore how pro-social behavior depends on the risk in these two dimensions of the investment as well as their interactions.

We consider this a realistic setting for a wide range of applications: Contributions to impure public goods (e.g., Andreoni, 1990; Cornes and Sandler, 1994; Kotchen, 2005; Chan and Kotchen, 2014) involve private and public benefits, yet this literature does typically not consider the impact of risks. Social investments often involve risks in private and social returns, yet they can interact in diverse ways:[2] Investors may get a fixed private return or get repaid depending on the success of the project. The public benefit may be positively correlated with the private benefits. For example, imagine environmental benefits from a green technology investment that realize together with the financial success. In other instances, the realization of public benefits may partly depend on exogenous factors, whereas the private return to the investor may be secured by financial and managerial skills. In this case the risks might be imperfectly correlated or independent. Conversely, a negative correlation between private and public returns may result when monetary investments are pledged, yet only are deducted if the project materializes.

In this paper, we investigate the importance of these channels through which risks impact voluntary contributions, i.e. investments. Inspired by the example of social investments, we compare different correlation structures of private and public returns. For this, we modify the well-established workhorse of a public good game by separating the returns of an individual contribution to herself vs. others.

First, we identify the effect of the *simultaneous* presence of risks in both private and public returns on contribution choices over time. When risks in the two dimensions are

[2]The trend to use "repayable finance to achieve a social as well as a financial return" has been widely recognized in Western societies (e.g., Warner, 2013, p.5). In the U.S., for example, social investments have seen an estimated growth of 33% from 2014 to 2016 alone and amounted to 8.72 trillion dollars at the beginning of 2016, thereby benefiting charities as well as social enterprises (see Voorhes and Humphreys, 2016).

positively correlated or independent, contributions decrease to an similar extent. This suggests that subjects' behavior might be particularly driven by the downside risk, i.e. by the possibility that investments may generate neither a private nor a public return. Conversely, an insurance effect arises when both returns are negatively correlated which stabilizes investments.

Second, we disentangle the importance of risk in either dimension. We find that, even though investments are slightly reduced when private returns are risky, investments particularly respond to risk in the public return dimension. Our findings thus suggest that a reduction of risk in the social domain is particularly important, both in presence and in absence of risk in private returns.

Third, we show that the treatment differences are particularly driven by risk-averse participants who more strongly react to the riskiness of returns, particularly in the private returns dimension.[3] and show that treatment differences are particularly driven by risk-averse types.

Our study is related to several different strands of the literature. In non-strategic two-person interactions, the impact of risk on giving decisions has been shown by, e.g. Dana et al. (2007), Andreoni and Bernheim (2009), Krawczyk and LeLec (2010) and Brock et al. (2013).[4] Within the literature on charitable donations, for example, Potters et al. (2005, 2007) and Sleesman and Conlon (2017) discuss the importance of revealing information about the charity's quality, i.e. about its ability to convert donations into impact. The impact of risks also has been investigated within public good environments: an early study by Dickinson (1998) shows reduced giving when returns from the public good become risky. Similarly, Levati and Morone (2009), Levati and Morone (2013), Stoddard et al. (2015), Björk et al. (2016) and Théroude and Zylbersztejn (2020) consider voluntary contributions with risky returns. While all these studies consider benefits from the public good that accrue equally to all subjects. Differently, Brennan et al. (2008)

[3]Several studies in economics compare risk attitudes when making a decision about own payoff vs. the payoff of another person or the payoff of the group. Evidence of the arising biases appears mixed. Harrison et al. (2013) find more risk-aversion in groups compared to individual decisions, other studies find mixed or insignificant results (e.g. Rockenbach et al., 2007; Baker et al., 2008) or results that depend on the type of risk (e.g. Shupp and Williams, 2008). Comparing risk attitudes about own payoff vs. the payoff of *one* other person, again evidence is mixed (e.g. Pahlke et al., 2015; Agranov et al., 2014; Eckel and Grossman, 2008b; Chakravarty et al., 2011; Faro and Rottenstreich, 2006). A meta-analysis by **?** including 49 studies in economics, psychology and medicine finds no overall difference between making decisions over risky prospects for oneself versus another person.

[4]Overall, their findings suggest a deterring impact of own and other's risk on prosocial behavior, although the evidence especially concerning risk for the other person is mixed. This effect appears to depend on which possible motivations drive behavior in the specific experimental design. Those include for example the distinction between concerns for ex ante vs. ex post payoff comparisons (Krawczyk and LeLec, 2010; Brock et al., 2013) or self-deceptive behavior (Dana et al., 2007; Andreoni and Bernheim, 2009; Exley, 2014).

consider an asymmetric setting in a 2-player game where the return from the public good is risky only for one contributor.

Perhaps closest to our study, Gangadharan and Nemes (2009) and Stoddard (2017) conduct lab experiments in which the return from the public account is risky for all group members or the return from the subject's private account, i.e. the amount she did not invest in the public good, is risky. In contrast, our experiment always keeps safe the non-invested amount, i.e. the private account, while introducing risks in the private and public components of the *return* from investment. We consider this to be a relevant risk structure for typical investment situations. By considering different variants of risky private and public returns, our setting thereby allows to decompose the reasons *why* individual contributions to public goods are typically lower when returns are risky.

The remainder of the paper is structured as follows: in section 4.2.1, we introduce the experimental treatments. The experimental procedure is reported in 5.2 and predictions are discussed in section 4.2.3. Results are presented in section 5.3, before we conclude in section 4.4.

4.2 Experimental Design

In all treatments, subjects play ten rounds of a modified public good game. The treatments vary the risk in private and public returns as described in section 4.2.1. Section 5.2 then details the procedure and implementation, before we outline the behavioral predictions in 4.2.3.

4.2.1 Treatments

In groups of four, subjects make symmetric and simultaneous decisions on how much of their private endowment to invest in a public account. Individuals' investments may generate both a private return to the player herself, as well as a return to other group members.

Specifically, an investment x_i into the public good by player i triggers a *private* payoff of rx_i to herself and a *public* return of hx_i for each of the group members. That is, while in the typical public good game private (marginal) returns coincide with the (marginal) returns to others ($r = h$), we separate these two dimensions. The design relates to Goeree et al. (2002) but additionally introduces risk in the respective dimensions.[5] This

[5] Goeree et al. (2002) employ this mechanism to disentangle altruistic motivation from noisy behavior in a public good game. They call the two returns from an individual's contribution an “internal return” for oneself and an “external return” from/to the other group members.

separation allows us to isolate the impact of risks in private and public returns.

The state-dependent payoff of an individual i within four-player variants of the public good game is thus given by:

$$\pi_i(s_r, s_h) = m - x_i + r(s_r)x_i + h(s_h)\sum_{j\neq i} x_j \tag{4.1}$$

Here, m is the initial endowment and x_i denotes the investment by individual i. s_r and s_h reflect the states of nature that determine the private and public return, $r(s_r)$ and $h(s_h)$ respectively.[6] Due to the symmetry of the game, each player benefits from the public return of investments by other players, generating the payoff component $h(s_h)\sum_{j\neq i} x_j$.

We denote the expected returns by $\bar{r} := \mathbb{E}[r(s_r)]$ and $\bar{h} := \mathbb{E}[h(s_r)]$. To satisfy the social dilemma, we assume that $\bar{r} < 1$, $\bar{r} + (n-1)\bar{h} > 1$: while it is socially desirable in expected payoff terms to invest the full endowment, this does not pay out at the individual level.

Our baseline treatment *NoRisk* is payoff-equivalent to the standard public good game with an MPCR of 0.5: each token invested by a player generates half a token to the player herself ($r = \bar{r} = 0.5$) as well as to each of the other players ($h = \bar{h} = 0.5$). With 4 players, an individual contribution to the public good is thus multiplied by 2 before being distributed among all group members. All other treatments implement identical expected returns in both dimensions ($\bar{h} = \bar{r} = 0.5$), but introduce risk in the public or the private return or in both.

For this, we allow for two different states $s_r, s_h \in \{0, 1\}$ and denote

$$r(s_r) = \begin{cases} r^H & \text{if } s_r = 1 \\ r^L & \text{if } s_r = 0 \end{cases} \qquad h(s_h) = \begin{cases} h^H & \text{if } s_h = 1 \\ h^L & \text{if } s_h = 0 \end{cases}$$

with $h^H > h^L$ and $r^H > r^L$. Again, we keep a symmetric structure in the experiment when choosing the parameters

$$r^H = h^H = 1, \qquad r^L = h^L = 0$$

and we assume that the states are equally likely, i.e. high or low returns result with a probability of 50%. In the main treatments, the random draws are executed at the group level, i.e. either all or none of the players of a group get a return in the respective dimension. This prevents concerns about ex-post inequality in the returns from the public

[6] The private return bears similarity to the functioning of a rebate in the charitable giving literature (e.g., Eckel and Grossman, 2006; Karlan and List, 2006). This literature does not consider risk.

good to influence contribution decisions.

Treatment	**Private Return**	**Public Return**
NoRisk	$\bar{r}$	$\bar{h}$
BothRisks (Ind,Pos,Neg)	r^H or r^L	h^H or h^L
PrivateRisk	r^H or r^L	$\bar{h}$
PublicRisk	$\bar{r}$	h^H or h^L

Table 4.1: Treatment Overview

With simultaneous risks in both dimensions, the overall riskiness of the investment and the effect on the investor's decision depend on the *interaction* of the two risks. We distinguish three possible cases: the risks can be independent ($BothRisks_{Ind}$), (perfectly) negatively correlated ($BothRisks_{Neg}$), or (perfectly) positively correlated ($BothRisks_{Pos}$).

Treatment $BothRisks_{Pos}$ is equivalent to a public good game with risky MPCR: own and public return are identical and coincide for all subjects. Given our parameter choice of $r^L = h^L = 0$, we can interpret this situation as the public good *not* being provided with a probability of 50%. $BothRisks_{Neg}$ resembles a situation where subjects pledge to invest a particular amount for an envisioned project. If the project materializes (with 50% chance), payments are enforced ($r^L = 0$) and a public return is generated ($h^H = 1$). If the project does not materialize, no public return results ($h^L = 0$), but the pledged amount is fully returned to the investor ($r^H = 1$). In addition, in $BothRisks_{Ind}$ we consider the case where both private and public returns are independent, simplifying a situation where the underlying factors that determine whether the investment is successful are independent for the private and the public return.

We further use the decomposed returns to isolate the extent to which risk in either return impacts investment decisions, i.e. we consider treatments *PrivateRisk* with risk *only* in the private return ($r^L = 0, r^H = 1, h = 0.5$), and *PublicRisk* with risk *only* in the public dimension ($r = 0.5, h^L = 0, h^H = 1$). We again chose to implement the random draws at the group level to capture a situation where the public project either fails or is successful, in which case it generates returns to all players.

While this appears a reasonable feature for real-life investments, it comes at a cost in our symmetric four-player public good game environment: when investments by player i result in successful giving to others, the investments of the other three players also generate a return to i. Thus, own final income of a player and the return of her own investment to others are positively correlated. As a robustness check, a final treatment $PublicRisk_{Ind.Level}$

implements independent individual random draws for each group member. By independently determining the return from the investment by each player, this treatment controls for (expected) wealth effects by breaking the positive correlation between own income and the success of own giving decisions.

4.2.2 Procedure

The experiment consists of two parts, a repeated public good game in the variants described in section 4.2.1 in Part 1 and two risk preference elicitation tasks in Part 2.

4.2.2.0.1 Part 1 In Part 1, participants make investment decisions over ten periods in a partner matching under one of the treatment conditions (between-subjects design). Each group consists of four players. In each round, a player receives an endowment of 100 Tokens in her private account, called "*Account A*", and decides how many of these to "*transfer*" into another account, "*Account B*". Account B then determines the returns to herself and the other group members. At the end of each round, feedback on the aggregate decisions by the other players in a subject's group is given, whereas random draws on the returns from investments are only drawn after all decisions have been made.[7] One round is randomly chosen for payment. Before the beginning of the experiment, subjects are asked to answer a set of control questions covering several possible situations in the experiment and, in order to create common knowledge of participants' understanding, they are only allowed to proceed after having answered all questions correctly.

4.2.2.0.2 Part 2 In Part 2, we use the simple risk elicitation task by Eckel and Grossman (2008a) and Dave et al. (2010) in which subjects choose one of six lotteries, summarized in the "outcome" columns of Table 4.2. All lotteries give payoffs A or B with a probability of 50% each. The last two columns of Table 4.2 were not shown to subjects but show that expected value and standard deviation of the lotteries increase from the top to the bottom. A very risk-averse individual should thus choose lottery 1, a risk-neutral person is predicted to choose lottery 5 or 6 with the highest expected value, and risk-seeking subjects may choose lottery 6. Participants are asked to make this choice twice in random order: one decision only matters for their own payoff (choice L_{own}) and the other

[7] By letting subjects play the game over ten periods, we allow for behavior to converge towards some equilibrium (e.g., if players show reciprocal behavior) and expect a typical downward trend of public good contributions. Informing players only about others' contributions prevents creating noise through different realizations of the random draws on the return from investments in each group in each round. Such noisy payoffs have been shown to possibly affect subjects' strategic learning (e.g., Bereby-Meyer and Roth, 2006), which we want to exclude as a possible explanation for treatment differences to *NoRisk*.

	Outcomes			
	A	B	EV	SD
1	56	56	56	0
2	48	72	60	12
3	40	88	64	24
4	32	104	68	36
5	24	120	72	48
6	4	140	72	68

Table 4.2: Lottery list in part 2. Each row is one lottery with equally likely outcomes A and B. EV denotes expected value and SD the standard deviation of the respective lottery, both were not shown to participants.

decision determines the lottery affecting the payoff of all members of their group (choice L_{group}). After all choices have been made, a random draw on the group level determines which of the two choices is relevant for payment.[8] For our discussion of results, we use the lottery choices for own payoff to classify subjects who choose 1 through 4 ($L_{own} \leq 4$) as risk-averse (RA) and those who choose 5 or 6 ($L_{own} \geq 5$) as non-risk-averse (NRA).

4.2.2.0.3 Implementation The experiment was conducted in the experimental laboratory of the Faculty of Business, Economics and Social Sciences, University of Hamburg, Germany, in 2015. We conducted 14 sessions with students from all departments of the University of Hamburg. The total number of participants is 336 with 48 subjects per treatment (24 per session). The experiment is implemented using ztree (Fischbacher, 2007). Tokens are converted into Euro at an exchange rate of 100 Tokens=5€. Total payoffs consist of a show-up fee (5€) plus the payoff from Part 1 plus the payoff from Part 2. Total average earnings are 15.63€. An English translation of the experimental instructions, originally in German, can be found in Appendix C.

4.2.3 Predictions

There is overwhelming evidence that individuals in public good games give positive amounts (e.g., Zelmer, 2003). This can be attributed to a combination of behavioral motivations like conditional cooperation (Chaudhuri, 2011) together with efficiency concerns or kindness (e.g., Rabin, 1993). These concerns essentially mean that a (conditional) cooperator cares not only about the impact of his actions on his own payoff, but also about the impact on the payoffs of others. The weight put on giving to others might depend

[8]In case the group choice is drawn, the choice of a randomly selected group member is implemented for the whole group.

on the observed behavior of others ($x_j, j \neq i$) and the kindness she infers from their level of contributions.[9] In Appendix B we provide a simple model capturing these ideas. The model additionally allows for diverse risk attitudes with respect to own payoff as well as to giving to others. Here, we will briefly discuss the reasoning for the treatment differences we expect to observe.

NoRisk ($r(s_r) = h(s_{h,i}) = 0.5$) provides identical incentives as a standard public good game with a homogeneous marginal return of 0.5. We expect to observe similar behavior as found in the literature as we do not expect contribution decisions to be impacted by the different framing that separates own and others' returns. While the other treatments introduce risky returns, expected returns are kept the same across treatments. Therefore, no treatment differences should result under risk-neutrality—unless they exhibit concerns for conditional cooperation (Prediction (i)). In this case, we expect them to react to lower contributions by others which makes their behavior similar to that of risk-averse types.

Risk attitudes can generate treatment differences and we present predictions for a person who is risk-averse with regard to own payoff in the following. In the presence of both risks, the correlation structure is decisive: positive correlation in $BothRisk_{Pos}$ is expected to lead to the smallest transfers, followed by independent draws ($BothRisk_{Ind}$). The negative correlation in $BothRisk_{Neg}$ is predicted to generate an insurance effect that leads to larger transfers than in $BothRisk_{Pos,\ Ind}$ and in *PublicRisk* (Predictions (ii) and (iv)).

Importantly, investments in $BothRisk_{Neg}$ can also be compared with the *NoRisk* case: Investments in $BothRisk_{Neg}$ do not lead to a change in income ($1 - r(s_r) = h(s_{h,i}) = 0$) with 50% chance, but alternatively result in a marginal effect on own and others' payoffs ($1 - r(s_r) = h(s_{h,i}) = 1$) that is twice as large as under *NoRisk*. To put it differently, in 50% of the cases, subjects end up with their initial endowment independent of their decisions. With 50% chance, though, the high returns are drawn and all payoffs are identical to those in the *NoRisk* treatment if the investments are half the ones chosen in *NoRisk*. Without changes in the perceived kindness, we would thus expect investments in $BothRisk_{Neg}$ to be half of those in *NoRisk* (Prediction (ii)).

Considering risk in only one dimension, *PublicRisk* should lead to smaller transfers than when only private risks are introduced in *PrivateRisk* (Prediction (iv)). Both single risk conditions can be expected to trigger larger transfers than when both risks are introduced

[9]People may also directly receive utility from giving to others. Altruistic preferences are seen as increasing in the monetary payoff to other subjects (Palfrey and Prisbre, 1997; Ledyard, 1995), with evidence gathered by, e.g., Anderson et al. (2011) and Goeree et al. (2002). Differently, warm-glow preferences sometimes refer to a situation where subjects get utility from the intention of contributing (Palfrey and Prisbre, 1997; Andreoni, 1989), not necessarily considering the success of giving.

in independent or positively correlated fashion (Predictions (iii) and (iv)).

Clearly, treatment differences depend on risk attitudes over both own and others' payoffs. It is therefore crucial to empirically elicit both measures.

4.3 Results

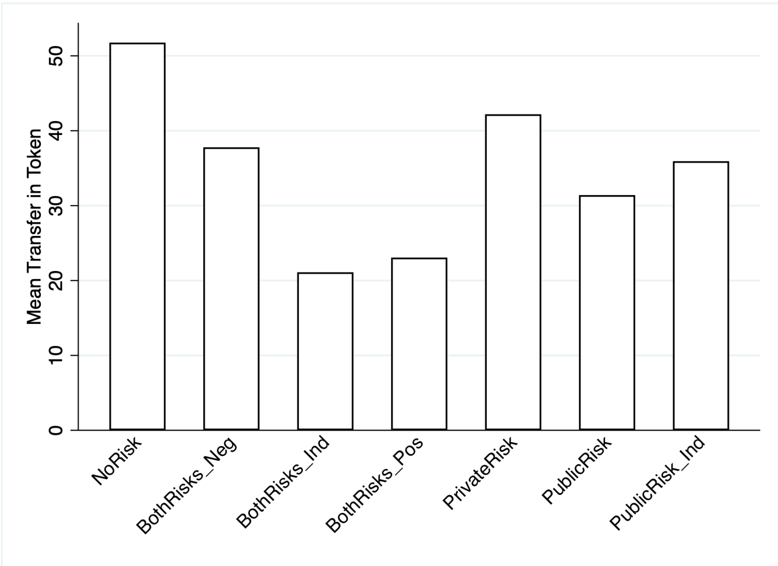


Figure 4.1: Mean transfers in each treatment (all periods)

The summary statistics for decisions in all treatments are given in Table 4.5. We report average decisions as well as the fraction of positive investments ($x > 0$) across all 10 periods, for the first period, for periods 1 through 5, as well as for periods 6 through 10. We further provide mean decisions separated by risk type. Following the neutral wording in the experimental instructions, the tables report the "transfer" decisions instead of "investments". The means across all periods are also reported in Figure 4.1.

4.3.0.0.1 Treatment Differences in Contributions. Investments are highest in the baseline treatment *NoRisk* (51.77 out of 100 tokens). The lowest average investments result if both own return and others' return are risky and positively correlated ($BothRisks_{Pos}$, 23.06 tokens) or independently drawn ($BothRisks_{Ind}$, 21.10 tokens). For both treatments, differences to *NoRisk* are significant (*Pos*: $p_{MW} = 0.011$, Mann-Whitney U test of the equality of distributions (MW), $p_{tt} = 0.045$, two-sample bootstrapped t-test (2-sided, tt), comparing group averages across all periods; *Ind*: $p_{MW} = 0.018$, $p_{tt} = 0.011$).[10] All treatment comparisons are summarized in Table 4.6.

[10]Throughout this section, we report p-values from both tests. We chose to report p-values from bootstrapped t-tests in addition to the non-parametric MWU tests, because the latter does not take the

Result 9. *Relative to non-risky returns, average investments are significantly lower when both private and public returns are risky and positively correlated or independently drawn.*

Result 9 is consistent with prediction (ii) as adding risks on both private and public returns is expected to reduce investments. The finding that investments are reduced when private and public returns are risky and positively correlated corresponds to earlier findings in the literature that risky returns may reduce giving in standard public good games (e.g., Dickinson, 1998; Gangadharan and Nemes, 2009). We obtain a similar level of investments when private and public returns are drawn independently. Consistent with predictions (ii) and (iv), investments in $BothRisk_{Neg}$ are (weakly) significantly larger (37.79) than in the case of positively or independently correlated risks (*Pos*: $p_{MW} = 0.024$, $p_{tt} = 0.206$, *Ind*: $p_{MW} = 0.038$, $p_{tt} = 0.035$). Here, the negative correlation of public and private returns essentially attenuates the downside risk.

Following prediction (ii), we finally compare total investments in $BothRisk_{Neg}$, where decisions correspond to a pledge to give, with half the amount in *NoRisk* which generates the exactly same payoffs with 50% chance. We find that investments tend to be larger than expected under negatively correlated risk (37.79 vs. 0.5*51.77, $p_{MW} = 0.106$, $p_{tt} = 0.113$).

When introducing either only private risk (*PrivateRisk*) or only public risk (*PublicRisk* and $PublicRisk_{Ind.Level}$), investments decrease relative to *NoRisk*, even though the differences are not statistically significant when averaging across all agents and all periods. However, considering only first period decisions—which allows for taking individual decisions as independent observations—investments are weakly significantly smaller under private risk ($p_{MW} = 0.088$, $p_{tt} = 0.082$) and significantly smaller under both public risk conditions than under *NoRisk* (for *PublicRisk*: $p_{MW} = 0.004$, $p_{tt} = 0.005$; for $PublicRisk_{Ind.Level}$: $p_{MW} = 0.002$, $p_{tt} = 0.001$), supporting predictions (ii) and (iii). In line with prediction (iv), the decline is stronger for risk in the public dimension.

Investment decisions in *PublicRisk* and $PublicRisk_{Ind.Level}$ do not differ significantly overall periods. This shows that it is not important if *all* subjects' giving has identical success.[11] This minor difference between the two public risk conditions also indicates that income risk from investments by others does not severely affect the own investment decision.[12]

cardinal information in the data into account. As a conservative measure, we discuss treatment differences while taking group means across all periods as one independent observation. The results are robust to using decisions in the first period which allows to take one decision per individual as an independent observation.

[11]This finding is consistent with our assumption of $u''' = 0$ which we used in order to derive parts of our predictions, i.e. to approximate giving decisions under *PublicRisk* and $BothRisk_{Ind}$.

[12]The comparison between these two treatments resonates with Fischbacher et al. (2014) and Théroude and Zylbersztejn (2020) who introduce asymmetries between contributors. In line with our finding,

Notably, adding positively correlated or independent public risk further reduces investments relative to a situation in which only private returns are risky (prediction (iii), *PrivateRisk* > *BothRisk*$_{Pos}$ ($p_{MW} = 0.038$, $p_{tt} = 0.15$), *PrivateRisk* > *BothRisk*$_{Ind}$ ($p_{MW} = 0.018$, $p_{tt} = 0.039$)), while investments under negatively correlated risk are almost indistinguishable from *PrivateRisk* only. At the same time, relative to a situation where only public returns are risky, adding private risk does not lead to statistically significant changes, i.e. we do not find support for prediction (iv) at the aggregate level.

Result 10. *Adding independent or positively correlated public risk to already existing private risk (further) reduces investments, while additional negatively correlated public risk does not change investments compared to a situation with only private risk.*

To shed further light into these results, Table 4.7 reports results from (individual) random effects models on the transfer decision with standard errors clustered at the group level,[13] as well as results from random effect probit models explaining positive giving. In both models, treatments are defined as binary variables equal to one if the respective treatment condition applies. *NoRisk* serves as the baseline. The regressions confirm the treatment effects we have identified above.

Table 4.7, column (1) shows that risk in each dimension, but particularly public risk, reduces investments. Interestingly, the negative effects of risk in private and public dimensions only almost add up when going to *BothRisk*$_{Ind}$ and *BothRisk*$_{Pos}$. Similar results are obtained for the participation decision (column (3) and (4) of Table 4.7): relative to *NoRisk*, risky public returns in *PublicRisk*, *BothRisk*$_{Ind}$ and *BothRisk*$_{Pos}$ reduce the share of subjects investing a positive amount.

4.3.0.0.2 Treatment Differences Over Time. Over the course of all ten periods, average investments exhibit a downward trend in all treatments (see Figure 4.2). However, the negative impact of independent or positively correlated public risk on investments relative to private risk (and relative to no risk) remains stable when only considering decisions in the last five periods. In the presence of risky private returns, additional risk on the impact of giving thus crucially impacts investment decisions even after several periods of interactions. To confirm these observations, Table 4.8 reports the same regressions as in Table 4.7 for decisions in period 6-10 only. While the treatment differences are smaller

Théroude and Zylbersztejn (2020) do not observe significant changes in behavior when random draws are made at the group level or at the individual level. However, in contrast to our finding (and much of the literature) they do not find a significant difference between the introduction of risk in the returns from the public good compared to a VCM under certainty.

[13]With 84 groups, the number of clusters is large enough to obtain reliable estimates. The results are robust when running tobit models instead.

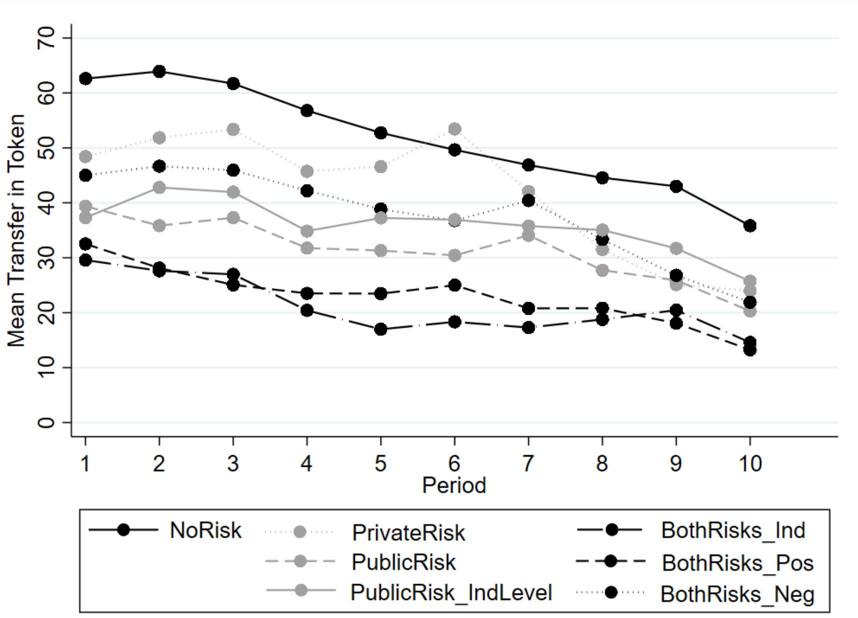


Figure 4.2: Transfers over all periods

in later periods, the overall pattern of treatment effects remains, except for the effect of *PublicRisk* which, while it is still large in magnitude, is no longer statistically significant.

4.3.0.0.3 Heterogeneity in Treatment Differences. To identify risk types, we begin by reviewing the two lottery choices in Part 2 of the experiment. Both choices over lotteries, reported in Table 4.3, show on average moderate risk aversion. We code 203 subjects as risk-averse (RA) and the remaining 133 as non-risk-averse (NRA). We use this classification to study the role of risk attitudes for investment decisions in the respective treatments.[14] While the lottery choices over own and group payoffs are significantly correlated, 54% of the participants switched between the two choices (see Table 4.4).[15]

Following our predictions, we allow for interactions between treatments and risk attitudes in the random effects models in Table 4.7 to gain more detailed insights into how

[14] As we were primarily interested in the investment behavior, we decided to conduct the risk-elicitation task in Part 2. While mean lottery choices vary slightly between treatment groups, this variation appears unsystematic and is unrelated to the degree of riskiness of the public good that participants were exposed to in Part 1.

[15] In line with much of the literature (e.g., Harrison et al., 2013), we do not observe a significant difference in average risk attitudes of subjects when acting on behalf of their group vs. only for themselves. Our predictions suggest to additionally separate risk attitudes w.r.t. own payoff vs. the payoff of others. In L_{group}, both those attitudes are mixed. However, subjects who act *more* risk-averse when deciding for the group than for their own payoff, can be identified as risk-averse w.r.t. the payoff of others. We therefore use the *difference* between the decisions in L_{own} and L_{group} to classify subjects as risk-averse w.r.t. other's payoff (RAother if $L_{group} < L_{own}$) or as non-risk-averse w.r.t. others' payoff (NRAother if $L_{group} \geq L_{own}$). Thus, 94 subjects are coded as RAother, while the remaining 242 are coded as NRAother. However, an additional inclusion of these variables in our analysis cannot explain decisions such that our analysis below concentrates on risk attitudes w.r.t. own payoff, i.e. on RA vs. NRA.

risk attitudes affect investment decisions. Risk attitudes are coded as binary variables as described above.

Controlling for risk types, we see in column (2) that the treatment differences discussed above are strengthened for risk-averse subjects. For non-risk-averse subjects, they are generally smaller and fully vanish in $BothRisks_{Neg}$ and *PrivateRisk*.[16] This is consistent with our Prediction (i). The differences in participation decisions (column (3) and (4) of Table 4.7) across treatments are robust for both NRA and RA types.

Overall, our data therefore shows that the detrimental effect of risk in returns is particularly driven by the risk in public returns, and that the negative effect of risk in each dimension alone adds up for risk-averse players when having risk in both dimensions in $BothRisk_{Ind}$ and $BothRisk_{Pos}$. In fact, transfers in these two treatments is very similar. One reason for this, that goes beyond our predictions, might be that the worst case in both treatments lets individuals obtain neither any private nor any public return, i.e. a return of zero.[17] In line with this interpretation but also consistent with our predictions, the negative correlation of public and private returns in $BothRisk_{Neg}$ essentially attenuates the downside risk and might provide an insurance effect. While these effects are consistent with our predictions derived from a model that allows for different risk attitudes over private and public payoffs, we do not find evidence supporting a need for this distinction. Rather, risk-aversion in the two dimensions is highly correlated.

4.4 Conclusions

We investigated how the riskiness of both private and public returns matters for investment decisions. Based on variants of public good games which allow us to disentangle the risk in *own* vs. *others'* returns from investments, we find that particularly the latter is detrimental for investments. The correlation between both risks matters: only if both risks are positively correlated or independent, investments are substantially negatively impacted, compared to a situation where only private returns are risky or no risks exist. The similarity of investments under both positively correlated and independent risk treatments suggests that people might not consider the whole distribution of risks, but instead focus on the downside risk of ending up with zero returns in both dimensions.

We can conclude that a reduction in risk in the success of giving to others, i.e. the

[16]The negative coefficient of NRA (though not significant) indicates that risk-aversion may be positively correlated with investments in the *NoRisk* treatment, in line with results by Freundt and Lange (2017) for giving in dictator games.

[17]However, we do not see any explanatory power in our measures of risk-aversion to explain these treatment differences.

return of the investment to others, is crucial to stabilize investments when a public component is present. This holds true in the absence as well as in the presence of private risk. With this, our findings further support models that assume utility being driven by the *success* in giving, rather than by the *act* of giving up own payoff alone, thereby extending existing evidence on the nature of social preferences (e.g., Anderson et al., 2011; Goeree et al., 2002; Palfrey and Prisbre, 1997; Hungerman and Ottoni-Wilhelm, 2018) to risky situations.

In our experiment we chose a rather extreme distribution of possible returns, i.e. returns of 0 and 1, which facilitated the derivation of predictions (see Appendix B) and greatly simplified participants' decisions in the lab. It remains to be seen in future research how robust the findings are to less extreme returns, particularly when a positive return is secured in any state of the world. However, we consider our setting to be informative of many applications: microlending, crowdinvesting, charitable donations, and environmental protection (e.g., abatement of emissions) may all come with a risk of complete failure to provide a promised public return. Our experiment indicates that reducing the risk in such situations may be crucial to attracting investments.

Appendix

Appendix A: Figures and Tables

	1	2	3	4	5	6	Mean
L_{own}	37	32	66	68	69	64	3.87
L_{group}	34	40	39	88	66	69	3.95

Table 4.3: Lottery choices in risk tasks for own and group payoff

	L_{own}						
	1	2	3	4	5	6	all
$L_{own} < L_{group}$	26	15	36	11	6		94
$L_{own} = L_{group}$	11	13	16	38	45	45	168
$L_{own} > L_{group}$		4	14	19	18	19	74

Table 4.4: Changes in choices for lottery for own vs. group payoff

Periods		all	1	1-5	6-10	all	all
Participants		all	all	all	all	RA	NRA
NoRisk	x	51.77	62.60	59.55	43.98	55.49	44.31
		(41.59)	(37.85)	(39.59)	(42.18)	(40.35)	(43.15)
	$x > 0$	0.79	0.92	0.85	0.75	0.81	0.76
	n	480	48	240	240	320	160
$BothRisks_{Ind}$	x	21.10	29.58	24.32	17.89	19.26	23.92
		(27.92)	(27.65)	(28.83)	(26.66)	(23.49)	(33.45)
	$x > 0$	0.58	0.71	0.61	0.56	0.62	0.53
	n	480	48	240	240	290	190
$BothRisks_{Pos}$	x	23.06	32.52	26.54	19.59	21.93	24.53
		(30.06)	(33.32)	(30.84)	(28.91)	(27.04)	(33.56)
	$x > 0$	0.67	0.69	0.70	0.65	0.71	0.62
	n	480	48	240	240	270	210
$BothRisks_{Neg}$	x	37.79	45.00	43.73	31.85	33.74	46.71
		(33.66)	(34.56)	(33.80)	(32.52)	(29.81)	(39.55)
	$x > 0$	0.77	0.79	0.80	0.74	0.78	0.74
	n	480	48	240	240	330	150
PrivateRisk	x	42.20	48.42	49.18	35.22	37.67	46.37
		(38.01)	(36.26)	(36.37)	(38.40)	(32.63)	(41.99)
	$x > 0$	0.75	0.87	0.84	0.68	0.79	0.72
	n	480	48	240	240	230	250
PublicRisk	x	31.40	39.38	35.12	27.69	30.96	31.93
		(34.54)	(36.11)	(34.46)	(34.28)	(33.98)	(35.25)
	$x > 0$	0.70	0.77	0.74	0.67	0.70	0.69
	n	480	48	240	240	260	220
$PublicRisk_{Ind.Level}$	x	35.93	37.31	38.83	33.03	35.75	36.34
		(33.58)	(31.86)	(33.62)	(33.37)	(32.99)	(34.96)
	$x > 0$	0.75	0.75	0.75	0.75	0.77	0.69
	n	480	48	240	240	330	150

Table 4.5: Summary statistics of transfers. (x = mean number of Taler transfered (with std. dev.) as well as fraction of positive transfers ($x > 0$), n = number of subjects)

	NoRisk	*BothRisks*$_{Ind}$	*BothRisks*$_{Pos}$	*BothRisks*$_{Neg}$	*PrivateRisk*	*PublicRisk*
BothRisks$_{Ind}$	$<$ $p_{MW} = 0.018$ $p_{tt} = 0.011$ *(prediction (ii))*					
BothRisks$_{Pos}$	$<$ $p_{MW} = 0.011$ $p_{tt} = 0.045$ *(prediction (ii))*	$>$ $p_{MW} = 0.863$ $p_{tt} = 0.818$ *(prediction (ii))*				
BothRisks$_{Neg}$	$<$ $p_{MW} = 0.356$ $p_{tt} = 0.231$ *(prediction (ii))*†	$>$ $p_{MW} = 0.038$ $p_{tt} = 0.035$ *(prediction (iv))*	$>$ $p_{MW} = 0.024$ $p_{tt} = 0.206$ *(prediction (ii))*			
PrivateRisk	$<$ $p_{MW} = 0.387$ $p_{tt} = 0.412$ *(pred. (ii,iii))*	$>$ $p_{MW} = 0.018$ $p_{tt} = 0.039$ *(prediction (iii))*	$>$ $p_{MW} = 0.038$ $p_{tt} = 0.150$ *(prediction (ii))*	$>$ $p_{MW} = 0.453$ $p_{tt} = 0.621$		
PublicRisk	$<$ $p_{MW} = 0.133$ $p_{tt} = 0.090$ *(prediction (ii))*	$>$ $p_{MW} = 0.166$ $p_{tt} = 0.148$ *(prediction (iv))*	$>$ $p =_{MW} 0.073$ $p_{tt} = 0.413$ *(prediction (ii))*	$<$ $p_{MW} = 0.273$ $p_{tt} = 0.434$ *(prediction (iv))*	$<$ $p_{MW} = 0.019$ $p_{tt} = 0.008$ *(prediction (iv))*	
PublicRisk$_{Ind.Level}$	$<$ $p_{MW} = 0.157$ $p_{tt} = 0.097$ *(rob. check)*					$>$ $p_{MW} = 0.854$ $p_{tt} = 0.936$ *(rob. check)*

Table 4.6: Summary of all treatment comparisons: Inequality signs indicate how the treatment on the LHS compares to the one in the top row. The p-values are taken from Mann Whitney U tests and from 2-sample bootstrapped t-tests based (2-sided) on group mean contributions for all periods. (i) to (vi) indicate the prediction guiding each comparison.

† For testing prediction (i) that *BothRisks*$_{Neg} = 0.5 *$ *NoRisk*, the p-values are $p_{MW} = 0.106$ and $p_{tt} = 0.113$.

	(1)	(2)	(3)	(4)
Dep.Var.	Transfer	Transfer	Participation	Participation
$BothRisks_{Ind}$	-30.66***	-36.24***	-1.17***	-1.09**
	(-3.11)	(-3.41)	(-3.20)	(-2.37)
$BothRisks_{Pos}$	-28.70***	-33.57***	-0.90**	-0.82*
	(-2.65)	(-3.11)	(-2.45)	(-1.74)
$BothRisks_{Neg}$	-13.98	-21.76**	-0.23	-0.18
	(-1.36)	(-2.06)	(-0.63)	(-0.41)
PrivateRisk	-9.56	-17.82	-0.48	-0.51
	(-0.89)	(-1.53)	(-1.31)	(-1.05)
PublicRisk	-20.36**	-24.53**	-0.69*	-0.78*
	(-2.00)	(-2.15)	(-1.90)	(-1.65)
$PublicRisk_{IndLevel}$	-15.83	-19.75*	-0.43	-0.49
	(-1.47)	(-1.71)	(-1.16)	(-1.08)
NRA		-11.18		-0.39
		(-1.31)		(-0.71)
NRA x $BothRisks_{Ind}$		15.85		-0.15
		(1.57)		(-0.19)
NRA x $BothRisks_{Pos}$		13.78		-0.08
		(1.34)		(-0.11)
NRA x $BothRisks_{Neg}$		24.16**		-0.16
		(2.02)		(-0.20)
NRA x *PrivateRisk*		19.88*		0.19
		(1.70)		(0.26)
NRA x *PublicRisk*		12.15		0.29
		(1.12)		(0.39)
NRA x $PublicRisk_{IndLevel}$		11.77		0.12
		(0.95)		(0.16)
Constant	51.77***	55.49***	1.42***	1.56***
	(5.84)	(5.78)	(5.33)	(4.73)
Observations	3,360	3,360	3,360	3,360
Number of subjects	336	336	336	336

Robust z-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.7: Random effect models on transfers x_i across all periods with standard errors clustered at the group level (column (1) and (2)). Random effects probit on participation ($x_i > 0$) across all periods (column (3) and (4)).

Dep. Var.	(1) Transfer	(2) Transfer	(3) Participation	(4) Participation
$BothRisks_{Ind}$	-26.10**	-31.22**	-1.21***	-1.20**
	(-2.34)	(-2.53)	(-2.89)	(-2.30)
$BothRisks_{Pos}$	-24.40**	-29.71**	-0.75*	-0.76
	(-2.09)	(-2.41)	(-1.84)	(-1.45)
$BothRisks_{Neg}$	-12.13	-19.30	-0.18	-0.15
	(-1.05)	(-1.56)	(-0.44)	(-0.29)
PrivateRisk	-8.76	-15.25	-0.51	-0.52
	(-0.73)	(-1.17)	(-1.25)	(-0.96)
PublicRisk	-16.29	-20.80	-0.63	-0.77
	(-1.41)	(-1.63)	(-1.53)	(-1.45)
$PublicRisk_{IndLevel}$	-10.95	-17.03	-0.10	-0.25
	(-0.94)	(-1.35)	(-0.24)	(-0.49)
NRA		-14.64*		-0.62
		(-1.96)		(-1.02)
NRA x $BothRisks_{Ind}$		15.26		0.07
		(1.64)		(0.08)
NRA x $BothRisks_{Pos}$		15.62*		0.16
		(1.71)		(0.20)
NRA x $BothRisks_{Neg}$		21.97*		-0.13
		(1.90)		(-0.15)
NRA x *PrivateRisk*		17.73		0.23
		(1.50)		(0.28)
NRA x *PublicRisk*		13.83		0.47
		(1.31)		(0.56)
NRA x $PublicRisk_{IndLevel}$		18.47*		0.37
		(1.68)		(0.42)
Constant	43.98***	48.86***	1.06***	1.28***
	(4.35)	(4.40)	(3.56)	(3.46)
Observations	1,680	1,680	1,680	1,680
Number of subjects	336	336	336	336

Robust z-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.8: Random effect models on transfers x_i across periods 6-10 with standard errors clustered at the group level (column (1) and (2)). Random effects probit on participation ($x_i > 0$) across periods 6-10 (column (3) and (4)).

Appendix B: Illustrating behavioral model

To guide our intuition on the behavior in our treatments, we formulate a simple model on conditional cooperation:

$$\mathbb{E}[U_i] = \mathbb{E}\left[u_i(m - x_i + r(s_r)x_i + \sum_{j\neq i} h(s_{h,j})x_j) + \kappa_i(\sum_{j\neq i} x_j)v_i(h(s_{h,i})x_i)\right].$$

A subject i's utility depends on her own payoff $m - x_i + r(s_r)x_i + \sum_j h(s_{h,j})x_j$ and additionally on her impact on the payoff of others $h(s_{h,i})x_i$, e.g. through a warm-glow sensation. For both utility components we allow for diverse risk attitudes, captured by u_i and v_i, respectively. Conditional cooperation motives are captured by κ_i which we assume to be increasing in the investments of others, i.e. in their intention to give to i.

The first order condition is given by

$$\mathbb{E}\left[-u_i'(\cdot)(1 - r(s_r)) + \kappa_i(\cdot)v_i'(\cdot)h(s_{h,i})\right] \quad \leq \quad 0$$

with equality for an interior solution where we assume that subjects do not invest all their income. It is obvious that for subjects that are risk-neutral in both dimensions, corner solutions of either no or full giving would result. In the following discussion, we focus on the comparative statics when agents are risk-averse in at least on dimension.

It is well-known that the qualitative impact of risk on decisions is typically ambiguous as it depends on the third derivative of the utility function. To illustrate this, consider the impact of risky public returns. For this, we have to compare $\mathbb{E}\left[v'(hx)h\right]$ and $\mathbb{E}[h]v'(\mathbb{E}[h]x)$. The curvature of the former is given by $v'''(hx)hx^2 + 2v''(hx)x$ which depends on both the second and the third derivative of v. As such, the comparison of $\mathbb{E}\left[v'(hx)h\right]$ and $\mathbb{E}[h]v'(\mathbb{E}[h]x)$ is generally ambiguous. However, if v displays risk aversion ($v'' < 0$) and prudence ($v''' > 0$), the investment is smaller in the presence of public risk than without risk.

The parameters in our experiment, however, are chosen to allow for cleaner predictions. Because $h^L = 0$, we obtain:

$$\mathbb{E}\left[v'(hx)h\right] = \pi^H h^H v'(h^H x) = \mathbb{E}[h]v'(h^H x) < \mathbb{E}[h]v'(\mathbb{E}[h]x)$$

as long as $v'' < 0$, i.e. with risk-aversion over the public return. Therefore, we expect such risk-aversion to decrease giving when the public return is risky.

Given our parameter settings, the first-order condition reduces to simple expressions for the respective treatments:

$$\begin{aligned}
&\textit{NoRisk} && -0.5u_i'(m - x_i + 0.5x_i + 0.5\sum_{j\neq i} x_j) + 0.5\kappa_i(\cdot)v_i'(0.5x_i) \leq 0 \\
&\textit{BothRisk}_{Pos} && -0.5u_i'(m - x_i) + 0.5\kappa_i(\cdot)v_i'(x_i) \leq 0 \\
&\textit{BothRisk}_{Neg} && -0.5u_i'(m - x_i + \sum_{j\neq i} x_j) + 0.5\kappa_i(\cdot)v_i'(x_i) \leq 0 \\
&\textit{BothRisk}_{Ind} && -0.25u_i'(m - x_i) - 0.25u_i'(m - x_i + \sum_{j\neq i} x_j) + 0.5\kappa_i(\cdot)v_i'(x_i) \leq 0 \\
&\textit{PrivateRisk} && -0.5u_i'(m - x_i + 0.5\sum_{j\neq i} x_j) + 0.5\kappa_i(\cdot)v_i'(0.5x_i) \leq 0 \\
&\textit{PublicRisk} && -0.25u_i'(m - x_i + 0.5x_i + \sum_{j\neq i} x_j) \\
& && \quad -0.25u_i'(m - x_i + 0.5x_i) + 0.5\kappa_i(\cdot)v_i'(x_i) \leq 0
\end{aligned}$$

The second-order conditions are automatically satisfied under risk-aversion. The first-order conditions allow to make predictions on the treatment differences in investment by player i, conditional on the decisions of other players $j \neq i$.

The first-order conditions reveal that, assuming a fixed κ_i, all subjects are predicted to give half the amount in $\textit{BothRisk}_{Neg}$ than in $\textit{NoRisk}$. For subjects that are risk-averse w.r.t. own payoff ($u'' < 0$), giving in $\textit{BothRisk}_{Pos}$ is smaller than in $\textit{BothRisk}_{Ind}$ than in $\textit{BothRisk}_{Neg}$.[18] Allowing κ_i to depend on perceived generosity might make these treatment differences even larger. Note that we assume κ_i to depend on $\sum_{\neq i} x_j$. This implies that, if no one else gives ($\sum_{\neq i} x_j = 0$), κ_i can be assumed to be very small, such that zero giving results in all treatments.

It is instructive to compare the treatment conditions under the assumption that subjects anticipate average giving to be at the same level as own giving, i.e. $\sum_{j\neq i} x_j = 3x_i$.

[18]For fixed κ_i – the first order condition in *BothRisk*$_{Neg}$ holds if investments are exactly half the ones that solve the conditions in *NoRisk*. If these reduced investments are perceived as less kind and reduce κ_i, we would predict that investments in *BothRisk*$_{Neg}$ are less than half of those in *NoRisk*. The relationship *BothRisk*$_{Neg}$ vs. *BothRisk*$_{Ind}$ and *BothRisk*$_{Pos}$ follows from $u_i'(m - x_i) \geq u_i'(m - x_i + \sum_{j\neq i} x_j)$. Similarly, the relationship of *BothRisk*$_{Pos}$ and *PrivateRisk* or *PublicRisk* follows from comparing the arguments of the Bernoulli function.

We rewrite the conditions above as:

$$\begin{aligned}
&\textit{NoRisk} && u_i'(m+x) \geq \kappa_i(\cdot)v_i'(0.5x)\\
&\textit{BothRisk}_{\textit{Pos}} && u_i'(m-x) \geq \kappa_i(\cdot)v_i'(x)\\
&\textit{BothRisk}_{\textit{Neg}} && u_i'(m+2x) \geq \kappa_i(\cdot)v_i'(x)\\
&\textit{BothRisk}_{\textit{Ind}} && 0.5u_i'(m-x) + 0.5u_i'(m+2x) \geq \kappa_i(\cdot)v_i'(x)\\
&\textit{PrivateRisk} && u_i'(m+0.5x) \geq \kappa_i(\cdot)v_i'(0.5x)\\
&\textit{PublicRisk} && 0.5u_i'(m+2.5x) + 0.5u_i'(m-0.5x) \geq \kappa_i(\cdot)v_i'(x) \leq 0
\end{aligned}$$

We first note that the left-hand side in $\textit{BothRisk}_{\textit{Ind}}$ is larger or smaller than $u_i'(m+0.5x)$ depending on $u_i''' > 0$ or $u_i''' < 0$. Similarly, the left-hand side in *PublicRisk* is larger or smaller than $u_i'(m+x)$ depending on $u_i''' > 0$ or $u_i''' < 0$. As we are agnostic about the prudence attitude (u_i'''), we approximate investments in these treatments by the levels given by

$$\begin{aligned}
&\textit{BothRisk}^*_{\textit{Ind}} && u_i'(m+0.5x) \geq \kappa_i(\cdot)v_i'(x)\\
&\textit{PublicRisk}^* && u_i'(m+x) \geq \kappa_i(\cdot)v_i'(x) \leq 0,
\end{aligned}$$

respectively. Under the assumption of fixed levels of κ, all subjects are predicted to give half the amount in $\textit{BothRisk}_{\textit{Neg}}$ than in NoRisk. For subjects that are risk-averse w.r.t. own payoff ($u'' < 0$), giving in $\textit{BothRisk}_{\textit{Pos}}$, is smaller than in $\textit{BothRisk}_{\textit{Ind}}$, than in $\textit{BothRisk}_{\textit{Neg}}$. Giving in *PublicRisk** is half the level of giving in *PrivateRisk*. Giving is maximal in *NoRisk*.

These statements follow since the first-order condition in $\textit{BothRisk}_{\textit{Neg}}$ holds if investments are exactly half the ones that solve the conditions in *NoRisk*. The same holds when comparing *PublicRisk** and *PrivateRisk*. The remaining statements follow directly from the arguments in the Bernoulli functions: $u_i'(m-x) \geq u_i'(m) \geq u_i'(m+x) \geq u_i'(m+2x)$ under the assumption of risk-aversion.

We can thus summarize our predicitions below:

(i) For risk-neutral subjects (in both dimensions) no treatment differences occur, unless they negatively reciprocate reduced giving by other (risk-averse) subjects—in which case the treatment comparisons are in line with those of risk-averse subjects.

(ii) For subjects that are risk-averse w.r.t. own payoff, investments are smallest in $\textit{BothRisk}_{\textit{Pos}}$ and largest in *NoRisk*. Giving in $\textit{BothRisk}_{\textit{Neg}}$ is half the amount given in $NoRisk$.

(iii) Giving in $PrivateRisk$ is smaller than under $NoRisk$ if the subject is risk-averse w.r.t. own payoff. It typically larger than giving under $BothRisk_{Ind}$.

(iv) For risk-averse types, giving in $PublicRisk$ is approximately half the amount under $PrivateRisk$ and between the amounts under $BothRisk_{Ind}$ and $BothRisk_{Neg}$.

Given these predictions, we expect the introduction of public risk to be more detrimental to investments than the introduction of private risks. In presence of both risks, the correlation structure is decisive: positive correlation is expected to lead to the least transfers, followed by independent draws. The negative correlation is predicted to generate an insurance effect that leads to larger transfers than under private or public risks alone.

Appendix C: Experimental instructions

Below we report an English translation of the instructions for treatment BothRisks$_{Pos}$. The original instructions have been in German. The parts that can differ between treatments are marked in italic.

Welcome to the Experimental laboratory and thank you for participating in this economic experiment.

Please switch off your phones during the entire experiment. Communication with other participants is not allowed and a violation of this rule will lead to an exclusion from the experiment as well as from all payments. If you have any questions during the experiment, please raise your hand, we will come to you.

Procedure

The experiment consists of two entirely independent parts. The instructions for the second part will be distributed and read out to you after the first part is over. The decisions you make in the first part are **not** relevant for the payoffs in the second part and the other way round. In the end, the earnings from part 1 and from part 2 will both be paid out to you.

Part 1

Payoffs

In part 1 you will make several decisions that determine your income as well as the income of other participants. The actual payoffs will partly depend on chance. As your decisions will determine the size of your earnings, it is important that you read the instructions carefully before making any decisions. If something is unclear to you, please do not hesitate to ask!

Your income in the experiment will be calculated in Taler. Taler will be converted into Euro with an exchange rate of

100 Taler=5 Euro.

Your total payoff consists of the sum of the payoffs from part 1 and part 2. In addition, you receive a show-up fee of 5 Euro for participating in the experiment. You will be paid out in cash immediately after the experiment is over. The other participants will not be able to see how much money you receive.

Procedure

All decisions will be **anonymous**, i.e. neither another participant nor the experimenter can match them to your identity.

The experiment consists of **10 rounds** in which you will be in the same decision situation. Before the beginning of the first round, you will be connected with three other participants that are chosen randomly into a group consisting of 4 people in total. None of the participants knows with whom they are matched into a group. During the entire experiment, i.e. in all 10 rounds, you stay in one group with the same participants. In each round all participants make the same decision. After the 10 rounds are over, you will see an information screen with the payoffs for each round. Out of the 10 rounds, one round will determine your payoff. Each round can be drawn with the same probability. This round will be determined by a random draw, for which one participants in the room will be chosen to draw one out of 10 cards with the numbers 1 to 10. Only the round that is drawn here will be paid out in the end.

Decision Situation

Before every decision, a brief description of the respective decision situation will appear on your screen. In case you have any questions, please raise your hand, we will come to you!

For each decision you make, you will be provided with **100 Taler** in your private account, **Account A**. Out of these 100 Taler you can transfer a chosen number of Taler into an **Account B**. You keep the remaining Taler in your private Account A. All participants face the **same** decision situation. The payoff of each participant consists of the following parts. You receive:

1. the number of Taler remaining in Account A.

2. a payoff from your transfer into Account B to **yourself**: For each Taler that you transferred into Account B you receive a payoff *of either 1 Taler or 0 Taler. Both events can happen with a probability of 50% and are determined by a random draw. Thus, when you transfer X Taler into Account B, you receive either a payoff of X Taler or of 0 Taler from Account B, depending on the outcome of the random draw.* At the same time you benefit from the transfers of the other three group members:

3. a payoff from the transfers of **the other three group members** into Account B: In addition to the payoff described in (2), you can get a payoff from the sum of the transfers of the other three group members into Account B, we call it $X_2 + X_3 + X_4$. *With a chance of 50% you get a payoff of $X_2 + X_3 + X_4$ Taler and with a chance of 50% you get a payoff of 0 Taler. The chance to receive this payoff depends on the same random draw as the chance to obtain part (2): Drawing the high payoff in (3) goes along with drawing the high payoff in (2), otherwise both parts of the overall payoff amount to 0 Taler.*

Thus, your overall payoff consist of the following three parts:

1. **The number of Taler in Account A: 100-X**

2. **The payoff from the own transfer X into Account B: *X or 0***

3. **The payoff from the transfers of the other group members into Account B: *($X_2 + X_3 + X_4$) or 0***

where X denotes the number of Taler, that you transferred into Account B and X_2, X_3, X_4 denote the transfers of the other three participants.

Total income in one round=(1)+(2)+(3)

This implies that your transfer into Account B also generates a payoff for the other members of your group: The fact that all participants face the same decision situation means, on the one hand, that you benefit from the transfers of your fellow group members into Account B as well as, on the other hand, that each other group member benefits from the Taler you transferred into account B. Your transfer into account B generates payoffs for the other group members in the same way as described above: *With a chance of 50%* ***all three group members*** *get a payoff of size X each and with a chance of 50% they get 0 Taler. Thus, in total 3*X Taler will be paid out to the other group members with a chance of 50%, 0 Taler otherwise.*
To illustrate, imagine the random draw as a coin toss:
There will be a coin toss and you receive -for example- the number of Taler transferred into Account B if head falls and 0 Taler if tail falls. *There will be only one coin toss. This coin toss determines the payoff of a participant from her own transfer in Account B and her payoff from the transfer of the others as well as the payoffs of the other three members from her transfer. The coin will be flipped once for each group (in each round) and the outcome holds for all group members. If head falls, each participant receivers the payoff from the own transfer and from the transfers of the others—and the others benefit from her transfer. If tail falls, all get 0 Taler from Account B.*
Above you can already see the computer screen with the decision situation as described above. There you type in the number of Taler you wish to transfer and then click on OK. After all participants have made their decision, in each round you can see how many Taler the other group members have transferred. After the end of the experiment, a random draw will be drawn by the computer and you will see the size of your payoffs from each round on your screen (out of which one will be paid out, see Procedure section).

Do you have any questions concerning the instructions? If not, we will now proceed with the control questions, that serve your understanding of the procedure of the experiment. As soon as all participants have answered all questions, the actual experiment begins.

Part 2

In part 2 of the experiment you choose between lotteries with outcomes of different sizes. On your screen you will see the following table:

	Payoff C	**Payoff D**
1	56	56
2	48	72
3	40	88
4	32	104
5	24	120
6	4	140

Each row in the table represents a lottery. Each lottery consists of two payoffs in Taler, payoff C and payoff D, that can each be drawn with a probability of 50%. The exchange rate Taler-Euro is the same as in part 1 of the experiment: 100 Taler= 5 Euro. The six lotteries differ only with respect to the possible outcomes C and D, the probability is 50:50 in each lottery. You will choose between lotteries 1 to 6 and what consequences your decision has, will be explained to you on your screen. There will be two different decision situations. A random draw determines which of those two decisions will be relevant for your earnings from part 2. Both decisions have the same chance to be drawn to determine your payment in the end.

In case you have any questions during the experiment, please raise your hand! If you do not have any questions now, we will now proceed with part 2.

Chapter 5

On the performance of green assets in financial markets—Evidence from a laboratory experiment[1]

Abstract

We investigate the financial performance of socially responsible investments in competition with conventional investments on financial markets. Setting up experimental asset markets for 'green', i.e. socially responsible, and non-green stocks and using a novel market design, we identify a causal impact of green investment opportunities on the development of prices and trading volumes. In particular, we test how speculation about future prices influences green market prices and to what extent subjects' behavior in the market is correlated with their generosity in individual donation decisions. We observe no price premium for green assets. On the contrary, the introduction of green investment opportunities increases prices for the non-green assets compared to markets with only conventional firms.

JEL Codes: G12, D40, D62

Keywords: socially responsible investments, experimental asset markets, ethical behavior in markets, social preferences

[1]This chapter is co-authored by Andreas Lange (University of Hamburg), Andreas Nicklisch (HTW Chur) and Stefan Palan (University of Graz).

5.1 Introduction

The economic importance of socially responsible investments is constantly increasing in many industries in Western societies. The Forum for Sustainable and Responsible Investment's annual report 2018 finds that "Sustainable, Responsible and Impact (SRI) Investing[2] in the United States continues to expand at a healthy pace. The total US-domiciled assets under management using SRI strategies grew from $8.7 trillion at the start of 2016 to $12.0 trillion at the start of 2018, an increase of 38 percent. This represents 26 percent – or 1 in 4 dollars – of the $46.6 trillion in total US assets under professional management." (US SIF Foundation, 2018, footnote added) Yet, sustainability and social responsibility do not come free of charge. They restrict the scope and the profitability of SRI funds to those investment opportunities that meet certain ethical and environmental standards. That is, since SRI funds consider a subset of all investment opportunities, their profitability is only at best as good as the performance of conventional funds. On the other hand, it may very well be that investors have an inherent preference for SRI investments causing the performance of green investments to increase despite these disadvantages. Consequently, there is an ongoing, controversial debate in the economic literature over the impact of socially responsible investments and the financial performance in the market place. Some empirical studies find a positive relationship between social and financial performance (e.g., Waddock and Graves, 1997) while others report no or a negative relationship (e.g., Wright and Ferris, 1997; McWilliams and Siegel, 2000).[3]

In our study, we add an important aspect to this controversy: data from laboratory experiment allow us to compare the market performance of SRI assets to conventional assets while controlling for differences in their fundamental values. In experimental double auction markets subjects can simultaneously trade assets of two firms over ten periods. In our baseline markets, both firms represent conventional—that is, non-SRI—investments (hereafter denoted *neutral* assets). In our treatment condition, one firm represents conventional investments, while the second firm represents investments that meet the criteria of SRI (hereafter denoted *green* assets). *Neutral* assets pay a private dividend to the investor per period, whereas the green asset does not pay the full private dividend to the investor,

[2]Socially responsible investments or, synonymously corporate social responsibility (CSR) can be seen as a company's active compliance with ethical and environmental standards that enhance the social welfare beyond the company's direct interests and the requirements of the law (Bénabou and Tirole, 2010).

[3]An overview of this literature is provided by McWilliams et al. (2006). Three meta analyses find a positive relationship between SRI and financial performance (Orlitzky et al., 2003, Allouche and Laroche, 2005, Margolis et al., 2009). However, Margolis et al. (2009) report the positive effect to be fairly small and other meta studies by Revelli and Viviani (2013) and Revelli and Viviani (2015) find no overall link between SRI and financial performance.

but splits the dividend evenly into a private payment to the investor and a public benefit, operationalized as a donation to a charity outside the laboratory. We compare the development of prices and trading volumes across markets in the baseline and in the treatment condition. Prior to the trading phase, we elicit participants' willingness to pay for two *neutral* assets in baseline, and for one *green* asset and one *neutral* asset in the treatment condition, respectively. This provides us with an indicator for subjects' preferences both for *neutral* and *green* assets in an individual decision situation. Based on these individual choices, we define investors who have *green* and *non-green* preferences and compare their trading behavior in the asset market. Additionally, we ask subjects for their beliefs about future market prices at the beginning of each trading period. This enables us to identify to what extend speculation drives a price premium for the *green* assets. Finally, we address the question whether individuals' demand for *green* donations in the market is positively or negatively related to their pro-social behavior outside the market. We achieve this by complementing trading in the market with individual donation decisions. Each participant is asked whether she wishes to donate part of her show-up fee to a charity before leaving the experiment. These features allow us to shed light on underlying mechanisms for SRI assets' under- or over-performance relative to their underlying fundamental value.

We do not find a price bubbles for assets with nominal values that are constant in expectation unless they are in competition with a *green* firm. In the presence of a *green* firm, prices for the *neutral* assets increase above the fundamental value in the second half of the trading periods. Thus, the introduction of assets with a costly positive externality leads to a spillover effect on the conventional, *neutral*, assets that pay a higher private dividend, leading to a price bubble for the latter. There is no price premium for *green* assets; instead their valuation is as high as the valuation of *neutral* assets in the homogeneous markets and close to the fundamental values. Overall, participants rather correctly anticipate market prices but, once controlling for prices and fundamental values, beliefs are not a strong predictor of bidding behavior. We observe underbidding for all types of assets during the first periods with bids coming close to fundamental values in the second half of the market periods. Overall, but predominantly in the later periods of the market interactions, subjects with stronger *green* preferences (classified according to individual decisions in the first part of the experiment) hold more assets, especially more *green* assets, while those with weaker *green* preferences hold more cash. Thus, preference types seem to matter for trading behavior. There is no evidence of a crowding out of individual pro-social behavior due to trading *green* assets on the market. On the contrary, *green* behavior on the market and individual donations are rather independent decisions, if anything they are weakly positively related. Furthermore, being exposed to mixed markets with *green* assets leads to a crowding-in of donors.

Our study extends and connects several streams of literature in business, finance and behavioral economics. A broad interdisciplinary literature discusses the financial performance of firms engaging in corporate social responsibility (CSR) activities and possible motives of investors to engage in such investments.[4] An empirical study by Apostolakis et al. (2016) reports that Dutch pension beneficiaries are willing to pay a premium for socially responsible portfolios. Renneboog et al. (2008) find a willingness to pay a "price for ethics" but cannot identify the exact cause for such a price premium.[5] Friede et al. (2015) review evidence from more than 2000 empirical studies on financial effects of SRI to conclude that about 90 % of studies find a non-negative relation between financial performance and SRI.

While previous papers such as Friede et al. (2015) study the link between investment decisions and the financial performance of firms who respect environmental, social and governance criteria (ESG) in markets, they do not address the impact that the presence of such firms have on *other* (non-ESG) firms on the same market (i.e., they are not considering the full extent of the entry of ESG firms into a financial market setting). We provide evidence that establishes a clean causal effect of SRI on market developments and furthermore helps recovering the underlying mechanisms for the demand for SRI. With this, this paper analyzes as the first one – to the best of our knowledge – how SRI perform on an asset market in which they compete with assets of a conventional firm. We provide evidence of the impact of SRI on the developments of market indicators, most notably prices and trading volumes.

Previous experiments on investors' motivations to engage in SRI focus typically on simple individual portfolio allocation decisions between funds that are framed as being more or less ethical (e.g., Barreda-Tarrazona et al. (2011), Consolandi et al. (2009)) or hypothetical choices (e.g., Glac (2009), Hofmann et al. (2008), Pasewark and Riley (2010), Hofmann et al. (2008)). Most of these studies find some evidence in favor of ethical concerns, however, they cannot exclude that these findings are at least partly influenced by experimenter demand effects since they don't use monetary incentives and the descriptions of the 'ethical conduct' of the (fictious) firms are normatively loaded. In contrast, our study employs a controlled laboratory experiment with neutral language to shed light

[4]A range of potential motives has been proposed in the literature, including intrinsic motivation, monetary incentives, and social and self-esteem concerns, delegated philantrophy on behalf of stakeholders (Bénabou and Tirole, 2010), internalized norms or beliefs about others' behavior (Nyborg et al., 2006). **?** study non-professional investors and show that their investment decisions are primarily driven by their own attitudes toward sustainability and the desire to generate a good feeling.

[5]That is, it is unclear whether it arises due to investors' beliefs that SRI firms offer a better financial performance or due to preferences for ethical investments.

on the demand for SRI. By measuring individual preferences for SRI and beliefs about market prices, we are able to not only establish a positive demand for SRI investments but to also propose *why* people might demand these types of investments in markets. The motives for holding and buying SRI on a financial market differ qualitatively from motives for pro-sociality in other domain of economic behavior. We mimic holding SRI assets on a financial market by designing the experiment such that investors derive utility from *holding* the green asset but not from having an *impact* on the provision of the public good by buying a green asset. By ruling out a preference for having an *impact* on the public good, we concentrate on motivations like self-image or identity-related motives, which can be modeled in a general way as a feeling of "warm-glow" from holding the asset (Andreoni, 1989, 1990). We consider such motives to be closest to the motivation that play a role for investing in green assets on a real financial market from the perspective of a (small) individual investor.[6]

With our study, we also contribute to better understanding how behavior on markets reflects individual pro-social preferences. In a seminal contribution, Falk and Szech (2013) show experimentally that participants' willingness to accept harm done to an uninvolved third party increases in a (double auction) market compared to individual decisions and also in a multilateral compared to a bilateral market interaction. Bartling et al. (2015) find a stable and robust preference for avoiding negative externalities on uninvolved third parties among Swiss subjects. Sutter et al. (2019) propose that prices might be unaffected by the presence of (negative) externalities, but that trading volumes might decline in double auction markets. Contrasting these earlier market designs, our asset markets provides the crucial feature of resell. That is, we allow participants to trade their assets over several periods. Hence, it may not longer the case that individual preferences for SRI dives the market performance of assets, but expectations regarding others' preferences for SRI investments. If one speculates that prices of SRI assets increase in later periods, it may be profitable to buy those assets in early periods of the experiment. Even if the vast majority of traders having no *preference* for holding these assets, it is therefore possible to observe a price premium for SRI assets when sufficiently many traders expect higher

[6] This is actually a common feature of other types of markets as well. Consider for example energy markets, where an individual household decides to buy green energy, knowing that this will not have a direct impact on CO_2 emissions. In several real-world settings, we might furthermore be interested in predicting the demand for sustainable products on different markets, ranging from green energy consumption to sustainable investments, using observational data that stems from individual choices that impact the provision of the public good. For example, when combining field data with behavioral measures such as individuals' charitable giving, or simple social preference survey measures as in Riedl and Smeets (2017), this issue can occur. In this study, we test to what extend individual giving decisions are nonetheless a good predictor for behavior in markets with this characteristic.

prices at which they can resell the asset.[7]

Finally, the hypothesis that pro-social behavior in one domain can crowd-out generosity in another domain has attracted some attention in the literature; findings so far have been mixed. Evidence for a positive relationship between holding SRI assets and generosity in individual giving decisions has been found, for example, by Riedl and Smeets (2017) who link individual investor administrative data to self-reported investment behavior elicited in a survey and to behavior in two simple risk and social preference experiments. On the other hand, experimental evidence by Engelmann et al. (2012) and Munro and Valente (2015) point at the opposite relationship, albeit not for financial markets. They investigate the demand for so-called 'impure public goods', i.e. bundled goods that generates a private *and* a public payoff (e.g., Cornes and Sandler, 1994, Kotchen, 2005, Chan and Kotchen, 2014).[8] In fact, by designing our *green* assets as paying a bundled dividend, part of which is donated for public benefit and part of which is a private payout to the investor, this is the first study to investigate the demand for impure public goods in competitive financial markets with trading opportunities over several periods.

Next, we outline the experimental design in detail in Section 5.2 together with the predictions. The experimental results are reported and discussed in Section 5.3. Section 5.4 concludes.

5.2 Experimental Design

Each experimental session consists of three parts. In part 1, we elicit subjects' individual willingness to pay for the neutral and – when available – the green asset. Subjects can trade on the asset market in part 2. In part 3, they are offered the opportunity to donate to the same charity that receives part of the green asset's dividend in part 1 and part 2.

[7]We identify those effects within a novel market design that has more realistic features compared to those previously used in the tradition of Smith (1982). The assets value changes stochastically between periods and changes are related to the asset's current value. A non-deterministic fundamental value path (that does not decline deterministically to zero) is better suited for exploring issues like under- and overreaction to news about the assets' future value. With these modifications of the classical paradigm, the experiment mimics real-world markets more closely, while being simple enough to ensure subjects' understanding.

[8]Previous lab and field experiments have found a willingness to pay a price premium for a public benefit bundled with a private consumption good like organic cotton (Casadesus-Masanell et al., 2009a), certified toilet paper (Björner et al., 2004), charity-linked products ((Frackenpohl and Pönitzsch, 2013) in the laboratory, and Elfenbein and McManus (2010) in the field) or electricity from renewable sources (Kotchen and Moore, 2007). In the latter study, the authors show furthermore that altruistic attitudes influence the demand for such goods.

5.2.1 Experimental Procedure

5.2.1.0.1 Treatments. Throughout our experiment, subjects encounter two different types of assets. The assets differ in the structure of the dividends they pay:

Definition 3 (Green Assets)**.** *Shares of green firm pay an overall dividend of 5%. Half of the dividend is immediately credited to the owner's cash account at the end of a period, while the other half is a public benefit that is paid to a charity.*

Definition 4 (Neutral Assets)**.** *Shares of the neutral firm pay a private dividend of 5%. Half of the dividend is immediately credited to the owner's cash account at the end of a period, while the other half accrues to a locked account that is added to the subject's wealth only at the end of the session, before payment.*[9]

In both treatments, the same number of shares from each of two the firms are traded in the market. The stochastic returns are perfectly positively correlated across the two firms to rule out diversification considerations. This is necessary in order to clearly identify the effects of *green* preferences (and beliefs) while excluding risk-sharing motives. The only difference between treatments lies in the different structure of the two firms' dividend payments.Our experiment is structured into three main parts, followed by a questionnaire and subject payment, as illustrated in Figure 5.1.

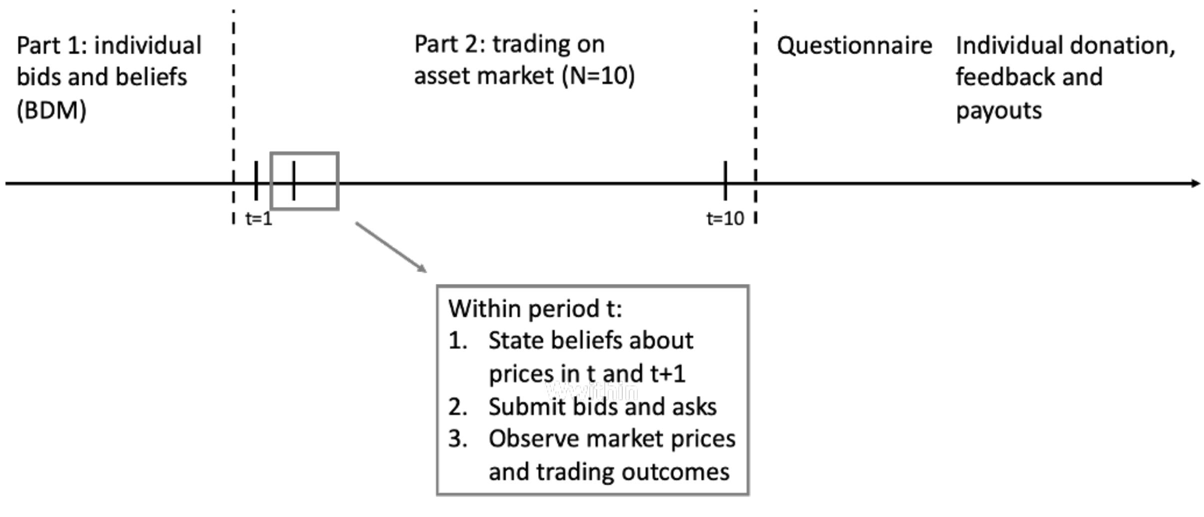


Figure 5.1: Timeline of an experimental session.

[9]Through this mechanism, the amount of cash that subjects receive from dividends and which is available for trading in subsequent periods is the same (in expectation) for neutral and green assets. This eliminates confounds due to differences in the cash/asset ratio between traders who hold equal amounts, but different types of, assets. See Palan (2013), Noussair and Tucker (2016), and the references therein for evidence on the effect of changes in the cash/asset ratio on mispricing.

5.2.1.0.2 Part 1. We start by eliciting each subject's individual willingness to pay for a ten period payoff stream of two *neutral* assets in *M_Baseline*, and of one *green* and one *neutral* asset in *M_Mixed.* Incentive compatibility is ensured by using a Becker et al. (1964) mechanism (hereafter BDM). The expected value of the *neutral* asset is the asset's value plus 10 times the dividend of 5%: $50 + 2.5 \cdot 10 = 75$ ECU. Likewise, the *green* asset's expected value that a risk neutral subject without green preferences (i.e., without a willingness to pay for donations to the charity) bid is $50 + 1.25 \cdot 10 = 62.5$ ECU. Since there is no interaction with other subjects in this first part, bids reveal the subjective willingness to pay for a *neutral* and for a *green* asset without any strategic incentives. Feedback concerning the outcome of the BDM mechanism is provided only at the end of the experiment. To obtain a measure for participants' beliefs about *other* participants' valuations of the *green* and the *neutral* asset, we elicit individual beliefs concerning the average willingness to pay of the other participants in the session. We incentivize this question using a linear scoring rule: a subject receives 1€ for the correct guess. For each ECU she deviates from the correct guess, 2 cents are subtracted from this 1€ until she misses the correct value by more than 50 ECU.

5.2.1.0.3 Part 2. In part 2, subjects trade assets in a double auction with ten other participants. At the beginning of each period, we ask participants to state their beliefs about the market prices of both assets one and two periods ahead, incentivized in the same way as in part 1. Then in each period of the market, subjects can submit an unlimited numbers of bids and asks for both assets (with the only restriction that prices of offers have to be at least as high as bids). After all subjects have finished submitting their bids and asks, the computer calculates the market-clearing price for the respective bids and asks submitted by the participants.[10] At the end of the ten periods, all assets that a subject holds are bought back at the assets' values in period ten.

5.2.1.0.4 Part 3. After part 2 is over, subjects are given the option to donate (part of) their show-up fee of 5€ to the charity 'Atmosfair', i.e. the same charity that also receives half of the dividend of *green* assets in parts 1 and 2.[11]

[10] The price determination algorithm follows the procedure used at the NASDAQ exchange, as outlined in Palan (2015). Specifically, the algorithm chooses the single price which (1) maximizes the feasible trading volume, (2) minimizes the absolute oversupply, and (3) accounts for a surplus of buy or sell volume. If all three of these criteria (which are ordered by priority) together fail to produce a unique auction price, the middle of the range of prices fulfilling the three criteria is chosen as the auction price.

[11] One might worry that the type of charity we chose for this experiment could affect the results we find on the nature of pro-social behavior. Picking a specific charity might well alter the *level* of demand for the charitable donation in all parts of the experiment (individuals' preferences for another charity could have been stronger or weaker). However, it is unlikely to affect the *differences* in pro-social behavior between

5.2.2 Market design

Groups of ten subjects trade assets for experimental currency units (ECU) over ten periods in an experimental, sealed, bid-ask double auction market. The asset's nominal value V starts out at $V_{t=0} = 50$ ECU at the beginning of the first period. At the end of each period, the nominal value grows by $r_t \in \{-25\%, -5\%, 10\%, 40\%\}$, where each possible value is equally likely. The average growth rate thus is 5%. At the same time, the asset pays a fixed dividend of 5% (also from its initial value in period one), such that the expected nominal value after dividend payment is constant. At the end of any period, the new nominal value thus is calculated as:

$$V_{t+1} = V_t \cdot (1 + r_{t+1} - div) \tag{5.1}$$

Here, *div* is the constant (relative) dividend of 5%. Since the expected after-dividend growth at time t is zero, i.e., $\mathrm{E}_t\left[r_{t+1} - div\right] = 0$, fundamental values are constant in expectation:

$$\mathrm{E}_t\left[V_{t+1}\right] = V_t \quad \forall t < T \tag{5.2}$$

The average period return of 5% is a risk-premium over the risk-free rate of return of $r_f = 0$ on subjects' ECU holdings.

This design yields nominal asset values (i.e., asset values excluding dividends) being *constant in expectation*, thus representing a martingale. The plots in Figure 5.2 show the nominal value paths used in the experiment.[12] The constant expected nominal value (from the perspective of $t = 0$) is drawn in gray. The nominal value paths are the 'correct' view of the fundamental value's development for subjects whose risk-aversion is just compensated by the risk-premium of 5%. Conversely, a risk-neutral investor should judge the asset's price in comparison to its fundamental value that includes the expected value of the remaining future dividend payments.

treatments or tasks. We do not expect measures of the correlation of individual pro-social behavior across tasks as well as measures of the impact of pro-social traders on our market indicators and the determinants of the demand for green assets in markets to be affected.

[12] To obtain the value paths presented in Figure 5.2, we randomly drew five paths using the four possible return realizations $r_t \in \{-25\%, -5\%, 10\%, 40\%\}$. For the remaining five paths, we inverted the first five paths as follows: We replaced each instance of a growth of -25% in an original path by a growth of 40%. We replaced each growth of -5% by 10%, 10% by -5%, and 40% by -25%. By replacing the original returns r_t using the alternative vector $\dot{r}_t \in \{40\%, 10\%, -5\%, -25\%\}$, we thus obtain a new price path which inverts the original price path, while retaining the same theoretical distribution, including the mean of $\bar{r} = 5\%$. This procedure follows the idea of Kirchler (2009). Pairing every market with higher-than-expected returns with a related market with lower-than-expected returns ameliorates the influence of specific individual realizations of the price paths.

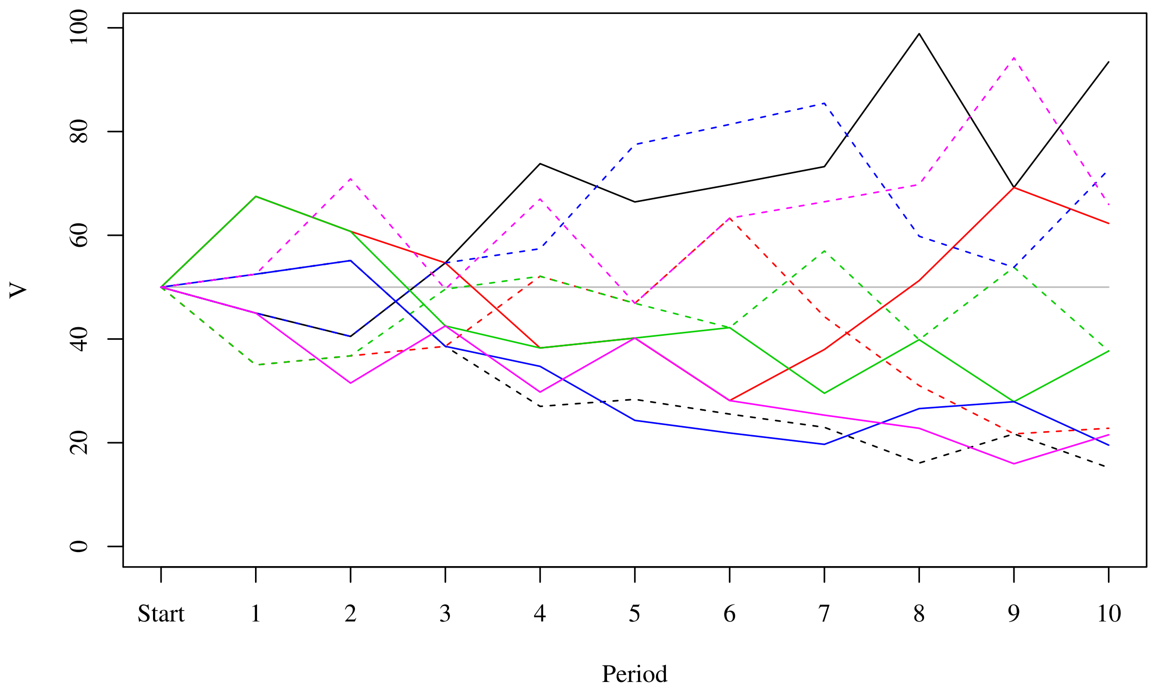


Figure 5.2: The ten nominal value paths used in the experiment. Five paths were drawn randomly (solid lines) and then inverted to create the other five (dashed lines).

Our design mimics more closely markets outside the lab than market paradigms like, for instance, Smith (1982) and Smith et al. (1988). The latter designs have been criticized for a lack of realism (e.g., Kirchler et al., 2012): first, percentage changes in asset value are more realistic than absolute ECUs changes.[13] Second, our design yields a non-deterministic fundamental value path, which is better suited for exploring issues like under- and overreaction to news about the assets' future value (while constant in expectation, the fundamental value is random and non-stationary). Third, our asset's fundamental value does not decline deterministically to zero as those used in studies based on the paradigm by Smith et al. (1988).[14] Despite these more realistic characteristics, the asset is designed to be simple enough to minimize possible confusion among our experimental subjects. Provided today's nominal value, the asset's expected future

[13]We choose the dividend to be 5% of the fund's value at the beginning of the period in order to have a round number (a dividend of 4.76% times the value at the end of the period − after growth − would achieve the same) that equals the average growth rate of the fund. Conceptually, our design mirrors a situation where the fund fixes the dividend to be paid at the end of the period at the beginning of a period.

[14]See Palan (2013), Powell and Shestakova (2016) and Nuzzo and Morone (2017) for a comprehensive overview of this literature.

nominal values equal today's nominal value and are constant. Thus, a subject's best forecast of the future nominal value is the current nominal value. Furthermore, we refrain from using a continuous interval of possible returns as in, for instance, Kirchler (2009). Instead, we stick to the design of Smith et al. (1988) of confronting subjects with only four possible, equi-probable period returns that remain unchanged for the entire duration of the experiment. Finally, the asset's dividend is certain and constant relative to the current nominal asset value.[15]

5.2.3 Implementation

We run 10 sessions with a total of 200 participants, 100 per treatment. All experimental sessions were conducted at the experimental laboratory of the School of Economics and Social Sciences at the University of Hamburg, Germany, in September and October 2015. Almost all participants were students (2% nonstudents) with various academic backgrounds, 56% were female and the median age was 24. We used GIMS 7.0.10 (Palan, 2015) and z-Tree 3.4.7 (Fischbacher, 2007) to program the experiment, and hroot (Bock et al., 2014) for recruiting. Each subject participated in only one treatment condition. The average payoff amounted to €18.52 and each session lasted approximately 120 minutes.

At the start of the experiment, subjects were given printed instructions for part 1.[16] The experimenter announced that there would be a second part, but the instructions for part 2 were only distributed after part 1 had been completed. All instructions were read out aloud by the experimenter to establish common knowledge among the subjects, and subjects were given the opportunity to ask questions (in private) before and during the session. Then subjects answered a set of control questions to ensure understanding. Before trading on the experimental market, subjects participated in three practice periods of trial order submission. During the practice periods we performed no order matching and allowed no interaction with other subjects. Rather, subjects interacted solely with the computer. After completing parts 1 and 2 of the experiment, subjects answered a standard socio-demographic questionnaire, augmented by some risk-related questions (Dohmen et al., 2011; Harrison et al., 2005), and were given the option to donate their show-up fee.

After finishing the experiment, each subject was paid in private and the experimenters executed the individual wire transfers to the charity 'Atmosfair'. We employed several measures to foster subjects' trust in the donation outside the lab (that were announced

[15] Smith et al. (2000) show that dividends play an important role in triggering mispricing. Porter and Smith (1995) find no difference in the effects of random and certain dividends on mispricing.

[16] English translations of the experimental instructions can be found in Appendix B.

in the instructions): we distributed a handout with the description of the charity that subjects were allowed to take with them to look up the charity's name after the experiment if they wish to do so. We transferred each subject's donation in their presence, while handing them out their cash payment (both in private). After all sessions had been completed, we furthermore sent an e-mail to all participants with a receipt of all donations made for this experiment.

5.2.4 Predictions

Our main research question concerns the impact on market developments of introducing a green investment opportunity. The demand for *green* assets in a market depends on individuals' willingness to pay and thus our predictions crucially depend on the assumptions we make regarding traders' preferences. We begin with stating the null hypothesis that traders are rational, risk-neutral, and self-interested, i.e., that they have no particular concern for the public component of the dividend of the *green* firm in *M_Mixed*. In this case, market prices of the assets of the *green* firm and the *neutral* firm in treatments *M_Mixed* and *M_Baseline* should closely follow the fundamental value paths. Note that the assumption of "no concern for the public component" may allow for the individual to have a preference for the *provision* of the public good nevertheless: Due to the way we design the public dividend, the public good will be provided irrespective of who owns shares of the green asset at the end of the experiment.

The alternative hypothesis states that traders have a willingness to pay for the *green* assets that exceeds the expected private payoff. If prices deviate from expected fundamental values, this can be explained by two sorts of discrimination in *M_Mixed*, namely taste-based discrimination and strategic considerations. First, a trader might have a *preference for holding assets* of a firm engaging in CSR, even though her action has no direct impact on the public good provided. This motivation can be represented as a utility from the act of buying and holding the green asset as in a model of 'warm-glow' (Andreoni, 1989, 1990). [17][18] Second, speculation, i.e., traders' expectations about *other* participants' willingness to pay for the *green* component, can lead to a price premium for the *green* asset. Even in the absence of *own* pro-social concerns, such expectations can lead to higher individual bids and consequently higher market prices—if a sufficiently large share

[17]Possible psychological underpinnings of such a feeling of warm-glow are a preference for maintaining a certain self-image, or identity concerns. Several experimental studies have shown that theses concerns can drive pro-social behavior (see for example Gneezy et al., 2012, Ariely et al., 2009, Andreoni and Bernheim, 2009, Bénabou and Tirole, 2006). Evidence in Teyssier et al. (2014) suggests that self-image concerns play a role in determining the demand for impure public goods (fair trade chocolate in their experiment)—next to social-image concerns.

[18]Social image and reputation concerns are excluded by the anonymity in our study.

of traders expect others to exhibit a positive willingness to pay for the socially responsible component.

Prediction 1 (M_Baseline)**.**

(i) H_0*: As both firms are identical in the unmixed market and the random draws of stock returns are perfectly positively correlated, there will be no difference in market prices between firms. With rational investors, opportunities for cross-arbitrage do not occur.*

Prediction 2 (Green Premium in M_Mixed)**.**

(ii) H_0*: Self-interested investors should not have a willingness to pay that exceeds the private expected value for the green assets.*

(iii) H_1*: The existence of traders with warm-glow preferences, and of traders who hold the belief that others have a willingness to pay for the green component, can ultimately lead to higher demand for the green assets in the market. Consequentially, market prices will exceed the respective expected fundamental values (identification of a "green premium").*

How will individual donations in Part 3 relate to trading in the asset market? A trader without pro-social concerns neither exhibits a preference for holding green assets, nor will she donate. She might, however, expect others to have a high willingness to pay for the green asset and therefore she might submit higher bids, even though she does not donate a positive amount in part 3. If a trader has a preference for contributing to the provision of the public good (rather than a warm-glow utility), this would not show in her trading behavior, but it could lead her to donate a positive amount in part 3. We expect to observe a positive correlation between behavior in the donation decision and on the market if participants have warm-glow preferences like self-image concerns. In this case, they are expected to do both, donate a positive amount and demand green assets.

Based on these arguments, we distinguish the following behavioral patterns across domains: (1) No green willingness to pay in the market and a positive donation indicates the importance of having an impact with one's pro-social action as a driver of ethical behavior. (2) A green willingness to pay but no donation hints at the importance of speculative motives, i.e., at beliefs about the future demand for green assets. (3) If there is neither a green willingness to pay in the market nor a donation, a trader's preferences can be represented by a classical selfish utility model. (4) Both motives—beliefs and green preferences—can play a role and can interact, which leads to a positive willingness to pay in the market and positive donations. Disentangling these will be the subject of a rather exploratory analysis.

Prediction 3 (Behavior across domains)**.**

(iii) We expect pro-social behavior to be either positively correlated or independent across market interactions and individual decisions.

The literature also advances an alternative hypothesis regarding social preferences, namely a "crowding out" of generosity by previous green behavior in the market. What type of preferences would lead to a crowding-out of donations after having traded green assets in the market, i.e., to a negative correlation between behavior in the two domains? This hypothesis requires additional assumptions on people's preferences. The most prominent argument put forward in the literature is so-called 'moral licensing' (see the meta-analysis by Blanken et al., 2015).[19] Trading-off one's moral behavior can lead a subject to donate less, the more *green* assets she bought and held in the market.

5.2.5 Empirical Identification

5.2.5.0.1 Market indicators (predictions 1 and 2). In order to assess the impact of green investment opportunities on the market, and to test prediction 1, we compare market prices and trading volumes between the two treatments, *M_Mixed* and *M_Baseline*. Our main prediction concerns market prices. We assess the impact of green investment opportunities in the development of market prices of the two types of assets and especially the extend of mispricing. To calculate relative prices, we correct for each periods fundamental value. The fundamental values are composed of the current period's base value plus the sum of the private dividends of all remaining periods. In addition to analyzing (relative) price trajectories, we compare the extend of mispricing of each type of asset using two measures that are commonly employed for measuring the occurrence of bubbles in the literature.[20] Both measures the average difference between the market price and the fundamental value in each period (Powell, 2016, 57). In our market with stochastic fundamental values, we conceptualize mispricing by measuring the distance of bids from *expected* fundamental values (which is by design always the last period's fundamental value). We calculate the geometric absolute deviation as a measure for mispricing:Note

[19] For example, Tiefenbeck et al. (2013) show that moral licensing may be able to explain shifts in the demand for different green goods: they find negative spillovers of environmentally friendly behavior in one domain on another, related domain, which offset the positive effect of an intervention to promote ethical behavior.

[20] For an overview see Stöckl et al. (2010).

that the deviations have a lower bound at 0% but no upper bound. In order to ensure numeraire independence, the GAD uses the natural log of the deviations to account for the asymmetry (see also Powell (2016)).

$$GAD = exp\left(\frac{1}{N}\sum_i \left| ln\left(\frac{price_i}{fv_i}\right)\right|\right) - 1$$

To ensure robustness of our results, we perform the same analysis with the mispricing measure proposed by Stöckl et al. (2010), the relative absolute deviation (RAD), and report the results in the appendix.

$$RAD = \frac{1}{N}\sum_i \left|\bar{price_i} - fv_i\right| / \left|\bar{fv}\right|$$

, where $price_i$ and fv_i are the periods price and fundamental value and $\bar{fv}$ is the average fundamental value of the market

5.2.5.0.2 Individual behavior in the market (prediction 3 and exploratory analysis). Our design allows for the identification of individuals' preferences and beliefs, based on which we perform an exploratory analysis of psychological mechanisms underlying the demand for *green* assets. The deviation of a subject's bid under the BDM from the *neutral* asset's fundamental value, i.e., the rational valuation of a risk neutral subject, serves as a measure for the subject's risk preferences. The additional deviation in the *green* bid (in percentage points from the fundamental value) captures a *green premium* the subject is willing to pay. Differences in the demand for the *green* and *neutral* assets might be driven by pro-social types of players who are willing to pay a green premium. We designed the experiment such that we can identify participants' willingness to pay for the *green* component in an individual decision *before* they are introduced to the market. We test to what extend a green premium in subjects' bids in part 1 predicts an individual green premium in the market. Our identification can be seen as a lower bound: out of all participants who exhibit a pro-social concern in the individual decision, only those whose concern is driven by motives that can be described as a warm-glow utility – which is not driven by a desire to have an impact on the public good – are expected to also show pro-social behavior in the market.

Figure 5.3 shows that bids for the *green* and the *neutral* assets (relative to fundamental values) in part 1 are positively correlated; the *green* relative bid increases by about .5 for a 1.0-point increase in the *neutral* relative bid. We define "*green*" and "*not green*" types

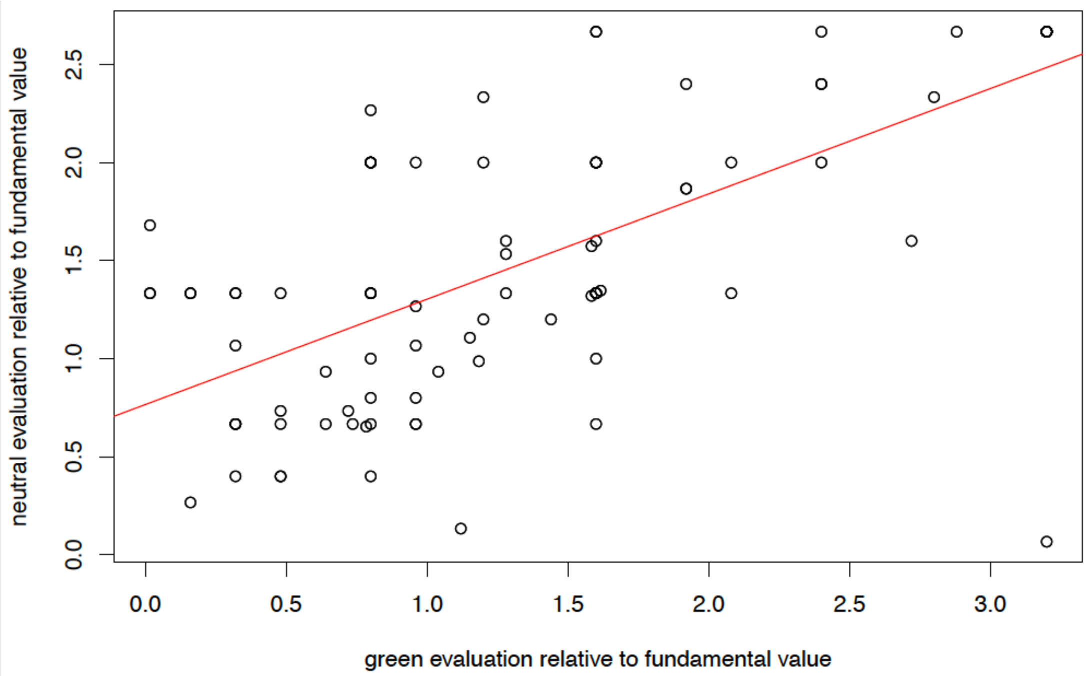


Figure 5.3: Valuations, i.e. bids for the green and neutral assets relative to their respective fundamental values in part 1 ($\mathrm{FV}_{nt} = 75, FV_{gr} = 62.5$).

of subjects based on the difference between these two variables. The resulting binary type variable is rather symmetrically distributed around zero and it has the advantage that general overbidding in part 1 cancels out and thus does not confound our analysis. We define

Definition 5 (Green_Type)**.** *We say that a participant is a* $Green\_Type = 1$ *if and only if* $Green\_Bid_i \div Green\_EV > Neutral\_Bid_i \div Neutral\_EV$.
A participant is of $Green\_Type = 0$ *if and only if* $Green\_Bid_i \div Green\_EV \leq Neutral\_Bid_i \div Neutral\_EV$.

We derive classifications for identifying more or less *green markets* from this type definition. A market is called *green* if the number of green types is higher than the median (and mean) of 5.5 out of 10 subjects per market.

Definition 6 (Green_Market)**.** $Green\_Market = 1$ *if the number of subjects with* $Green\_Type = 1$ *in this market is larger than the median number. Otherwise* $Green\_Market = 0$.

An alternative classification takes the *strength* of the green preferences subjects exhibit in part 1 into account and serves as a robustness check: According to this definition, a market is called a *green market* if the average difference in bids for the green and neutral assets in part 1 is larger than zero (i.e. if $\sum(Green\_bid/62.5 - Neutral\_bid/75))/10 > 0$).

In addition to preferences, we hypothesized that speculation, i.e., beliefs about future market developments, drives the demand for *green* assets. When the trajectory of market prices over the ten periods follows that of the (expected) fundamental value, beliefs may not be able to explain much of the variance in market prices after controlling for fundamental value. However, if prices *deviate* from this path, we explore if beliefs are able to explain the deviations. On the individual level, we explore to what extent beliefs matter for the bids and asks individuals submit for the *green* and the *neutral* assets. We calculate a subject's valuation of, say, the *green* asset, as the average of her maximum willingness to pay (WTP, the highest bid she submits) and her minimum willingness to accept (WTA, the lowest ask she submits) in a given period.[21] In addition to investigating the relationship between participants bids and beliefs in a given period, we analyze to what extent her optimism or pessimism about the *trend* of the market prices influence bids. Subjects predict future market prices of each asset for periods $t+1$ and $t+2$ in each period t, $Belief\_MP\_t+1 >$ and $Belief\_MP\_t+2$. Based on these two incentivized belief measures, we define an individual as an *Optimist* if she believes the market price to increase:

Definition 7 (Optimist)**.** *We call a subject an Optimist* $= 1$ *if and only if* $Belief\_MP\_t+2 > Belief\_MP\_t+1$*. A subject is of the type Optimist* $= 0$ *if* $Belief\_MP\_t+2 \leq Belief\_MP\_t+1$*.*

5.3 Results

5.3.1 Individual Bids in Part 1

While the rational valuation of the *neutral* asset by a risk-neutral subject is 75 ECU, the average bid is 116.25 ECU in *M_Mixed*. Correspondingly, for the *green* asset with a risk-neutral total dividend value of 62.5 ECU, the average bid is 91.33 ECU. The difference between bids for the *neutral* and the *green* asset in *M_Mixed* is highly significant ($p < .001$, 2-sided Wilcoxon signed rank test of equality of distributions (WSR), t-test, bootstrapped t-test of equality of means (BTT)). The mean bid for the *neutral* asset in *M_Mixed* is insignificantly higher than the mean bids in the *Baseline*

[21]We refrain from weighting the WTP or WTA by the number of assets a subject demands or offers at this price because by doing so we would weight different periods differently and we prefer to keep this index equal across all periods of the market interaction.

treatment with 105.5 ECU and the median bids are the same (100 ECU as opposed to 80 ECU for the green asset, see Figure 5.4.[22] Relative to each asset's fundamental value, participants' bids under the BDM mechanism show substantial overbidding for both types of assets to an equal degree: Subjects bid on average 155% of the *neutral* asset's fundamental value and 146.13% of the *green* asset's fundamental value. Taking into account the general overbidding, average bids in the individual decisions in part 1 do not exhibit a 'green premium', i.e., a willingness to pay for the green component.

Participants' beliefs about other participants' bids show that they systematically underestimate others' bids, stating average beliefs closer to (but still higher than) the expected values. Figure 5.4 also shows that beliefs about others' bids are less dispersed than actual bids, especially in the baseline treatment. While estimates are smaller than actual bids for *green* and for *neutral* assets in *M_Mixed*, the difference is only statistically significant for the *neutral* assets ($p < .01$ WSR, BTT, t-test).[23] In *M_Baseline*, average bids are underestimated by a similar magnitude as in *M_Mixed* ($p < .05$ (BTT, t-test), $p < .1$ WSR), i.e., bids are quite consistently underestimated by about 10 ECU on average. Compared to fundamental values, subjects expect others to overbid for each asset by roughly the same amount (*neutral*: 139.03% of FV, *green*: 130.70% of FV). They thus partly predict the substantial overbidding for both assets and correctly predict the lower demand for the *green* asset in *M_Mixed*.

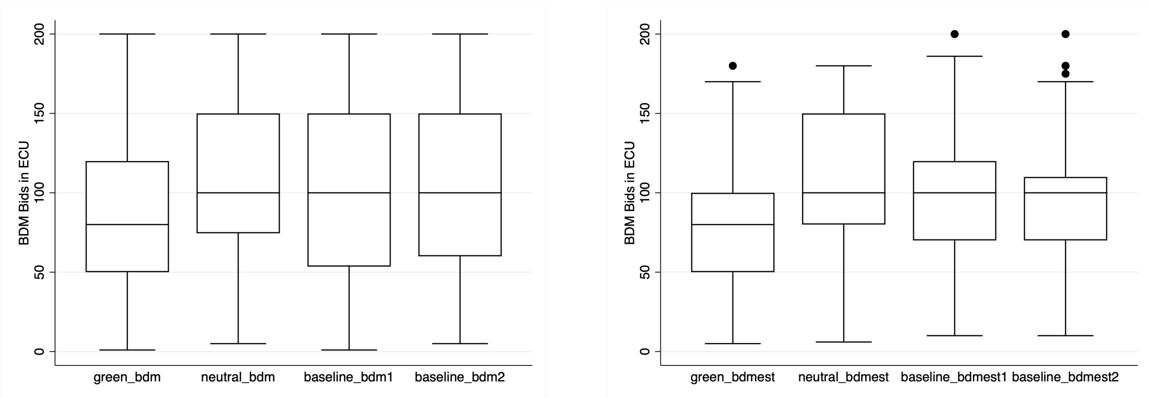


Figure 5.4: Bids in ECU and beliefs about other participants' average bids, BDM mechanism, part 1.

[22]As expected, the differences in average bids and estimates between assets of the two equal firms in *M_Baseline* are minor and statistically insignificant. We therefore report the pooled variables in the text and in Figure 5.4.

[23]The difference for *green* is insignificant according to WSR and significant at the 5% level according to t-test and BTT.

5.3.2 Trading in the Asset Market

5.3.2.1 Market Developments

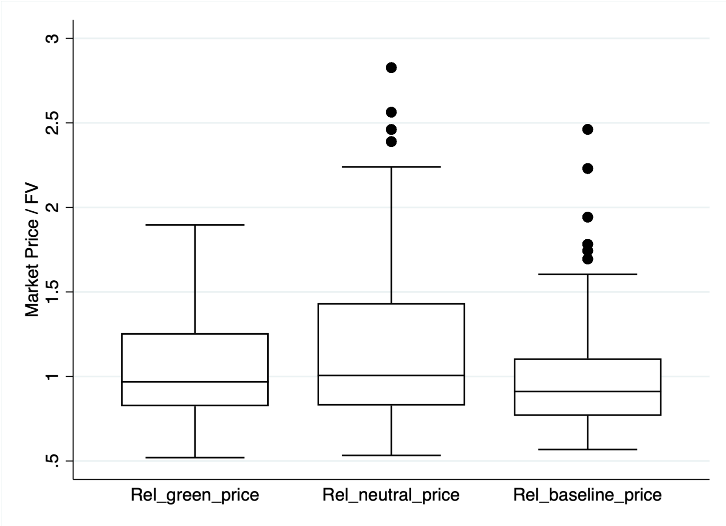


Figure 5.5: Mean relative prices (Price/FV) over all periods, by asset type and treatment.

5.3.2.1.1 Market prices. Taking both types of assets together, overall prices on the mixed and the homogeneous markets are not significantly different (MWU, $p = 0.13$).[24] Averaged over the ten markets in *M_Mixed* and over all ten periods, we find that market prices for the *green* and the *neutral* assets differ significantly ($p < .001$ WSR, t-test, BTT), with the *green* market prices being on average lower (52.92 vs 63.29 ECU). We find the same result comparing prices relative to fundamental values (WSR, p=0.002). Figure 5.5 furthermore shows that relative *green* prices in *M_Mixed* are not significantly different from prices in *M_Baseline* (MWU, $p = 0.7$); the same holds for absolute prices. The *neutral* assets on the other hand are overall traded for a higher average market price in *M_Mixed* than the same assets in the homogeneous market *M_Baseline* (mean in *M_Baseline*=54.99 ECU, $p < .001$ MWU, t-test, BTT). This suggests that in the market with two different types of firms, the one with the higher private dividend becomes relatively more attractive, outperforming the exact same assets on homogeneous markets. It also suggests that, even though the private dividend is smaller, the *green* prices reached in the mixed markets are similar to those of the *neutral* assets in *M_Baseline.* Thus, demand for the *green* asset does not disappear in the market: despite their lower private payoff, they perform as well as the *neutral* ones on a market without *green* firms.

[24]Notice that we present pooled results for the two assets in *M_Baseline* since, as expected, there is no significant difference between the two assets' prices ($p > 0.1$, two-sided Kolmogorov-Smirnov test).

Over the course of the ten period, *green* prices track *neutral* baseline prices relatively closely while the *neutral* prices in *M_Mixed* rise in the second half of the periods, see Figure 5.6. While *green* and *neutral* baseline prices reflect the fundamental values of the respective assets, *neutral* prices in *M_Mixed* clearly rise above the rational valuation and exhibit a bubble. The finding that *baseline* and *green* prices do not deviate substantially from fundamental values over time in an asset market with constant nominal value resonates with Kirchler et al. (2012), who find that constant fundamental values lead to significantly fewer and smaller bubbles compared to declining fundamental values as used by Smith et al. (1988). Our results indicate that this partly holds also for paths that are non-stationary in expectation.

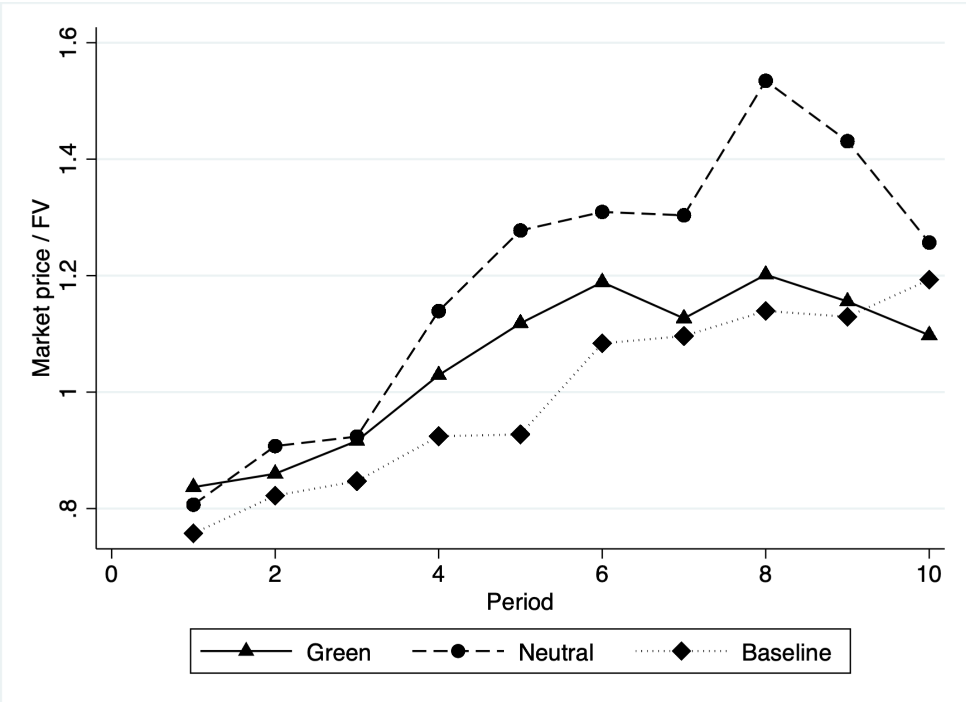


Figure 5.6: Development of relative market prices (price/FV) over time by treatment and type of asset.

5.3.2.1.2 Mispricing indicator. Table 5.5 summarizes the geometric absolute deviation (Powell, 2016, 59) for both treatments.[25] This measure summarizes mispricing over all periods for each type of asset and each type of market. The mispricing indicator is significantly larger for the *neutral* than for the *green* assets in *M_Mixed* ($p < .05$ WSR, t-test, $p < .1$ BTT), confirming our previous finding of a *neutral* bubble. With a

[25]The neutral market price in market 101 in period 2 has been excluded for all analyses in this section since it is a stark outlier and confounds results.

value of .26, *green* prices deviate on average over all 10 periods by about 26% from the fundamental values, compared to 37% for *neutral* assets. Mispricing is overall higher in the mixed markets due to the price bubble for the *neutral* assets, but the difference to mispricing in the homogeneous markets is not statistically significant. Again, *green* mispricing is at a similar level as overall mispricing in the baseline markets. As a robustness check, we provide the same analysis using the mispricing measure by (Stöckl et al., 2010), the relative absolute deviation (RAD), in the appendix. All results hold when using this measure instead of the GAD (here, the difference between green and neutral mispricing is not statistically significant according to BTT).

	M_Mixed		M_Baseline
	Green Asset	Neutral Asset	Baseline Assets
GAD	.258	.370	.278

Table 5.1: Mispricing measures for each asset and treatment: absolute geometric deviation over the N=10 periods.

Result 11 (Market prices)**.**

1. ***M_Baseline.*** *We find no significant differences between prices of the two assets in markets with two equal firms, and no price bubbles with nominal values that are constant in expectation.*

2. ***M_Mixed—Price premium.*** *There is no price premium for green assets in the mixed markets. However, the demand for green investments does not disappear in the market as green prices remain as high as market prices in the homogeneous baseline markets with only neutral assets.*

3. ***M_Mixed—Spillover effects.*** *In the presence of a green firm, prices for the neutral assets increase above the fundamental value in the second half of the trading periods, leading to a bubble with higher market prices than those observed for the green assets, and than those observed in the homogeneous baseline markets.*

5.3.2.1.3 Trading volumes. Contrary to market prices, market activity seems to increase over time in the mixed market: the overall trading volume amounts to an average of 47.5 assets traded per period as compared to 45.0 in *M_Baseline*, which is a statistically significant difference ($p < .001$ MWU, t-test, BTT). Trading volumes of *green* assets are on average somewhat lower than for *neutral* assets in *M_Mixed*, however, the differences are less strong than for market prices (46.1 versus 48.9, $p \leq .05$ WSR, t-test, BTT), see also Figure 5.7. These observations clearly contradict the suggestion by Sutter et al. (2016) to "look for the morals in markets" in trading volumes rather than in market prices. Figure 5.7 rather indicates that subjects seem to trade both assets *simultaneously*.

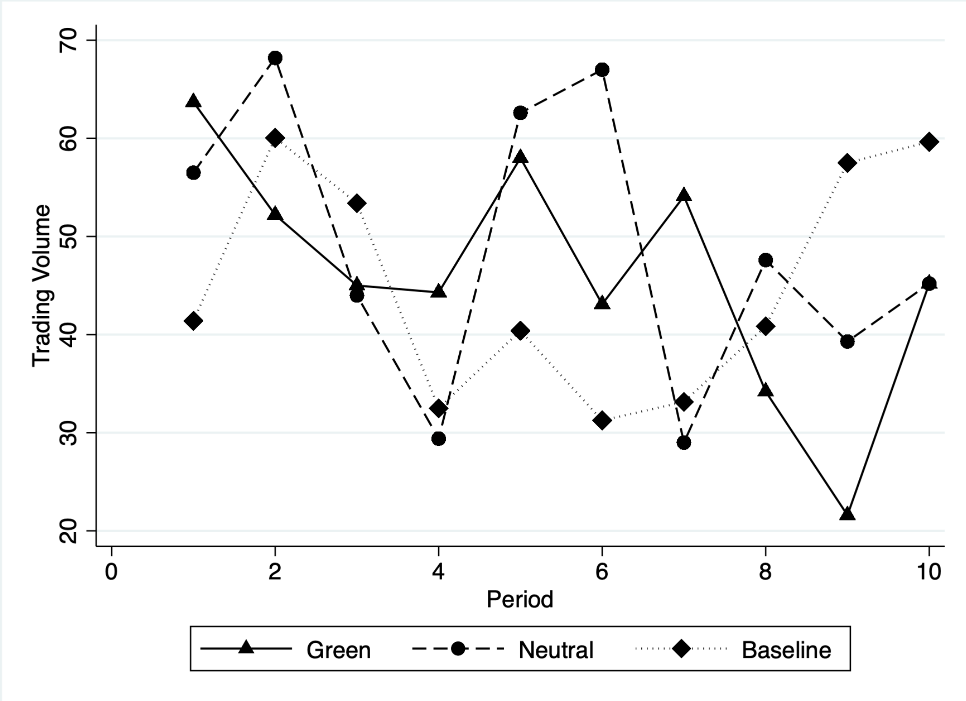


Figure 5.7: Trading volume (average number of assets traded in a market) over time by type of asset and treatment.

Result 12 (Trading volumes)**.**

1. *Trading volumes are overall significantly higher in the mixed markets compared to the homogeneous baseline markets. In M_Mixed, the difference between trading volumes of green and neutral assets is rather small with slightly less green assets being traded.*

We will now take advantage of the different fundamental value paths generated in our experiment to better understand why and how mispricing occurs or does not occur. Figures 5.8 and 5.9 plot the market prices of both types of assets together with the market's fundamental values for each of the ten markets in both treatments. Overall, the price and fundamental value paths indicate that, when mispricing occurs, it is rather due to the *stickiness* of prices than due to an (upward) deviation of prices (except for a few cases such as the neutral price trend in market 901, figure 5.8). In other words, prices tend to be overall more flat and only follow the developments of fundamental values with some inertia and to a lesser extend. This is in line with previous findings from experimental asset markets: In traditional market designs where the fundamental value

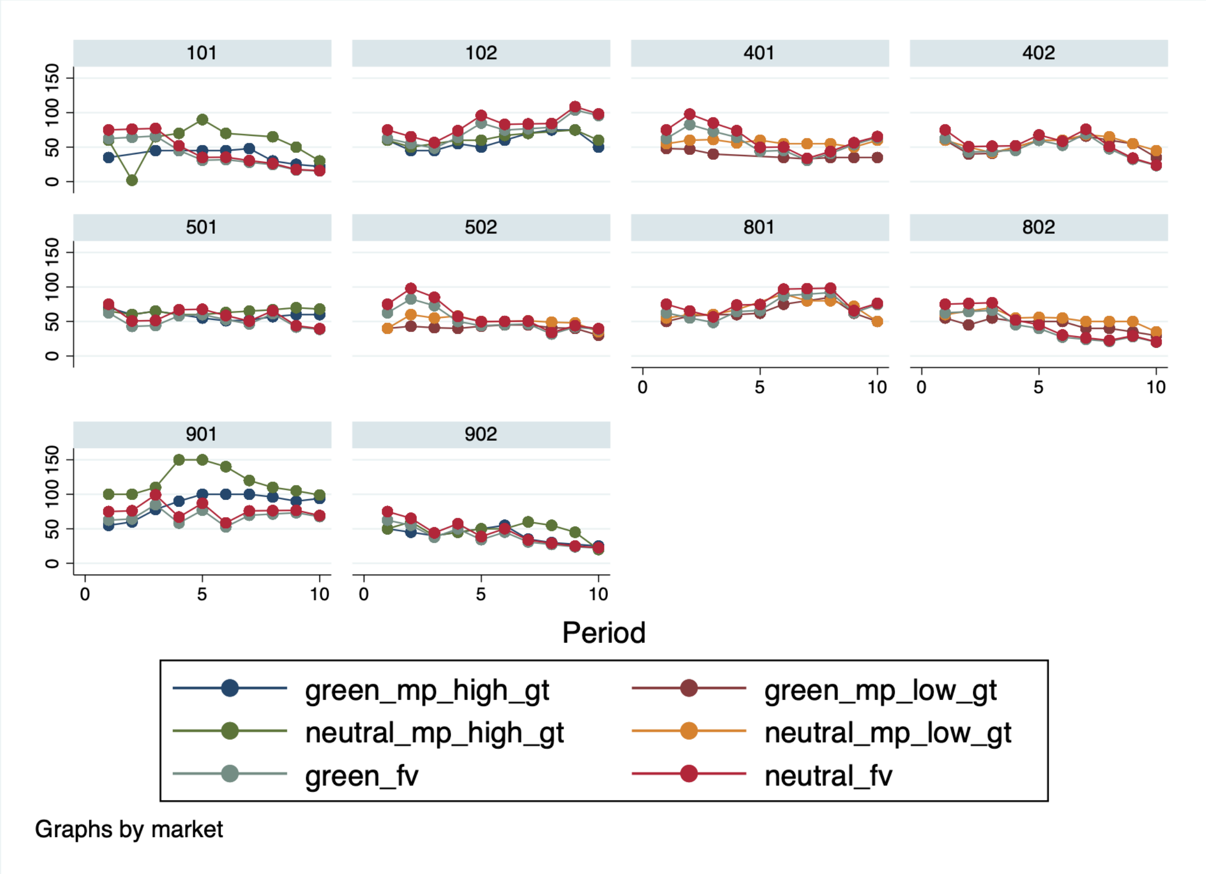


Figure 5.8: Green and neutral market prices and fundamental values in each mixed market, by high and low number of green types.

declines deterministically to zero, bubbles occur when prices do not decline as much, i.e. when prices are more sticky. In markets with constant fundamental values, usually no bubbles occur (for an overview see (Palan, 2013)). This is exactly in line with our general observation that mispricing occurs when prices are sticky and exhibit a less pronounced upward or downward movement.

5.3.2.2 Bids and Beliefs in the market

5.3.2.2.1 Beliefs. In *M_Mixed*, subjects' stated beliefs over market prices in the subsequent period show that they anticipate the differences in demand between the two types of assets: Expected market prices are on average lower for *green* than for *neutral* assets (mean beliefs are 51.2 versus 62.0, median beliefs are 50 vs 55 ($p < .001$ WSR, t-test, BTT)).[26] As for actual market prices, expected prices for the *neutral* assets in *M_Mixed*

[26]Three participants had to be excluded from this analysis as they stated clearly not serious beliefs such as 99999.

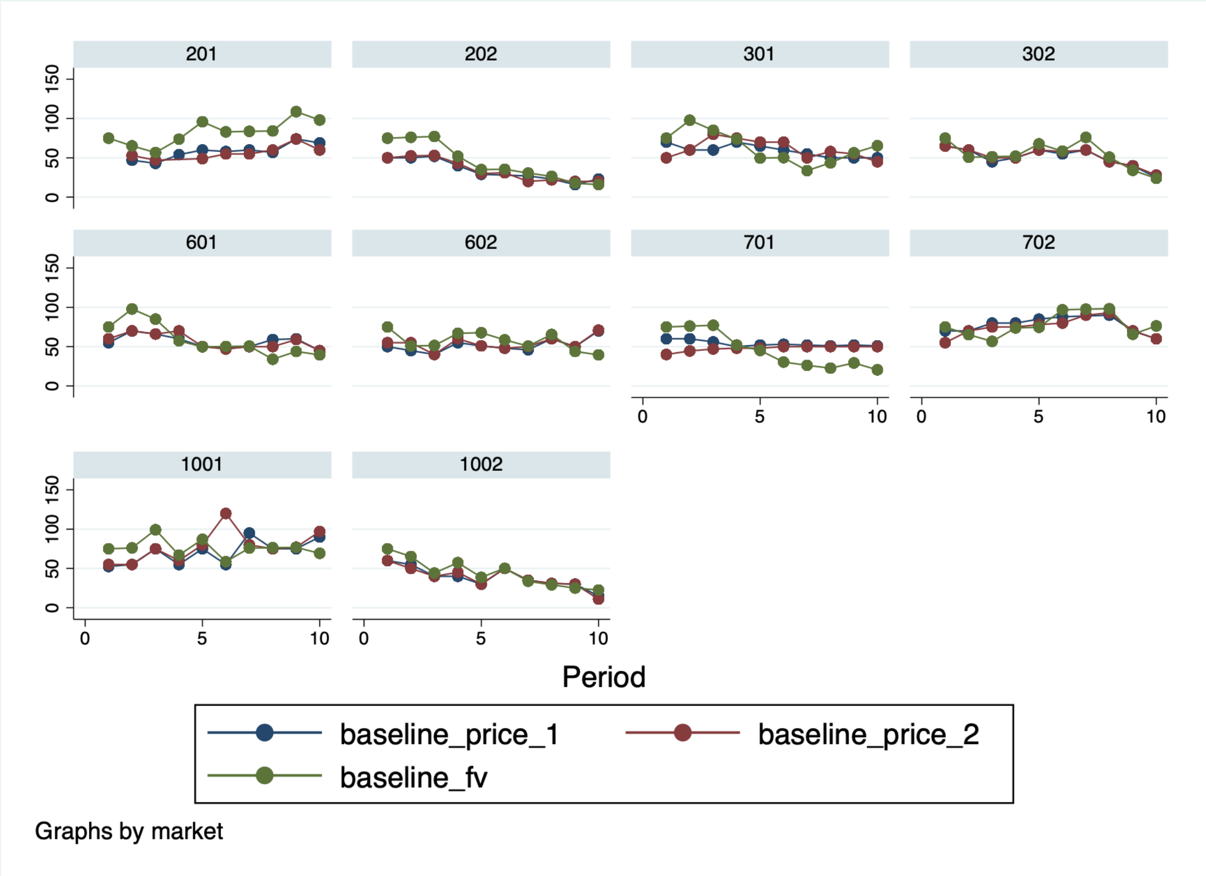


Figure 5.9: Market prices of both neutral assets and fundamental values in each baseline market.

are higher than expected prices for the same type of asset in the homogeneous baseline markets (mean=55.15, median=51, $p < .001$ MWU, t-test, BTT). *Green* beliefs are even significantly lower than beliefs about the performance of *neutral* assets in the baseline markets ($p < .001$ MWU, t-test, BTT) even though the magnitude of the difference is very small. When correcting for fundamental values, that take into account that the private dividend is smaller for the *green* assets, the relative *neutral* belief (aver belief/fundamental value=1.15) is still significantly larger than both, the *green* and baseline relative beliefs ($p < .001$ MWU, t-test, BTT for both), but the difference between the *green* and baseline beliefs over all periods and markets reverses with beliefs relative to fundamental values being somewhat lower in *M_Baseline* (1.04 versus 0.98, $p < .05$ MWU, t-test, BTT for both). Do these beliefs matter for trading behavior? Prices should follow (expected) fundamental values, but we hypothesize that beliefs might be able to explain deviations from fundamental value paths. Indeed, while we cannot establish a causal link, we clearly observe that prices and beliefs deviate jointly from fundamental values, if they do so, and that overall beliefs and market prices are remarkably close in most markets (see figures

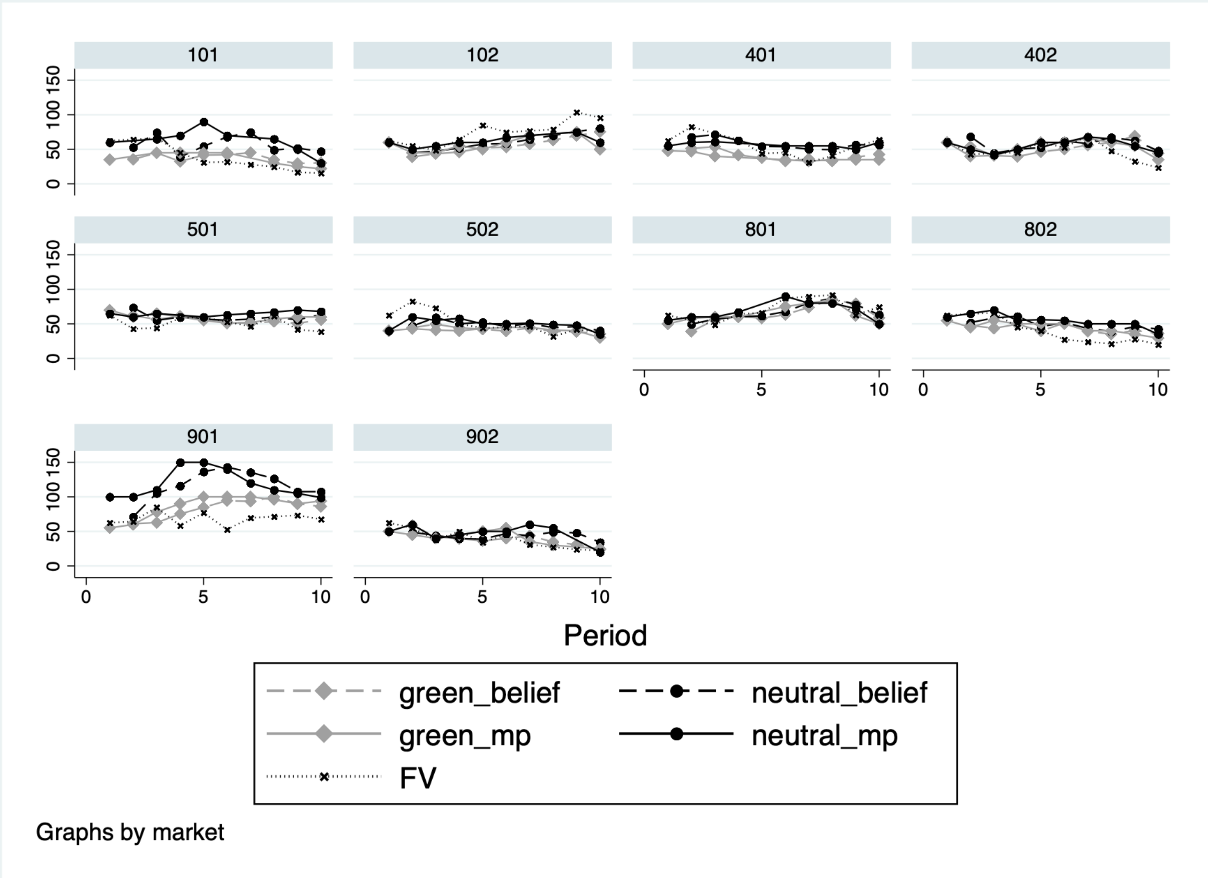


Figure 5.10: Development of beliefs (stated in t-1 about prices in period t), market prices and fundamental values over time in *M_Mixed*

5.10 and 5.11).

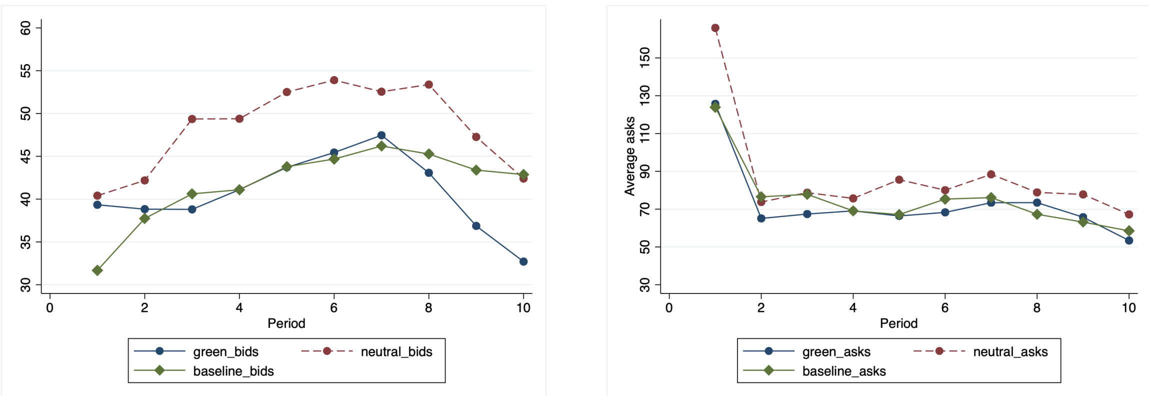


Figure 5.12: Bids (lhs) and asks (rhs) over time.

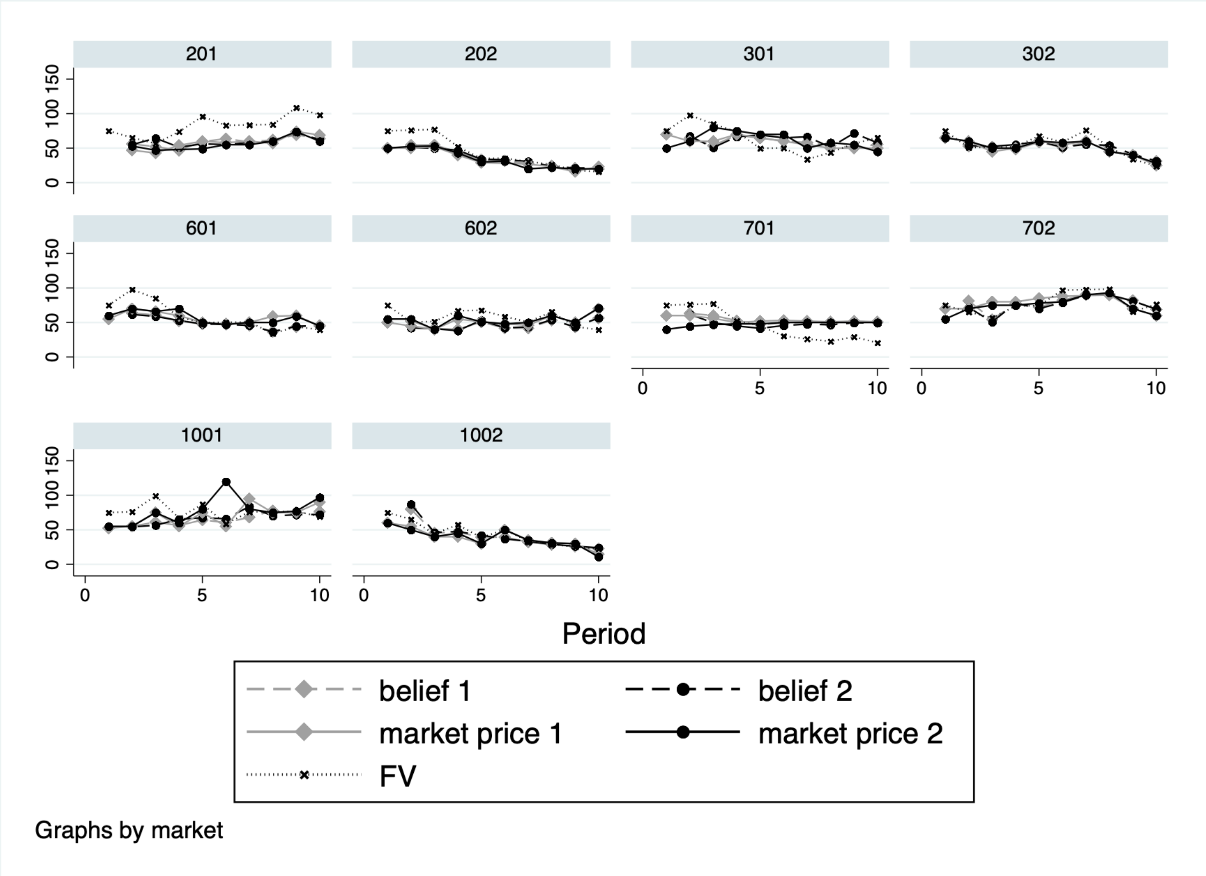


Figure 5.11: Development of beliefs (stated in t-1 about prices in period t), market prices and fundamental values over time in *M_Baseline*

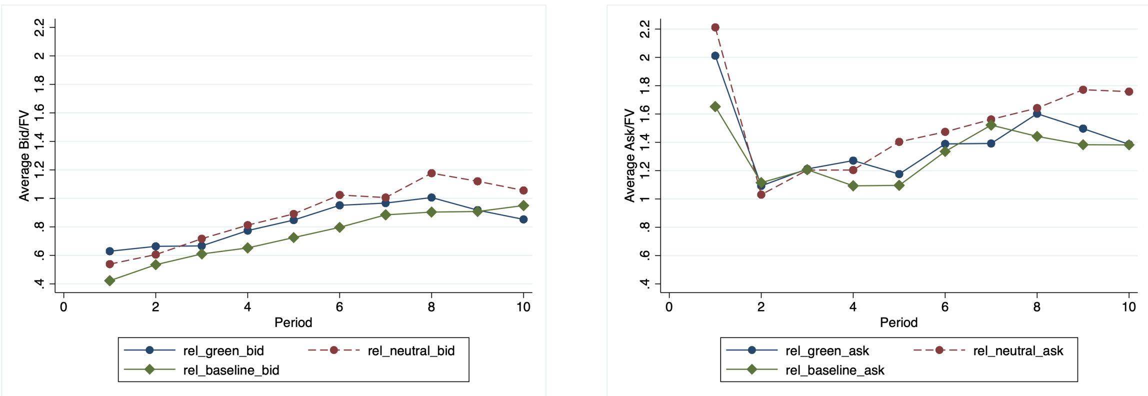


Figure 5.13: Bids (lhs) and asks (rhs) relative to fundamental values over time.

5.3.2.2.2 Do optimists trade differently from non-optimists? For all types of assets and markets, *optimists* hold significantly more assets in their portfolio compared to traders who expect market prices to stay the same or fall (on average 51 versus 65 *green* assets, 56 versus 65 *neutral* assets, 57 versus 64 baseline assets; green: $p < .001$, neutral/baseline: $p < .05$ MWU, t-text, BTT). Overall, *optimists* also tend to submit higher bids and asks in *M_Mixed*, but only for asks the difference is (weakly) significant (green: $p < .05$, neutral: $p < .1$ MWU, t-text, BTT), in *M_Baseline*, there is no difference in behavior of the two types of traders. Thus, overall, but especially in *M_Mixed* and for *green* assets, the belief that market prices will increase in the future is positively related to holding green assets and bidding/asking for higher prices.

Result 13 (Beliefs about market prices)**.**

1. *Overall, participants rather correctly anticipate market prices as well as the differences between prices of the two types of assets. Optimistic beliefs that prices will increase over the next periods lead traders to hold more assets.*

5.3.2.2.3 Bids and asks. To better understand this observation, we will now investigate the bids and asks subjects submit in each period. Individual bids for *green* and *neutral* assets in *M_Mixed* differ substantially. On average, subjects bid 41.01 ECU for the *green* and 48.38 ECU for the *neutral* asset (median bids are 40 and 45), which is a statistically and economically significant difference (p$\leq$.01 WSR, t-test, BTT). Mean bids in *M_Baseline* are close to bids for the *green* assets in magnitude (mean=41.83 ECU, median=44.16) and significantly lower than mean *neutral* bids in *M_Mixed* ($p < .001$ MWU, t-test, BTT). *Neutral* asks are also substantially higher than *green* asks and than asks submitted in *M_Baseline* ($p < .001$ WSR, t-test, BTT for *neutral* versus *green*, $p < .05$ WSR, t-test, BTT for *neutral* versus baseline). However, note that for asks the distributions are rather skewed with mean neutral asks of 88.69 ECU (median=70) compared to mean green asks of 73.85 ECU (median=60) due to some high outliers. Asks in the baseline markets are 76.92 ECU (median=66.75). Comparing bids and asks relative to each periods fundamental value confirms the above observations while correcting for the smaller remaining financial dividend of the *green* assets. For this reason, relative *green* bids and asks are now higher than bids and asks in the baseline markets (bids: $p < .001$ MWU, t-test, BTT; asks: $p < .05$ MWU, insignificant according to t-test, BTT). Average *neutral* relative bids and asks are highest (bids/asks compared to baseline: $p < .001$ MWU, t-test, BTT; compared to *green*: $p < .05$ MWU, t-test, BTT). Figures 5.12 and 5.13 show the development of individuals' (relative) bids and asks over all periods, separated by asset type and by treatment. Figure 5.12 demonstrates how the magnitude of *neutral* bids deviates substantially from *green* and baseline bids, except for the first and last periods. All three types of assets exhibit average bids that increase over the first 2/3 of periods and then decline again. Accounting for fundamental values that include the value of the dividends to be paid in all remaining periods in figure 5.13 shows that subjects start by overall undervaluing assets (independent of their type) to then increase bids over time until average bids are very close to fundamental values. This suggests that the substantial overbidding we observed for the individual decisions under the BDM mechanism in part 1 might be corrected in the market. Trends for (relative) asks are similar but less clear, especially due to some stark outliers in the first period, which may reflect subject's lack of experience with this type of setup. Since it seems that, overall, subjects had a harder time submitting asks, we will mainly focus on bids for the subsequent analysis.

Result 14 (Bidding behavior)**.**

1. *We observe underbidding for all types of assets during the first periods. Bids come close to fundamental values in the second half of the market periods. In M_Mixed, bids for the neutral assets are significantly higher than bids for the green assets.*

5.3.2.2.4 Who holds the green assets? Subjects who are classified as *Green_Type* = 1 according to their bids in part 1 hold substantially more *green* assets than *Green_Type* = 0 subjects ($p < .001$ MWU, t-test, BTT), with the difference becoming larger in later periods, see figure 5.14. They also hold overall more *neutral* assets but the difference is smaller, more stable over time and not statistically significant (figure 5.15). *Green_Type* = 0 subjects, on the other hand, hold more cash ($p < .01$ MWU, t-test, BTT)—in particular during the later periods of the market they sell more assets and hold more cash in their portfolios, see figure 5.16.

For the *magnitude* of individuals' bids and asks we find no significant differences between *green* and *notgreen_types* (even though the *green_types*'s bids are slightly higher and their asks slightly lower for both assets), see also figure 5.17. Thus, subjects with a preference for the green component buy more green assets and keep them in their portfolio instead of exchanging them for cash, but their preference does not become visible in the size of their bids and asks.

5.3.2.2.5 Do beliefs matter for bidding behavior? From figures 5.10 and 5.11 we observed that prices and beliefs seem to follow each other closely in most markets and, moreover, that they often seem to jointly deviate from fundamental values if they do so at all. The experimental design does not allow us to establish a causal link, but to look into individual behavior more closely nonetheless. First, we establish that subjects defined as *Green_Type* = 1 do not hold systematically different beliefs. Overall, *green*

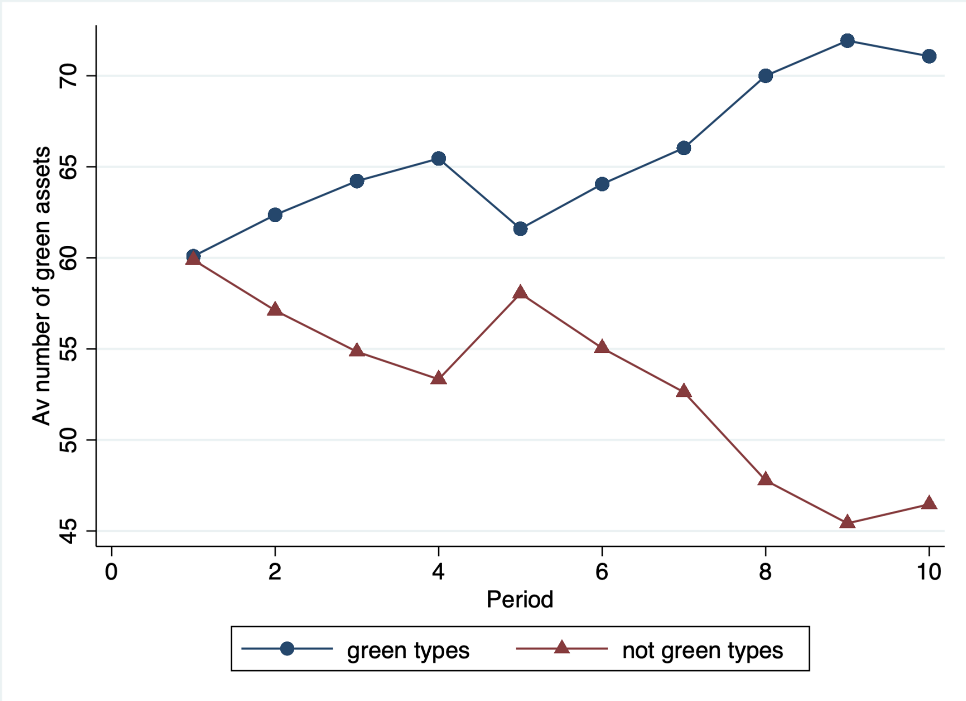


Figure 5.14: Average number of green assets held by subjects, by *green_type*.

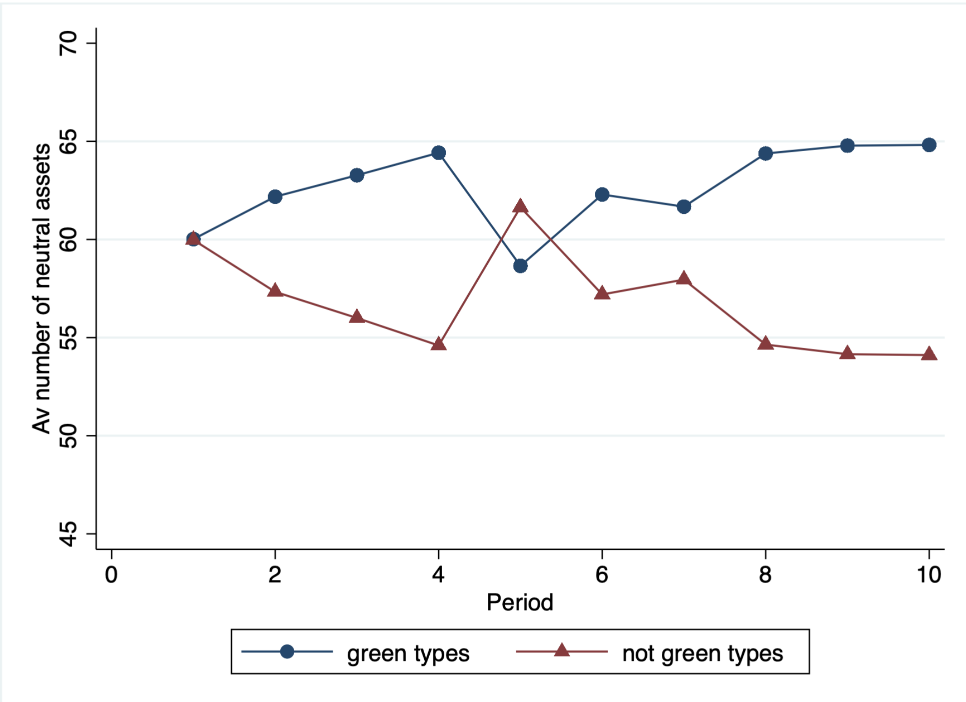


Figure 5.15: Average number of neutral assets held by subjects, by *green_type*.

and *neutral* beliefs about prices in $t+1$ (divided by period t's fundamental values) do not differ significantly between both types of players (among $Green\_Type = 1$ subjects both beliefs are slightly higher, the difference being only statistically significant for *green* beliefs according to WSR, not according to t-tests). The different types of traders seem to

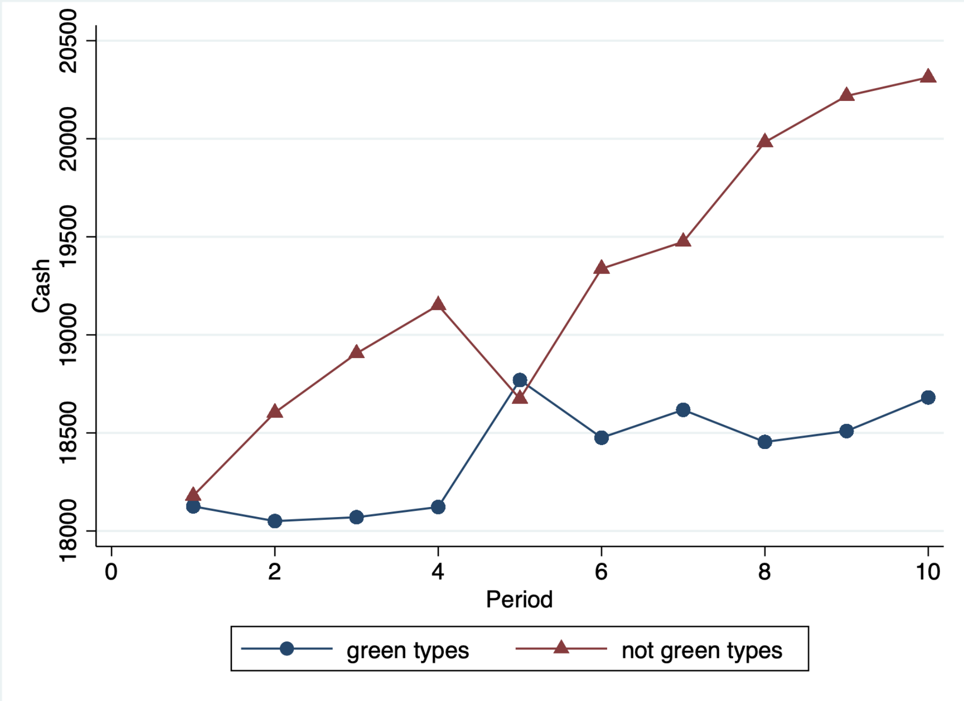


Figure 5.16: Cash in ECU held by subjects, by *green_type*.

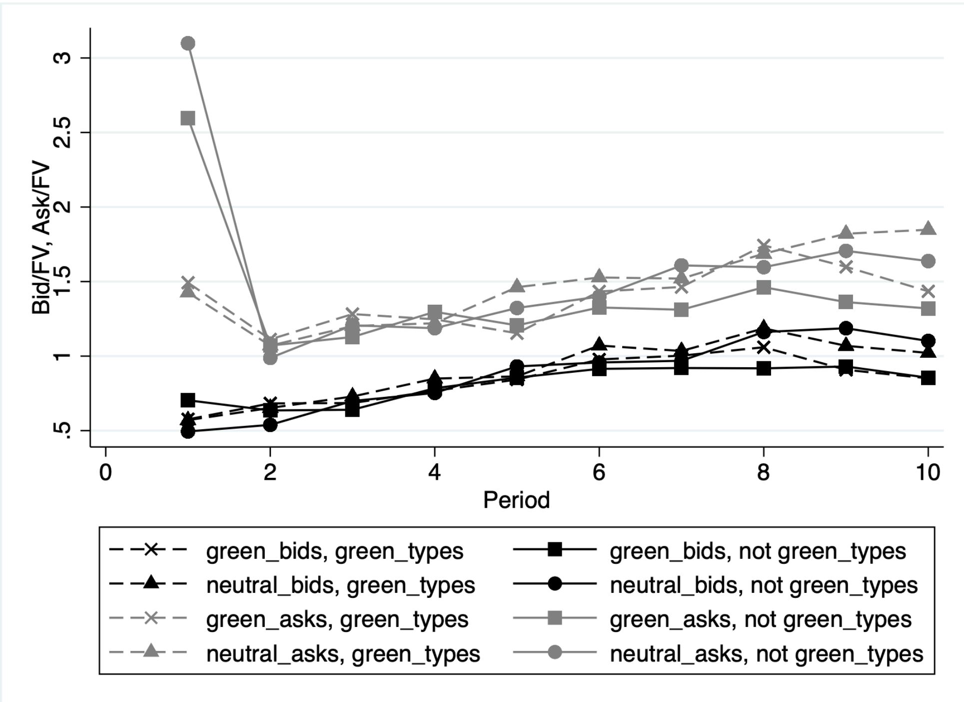


Figure 5.17: Bids and asks relative to FV in each period by *green_type*.

not hold systematically different beliefs. This is confirmed when looking at the differences in beliefs in period 1 prior to any market experiences. Here, *green* beliefs are higher among *green_types* (55.16 vs 52.48), hinting at a correlation between own preferences and beliefs, however, this difference is only significant according to MWU ($p<.01$)), not according to t-test/BTT.

Result 15 (Green Types)**.**

1. *Overall, but predominantly in the later periods of the market interaction, Green_Types hold more assets, especially more green assets, while not Green_Types hold more cash.*

2. *Green_Types do not hold significantly different beliefs from not Green_Types.*

How strongly do subjects' beliefs influence their bidding decisions? To answer this question, we can exploit the timeline of the experiment to better understand how bids and beliefs relate. If bids are strongly correlated with past period's beliefs, i.e. the belief submitted in $t-1$ about the market price in period t, this suggests that subjects base their bid on what they believe the market price in this period to be. If bids are rather correlated with beliefs submitted at the end of period t about prices in $t+1$, this suggests that subjects form their beliefs based on their preferences for each asset and their own trading behavior. The data provide some evidence for both cases, however, the correlations with current period's beliefs about future prices are of larger magnitude. Thus, support for the second interpretation is somewhat stronger. We calculate Pearson's correlation coefficients as well as Spearman's rho for relative beliefs and relative bids (thereby controlling for the common fundamental values) for both types of assets in *M_Mixed*.[27] Both correlations are somewhat stronger for *neutral* assets compared to *green* assets. Both, *green_types* and *notgreen_types* show somewhat higher correlations between bids and beliefs for bids for the *neutral* assets (this holds for lagged and same period beliefs). Interestingly, the correlation between bids and lagged beliefs, indicating an influence of beliefs on bids, is stronger among *notgreen_types*, while the correlation between bids and beliefs submitted in the same period about future prices tends to be larger among *green_types*. Over time, the strength of correlations between bids and beliefs fluctuates somewhat between periods, but there are no clear time trends observable (figures omitted to save space).

Intuitively, while $Green\_Type = 1$ subjects might buy and hold *green* assets out of a preference for the *green* dividend, among $Green\_Type = 0$ subjects, only beliefs can be expected to drive their demand for *green* investments. In other words, we suspect them to only buy green assets if the expect to be able to profitably sell them in future periods. Our results partly support this intuition, but future research will have to establish if this relationship holds true.

[27]Green same period: corr. coef.=0.47, rho=0.57, $p < .001$; Neutral same period: corr. coef.=0.58, rho=0.62, $p < .001$; Green previous period: corr. coef.=0.23, rho=0.28, p.001; Neutral previous period: corr. coef.=0.33, rho=0.38, $p < .001$. For correlations with lagged beliefs, the first period is dropped.

	(1) Gr_Bid	(2) Gr_Bid	(3) Gr_Bid	(4) Gr_Bid
Gr_BDM	0.059**	0.054*	0.071***	0.066***
	(0.022)	(0.025)	(0.015)	(0.016)
Gr_Belief	0.088			
	(0.064)			
Rel_gr_price	30.53***	32.62***		
	(6.915)	(9.578)		
Gr_Vol	-0.006	-0.008		
	(0.018)	(0.019)		
Gr_FV	0.765***	0.828***		
	(0.132)	(0.152)		
Gr_Optimist		4.242*		
		(2.216)		
Gr_Belief_lag			0.026	
			(0.042)	
Rel_gr_price_lag			27.06***	26.47***
			(8.057)	(9.038)
Gr_Vol_lag			0.008	0.007
			(0.017)	(0.021)
Gr_FV_lag			0.708***	0.714***
			(0.114)	(0.147)
Gr_Optimist_lag				-0.447
				(1.902)
_cons	-41.75***	-44.18**	-33.62***	-31.11**
	(12.270)	(16.380)	(12.096)	(15.396)
N	549	506	481	439

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5.2: Bids for the green assets explained by BDM bids in part 1, beliefs, market prices, trading volumes and fundamental values. OLS and random effects estimations. Standard errors clustered at the market level.

	(1) Nt_Bid	(2) Nt_Bid	(3) Nt_Bid	(4) Nt_Bid
Nt_BDM	0.046***	0.038*	0.057***	0.045***
	(0.014)	(0.018)	(0.015)	(0.014)
Nt_Belief	0.156			
	(0.124)			
Rel_nt_price	27.79**	33.80**		
	(11.760)	(13.329)		
Nt_Vol	-0.01	0.013		
	(0.026)	(0.026)		
Nt_FV	0.765***	0.931***		
	(0.230)	(0.284)		
Nt_Optimist		4.773		
		(3.347)		
Nt_Belief_lag			0.190*	
			(0.110)	
Rel_nt_price_lag			16.64**	20.48
			(7.730)	(13.148)
Nt_Vol_lag			-0.019	-0.02
			(0.025)	(0.031)
Nt_FV_lag			0.562***	0.660**
			(0.148)	(0.285)
Nt_Optimist_lag				-3.992
				(2.484)
_cons	-44.72*	-55.49*	-20.69	-16.19
	(21.992)	(27.774)	(16.327)	(27.165)
N	549	507	481	440

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5.3: Bids for the neutral assets explained by BDM bids in part 1, beliefs, market prices, trading volumes and fundamental values. OLS and random effects estimations. Standard errors clustered at the market level.

	(1) Bl_Bid	(2) Bl_Bid	(3) Bl_Bid	(4) Bl_Bid
Bl_BDM	0.027	0.025	0.010	0.008
	(0.033)	(0.035)	(0.043)	(0.045)
Bl_Belief	-0.003*			
	(0.002)			
Rel_bl_price	30.58***	33.67***		
	(6.744)	(7.649)		
Bl_Vol	0.017	0.011		
	(0.039)	(0.038)		
Bl_FV	0.653***	0.665***		
	(0.067)	(0.075)		
Bl_Optimist		0.032		
		(2.487)		
Bl_Belief_lag			-0.000	
			(0.000)	
Rel_bl _price_lag			24.57***	26.24***
			(5.009)	(5.802)
Bl_Vol_lag			-0.010	-0.023
			(0.020)	(0.024)
Bl_FV_lag			0.530***	0.529***
			(0.069)	(0.082)
Bl_Optimist_lag				0.090
				(1.030)
_cons	-32.22**	-35.93**	-15.42	-16.24
	(9.979)	(11.882)	(9.396)	(10.562)
N	716	662	641	587

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5.4: Bids for the baseline assets explained by BDM bids in part 1, beliefs, market prices, trading volumes and fundamental values. OLS and random effects estimations. Standard errors clustered at the market level.

5.3.2.2.6 Regression analyses A regression analysis supports the above observations: Using OLS and random effects regressions with standard errors clustered at the market level, we assess how individual bids are correlated with beliefs about future market prices, risk preferences and *green* preferences as measured by bids in part 1 as well as by market indicators. In each table 5.2, 5.3, and 5.4, columns (1) and (2) present results from OLS regressions in which variables have been aggregated over all 10 periods, while columns (3) and (4) use random effects models with lagged fundamental values, trading volumes, relative prices and beliefs as covariates (excluding period 1). For *green* and *neutral* assets in *M_Mixed*, we see a small but statistically significant correlation between BDM bids and market bids. Prices and fundamental values can largely explain the variance in bidding behavior; beliefs play a minor role after controlling for these two variables. This relationship holds true for aggregate variables in estimations (1) and (2) as well as for observed prices and fundamental value realizations from previous periods in models (3) and (4) (only for *neutral* assets, lagged prices are no longer significant in (4)). This confirms our initial intuition that bids and therefor market prices are relatively accurate and overall do not part substantially from fundamental values, such that beliefs have relatively little additional explanatory power (see also figures 5.10 and 5.11). In the baseline markets, we observe the same phenomena, with the exception that BDM bids are not significantly correlated with bids in the market. Trading volumes play overall no role in explaining bids. A word of caution when interpreting these findings may be warranted since the number of clusters in all estimations is quite small.

Result 16 (Bids and beliefs)**.**

1. *Bids and beliefs are significantly positively correlated but once controlling for prices and fundamental values, beliefs do not have substantial explanatory power for bidding behavior.*

5.3.2.2.7 Does the presence of green types affect market-level measures? Suggestive evidence for the importance of *green* traders on a market can be gathered from dividing out 10 markets in *M_Mixed* into more and less *green* markets as done in section 5.2.5. Markets defined as $Green\_Market = 1$, do not differ systematically with respect to relative *green* market prices and trading volumes. We merely observe that *neutral* prices are higher and *neutral* trading volumes are lower, meaning that on *greener* markets, on average, less *neutral* assets are being traded at higher prices. Using our alternative classification based on the average strength of the green preferences of the traders on a given market rather than the binary type categorization uses more of the information we have about the preferences of market participants. Figure 5.18 confirms that trading volumes to not vary systematically with the type of markets. It suggests, however, that both relative prices, *green* and *neutral*, are somewhat higher the *greener* that market. This hints at an interesting relationship between the prevalence and intensity of *green* preferences among traders on a market and asset pricing. However, with only 10 markets, we can merely provide anecdotal evidence that calls for a thorough analysis in future studies.

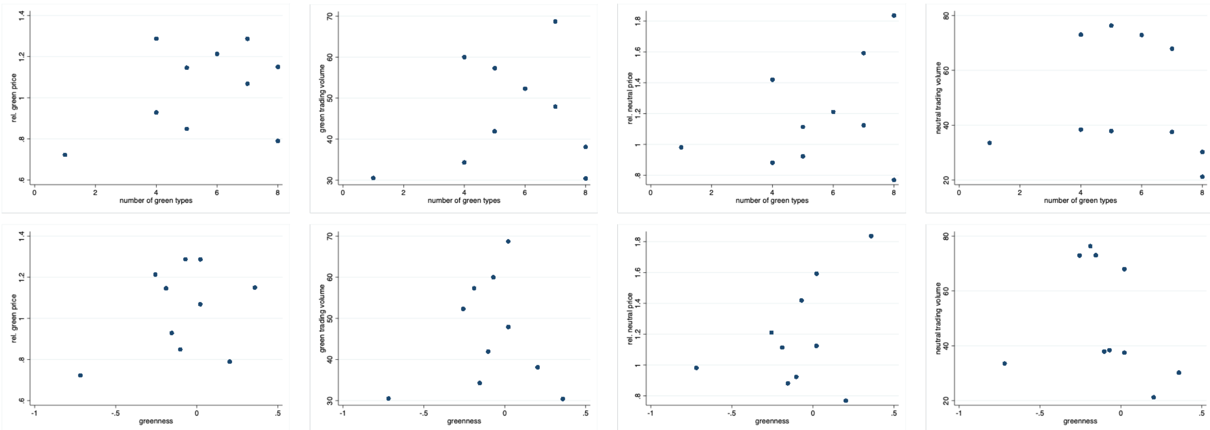


Figure 5.18: Average green and neutral market prices and trading volumes over all periods. By number of green types on a market and strength of the green preference by traders on a market (M_Mixed).

5.3.3 Individual Donation Decisions

On average, subjects donate €1 (median = €0.44) of their show-up fee of €5 at the end of the experiment. 46.5% of the participants donated zero and the maximum donation was €5, chosen by 10 participants (5%).

How does market experience impact the decision to donate? This question builds up on, e.g., Herz and Taubinsky (2018), who show that fairness judgments are shaped by previous market experiences in ultimatum games. We observe slightly higher average donations in *M_Mixed* than in the baseline treatment, (€1.04 compared to €0.97, not statistically significant (*t*-test, BTT)). Median giving differs quite substantially between subjects in *M_Mixed* (€.66) and *M_Baseline* (€0, $p = .031$, 2-sided MWU). More participants in *M_Baseline* decide to donate zero ECU, namely 58% versus 35% in *M_Mixed* ($p < 0.01$ Fisher's exact test). In our asset market setting, it thus seems as if exposure to the green asset during the market phase increases participants' willingness to donate a positive amount, even though the amounts given are small such that they do not significantly affect mean giving. Therefore, despite the steep increase in the number of donors, the total revenue of the charity is only higher by $104.23 - 96.58 = 7.65$ ECU.

Second, we ask if the donation decision is a complement or a substitute for *green* behavior in the asset market? Do we observe consistent green and non-green behavior across domains or are decisions rather independent (prediction 3)? The number of *green* assets a subject held over the 10 periods of trading in the market is weakly positively correlated with the size of her donation (Pearson correlation coefficient (PCC) of .078, Spearman's rho=.17, Kendall's tau-b=.12, $p < .1$ for the latter two, $p > .1$ for the first). Comparing subgroups of participants, those who held more *green* assets than the median number donate on average €1.2 (median=1) compared to €.9 (median=.4) donated by those who held less than the median number of *green* assets in the market (p<.05 MWU). Subjects who held on average over all periods more *green* than *neutral* assets in their portfolio donate on average €1.3 (median=1) compared to €.8 (median=.3) for those who held (weakly) more *neutral* than *green* assets (p=.014, MWU). Thus, holding overall more *green* assets in one's portfolio on the market is (weakly) positively associated with donating more.

Mean individual (relative) *green* bids and asks are not significantly correlated with the size of a person's donation. Participants who submitted higher bids for the *green* than for the *neutral* asset (averaged over all ten periods and relative to fundamental values) do not donate substantially different amounts compared to those who submitted lower *green* than *neutral* bids.

Result 17 (Pro-social behavior across domains)**.** *We do not find evidence of a crowding out of individual pro-social behavior due to trading green assets on the market. On the contrary, green behavior on the market and donations are rather independent decisions, if anything they are weakly positively related. Furthermore, being exposed to mixed markets with green assets leads to a crowding-in of donors.*

5.4 Conclusion

In this paper, we analyze how socially responsible investment opportunities perform in an asset market in which they compete with assets of a conventional firm. We set up an experimental asset markets in which participants can trade assets of two firms with stochastic fundamental values that are constant in expectation over the course of ten periods in a double auction. In the baseline markets both firms offer conventional assets with private dividends paid to investors whereas in the mixed markets, a conventional firm is in competition with a *green* firm whose assets pay out bundled dividends that benefit the investor and a public good.

We do not find a price bubbles for assets with nominal values that are constant in expectation unless they are in competition with a *green* firm. In the presence of a *green* firm, prices for the *neutral* assets increase above the fundamental value in the second half of the trading periods. Thus, the introduction of assets with a costly positive externality leads to a spillover effect on the conventional, *neutral*, assets that pay a higher private dividend, leading to a price bubble for the latter. There is no price premium for *green* assets; instead their valuation is as high as the valuation of *neutral* assets in the homogeneous markets and close to the fundamental values. Market activity, i.e. trading volume, is overall significantly higher in the mixed markets in which a *green* and a *neutral* firm compete compared to the homogeneous baseline markets.

Our design furthermore allows to explore individual behavior and trading motives in the market. Overall, participants rather correctly anticipate market prices as well as the differences between prices of the two types of assets in M_Mixed. Once controlling for prices and fundamental values, beliefs are, however, not a strong predictor of bidding behavior. We observe underbidding for all types of assets during the first periods with bids coming close to fundamental values in the second half of the market periods. In M_Mixed, average bids for the *neutral* assets are significantly higher than bids for the *green* assets, leader to higher markets prices for the first. Analysing individual behavior more closely, we find systematic differences between subjects classified as $Green\_Type = 1$ or $Green\_Type = 0$ according to their bids in part 1. Overall, but predominantly in the later periods of the market interactions, $Green\_Types$ hold more assets, especially more *green* assets, while those with weaker *green* preferences hold more cash. Thus, preference types seem to matter for trading behavior.

There is no evidence of a crowding out of individual pro-social behavior due to trading *green* assets on the market. On the contrary, *green* behavior on the market and individual donations are rather independent decisions, if anything they are weakly positively related. Furthermore, being exposed to mixed markets with *green* assets leads to a crowding-in of donors.

The results add to mixed previous findings in empirical and experimental studies on the financial performance of SRI. Our experimental design overcomes several identification issues from earlier studies and investigates the financial performance on an experimental asset market in competition with a conventional firm. By investigating individual behavior in addition to market developments, our study furthermore explores the mechanisms behind the demand for *green* investments and thereby complements previous work on how pro-social behavior plays out on markets. Our exploratory analysis of the impact of individual *green* preferences and beliefs about future market prices in competitive financial markets reveals very interesting relationships. Since these findings are mainly correlational and of a rather suggestive nature, we see them as exciting avenues for future research and hope that future studies will show their robustness as well as identify causal links behind them.

Appendix

A: Additional Analyses: Robustness of mispricing

As a robustness check, we rerun the same tests as in section 5.3.2 using the mispricing measure proposed by (Stöckl et al., 2010), the relative absolute deviation (RAD).

$$RAD = \frac{1}{N} \sum_i \left| \bar{price_i} - fv_i \right| / \left| \bar{fv} \right|$$

, where $price_i$ and fv_i are the periods price and fundamental value and $\bar{fv}$ is the average fundamental value of the market. We reproduce the finding that neutral mispricing is (weakly) significantly higher than green mispricing with a RAD of .33 compared to .23 (this difference is weakly statistically significant according to WSR (p=.059) and *t*-text (p=.069), but not according to BTT (p>.1). As before, green mispricing is highly similar and statistically indistinguishable from what we observe in the baseline market. As a result, again, the overall difference in mispricing between the mixed and the homogeneous markets is minor and not significant.

	M_Mixed		M_Baseline
	Green Asset	Neutral Asset	Baseline Assets
GAD	.233	.332	.247

Table 5.5: Mispricing measures for each asset and treatment: relative absolute deviation (RAD) over all N=10 periods. The two assets in M_Baseline are summarized here as the values are highly similar.

B: Instructions

In the following you find an English version of the original instructions given to participants in German. The instructions are those of the mixed market treatment, the parts that are changed in the baseline treatment are marked in italic.

Instructions

Thank you very much for your participation in this economic experiment.

Please note that communication with other participants is not allowed. If you do not follow this rule you will be excluded from the experiment as well as from all payments.

Procedure

All decisions will be made anonymously, i.e. neither another participant nor the experimenter can match them with your person. Also payments are anonymous, i.e. no participant gets to know the payment of another one.
In case you have a question during the experiment please raise your hand, we will come to your place.
The experiment consists of 3 parts. In the following we will describe part 1 of the experiment. the instructions for part 2 and 3 will be distributed and read to you once the previous part is over. The decisions you make in one part of the experiment have no consequences for the payoffs in another part of the experiment.

Payoffs

In the end, your payoffs from part 1, 2 and 3 will be added up. *During the course of the experiment there will be -beside your own payments- payments to the nonprofit organisation Atmosfair. Atmosfair is a NGO in Bonn engaged in climate protection by compensating CO_2 emissions by renewable energies. You will find further information on Atmosfair on an extra sheet attached to the instructions.*
During the while experiment we calculate all earnings in Taler with an exchange rate of 1 Euro = 1800 Taler.

Part 1

In this experiment, there are 2 fictitious firms, we will call them Nanpal and Grenik. In part 1 of the experiment you can buy assets of those two firms.

Description of the assets

In the beginning, every Nanpal asset and every Grenik asset has a base value of 50 Taler. The value of a Nanpal asset and of a Grenik asset changes randomly: The

computer randomly decides whether the value decreases by 25%, decreases by 5%, increases by 10% or increases by 40%. All four changes are equally likely. The changes in value is the same for each piece and for each asset (Grenik or Nanpal)! At the same time, each asset pays a dividend of 5% of the value of the asset. Due to this dividend payment, the value of the asset diminishes by exactly this amount. This means that the overall change in value of an asset consist of the change in value and the value reduction of 5% due to the dividend payment. as the average change in value is +5% and the asset pays exactly 5% dividend the value remains constant on average.
The value of an asset changes 10 times in a row. The change always happens from the assets current value. It is independent of the previous or subsequent changes. each time, when the value changes, thus 10 times overall, the owner of the asset gets the dividend of 5%. As the value of the asset changes each time, the the absolute magnitude of the dividend in Taler changes aswell. In case you buy an asset you will get the sum of the 10 dividend payment in the end of part 1.
In the appendix of the instructions you will find a table with examples how the value of a Nanpal or of a Grenik asset can change over time. Please not that those are merely examples that have been drawn by a random process and not actual value changes that you will observe in part 1!

The dividend payments of the two firms *differ*:
The firm Nanpal pays out he full 5% dividend to you.
The firm Grenik pays out half of the 5% dividend to you. The other half, 2.5% of the value before the value change, will be paid into an extra account and transferred to the nonprofit organisation Atmosfair at the end of the experiment. In case of a transfer the money will be transferred online for each participant separately and in your presence.

Decision situation in part 1

In the beginning of part 1 you can make an offer to buy a parcel of exactly 30 Nanpal assets, or to buy a parcel of exactly 30 Grenik assets respectively. With this offer you state how much your are willing to pay for one Grenik asset, or for one Nanpal asset. and the payoffs it generates. You state that you are willing to buy a Nanpal asset, or a Grenik asset, to exactly the price offered or a lower price.

In order to do this, you will get 6000 Taler, i.e. you can bid between 0 and 200 Taler for one asset. Please insert the two bids in the respective fields on the screen and then confirm your choice by clicking OK. After that you cannot modify your bids anymore. In the next step, the computer randomly selects one of the two assets that will be relevant for your payment! Each asset can be drawn with the same probability. The actual price of the chosen asset is randomly determined by the computer. This happens by means of a random draw that draws any oprice between 0 and 200 with an equal likelihood. Whether you get the asset depends on wehther your bid has been larger or smaller than the randomly drawn price: If the price is larger than you bid you do not get the asset. If the randomly drawn price is smaller than or equally high as your bid, you buy 30 pieces of the asset at the randomly drawn price. Thus, your bid does not determine the price you have to pay but rather an upper limit until which you are willing to buy the asset. Therefore is is the best strategy for you to bid exactly the amount that you are still willing to pay for one piece of the asset.

Payoff from part 1

In case you do not purchase the asset: 6000 Taler
In case you purchase the asset: 6000 - (price*30) + 30*ten dividends of the respective asset + 30*future value of the asset
Please not that the decisions in part 1 have no relevance for part 2 of the experiment.

Estimations in part 1

In addition we ask you to estimate which bids the other participants in the room have submitted on average.
In case you correctly guess the average bid for an asset you receive 1€(1800 Taler) for this estimation. For each Taler that your estimation deviates from the correct amount, 2 cent (36 Taler) will be subtracted from this 1€(1800 Taler). Your payment is higher, the closer you are to the actual average value of the bids of the other players.
Again, one of the assets will be chosen randomly for payment. the result of your estimation will be paid out to you in addition to your other payments at the end of the experiment.

Do you have questions concerning the instructions? If not, the first part of the experiment will begin now with short comprehension questions to make sure that you understand the procedure. As soon as all participants answered them correctly, the actual experiment will start.

Part 2

In part 2 of the experiment, you can trade assets of the fictitious firms Nanpal and grenik in an asset market, i.e. buy and sell them. There are 10 participants in one market. During the whole experiment you will trade with the same nine other participants.
Each participant receives 60 Grenik assets and 60 Nanpal assets at the beginning of part 2. In addition each participants is endowed with 18000 Taler. At the beginning of period 1 each Nanpal asset and each Grenik asset has a base value of 50 Taler. The value of a Nanpal asset and of a Grenik asset changes randomly: A random draw by the computer determines if the value of an asset decreases by 20%, decreases by 5%, increases by 10% or increases by 40%. All four changes are equally likely. Thus, on average the value of an asset increases with each value change by 5%. The change is the same for each piece and for each asset (Grenik and Nanpal)! If, for example, the value of a Nanpal asset increases by 10% in a given period, then the value of a Grenik asset also increases by 10% in that period. At the same time each asset pays a dividend of 5% of the value of the asset. Due to the dividend payment the value of that asset diminishes by that amount. Thus, the overall value change of an asset consists of the percentage change of its value minus the 5% dividend payment! As the average change in value is 5% and the asset pays exactly a 5% dividend the value remains constant on average.

Overall there are 10 trading periods. Each period is built up in the exact same way. Between those trading periods the value of a Nanpal asset and of a Grenik asset changes in the way described above. Each change happens based on the current value in that period. It is independent of previous or subsequent changes. In each period the 5% dividend is booked to the account of the owner of the asset. Again, the absolute magnitude of the dividend in Taler changes with the value of the asset. The person who own the asset after the 1th trading period will be paid its future value. In addition, the sum of the dividends from each period will be paid out.

In the appendix of the instructions you will find a table with examples how the value of a Nanpal or of a Grenik asset can change over time. Please not that those are merely examples that have been drawn by a random process and not actual value changes that you will observe in part 2!

The dividend payments of the two firms *differ*:

The firm Nanpal pays out half of the 5% dividend to you at the end of a period. The other half, 2.5% of the value at the beginning of a period, will be paid into a locked account sand paid out to you only at the end of the experiment. Those Taler will not be available for trading.

The firm Grenik pays out half of the 5% dividend to you. The other half, 2.5% of the value at the beginning of a period, will be paid into an extra account and transferred to the nonprofit organisation Atmosfair at the end of the experiment.

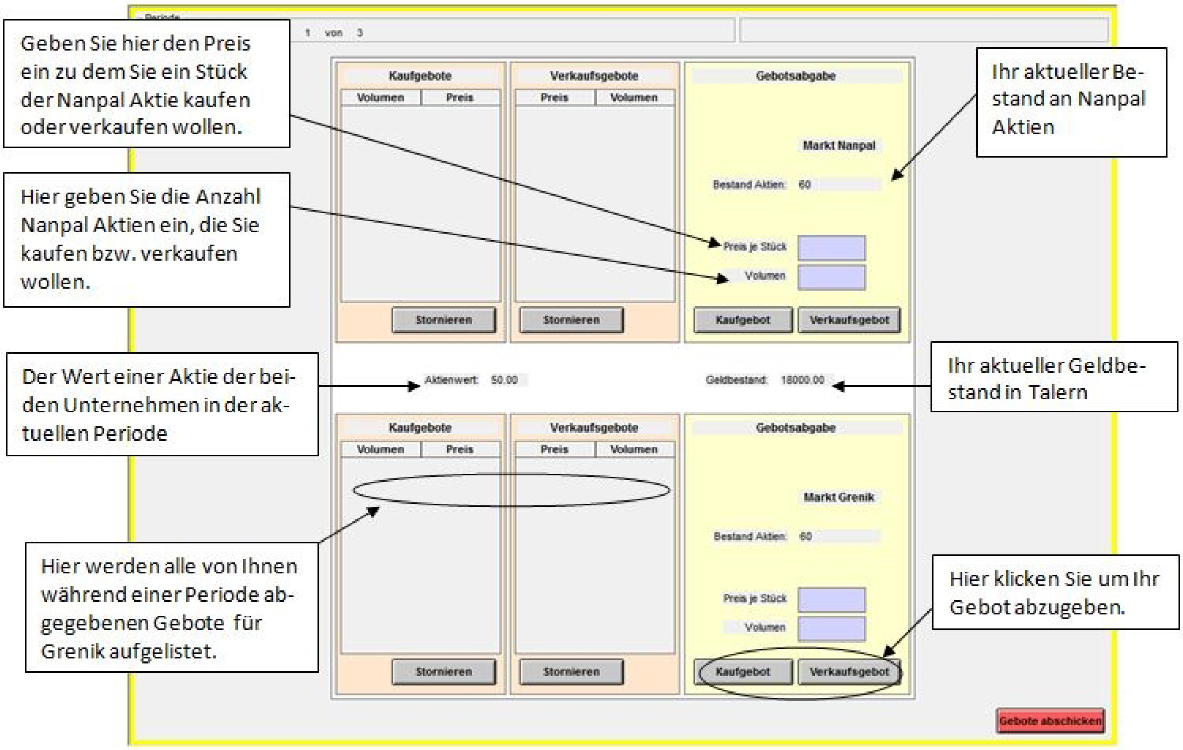


Figure 5.19: Screenshot from the instructions

How do I trade on the market?

In each trading period you have the possibility to buy and sell assets.

In order to buy assets you need money. In the middle of your screen you will see your money holdings in Taler in the white area. You can use these to trade. In order to sell assets you need assets. Your current stock of Nanpal assets and of Grenik assets will each be shown to you in the two yellow areas on the right side of the screen (see figure 1).

I. Sell: You can offer assets from your depot for sale. In order to do this, enter the minimum price per piece you want to receive for one asset into the field “Price per Piece”.

Please note, that this is the price per piece, i.e. for one asset. Also you enter the number of assets you want to sell in the field “Volume”. You can offer assets of the firm Grenik and of the firm Nanpal for sale at the same time.
II. Buy: You can make an offer to buy assets. In order to do this, enter the maximum price per piece you are willing to pay for an asset in the field “price per piece”. Please note, that this is the price per piece, i.e. for one asset. Also you enter the number of assets you want to buy in the field “Volume”. You can buy assets of the firm Grenik and of the firm Nanpal at the same time.
III. Delete buy and sell offers: you can take an offer back by clicking on it in the list of buy and sale offers and then click on the button “cancel”. During the course of a period you can make several buy and sale offers with different prices. They will all be listed in the respective column “Buy offers” or “Sale offers”, see figure 1. Only after you clicked the button “Send” in the bottom right corner the offers will be send and you cannot reverse them anymore. Please not the following trading restriction: The prices of your sale offers have to be above the prices of your buy offers.
Trade: At the end of each period, the computer calculates the two market prices of the period (one market price for Grenik and one for Nanpal). The market price is the price for which the largest amount of assets of a firm can be traded according to the buy and sell offers of all ten market participants. Depending on the offers of the participants the market prices of Nanpal and of Grenik assets can differ. All transactions take place to a single market price for Grenik assets, and for a single market price for Nanpal assets. In case trade takes place, the amount of Taler corresponding to the market price, multiplied by the number of assets sold, will be added the account of the seller and the number of traded assets will be subtracted from the depot of the seller. At the same time, the amount of Taler corresponding to the market price, multiplied by the number of assets bought, will be subtracted from the account of the buyer and the number of traded assets will be added to the depot of the buyer.

Estimations at the beginning of each period

Before the beginning of each period we ask you to estimate the market price of a Nanpal asset and the market price of a Grenik asset in the current and the subsequent period. In the end the computer will randomly draw one of the assets and one period for payment.
In case you have correctly guessed the market price of the chosen asset in the chosen period, you receive 1€(1800 Taler) for this estimation. For each Taler that your

estimation deviates from the correct value, 2 cent (36 Taler) will be subtracted from this amount. Thus, your payment is higher, the closer you are to the actual price. If your estimation deviates more than 50 Taler from the actual price you will receive no payment for your guess.
The payment will be paid to you in addition to your other payments at the end of the experiment.

At the end of each period you will see a payoff screen. This screen displays all information concerning your current stock of money, your current stock of assets, the value and the dividends.

Practice periods

We will now conduct three practice periods. Please use those practice periods to get accustomed to the trading platform! You will have 100 assets in your depot and 5000 Taler in your account. Transactions during the practice periods have no consequences for your payoffs!
Do you have any questions about the experiment?

Part 3

In addition to your payoffs from the experiment you receive the 5€show-up fee. You now have the possibility to transfer any desired amount out of those 5€to the nonprofit organisation Atmosfair. You can transfer 0€, you can transfer the whole 5€or any chosen amount between 0 and 5. The share you want to transfer will be transferred together with the extra dividend from parts 1 and 2 online at the end of the experiment.

5..0.1 Description of Atmosfair

Note that the below text is a direct translation of the German original document subjects received as part of their instructions for the epxeriment.

Atmosfair is a charitable organization based in Bonn. Among other activities, it is active in climate protection by working on compensation of greenhouse gases by renewable energy. By means of the money wired them, this organization promotes renewable energy, especially in developing countries. This way, Atmosfair saves CO_2 which would otherwise have been created through the use of fossil energy in these countries. According to their emissions calculator, a donation of for example €10 leads to a reduction of CO_2 emissions by approximately 430kg. At the same time, the population in these countries profits from getting access to clean and reliable energy. Atmosfair climat protection projects not only mitigate CO_2 emissions, but also promote sustainable development and the fight against poverty.

Out of the many supported by Atmosfair, for this experiment we have chosen the following project to receive the money transfer: Atmosfair offers financial support and logistical help in providing renewable energy systems for inhabitants of the village of Chispani in Nepal, approximately 40km north of the capital Kathmandu. The technologies implemented include solar lamps, solar panels, solar water heatrs, small-scale bio gas reactors and efficient ovens.

In case of a money transfer to Atmosfair, the money will be wired for every subject separately once the experiment has ended. This takes place using electronic money transfer in the course of the cash payment in your presence. The money transfer will make use of the online donation platform betterplace.org, which transmits 100% of the money to Atmosfair. A few days after the experiment, you will receive an e-mail from the lab team. This e-mail will contain a link which will allow you to view the original receipts for the money transfers we made. If you would like to obtain further information on Atmosfair after the experiment, you can find it under https://www.atmosfair.de/ and under www.betterplace.org/de/.

References

Agranov, M., A. Bisin, and A. Schotter (2014). An experimental study of the impact of competition for other people's money: the portfolio manager market. *Experimental Economics 17*(4), 564–585.

Akerlof, R. (2017). The importance of legitimacy. *The World Bank Economic Review 30*, 157–165.

Allouche, J. and P. Laroche (2005). A meta-analytical investigation of the relationship between corporate social and financial performance. *Revue de gestion des ressources humaines* (57), 18.

Almås, I., A. W. Cappelen, and B. Tungodden (2017). Cutthroat capitalism versus cuddly socialism: Are americans more meritocratic and efficiency-seeking than scandinavians? HCEO Working Paper 2017-003.

Anderson, L. R., F. J. DiTraglia, and J. R. Gerlach (2011). Measuring altruism in a public goods experiment: A comparison of u.s. and czech subjects. *Experimental Economics 14*, 426–437.

Andreoni, J. (1989). Giving with impure altruism: Applications to charity and ricardian equivalence. *Journal of Political Economy 97*(6), 1447–1458.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal 100*(401), 464–477.

Andreoni, J. and B. D. Bernheim (2009). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica 77*(5), 1607–1636.

Apostolakis, G., F. Kraanen, and G. van Dijk (2016). Examining pension beneficiaries willingness to pay for a socially responsible and impact investment portfolio: A case study in the dutch healthcare sector. *11*, 27–43.

Ariely, D., A. Bracha, and S. Meier (2009). Doing good or doing well? image motivation and monetary incentives in behaving prosocially. *American Economic Review 99*, 544–555.

Baker, R. J., S. K. Laury, and A. W. Williams (2008). Comparing small-group and individual behavior in lottery-choice experiments. *Southern Economic Journal 75*(2), 367–382.

Bardhan, P. (2000). Irrigation and cooperation: An empirical analysis of 48 irrigation communitiesin south india. *Economic Development and Cultural Change 48*(4), 847–865.

Bardsley, N. and P. G. Moffatt (2007). The experimetrics of public goods: Inferring motivations from contributions. *Theory and Decision 62*(2), 161–193.

Barreda-Tarrazona, I., J. C. Matallin-Saez, and R. Balaguer-Franch (2011). Measuring investors' socially responsible preferences in mutual funds. *Journal of Business Ethics 103*, 305–330.

Bartling, B., E. Fehr, and H. Herz (2014). The intrinsic value of decision rights. *Econometrica 82*(6), 2005–2039.

Bartling, B., M. Grieder, and C. Zehnder (2015). Does competition justify inequality? *University of Zurich Working Paper Series 158.*

Bartling, B., R. A. Weber, and L. Yao (2015). Do markets erode social responsibility? *Quarterly Journal of Economics*, 219–266.

Battigalli, P. and M. Dufwenberg (2009). Dynamic psychological games. *Journal of Economic Theory 144*(1), 1–35.

Becker, G., M. DeGroot, and J. Marschak (1964). Measuring utility by a single-response sequential method. *Behavioral Science 9*, 226–232.

Becker, G. S. (1957/2010). *The economics of discrimination.* University of Chicago press.

Becker, G. S. (1981/2009). *A Treatise on the Family.* Harvard University Press.

Bénabou, R. and J. Tirole (2006). Incentives and prosocial behavior. *The American Economic Review 96*(5), 1652–1678.

Bénabou, R. and J. Tirole (2010). Individual and corporate social responsibility. *Economica 77*(305), 1–19.

Bénabou, R. and J. Tirole (2012). Laws and norms. IZA Discussion Paper No. 6290.

Bereby-Meyer, Y. and A. E. Roth (2006). The speed of learning in noisy games: Partial reinforcement and the sustainability of cooperation. *American Economic Review 96*(4), 1029–1042.

Berg, J., J. Dickhaut, and K. McCabe (1995). Trust, reciprocity, and social history. *Games and Economic Behavior 10*(1), 122–142.

Berman, E., M. Callen, C. Gibson, and J. D. Long (2014). Election fairness and government legitimacy in afghanistan. NBER Working Paper No. 19949.

Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy 102*(5), 8.

Besley, T., A. Jensen, and T. Persson (2015). Norms, enforcement, and tax evasion. CEPR Working paper No. DP10372.

Björk, L., M. G. Kocher, P. Martinsson, and N. Pham-Khanh (2016). Cooperation under risk and ambiguity. *Working Paper*.

Björner, T. B., L. G. Hansen, and C. S. Russell (2004). Environmental labeling and consumers' choice –an empirical analysis of the effect of the nordic swan. *Journal of Environmental Economics and Management 47*(3), 411–434.

Black, S. E. and L. M. Lynch (2001). How to compete: The impact of workplace practices and information technology on productivity. *Review of Economics and Statistics 83*(3), 434–445.

Blanken, I., N. van de Ven, and M. Zeelenberg (2015). A meta-analytic review of moral licensing. *Personality and Social Psychology Bulletin 41*(4), 540–558.

Bock, O., I. Baetge, and A. Nicklisch (2014). hroot: Hamburg registration and organization online tool. *European Economic Review 71*, 117–120.

Bolton, G. E. (1991). A comparative model of bargaining: Theory and evidence. *The American Economic Review*, 1096–1136.

Bolton, G. E. and A. Ockenfels (2000). Erc: A theory of equity, reciprocity, and competition. *The American Economic Review 90*(1), 166–193.

Bone, J., J. Hey, and J. Suckling (2004). A simple risk-sharing experiment. *Journal of Risk and Uncertainty 28*(1), 23–38.

Bonin, J. P., D. C. Jones, and L. Putterman (1993). Theoretical and empirical studies of producer cooperatives: Will ever the twain meet? *Journal of Economic Literature 31*(3), 1290–1320.

Brennan, G., L. G. Gonzalez, W. Güth, and M. V. Levati (2008). Attitudes toward private and collective risks in individual and strategic choice situations. *Journal of Economic Behavior & Organization 67*(1), 253–262.

Brock, J., A. Lange, and E. Y. Ozbay (2013). Dictating the risk -experimental evidence on giving in risky environments. *American Economic Review 103*(1), 415–437.

Brusco, V., M. Nazareno, and S. C. Stokes (2004). Vote buying in argentina. *Latin American Research Review 39*(2), 66–88.

Camerer, C. F. (2003). *Behavioral Game Theory -Experiments in Strategic Interaction.* Princeton, New Jersey: Princeton University Press.

Cappelen, A. W., A. D. Hole, E. Ø. Sørensen, and B. Tungodden (2007). The pluralism of fairness ideals: An experimental approach. *The American Economic Review 97*(3), 818–827.

Cappelen, A. W., J. Konow, E. Sørensen, and B. Tungodden (2013). Just luck: An experimental study of risk-taking and fairness. *American Economic Review 103*(4), 1398–1413.

Casadesus-Masanell, R., M. Crooke, F. Reinhardt, and V. Vasishth (2009a). Households' willingness to pay for "green" goods: evidence from patagonia's introduction of organic cotton sportswear. *Journal of Economics and Management Strategy 18*(1), 203–233.

Casadesus-Masanell, R., M. Crooke, F. Reinhardt, and V. Vasishth (2009b). Households' willingness to pay for green goods: evidence from patagonia's introduction of organic cotton sportswear. *Journal of Economics & Management Strategy 18*(1), 203–233.

Chakravarty, S., G. W. Harrison, E. E. Haruvy, and E. E. Rutstroem (2011). Are you risk averse over other people's money? *Southern Economic Journal 77*(4), 901–913.

Chan, N. W. and M. J. Kotchen (2014). A generalized impure public good and linear characteristics model of green consumption. *Ressource and Energy Economics 37*, 1–16.

Charness, G. and U. Gneezy (2010). Portfolio choice and risk attitudes: An experiment. *Economic Inquiry 48*(1), 133–146.

Charness, G. and U. Gneezy (2012). Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization 83*(1), 50–58.

Charness, G. and M. Rabin (2002). Understanding social preferences with simple tests. *Quarterly Journal of Economics 117*(3), 817–869.

Charness, G. and A. Viceisza (2012). Comprehension and risk elicitation in the field: Evidence from rural senegal. *University of California at Santa Barbara, Economics Working Paper Series qt5512d150*.

Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics 14*(1), 47–83.

Consolandi, C., A. Innocenti, and A. Vercelli (2009). Csr, rationality and the ethical preferences of investors in a laboratory experiment. *Research in Economics 63*, 242–252.

Cooper, R., D. V. DeJong, R. Forsythe, and T. W. Ross (1996). Cooperation without reputation: experimental evidence from prisoner's dilemma games. *Games and Economic Behavior 12*(2), 187–218.

Cornes, R. and T. Sandler (1994). The comparative static properties of the impure public good model. *Journal of Public Economics 54*, 403–421.

Costa-Gomes, M. A., S. Huck, and G. Weizsäcker (2014). Beliefs and actions in the trust game: Creating instrumental variables to estimate the causal effect. *Games and Economic Behavior 88*, 298–309.

Cottrell, D., M. C. Herron, and S. J. Westwood (2018). An exploration of donald trump's allegations of massive voter fraud in the 2016 general election. *Electoral Studies 51*, 123–142.

Dal Bó, E., P. Dal Bó, and E. Eyster (2017). The demand for bad policy when voters underappreciate equilibrium effects. *The Review of Economic Studies 85*(2), 964–998.

Dal Bó, P. (2014). Experimental evidence on the workings of democratic institutions. In S. Gallani and I. Sened (Eds.), *Institutions, Property Rights, and Economic Growth: The Legacy of Douglass North*. New York: Cambridge University Press.

Dal Bó, P., A. Foster, and K. Kamei (2019). The democracy effect: a weights-based identification strategy. Technical report, National Bureau of Economic Research.

Dal Bó, P., A. Foster, and L. Putterman (2010). Institutions and behavior: Experimental evidence on the effects of democracy. *American Economic Review 100*, 2205–2229.

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.

Dave, C., C. C. Eckel, C. A. Johnson, and C. Rojas (2010). Eliciting risk preferences: When is simple better? *Journal of Risk and Uncertainty 41*, 219–243.

Dawes, R. M. and R. H. Thaler (1988). Anomalies: cooperation. *The Journal of Economic Perspectives 2*(3), 187–197.

Dickinson, D. (1998). The voluntary contributions mechanism with uncertain group payoffs. *Journal of Economic Behavior and Organization 35*, 517–533.

Dickson, E. S., S. C. Gordon, and G. A. Huber (2015). Institutional sources of legitimate authority: An experimental investigation. *American Journal of Political Science 59*(1), 109–127.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic 9*(3), 522–550.

Dufwenberg, M. and G. Kirchsteiger (2004). A theory of sequential reciprocity. *Games and Economic Behavior 47*(2), 268–298.

Dwenger, N., H. Kleven, I. Rasul, and J. Rincke (2016). Extrinsic and intrinsic motivations for tax compliance: Evidence from a field experiment in germany. *American Economic Journal: Economic Policy 8*(3), 203–232.

Eckel, C. C. and P. J. Grossman (1996). Altruism in anonymous dictator games. *Games and Economic Behavior 16*, 181–191.

Eckel, C. C. and P. J. Grossman (2003). Rebate versus matching: does how we subsidize charitable contributions matter? *Journal of Public Economics 87*(3), 681–701.

Eckel, C. C. and P. J. Grossman (2006). Subsidizing charitable giving with rebates or matching: Further laboratory evidence. *Southern Economic Journal 72*(4), 794–807.

Eckel, C. C. and P. J. Grossman (2008a). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior and Organization 68*, 1–17.

Eckel, C. C. and P. J. Grossman (2008b). Subsidizing charitable contributions: a natural feld experiment comparing matching and rebate subsidies. *Exprimental Economics 11*, 234–252.

Elfenbein, D. W. and B. McManus (2010). A greater price for a greater good? evidence that consumers pay more for charity-linked products. *American Economic Journal: Economic Policy 2*(2), 28–60.

Engel, C. (2011). Dictator games: A meta study. *Working Paper*.

Engel, C. and P. G. Moffatt (2014). dhreg, xtdhreg, and bootdhreg: Commands to implement double-hurdle regression. *The Stata Journal 14*(4), 778–797.

Engelmann, D., A. Munro, and M. Valente (2012). On the behavioural relevance of optional and mandatory impure public goods: Results from a laboratory experiment. *GRIPS Discussion Papers 11-17, National Graduate Institute for Policy Studies*.

Enikolopov, R., V. Korovkin, M. Petrova, K. Sonin, and A. Zakharov (2013). Field experiment estimate of electoral fraud in russian parliamentary elections. *Proceedings of the National Academy of Sciences 110*(2), 448–452.

Ertan, A., T. Page, and L. Putterman (2009). Who to punish? individual decisions and majorityrule in mitigating the free rider problem. *European Economic Review 53*(5), 495–511.

Exley, C. L. (2014). Excusing selfishness in charitable giving: The role of risk. *Working Paper*.

Falk, A. and U. Fischbacher (2006). A theory of reciprocity. *Games and economic behavior 54*(2), 293–315.

Falk, A. and N. Szech (2013). Morals and markets. *Science 340*(6133), 707–711.

Faro, D. and Y. Rottenstreich (2006). Affect, empathy, and regressive mispredictions of others' preferences under risk. *Management Science 52*(4), 529–541.

Fehr, E., H. Herz, and T. Wilkening (2013). The lure of authority: Motivation and incentive effects of power. *American Economic Review 103*(4), 1325–59.

Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 817–868.

Feld, L. P. and J.-R. Tyran (2002). Tax evasion and voting: An experimental analysis. *Kyklos 55*(2), 197–221.

Fischbacher, U. (2007). Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*(2), 171–178.

Fischbacher, U., S. Schudy, and S. Teyssier (2014). Heterogeneous reactions to heterogeneity in returns from public goods. *Social Choice and Welfare 43*, 195–217.

Fjeldstad, O.-H., C. Kagoma, E. Mdee, and I. Hoem Sjursen (2018). The customer is king: Evidence on vat compliance in tanzania.

Forsythe, R., J. L. Horowitz, N. E. Savin, and M. Sefton (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior 6*(3), 347–369.

Frackenpohl, G. and G. Pönitzsch (2013). Bundling public with private goods. *Bonn Econ Discussion Papers No 05/2013*.

Freundt, J. and A. Lange (2017). On the determinants of giving under risk. *Journal of Economic Behavior & Organization 142*, 24–31.

Frey, B. S. (1997). A constitution for knaves crowds out civic virtues. *The Economic Journal 107*(443), 1043–1053.

Frey, B. S., M. Benz, and A. Stutzer (2004). Introducing procedural utility: Not only what, but also how matters. *Journal of Institutional and Theoretical Economics 160*, 377–401.

Friede, G., T. Busch, and A. Bassen (2015). Esg and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment 5*(4), 210–233.

Fudenberg, D. and D. K. Levine (2012). Fairness, risk preferences and independence: Impossibility theorems. *Journal of Economic Behavior and Organization 81*(2), 606–612.

Gangadharan, L. and V. Nemes (2009). Experimental analysis of risk and uncertainty in provisioning private and public goods. *Economic Inquiry 47*(1), 146–164.

Glac, K. (2009). Understanding socially responsible investing: The effect of decision frames and trade-off options. *Journal of Business Ethics 87*(1), 41–55.

Gneezy, A., U. Gneezy, G. Riener, and L. D. Nelson (2012). Pay-what-you-want, identity, and self-signaling in markets. *Proceedings of the National Academy of Sciences 109*(19), 7236–7240.

Gneezy, U., E. A. Keenan, and A. Gneezy (2014). Avoiding overhead aversion in charity. *Science 346*(6209), 632–635.

Gneezy, U. and J. Potters (1997). An experiment on risk taking and evaluation periods. *Quarterly Journal of Economics 112*(2), 631–645.

Goeree, J. K., C. A. Holt, and S. K. Laury (2002). Private costs and public benefits: Unraveling the effects of altruism and noisy behavior. *Journal of Public Economics 83*(2), 255–276.

Gonzalez-Ocantos, E., C. K. De Jonge, C. Meléndez, J. Osorio, and D. W. Nickerson (2012). Vote buying and social desirability bias: Experimental evidence from nicaragua. *American Journal of Political Science 56*(1), 202–217.

Gosling, S. D., P. J. Rentfrow, and W. B. Swann (2003). A very brief measure of the big-five personality domains. *Journal of Research in Personality 37*(6), 504–528.

Grossman, G. and D. Baldassarri (2012). The impact of elections on cooperation: Evidence from a lab-in-the-field experiment in uganda. *American Journal of Political Science 56*(4), 964–985.

Grossman, P. J. and C. C. Eckel (1996). Altruism in anonymous dictator games. *Games and Economic Behavior 16*, 181–191.

Gueth, W., V. M. Levati, and M. Ploner (2008). On the social dimension of time and risk preferences: An experimental study. *Economic Inquiry 46*(2), 261–272.

Gueth, W., R. Schmittberger, and B. Schwarze (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization 3*(4), 367–388.

Harms, M. (2016). Crowdinvesting deutschland marktreport 2016.

Harrison, G. W., M. I. Lau, E. E. RutstrÃ¶m, and M. B. Sullivan (2005). Eliciting risk and time preferences using field experiments: some methodological issues. *Research in Experimental Economics 10*, 125–218.

Harrison, G. W., M. I. Lau, E. E. Rutström, and M. Tarazona-Gomez (2013). Preferences over social risk. *Oxford Economic Papers 65*, 25–46.

Herz, H. and D. Taubinsky (2018). What makes a price fair? an experimental study of transaction experience and endogenous fairness views. *Journal of the European Economic Association 16*(2), 316–352.

Hofmann, E., E. Hoelzl, and E. Kirchler (2008). A comparison of models describing the impact of moral decision making on investment decisions. *Journal of Business Ethics 82*(1), 171–187.

Hungerman, D. M. and M. Ottoni-Wilhelm (2018). Impure impact giving: Theory and evidence. Technical report, National Bureau of Economic Research.

Kahneman, D., J. L. Knetsch, and R. Thaler (1986a). Fairness as a constraint on profit seeking: Entitlements in the market. *The American economic review*, 728–741.

Kahneman, D., J. L. Knetsch, and R. H. Thaler (1986b). Fairness and the assumptions of economics. *Journal of business*, S285–S300.

Karlan, D. and J. A. List (2006). Does price matter in charitable giving? evidence from a large scale natural field experiment. *NBER Working Paper Series*.

Kirchler, M. (2009). Underreaction to fundamental information and asymmetry in mispricing between bullish and bearish markets. an experimental study. *Journal of Economic Dynamics and Control 33*, 491–506.

Kirchler, M., J. Huber, M. Stefan, and M. Sutter (2015). Market design and moral behavior. *Management Science*.

Kirchler, M., J. Huber, and T. Stöckl (2012). Thar she bursts - reducing confusion reduces bubbles. *The American Economic Review 102*(2), 865–883.

Kotchen, M. J. (2005). Impure public goods and the comparative statics of environmentally friendly consumption. *Journal of Environmental Economics and Management 49*, 281–300.

Kotchen, M. J. and M. R. Moore (2007). Private provision of environmental public goods: Household participation in green-electricity programs. *Journal of Environmental Economics and Management 53*, 1–16.

Krawczyk, M. and F. LeLec (2010). 'give me a chance!' an experiment in social decision under risk. *Experimental Economics 13*, 500–511.

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

Lai, C.-Y., A. Lange, J. A. List, and M. K. Price (2017). The business of business is business: why (some) firms should provide public goods when they sell private goods. *NBER Working Paper Series*.

Landry, C. E., A. Lange, J. A. List, M. K. Price, and N. G. Rupp (2006). Towards an understanding of the economics of charity: Evidence from a field experiment. *Quarterly Journal of Economics 121*(2), 747–782.

Lange, A., J. A. List, and M. K. Price (2007). Using lotteries to finance public goods: Theory and experimental evidence. *International Economic Review 48*(3), 901–927.

Lange, A., C. Schwirplies, and A. Ziegler (2017). On the interrelation between the consumption of impure public goods and the provision of direct donations: Theory and empirical evidence. *Resource and Energy Economics 47*, 72–88.

LeDuc, L., R. G. Niemi, and P. Norris (Eds.) (2014). *Comparing Democracies*, Volume 4. London: SAGE Publications.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. H. Kagel and A. E. Roth (Eds.), *Handbook of Experimental Economics*, pp. 111–181. Princeton University Press.

Lehner, O. M. (2013). Crowdfunding social ventures: a model and research agenda. *Venture Capital: An International Journal of Entrepreneurial Finance 15*(4), 289–311.

Levati, M. V. and A. Morone (2009). Voluntary contributions with risky and uncertain marginal returns: the importance of the minimal value. *Jena Economic Research Papers 2009-062*.

Levati, M. V. and A. Morone (2013). Voluntary contributions with risky and uncertain marginal returns: the importance of the parameter values. *Journal of Public Economic Theory 15*(5), 736–744.

List, J. A., A. M. Shaikh, and Y. Xu (2016). Multiple hypothesis testing in experimental economics. Technical report, National Bureau of Economic Research.

Margolis, J. D., H. A. Elfenbein, and J. P. Walsh (2009). Does it pay to be good... and does it matter? a meta-analysis of the relationship between corporate social and financial performance.

McWilliams, A. and D. Siegel (2000). Research notes and communications. corporate social responsibility and financial performance: correlation or misspecification? *Strategic management journal 21*(5), 603–609.

McWilliams, A., D. S. Siegel, and P. M. Wright (2006). Corporate social responsibility: Strategic implications*. *Journal of management studies 43*(1), 1–18.

Munro, A. and M. Valente (2015). Green goods: Are they good or bad news for the environment? evidence from a laboratory experiment on impure public goods. *Environmental and Resource Economics*, 1–19.

Norris, P. (2014). Electoral integrity and political legitimacy. In L. LeDuc, R. G. Niemi, and P. Norris (Eds.), *Comparing Democracies* (4th ed.)., pp. 150–172. SAGE Publications Ltd.

Noussair, C. N. and S. Tucker (2016). Cash inflows and bubbles in asset markets with constant fundamental values. *Economic Inquiry 54*(3), 1596–1606.

Nuzzo, S. and A. Morone (2017). Asset markets in the lab: A literature review. *Journal of Behavioral and Experimental Finance 13*, 42–50.

Nyborg, K., R. B. Howarth, and K. A. Brekke (2006). Green consumers and public policy: On socially contingent moral motivation. *Resource and Energy Economics 28*, 351–366.

Orlitzky, M., F. L. Schmidt, and S. L. Rynes (2003). Corporate social and financial performance: A meta-analysis. *Organization studies 24*(3), 403–441.

Pahlke, J., S. Strasser, and F. M. Vieider (2015). Responsibility effects in decision making under risk. *Journal of Risk and Uncertainty 51*, 125–146.

Palan, S. (2013, July). A review of bubbles and crashes in experimental asset markets. *Journal of Economic Surveys 27*(3), 570–588.

Palan, S. (2015). Gims - software for asset market experiments. *Journal of Behavioral and Experimental Finance 5*, 1–14.

Palfrey, T. R. and J. E. Prisbre (1997). Anomalous behavior in public goods experiments: How much and why? *American Economic Review 87*(5), 829–846.

Pasewark, W. R. and M. E. Riley (2010). Itâsa matter of principle: the role of personal values in investment decisions. *Journal of business ethics 93*(2), 237–253.

Porter, D. P. and V. L. Smith (1995, October). Futures contracting and dividend uncertainty in experimental asset markets. *The Journal of Business 68*(4), 509–541.

Potters, J., M. Sefton, and L. Vesterlund (2005). After you - endogenous sequencing in voluntary contribution games. *Journal of public Economics 89*(8), 1399–1419.

Potters, J., M. Sefton, and L. Vesterlund (2007). Leading-by-example and signaling in voluntary contribution games: an experimental study. *Economic Theory 33*(1), 169–182.

Powell, O. (2016). Numeraire independence and the measurement of mispricing in experimental asset markets. *Journal of Behavioral and Experimental Finance 9*, 56–62.

Powell, O. and N. Shestakova (2016). Experimental asset markets: A survey of recent developments. *Journal of Behavioral and Experimental Finance 12*, 14–22.

Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American economic review*, 1281–1302.

Renneboog, L., J. T. Horst, and C. Zhang (2008). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance 14*, 302–322.

Revelli, C. and J.-L. Viviani (2013). The link between sri and financial performance: Effects and moderators. *International Management 17*(2), 105–122.

Revelli, C. and J.-L. Viviani (2015). Financial performance of socially responsible investing (sri): what have we learned? a meta-analysis. *Business Ethics: A European Review 24*(2), 158–185.

Riedl, A. and P. Smeets (2017). Why do investors hold socially responsible mutual funds? *The Journal of Finance 72*(6), 2505–2550.

Rockenbach, B., A. Sadrieh, and B. Mathauschek (2007). Teams take the better risks. *Journal of Economic Behavior and Organisation 63*, 412–422.

Romano, J. P. and M. Wolf (2005). Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association 100*(469), 94–108.

Saito, K. (2013). Social preferences under risk: Equality of opportunity versus equality of outcome. *American Economic Review 103*(7), 3084–3101.

Scharf, K. and S. Smith (2010). The price elasticity of charitable giving: Does the form of tax relief matter? *Working Paper*.

Schlag, K. H. and J. Tremewan (2016). Simple belief elicitation. *mimeo*.

Schmitz, C. et al. (2012). Limesurvey: An open source survey tool. LimeSurvey Project Hamburg, Germany. URL http://www. limesurvey. org.

Shupp, R. S. and A. W. Williams (2008). Risk preference differentials of small groups and individuals. *The Economic Journal 118*, 258–283.

Sleesman, D. J. and D. E. Conlon (2017). Encouraging prosocial decisions: The role of fairness salience and uncertainty. *Journal of Behavioral Decision Making 30*(2), 502–515.

Smith (1982). Markets as economizers of information: Experimental examination of the 'hayek hypothesis'. *Economic Inquiry 20*, 165–179.

Smith, V., M. van Boening, and C. P. Wellford (2000). Dividend timing and behavior in laboratory asset markets. *Economic Theory 16*, 567–583.

Smith, V. L., G. L. Suchanek, and A. W. Williams (1988, September). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica 56*(5), 1119–1151.

Stöckl, T., J. Huber, and M. Kirchler (2010). Bubble measures in experimental asset markets. *Experimental Economics 13*(3), 284–298.

Stöckl, T., J. Huber, and M. Kirchler (2010). Bubble measures in experimental asset markets. *Experimental Economics 13*, 284–298.

Stoddard, B. (2017). Risk in payoff-equivalent appropriation and provision games. *Journal of behavioral and experimental economics 69*, 78–82.

Stoddard, B. V. et al. (2015). Probabilistic production of a public good. *Economics Bulletin 35*(1), 37–52.

Sutter, M., S. Haigner, and M. G. Kocher (2010). Choosing the carrot or the stick? endogenous institutional choice in social dilemma situations. *The Review of Economic Studies 77*(4), 1540–1566.

Sutter, M., J. Huber, M. Kirchler, and M. Stefan (2019). Where to look for the morals in markets? *Experimental Economics*, 1–23.

Sutter, M., M. K. Jürgen Huber, and M. Stefan (2016). Where to look for the morals in markets? *IZA Discussion Paper No. 10105*.

Teyssier, S., F. Etilé, and P. Combris (2014). Social-and self-image concerns in fair-trade consumption. *European Review of Agricultural Economics 42*(4), 579–606.

Théroude, V. and A. Zylbersztejn (2020). Cooperation in a risky world. *Journal of Public Economic Theory 22*(2), 388–407.

Tiefenbeck, V., T. Staake, K. Roth, and O. Sachs (2013). For better or for worse? empirical evidence of moral licensing in a behavioral energy conservation campaign. *Energy Policy 57*, 160–171.

Tyler, T. (1990). *Why People Obey Rules.* Yale University Press.

Tyler, T. R. (2006). Psychological perspectives on legitimacy and legitimation. *Annual Review of Psychology 57*, 375–400.

Tyran, J.-R. and L. P. Feld (2006). Achieving compliance when legal sanctions are non-deterrent. *The Scandinavian Journal of Economics 108*(1), 135–156.

UK Electoral Commision (2018). Analysis of cases of alleged electoral fraud in the uk in 2017. Technical report (Summary of data recorded by police forces).

US SIF Foundation (2018). Report on us sustainable, responsible and impact investing trends 2018 executive summary.

Voorhes, M. and J. Humphreys (2016). Report on us sustainable, responsible and impact investing trends 2016. *Technical report US SIF Foundation*.

Waddock, S. A. and S. B. Graves (1997). The corporate social performance-financial performance link. *Strategic management journal 18*(4), 303–319.

Warner, M. E. (2013). Private finance for public goods: social impact bonds. *Journal of Economic Policy Reform 16*(4), 303–319.

Weber, M. (1978). *Economy and Society*. Berkeley: University of California Press.

Wright, P. and S. P. Ferris (1997). Agency conflict and corporate strategy: The effect of divestment on corporate value. *Strategic Management Journal 18*(1), 77–83.

WVS (2014). *WORLD VALUES SURVEY Wave 6 2010-2014 OFFICIAL AGGREGATE v.20150418*. World Values Survey Association (www.worldvaluessurvey.org). Aggregate File Producer: Asep/JDS, Madrid SPAIN.

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics 6*, 299–310.

Zwick, T. (2004). Employee participation and productivity. *Labour Economics 11*(6), 715–740.

Appendix A

Summaries

Chapter 2: *Manipulated Votes and Rule Compliance*

We design an online experiment with large voter groups to study how vote buying and partial disenfranchisement of the electorate during a referendum affects compliance with elected rules of redistribution. To our knowledge, this is the first experimental paper to study whether the well-documented positive behavioral effects of democratic institutions are sensitive to electoral manipulation. We establish a strong causal effect of manipulative interventions on compliance with rules promoting redistribution: When votes have been bought or parts of the electorate been excluded from the ballot, subjects comply significantly less with elected rules that ask them to share their income with other members of the experimental society. Analyzing beliefs, we find no evidence that treatment effects are driven by strategic concerns regarding the behavior of other subjects. Rather, subjects seem to react intrinsically to violations of the idea of inclusive and unbiased elections. Treatment effects are found mainly among individuals who—in a questionnaire that is presented as an unrelated survey two weeks after the experiment—indicate a high valuation of democratic institutions and little justifiability for bribes and (political) lobbying in the real world.

In diesem Online-Experiment analysieren wir wie Wahlmanipulation in Form von Stimmenkauf und Ausschluss von Teilen der Wählerschaft in einem Referendum beeinflusst, inwiefern gewählte Regeln zur Umverteilung von Einkommen befolgt werden. Hiermit testen wir, ob die empirisch gut dokumentierten positiven Verhaltens-Effekte von demokratischen Institutionen durch Wahlmanipulation beeinflusst werden. Wir identifizieren einen starken kausalen Effekt von Manipulation auf die Bereitschaft der Experimentteilnehmer Regeln, die Umverteilung fordern, zu befolgen: Wenn Stimmen gekauft wurden oder ein Teil der Wähler von der Wahl ausgeschlossen wurde, befolgen weniger der Teilnehmer gewählte Regeln, die von ihnen fordern, ihr Einkommen mit anderen Mitgliedern der Gesellschaft (im Experiment) zu teilen—im Vergleich zu durch faire Wahlen eingesetzten Regeln. Wir finden keine empirischen Hinweise darauf, dass die Erwartungen der Teilnehmer über das Verhalten anderer Teilnehmer deren Entscheidung eine Regel, die Umverteilung von Einkommen fordert, zu befolgen beeinflusst. Im Gegenteil scheint es, dass die Unterschiede zwischen den Treatments mit und ohne Manipulation stark davon bestimmt werden, dass die Teilnehmer direkt auf die Verletzung der demokratischen Prinzipien von repräsentativen Wahlen reagieren. Dies wird durch die Beobachtung unterstützt, dass wir die Treatment Unterschiede hauptsächlich unter den Teilnehmern finden, die in einem separaten Fragebogen angegeben haben, demokratische Institutionen sehr wert zu schätzen und Bestechungsgelder sowie Lobbying nicht vertretbar zu finden.

Chapter 3: *Investments in Impure Public Goods*

This paper experimentally investigates to what extend the type of risk inherent in social investments influences their attractiveness for investors. In particular, we analyze how risk in the provision of the public benefit and in the financial return to the investor each affect investment decisions separately and how, in addition, their correlation influences investments when both risks are simultaneously present. The results show that the reaction to risk in the private return to the investor and in the public benefit (that is paid to a charity) crucially depends on (1) the correlation of the co-existent risks and (2) the type of the investor. Identifying heterogeneous treatment effects shows a particularly strong reaction of pro-social and risk averse participants to co-existent risks when random draws are independent. It furthermore suggests that less inherently pro-social and less risk averse participants can be attracted to invest in risky impure public goods. The findings not only inform social investments such as crowdinvestments and microlending but also theories of giving under risk: the data rather support models that assume donors (also) care about the impact of their donation on the public good.

Wir untersuchen mithilfe eines Experiments inwiefern die Art des Risikos in ethischen Investitionen ('social investments') deren Attraktivität für Investoren beeinflusst. Genauer gesagt analysieren wir, wie Risiko in der ethischen Komponente des Investments (d.h. in der gemeinnützigen Auszahlung) und Risiko in der Rendite an den Investor jeweils

Investitionen beeinflussen. Darüber hinaus testen wir, wie die Art der Korrelation dieser beiden Risiken Investitionen beeinflusst wenn beide gleichzeitig auftreten. Die Resultate des Experiments zeigen, dass die Reaktionen auf Risiko in der privaten Rendite an den Investor und in der gemeinnützigen Rendite (modelliert als Zahlung an eine wohltätige Organisation) von (1) der Korrelation der co-existierenden Risiken und (2) dem Typ von Investor abhängen. Die Identifizierung von heterogenen Treatment-Effekten zeigt eine besonders starke negative Reaktion von denjenigen Investoren, die wir als sozial und risikoscheu charakterisieren—und dies umso mehr, wenn die beiden Risiken voneinander unabhängig sind. Die Analyse suggeriert darüber hinaus, dass weniger soziale und gleichzeitig risikofreudigere Investoren zu risikobehafteten Investitionen, die eine ethische Komponente beinhalten ('impure public goods'), motiviert werden können. Diese Ergebnisse sind nicht nur für 'social investments' wie beispielsweise crowdinvestments oder microlending von Interesse, sondern bereichern auch ökonomische Theorien über pro-soziales Verhalten: Die Daten aus dem Experiment unterstützen Modelle, die annehmen, dass es Spendern wichtig ist, dass ihre Spende tatsächlich eine Auswirkung auf die Bereitstellung des öffentlichen Gutes hat.

Chapter 4: *On the voluntary provision of public goods under risk: the role of risk in private or public returns*

We use a laboratory experiment to disentangle the impact of risk in the private and public dimension of social investments. Based on variants of a public good game, we separate the return of a subject's investment for herself vs. the return to others. We identify a detrimental effect of risk when returns in both dimensions are risky and positively correlated or independent. A negative correlation limits the downside risk and leads to more stable investments. Disentangling the impact of risk in the two dimensions, we find that investments particularly respond to the risk in the public return dimension.

Wir nutzen ein Laborexperiment um die Rolle von Risiko in der privaten und der öffentlichen Dividende von Investitionen von Gruppen von Individuen in öffentliche Güter zu untersuchen. Das Experiment basiert auf 'public good games', in denen wir die Rendite, die aus dem Beitrag eines Gruppenmitgliedes stammt, in zwei Teile trennen: den Teil, der an das beitragende Gruppenmitglied zurückgezahlt wird und den Teil, der an die anderen

Gruppenmitglieder ausgezahlt wird. Risiken wirken sich insbesondere dann negativ auf die Höhe der durchschnittlichen Beiträge zum öffentlichen Gut aus, wenn sie in beiden Komponenten der Rendite präsent sind und dabei entweder unabhängig voneinander oder positiv korreliert sind. Negative Korrelation hingegen stabilisiert die Höhe der Beiträge, da sie das Ausfallrisiko eliminiert. Wenn wir die Auswirkungen des Risikos in beiden Komponenten trennen, wird sichtbar, dass Beitragszahlungen vor allem auf das Risiko in der Auszahlung an die anderen Gruppenmitglieder reagieren.

Chapter 5: *On the performance of green assets in financial markets—Evidence from a laboratory experiment*

We investigate the financial performance of socially responsible investments in competition with conventional investments on financial markets. Setting up experimental asset markets for 'green', i.e. socially responsible, and non-green stocks and using a novel market design, we identify the impact of green investment opportunities on the development of prices and trading volumes. In particular, we test how speculation about future prices influences green market prices and to what extend subjects' behavior in the market is correlated with their generosity in individual donation decisions. We observe no price premium for the green assets. On the contrary, the introduction of green investment opportunities increases prices for the non-green assets compared to markets with only conventional firms.

Wir untersuchen die Wertentwicklung von ethischen Investitionen ('socially responsible investments' (SRI)) die auf dem Finanzmarkt mit konventionellen Investitionen im Wettbewerb stehen. Indem wir experimentelle Aktienmärkte für 'grüne' (d.h. ethische) und nicht-grüne Aktien aufsetzen, können wir die Auswirkungen von grünen Investitionsmöglichkeiten auf die Entwicklung von Preisen und Handelsvolumen auf dem Markt identifizieren. Insbesondere testen wir wie Spekulation über zukünftige Preise die grünen Marktpreise beeinflusst und inwiefern das Verhalten der Teilnehmer im Markt mit deren Spendenverhalten außerhalb des Marktes zusammenhängt. Wir finden keine höhere Zahlungsbereitschaft für die grünen Aktien mit einer positiven Externalität, die über den erwarteten Wert der Aktie hinausgeht. Im Gegenteil führen die grünen Investitionsmöglichkeiten auf einem Markt zu einer Erhöhung der Preise zu denen die Aktien der konventionellen Firma gehandelt werden.