Corrupted Votes and Rule Compliance[*]

Arno Apffelstaedt[†]   Jana Freundt[‡]
U of Cologne   UPenn

14th May 2018

Abstract

We study—using an online experiment with international subjects—how compliance with elected rules of conduct is affected by having experienced an election in which (1) subjects are asked for money to make their vote count, (2) subjects are offered money for voting differently, or (3) subjects with low household income are excluded from the ballot. We find strong and significant reductions in compliance rates across the population after such "corrupt elections", but only if elected rules ask subjects to behave prosocially. Treatment effects seem to be driven by intrinsic concerns about procedural aspects of the election mechanism and are prevalent mainly among individuals who—in a questionnaire that is presented as an unrelated survey two weeks after the experiment—express high value for democratic institutions and low value for bribing and (political) lobbying in the real world.

Keywords: Endogenous Institutions, Corruption, Rule Compliance

JEL Codes: D72, D91, C92

1 Introduction

An influential stream of papers in public and political economics suggests that democratic institutions may affect behavior.[1] Frey (1997), for example, finds that tax compliance is higher in Swiss cantons that see more democratic participation. Bardhan (2000) shows that South Indian farmers are more likely to follow irrigation rules if they partake in crafting them. Experimentally, Tyran and Feld (2006), Ertan, Page and Putterman (2009) and Sutter, Haigner and Kocher (2010), among others, demonstrate that punishments and rewards have greater impact on contributions to a public good when they are implemented by majority vote rather than exogenously by a computer. Dal Bó, Foster and Putterman (2010) provide experimental evidence of a similar 'democracy effect' in co-ordination games.[2] A conclusion that can be drawn from this literature is that giving citizens decision rights through elections and referenda can bring important efficiency gains to societies.

In many countries, however, promises of "free and fair" elections are undermined by practices ranging from systematic vote buying to arguably unintentional disfranchisement of poor voters.[3] Similar to how the introduction of a democratic procedure can generate positive behavioral responses, perceived malpractice and voter manipulation during elections may lead to negative behavioral consequences. In this essay, we test this hypothesis using a novel online experiment. The experiment studies how vote buying and voter disenfranchisement during a referendum affects the willingness of individuals to comply with elected rules asking them to behave pro-socially (to redistribute income) and with elected rules asking them to behave selfishly (to not redistribute). To our knowledge, this is the first experimen-

---

[1]There is a related literature in organizational economics that studies the value of "democratic" decision making mechanism within firms. Bonin, Jones and Putterman (1993), Black and Lynch (2001) and Zwick (2004), for example, provide empirical support that employee participation is associated with increased worker productivity. On a general account, Bartling, Fehr and Herz (2014) are able to demonstrate experimentally that many people yield intrinsic value from decision rights.

[2]This list of studies is not meant to be exhaustive. See, e.g., Dal Bó (2014) for further studies.

[3]In a survey study in Argentina from 2002, for example, 35% of respondents reported to have observed the distribution of gifts by political parties in their neighborhoods during election campaigns and 12% of low-income respondents reported to have received something from a political party or candidate (Brusco, Nazareno and Stokes, 2004, pp. 69-70). According to a list experiment by Gonzalez-Ocantos et al. (2012) (a technique that usually assures to minimize social desirability biases in sensitive survey questions) more than 24% of registered voters reported to have been offered some sort of gift for their vote after the 2008 Nicaraguan municipal election. Examples for arguably unintentional voter disenfranchisement are restrictive ID laws (De Alth, 2009) or felon disfranchisement (Manza and Uggen, 2008) in some states of the US. In 2017 alone, allegations of voter fraud have led to violent demonstrations in Turkey, Venezuela, Indonesia and the US, among other countries. A systematic, world-wide analysis of electoral malpractices and survey-based evidence of voters' expressed dissatisfaction with biased electoral procedures can be found, for instance, in Norris (2014).

tal study on whether the well-documented positive behavioral effects of democratic institutions are sensitive to electoral malpractice. In comparison to earlier studies on 'democracy effects', our experiment allows us to say more about the psychological mechanisms underlying behavior and treatment effects. We establish a strong negative (intrinsic) effect of electoral malpractices on compliance with pro-social rules: 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 redistribute. Maybe surprisingly, we find no significant treatment effects on compliance with selfish rules.

We study redistribution choices in experimental societies made up of 100 individual subjects. Subjects are recruited online via the platform Prolific.ac.[4] The experiment revolves around the decision of whether one should redistribute income earned through luck to another member of the society who was unlucky (i.e., did not receive any income). We implement this decision with a binary one-shot dictator game: Each subject in the society has to decide conditional on receiving income whether she wants to $Give_i \in \{0, 1\}$ thirty percent of her income to a randomly matched person $j \neq i$ who did not receive income. Before subjects decide whether to redistribute, there is a referendum on the right "code of conduct." Each subject can vote for a (society-wide) code that promotes giving (*Rule:Give*) or for a code that promotes non-giving (*Rule:Don't*). After the referendum, subjects decide (individually and anonymously) whether they want to $Give_i|Rule:Give \in \{0, 1\}$ conditional on *Rule:Give* being elected and whether they want to $Give_i|Rule:Don't \in \{0, 1\}$ conditional on *Rule:Don't* being elected. We are interested in how voluntary compliance with each of the two rules depends on electoral malpractice (in the form of vote buying or partial disenfranchisement) being present during the referendum.[5] The hypothesis guiding our analysis is that compliance with both rules should be *lower* in societies that experience malpractice during the referendum compared to the levels of compliance observed in a society that did not experience electoral malpractice.

Using different treatment groups (each consisting of a society with 100 subjects), we introduce interventions to the referendum that may either lead to some voters being excluded from the ballot (= partial disenfranchisement) or to some votes

---

[4]Prolific.ac has a subject pool of about 40.000 people and administers recruiting and payment. The Prolific.ac subject pool consists of individuals out of whom 60% are male, 26% are students, 85% speak English as a first language, roughly 60% have the UK nationality and 25% the US nationality. The remaining subjects have all kinds of different nationalities. The median age is 27. Education levels vary from no formal education (3%), college education (41%), undergraduate (33%) or graduate (18%) education to doctoral degrees (4%). See `https://www.prolific.ac/demographics` (accessed November 11th, 2017).

[5]Complying with the elected code of conduct is entirely voluntary: There is no formal punishment involved with deviation. There is also no possibility for other subjects to punish the choice of individual i.

not being representative of the true opinion of their issuer (= vote buying). Our interventions are the introduction of a small voting fee (the votes of subjects who do not pay are not counted towards the referendum), monetary offers to all subjects if they vote for the rule opposite to their first choice (vote buying), and an exclusion of all subjects from the ballot whose household income is below a certain threshold (GBP 40,000). A baseline treatment in which the votes of all 100 subjects are counted in an unbiased way serves as the comparison.

We choose to study behavior in one-shot dictator games primarily for two reasons. The first reason is that non-binding rules in this domain should mainly work by their normative appeal. In particular, (classical) co-ordination issues as well as punishment concerns that exist in other games should not play a role in this setting.[6] This makes dictator games particularly well suited for the analysis of whether procedural changes in how an election is conducted affect the intrinsic motivation of subjects to follow rules.[7] For reasons we discuss in the next paragraph we hypothesize that rules should have higher normative appeal when they were selected in an inclusive and unbiased way, that is, with a referendum that did not involve vote buying or disenfranchisement. The second reason is that we aim to create an experimental situation in which people disagree about the "right" code of conduct and hence, potentially, vote for different rules. Note, importantly, that there is no efficiency-dominant rule. *Rule:Give* and *Rule:Don't* differ only in their distributive nature. Earlier studies have shown that people differ in their judgements regarding whether income received through luck should be redistributed, see, in particular, Cappelen et al. (2007) and Almås, Cappelen and Tungodden (2017). Our setup allows us to study behavior under rules that promote "egalitarian" values (*Rule:Give*) and rules that promote "libertarian" values (*Rule:Don't*).[8]

Finer details of our experimental design are meant to identify the psychological determinants of behavior that underlie rule-compliance and treatment effects. Research in psychology and behavioral economics suggests that procedural aspects of

---

[6]Earlier experiments on the behavioral effects of democratic elections have primarily looked at repeated public good games, trust games, and co-ordination games, see e.g., Tyran and Feld (2006) and Dal Bó, Foster and Putterman (2010). In those games, expectations about the behavior of other subjects are likely to play a more important role than they do in a dictator game. While there are no classical co-ordination incentives in one-shot dictator games—conditional on being a dictator, the strategies of other agents cannot influence a subject's monetary payoff—there might be "psychological" co-ordination incentives arising from the wish to align one's behavior with what others do or value. Our experiment is designed to test for such incentives, see the next paragraph.

[7]Dictator games have been chosen in earlier studies for similar reasons, see, for example, Krupka and Weber (2013), albeit not to our knowledge in studies on the effects of democracy on behavior. Note also that dictator games, in comparison to other interesting games in which rule-compliance is key—for example, games used to study cheating or lying behavior (Fischbacher and Föllmi-Heusi, 2013; Gächter and Schulz, 2016)—, do not entail the possibility that with non-compliance a subject can punish the *experimenter* for a procedure she perceives as unfair.

[8]Our use of the words "egalitarian" and "libertarian" follows Almås, Cappelen and Tungodden (2017).

decision making can affect preferences directly. In particular, people seem to care about the "fairness" of decision making processes (see, e.g., Tyler, 1990; Frey, Benz and Stutzer, 2004; Cappelen et al., 2013) as well as about personally partaking in them (see, e.g., Bonin, Jones and Putterman, 1993; Bardhan, 2000; Bartling, Fehr and Herz, 2014). Vote buying and partial disenfranchisement during elections is certain to affect preferences on the latter domain. Intuitively, preferences concerning the fairness of the decision making process should also be affected. The view that procedural concerns may lower the normative appeal of elected rules and thus, directly affect the willingness of people to comply is related to theories of "legitimate authority" (Weber, 1978; Tyler, 2006; Akerlof, 2017). We control for three aspects that might affect a subject's decision to comply with rules in the dictator game apart from such concerns: (1) her preferences regarding the "right" code of conduct, (2) her behavior in the absence of a rule, and (3) her beliefs about the behavior of other subjects. To control for (1), we introduce our treatment interventions only after all subjects have stated a preference for the rule (*Rule:Give* or *Rule:Don't*) they want to vote for. This allows us to control for the unbiased vote of a subject in all treatments—even if this vote might not count towards the final referendum.[9] We control for (2) by introducing a prior round of the dictator game to our experiment in which subjects decide whether to $Give_i \in \{0, 1\}$ without knowing that there will be a second round that includes the election of a code of conduct. This allows us to identify a subject as a "natural" giver or non-giver—a categorization that turns out to play an important role in our analysis. Instead of giving subjects information about the behavior of other participants in this round—which might induce undesired punishment behavior in the second round of the dictator game following the referendum—, we present them with partial information about redistribution choices in our experimental pilot. By varying this information randomly on a subject-by-subject basis, we generate exogenous variance in the beliefs about the behavior of other subjects. This helps us to *causally* identify (3): The role of others in guiding behavior.[10] Beliefs about the voting and compliance behavior of other subjects as well as beliefs about the impact of manipulative interventions on the referendum outcome are elicited (in an incentive compatible way) from every subject at the end of the experiment. Our main finding regarding the psychological determinants of behavior is that beliefs about the behavior of other subjects seem to play little to no role in explaining our treatment effects. Rather, subjects seem to react intrinsically to violations of the democratic ideal that elections should be

---

[9]This control follows the identification procedure introduced by Dal Bó, Foster and Putterman (2010).

[10]For example, we can use variance in the information we give subjects after round 1 of the dictator game to instrument for variance in beliefs about the behavior of other subjects in round 2.

inclusive and unbiased.

How justifiable is...?
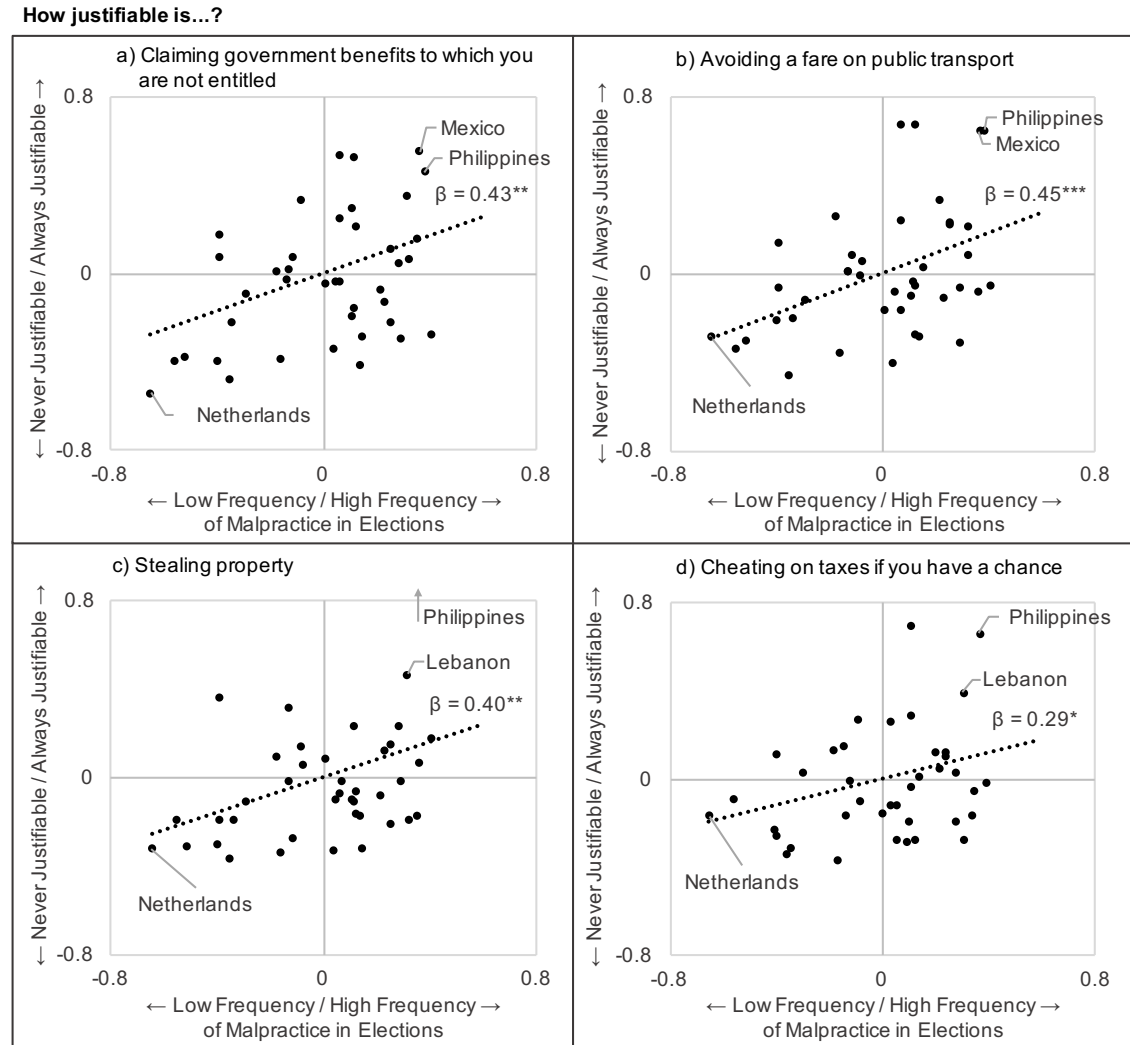
Figure 1: Country-level correlations between citizens' perceived frequency of malpractice in elections and their statements about the justifiability of violating rules and laws. Source: Country averages calculated from the WVS (2014). The figures plot the average answers in a country to questions V198-V201 against an index of perceived malpractice in elections. This index is calculated from the average of 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). We have normalized the data to show relative deviations from the average across all countries. For example, in panel d), Lebanon's data point is $(0.30, 0.38)$ meaning that is has a 30% higher measure of perceived malpractice and 38% higher measure of justifiability for tax cheating than the average country in our sample. The β-coefficients are from univariate OLS regressions without intercept: $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$ assuming OLS standard errors.

We complement our experiment with an extensive questionnaire on subjects' standpoints regarding various political issues such as redistribution, corruption, democratic values, and personal trust in institutions. To prevent the risk of spillovers from exposure to different treatments to questionnaire answers, the ques-

tionnaire is presented as an unrelated survey (using a different design and researcher profile) and is send to the same people about two weeks after they participated in the experiment. We use the questionnaire to study whether self-reported standpoints on the value of democratic institutions correlate with reactions to electoral manipulation in the experiment. Indeed, we find that our treatment effects are mainly driven by subjects who self-report to have a high valuation for democratic institutions.

Indicative evidence for the hypothesis that electoral malpractice affects the willingness of people to comply with social rules and laws can also be found in observational data. In answers gathered from the World Values Survey (see Figure 1) the level of electoral malpractice perceived in a country is positively correlated with individual judgments regarding the justifiability of breaking rules, ranging from wrongfully claiming government benefits to cheating on taxes. However, because the level of malpractice is difficult to randomize in real elections, causality is hard to establish in the field. Where this is possible, researchers then generally have to rely on surveys to measure aggregate effects on behavior.[11] Individual level behavioral measures of voluntary rule-compliance are almost impossible to come by due to the difficulty to control for formal and informal deterrence measures that are in place in the field. An additional comparative advantage to using real world data is that our experimental framework enables us to study the psychological mechanisms driving treatment effects.

By relying on direct instead of indirect behavioral measures of support and dissatisfaction among citizens, political scientists have mostly taken a different approach towards assessing people's acceptance of elected institutions. Extensive survey studies of whether electoral malpractice undermines citizens' expressed support for institutions is provided by Norris (2014). An experimental approach to eliciting such direct support is taken by, for example, Dickson, Gordon and Huber (2015), who measure the legitimacy of an institution by observing whether participants help or hinder an authority in punishing free-riders in a public good game. We are not aware of an experimental study that is trying to test what we are after.

The remainder of the chapter is structured as follows. Section 2 presents the experimental design in detail together with the predictions and identification strategy. Sections 3 and 4 present our results: We first estimate the average effect of

---

[11]For example, Berman et al. (2014) sent letters to a random sample of Afghan polling stations announcing that researchers would photograph election results and that these photographs would later be compared to certified results. This threat of control seems to have reduced election fraud (see also Callen and Long, 2015). The authors rely on a post-election survey to measure the effect of this treatment on attitudes towards government, of which "the willingness to report insurgent behavior to security forces" is the measure closest to what we are after. They find that sending a letter increases this willingness by 2.5 to 3 percentage points, which is statistically significant and supports our hypothesis.

vote manipulation on compliance rates and then study determinants of individual rule compliance. We conclude in section 5. Experimental instructions, screenshots, and the questionnaire can be found in the appendix to this chapter.

2 Experimental Design

The design of our online experiment is based on a referendum among 100 subjects on the preferred "code of conduct" regarding behavior in a dictator game. For each treatment, 100 subjects participate in a lottery that has one of them winning GBP 100. They are informed that the computer will unequally distribute lottery tickets among the 100 participants: 50 subjects will be "receivers" who get 10 lottery tickets each, while the remaining 50 subjects will be "non-receivers" and get no tickets. One of the 500 distributed lottery tickets is the winning ticket. We use this set-up to construct a dictator game with role uncertainty: Before learning whether one is a receiver or a non-receiver of tickets, each subject is asked to (privately) decide whether—in case of being a receiver—she wants to give three out of ten lottery tickets to a randomly selected non-receiver.[12] In other words, each subject decides whether she wants to redistribute chances to win that she received through luck to another participant who was unlucky. In each session, we implement two rounds of this dictator game. Round 1 is a simple individual decision, the choice of individual i in this part is coded $Give_i \in \{0, 1\}$. In round 2, before subjects play the the dictator game again, they hold a referendum on a "code of conduct" for the whole group of 100 subjects. All subjects vote (privately) for either *Rule:Give* ("everybody should give"), $Vote_i = 1$, or for *Rule:Don't* ("everybody should not give"), $Vote_i = 0$. After the referendum, each individual decides privately whether she wants to $Give_i|Rule:Give \in \{0, 1\}$ conditional on *Rule:Give* being elected and whether she wants to $Give_i|Rule:Don't \in \{0, 1\}$ conditional on *Rule:Don't* being elected. There is no (monetary) punishment involved in not following the elected rule.

Treatments differ in whether or not there is malpractice during the referendum and, if there is malpractice, in the form of malpractice introduced. We introduce treatment interventions *after* subjects have voted, but *before* they take decisions $Give_i|Rule:Give$ and $Give_i|Rule:Don't$. The baseline treatment (*T_Baseline*) implements a simple majority vote. After voting, subjects are informed that "the rule that receives more votes in total will be implemented as the code of conduct." The other

---

[12]Subjects are informed that in the case of being a receiver (50% probability), their decision is automatically implemented and determines the number of lottery tickets for them and for one random other. They are also informed that in the case of being a non-receiver (50% probability), their decision does not play a role for the distribution of lottery tickets.

three treatments allow for the possibility that either, some votes are not counted towards the majority vote, or that the final votes may have been manipulated. In *T_Pay4Vote*, after voting, subjects see a screen that asks them to pay GBP 0.20 to make their vote count and informs them that the code of conduct will be selected by majority vote among those subjects who accepted to pay. In *T_Bribe*, subjects see a screen that offers them a bonus payment of GBP 0.20 if they reverse their vote and informs them that the code of conduct will be selected by majority vote after each subject has decided to either accept or reject this offer. Finally, in *T_ExcludePoor*, subjects are informed that the code of conduct will be selected by majority vote among subjects with an annual household income above GBP 40,000. They are also informed whether this means that their personal vote is counted or not.[13] The prediction guiding our analysis is:

Prediction 1 (Malpractice Effect). *The manipulation of electoral processes lowers voluntary compliance with elected rules:*

(a) $E(Give_i|Rule:Give, Malpractice = 1) - E(Give_i|Rule:Give, Malpractice = 0) < 0$

(b) $E(Give_i|Rule:Don't, Malpractice = 1) - E(Give_i|Rule:Don't, Malpractice = 0) > 0$

In our experiment, Malpractice = 1 if individual i is in treatment T_Pay4Vote, T_Bribe, or T_ExcludePoor, and Malpractice = 0 if individual i is in treatment T_Baseline.

2.1 Theoretical Framework

To fix ideas, consider the following simple theoretical framework.[14] 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 the utility of individual i when deciding to give or not give, respectively. Individual i then chooses to give if and only if

$$\Delta u_i(Give) := u_i(1) - u_i(0) \geq 0.$$

Classical economic theory would predict that $\Delta u_i(Give)$ is negative. A positive $\Delta u_i(Give)$ may reflect social preferences of individual i or "warm glow".[15] People

---

[13]To identify a subject as having a household income above or below GBP 40,000, we use self-declared information provided to us (with consent of the participants) by the online-platform *Prolific.ac*.

[14]We provide a framework regarding the effect of our treatments on giving behavior. We extend this framework to cover voting behavior in the appendix.

[15]Typical examples in standard dictator games would be Fehr and Schmidt (1999), Bolton and Ockenfels (2000) and Andreoni (1989, 1990). Note however that due to individual i distributing lottery tickets, these theories can explain positive giving rates in our setting only if endowments are understood in an *ex ante* sense, that is, under the assumption that individual i has preferences

might also want to align their behavior with anticipated giving behavior of others, driven by preferences for conformity (Bernheim, 1994; Bénabou and Tirole, 2012) or positive reciprocity (Fehr and Gächter, 2000). We will call those who give *Givers* and those do not give *Non-Givers* throughout the analysis. Let $\Delta u_i(Give)$ be distributed in the population with cumulative density function $F[\cdot]$. In the absence of a rule, the share of *Givers* in the population is then given by $1 - F[0]$ as illustrated in Figure 2, panel a), below.

Now consider the case in which there exists a democratically elected code of conduct that either promotes giving, *Rule:Give*, or promotes non-giving, *Rule:Don't*. Theories of "legitimate authority" (e.g., Weber, 1978; Tyler, 2006; Akerlof, 2017) suggest that if a rule has come into force by a fair procedure, "people feel that they ought to defer [its] decisions and rules, following them voluntarily out of obligation rather than out of fear of punishment or anticipation of reward." (Tyler, 2006, p.375). This is in line with earlier literature in psychology and behavioral economics which suggests that procedural aspects of decision making affect preferences directly (Tyler, 1990; Frey, Benz and Stutzer, 2004; Cappelen et al., 2013; Bartling, Fehr and Herz, 2014, among others). If people care to align their behavior with others, elected rules might change behavior because they provide a signal about what others do and value (Basu, 2015; Akerlof, 2016). Earlier experiments (e.g., Tyran and Feld, 2006; Sutter, Haigner and Kocher, 2010; Dal Bó, Foster and Putterman, 2010) confirm that endogenously elected institutions have the power to change behavior, but do not disentangle the psychological reasons why. Our experiment is designed to provide more insights into the psychological mechanism. For the theoretical framework, we shall take a "reduced form" approach: Assume that complying with a democratically elected rule adds fixed utility $\bar{u}^B \geq 0$ to $u_i(0)$ or $u_i(1)$, respectively. It then follows that individual i chooses to give iff

$$\Delta u_i(Give) \geq \begin{cases} -\bar{u}^B & \text{under } Rule\text{:}Give, \\ +\bar{u}^B & \text{under } Rule\text{:}Don't. \end{cases}$$

Compared to the case without a code, the share of givers in the population increases or decreases, see Figure 2, panels b) and c). Note that the rule should only affect behavior of those individuals who in the absence of a code would have chosen the opposite action. For instance, a democratically elected code that promotes giving (*Rule:Give*) may convince a *Non-Giver* to give, but will leave the behavior of a "natural" *Giver* unaffected.

How is rule compliance affected by attempts to disenfranchise or manipulate

---

over the distribution of winning probabilities. Saito (2013), for example, offers a model that introduces such preferences.
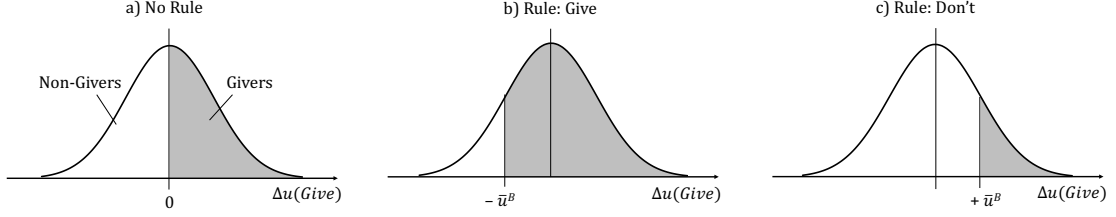
Figure 2: Theory: Share of Givers with and without rules

voters during the election of a code? Again, we take a simple reduced form approach and assume that our interventions lower the utility to follow the elected rule by a constant $\Delta \bar{u}^M > 0$. This is line with both theoretical explanations laid out above: When the elected code does not represent the true preferences of all voters, this might affect the intrinsic motivation of a subject to follow the rule. It will also introduce noise into the signaling process of underlying values. In both cases, malpractice lowers the incentives to follow a given code: Individual i chooses to give iff

$$\Delta u_i(Give) \geq \begin{cases} -(\bar{u}^B - \Delta \bar{u}^M) & \text{under } Rule{:}Give, \\ +(\bar{u}^B - \Delta \bar{u}^M) & \text{under } Rule{:}Don't. \end{cases}$$

First and foremost, manipulating or disenfranchising voters thus leads people to revert back to their individually preferred behavior: As $\Delta \bar{u}^M$ increases, a lower share of *Non-Givers* will follow *Rule:Give*, see Figure 3, panel b). Similarly, a lower share of *Givers* will be willing to follow *Rule:Don't*, Figure 3, panel c). As $\Delta \bar{u}^M$ becomes sufficiently large such that $\bar{u}^M - \Delta \bar{u}^M$ turns negative, people may even turn against rules that match their "natural" 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.
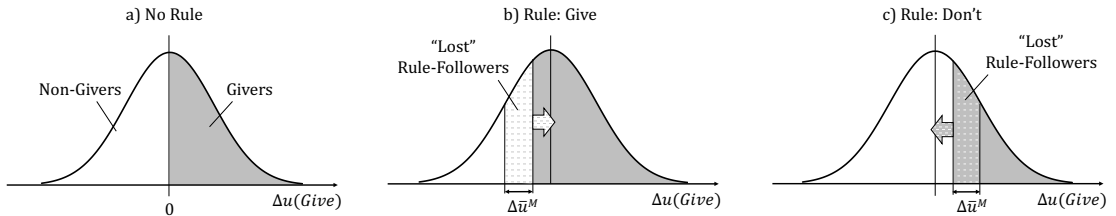
Figure 3: Theory: Effects of Interventions (Electoral Malpractice) on Rule-Compliance

11

2.2 Experimental Procedures

We will now detail all steps of an experimental session. For each treatment, 100 individual subjects are recruited on the online platform Prolific.ac with a small, fixed base payment and the prospect that one of 100 participants will win GBP 100. Before a participant starts the experiment, she receives detailed instructions on how the lottery tickets will be distributed (see Appendix D). Control questions at the end of each screen have to be answered correctly in order to proceed with the experiment.[16] Participants are informed that there are two rounds but they only learn about the referendum that will take place in round 2 after having completed round 1. One round is randomly drawn to determine the final distribution of lottery tickets. All decisions are taken anonymously.
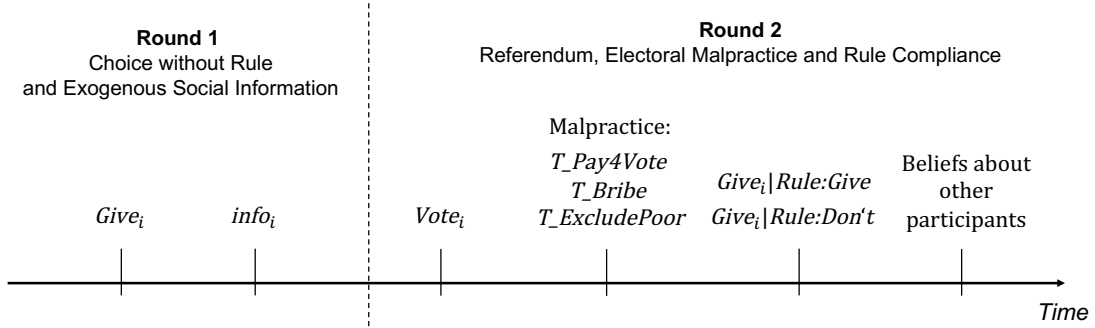
Figure 4: Timeline of Experimental Session

Timeline of Experimental Session. In round 1, each subject plays the dictator game ($Give_i \in \{0,1\}$) individually. After the decision, subjects do not receive feedback about the giving behavior in their cohort. Instead, we show each subject exogenous information on the giving decisions of five participants from an earlier session. An independent random draw determines if a subjects sees a sample where two out of five participants chose to give ($info_i = 2$) or one where four out of five participants chose to give ($info_i = 4$).

Participants then move to round 2, where they are informed that in this round, there will be a code of conduct for behavior in the dictator game. Every subject votes ($Vote_i \in \{0,1\}$) on whether she prefers to have a code of conduct for all 100 subjects that says "give" (*Rule:Give*) or one that says "don't give" (*Rule:Don't*). Treatments vary between subjects and are introduced after the vote. In *T_Pay4Vote*, each participant now decides whether she wants to pay GBP 0.20 to make her vote count. In *T_Bribe*, each participant decides whether she wants to accept GBP

---

[16]We observe the number of times an individual tried to proceed without having answered all questions correctly. The number of such mistakes is generally small and has no explanatory power for our results.

0.20 and reverse her original vote. In *T_Baseline* and *T_ExcludePoor*, subjects are simply informed about the vote aggregation process. Subjects in all treatments are informed that the 99 other participants see the same information, but are *not* informed about the number of votes being excluded our manipulated by these interventions. Following the referendum, each individual i decides whether she wants to $Give_i|Rule{:}Give \in \{0,1\}$ conditional on *Rule:Give* being elected and whether she wants to $Give_i|Rule{:}Don't \in \{0,1\}$ conditional on *Rule:Don't* being elected (strategy method). Round 2 ends with an incentivized elicitation of beliefs about the choices of the other 99 participants in their session. After all participants have finished the experiment, random draws are executed, subjects are matched into pairs and decisions are being implemented. Subjects receive all payments and an e-mail with a summary of the outcomes within two days after the experiment. Figure 4 summarizes the timeline of an experimental session.

Belief Elicitation. In all treatments, at the end of round 2, we ask participants to state their beliefs about how many of the other 99 group members (a) follow *Rule:Give* (b) follow *Rule:Don't* and (c) vote for *Rule:Give*. We incentivize truth telling by letting subjects indicate a bracket (0-9 subjects, 10-19 subjects...,..., 90-99 subjects) and paying them GBP 0.50 for each question where the true number of subjects falls into this bracket (see Schlag and Tremewan, 2016, for a discussion of this method). In *T_Pay4Vote*, *T_Bribe* and *T_ExcludePoor*, we also elicit beliefs about the impact of the intervention on final voting outcomes. In *T_ExcludePoor*, we ask participants to state their belief about the share of votes for *Rule:Give* separately for the high income (income > GBP 40,000) and for the low income participants (income ≤ GBP 40,000). In *T_Pay4Vote* we ask participants to state their beliefs about the share of *Rule:Give*-voters who pay for their vote and, separately, about the share of *Rule:Don't*-voters who pay for their vote. We do the same regarding the beliefs about the share of participants who accept the bribe in *T_Bribe*. Truth telling is incentivized in the same way as before, with subjects now indicating a bracket between 0-9% and 90-99%.

Post-Experimental Questionnaire. We conduct a post-experimental questionnaire to complement the standard background information on subjects we can access via Prolific.ac. The questionnaire is presented as an unrelated survey (using a different visual design and researcher profile) and is send to the same people about two weeks after they participated in the experiment. These measures are meant to minimize the risk of spillovers from decision in the experiment and especially from exposure to the different treatments to questionnaire answers. We ask participants about their standpoints on various political issues such as redistribution, corrup-

tion, democratic values, and personal trust in institutions. Most of the questions are either directly taken or adapted from questions featuring in the 6th wave of the World Value Survey (WVS, 2014). Additionally, we assess personality characteristics such as risk preferences (self-reported and hypothetical lottery choice), trust, and the Big Five personality traits. The questions and answer format (7 point Likert scale) of the very short version of the Big Five are taken from Gosling, Rentfrow and Swann (2003). The full list of questions can be found in the appendix.

2.3 Empirical Strategy

To identify the impact of our interventions ($T\_Pay4Vote$, T_Bribe, or $T\_ExcludePoor$) on compliance, we cannot rely on comparing compliance rates in these treatments with the compliance rate in $T\_Baseline$. Even though treatments are randomly assigned, treatment groups might differ in the ex-ante motivation of the average individual to follow a given rule. This can affect compliance levels and potentially hide or exaggerate treatment effects: Individual i may be more likely to follow a rule in the case that the rule corresponds to her individually preferred behavior or in the case that it corresponds to what she believes is the correct "societal" or "ought" behavior. We identify and control for these two motives by controlling for the type of an individual as indicated by her round 1 choice $Give_i \in \{0, 1\}$ and her $Vote_i \in \{0, 1\}$, indicating her preferred societal rule. Because treatment interventions are introduced after the votes are submitted in round 2, both variables are unbiased by the interventions. This identification is very close to the approach suggested by Dal Bó, Foster and Putterman (2010). Similar to them, we can estimate treatment effects on the type-level by conditioning on $Give_i \in \{0, 1\}$, $Vote_i \in \{0, 1\}$, or both. We go one step further and use the distribution of types in our experimental sample to estimate *average* treatment effects on the population level. Because there is no punishment associated with violating a rule, residual treatment differences measure to what extend the willingness to follow rules depends on the election process.

2.4 Implementation

The experiment is implemented online using a subject pool of (non-representative) international participants on the platform Prolific.ac based in Oxford, UK.[17] We programed the experiment using the software *LimeSurvey*, screenshots can be found in the appendix. All sessions were run in February and March 2017 on Tuesday, Wednesday or Thursday afternoons in order to keep the external circumstances

---

[17]https://prolific.ac

as similar as possible between treatments. Registered participants have a unique Prolific-ID that is used to identify subjects, to prevent repeated participation and to process payments. 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 1.60 for completing the study.[18] With each session taking roughly 15 minutes to complete, this base payment translates into an hourly wage of GBP 6.40. Additional payments are announced during the course of the experiment. For completing the 10 minute post-experimental questionnaire, subjects receive a compensation of GBP 1. The follow-up-rate is close to 100 per cent.[19] In addition, subjects' unique Prolific-ID allows us to access an extensive set of self-reported socio-demographic data including gender, nationality and income (see table 1). All information is provided voluntarily by the subjects but we required that only those who had filled out information on their gender and nationality were eligible for our study. For treatment *T_ExcludePoor* we also required that participants had filled out information on their annual household income (to make our intervention possible). To have a balanced sample in this particular treatment, we invited 50 participants with a stated household income above GBP 40,000—whose vote is counted in the election—and 50 participants with a stated household weakly below GBP 40,000—whose vote is *not* counted.[20] Table 1 shows a summary of sample demographics. With a mean age of 31, almost two thirds of the participants not being students and about one third having a non-Western nationality, our population sample differs in several respects from the typical subject pool at Western university labs.

	Age	Female	Western	Student	Unemployed	UGrad	Inc < 40K
Mean	31	0.42	.68	.36	.17	.58	.61
Std.Dev.	10.7						
Observations	394	400	400	400	400	390	321

Table 1: Participant Demographics. *Western* = 1 if Nationality is Western Europe, Australia, Canada, New Zealand, US. *Student* = 1 if participant is student at the moment of taking part. *UGrad* = 1 if highest education is at least undergraduate degree (BA/BSc/other). *Inc < 40K* if self-reported yearly household income is below GBP 40,000.

---

[18]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.

[19]Of 400 subjects, 387 filled out the questionnaire.

[20]Individuals registered on *Prolific.ac* can access a list of active studies for which they are eligible and can participate in. They are *not* informed about the criteria used to pre-select "eligible" participants. For example, in treatment *T_ExcludePoor*, they do *not* know that eligibility is based on stated household income.

3 Treatment Effects

To set the stage for the analysis of treatment effects, we begin by providing summary statistics of choices that precede the compliance decisions of subjects as well as of the impact of our interventions on the voting outcome. We also provide an overview of subjects' beliefs about the behavior of other participants in their group. The overall giving rate in round 1—that is, the share of subjects choosing $Give_i = 1$—is 61% (245/400).[21] Almost all of those who choose to give in round 1 also vote for *Rule:Give* in the beginning of round 2 (93%). Among those who do not give in round 1, a significant majority of 59% vote for *Rule:Don't*. Overall, 73% vote in favor of *Rule:Give*, making it the preferred rule in every session. 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_Bribe* accept to reverse their vote for the payment, and, by design, 50% of voters are excluded due to a low household income in *T_ExcludePoor*, see also Figure 5. Intuitively, excluding a substantial fraction of voters can affect the voting outcome. We measure "outcome bias" as the absolute value of the difference between the share of votes for *Rule:Give* before and after the intervention. While a large share of participants lose their voice, this has a relatively small effect on the voting outcome, see the right panel of Figure 5. In *T_Pay4Vote* the bias is in favor of *Rule:Give* (+5 percentage points), while in *T_Bribe* and *T_ExcludePoor* the bias is in favor of *Rule:Don't* (+11 and +3 percentage points, respectively). Beliefs about the impact of the treatment intervention (elicited at the end of the experiment) show that the large majority of subjects expected the interventions to lead to a considerable bias in the voting outcome (right panel of Figure 5).

Figure 6 shows the distributions of subjects' beliefs about the voting behavior and rule compliance of other participants in their session. From the histograms in the top panels we can see that beliefs are very heterogeneous. The median answer bracket regarding the question of how many of the other 99 participants voted for *Rule:Give* (panel a) is 50-59. This and the observation that the number of subjects stating extreme beliefs (0-9 or 90-99) is small gives us confidence that most subjects believed each of the two rules to have positive probability of being selected in the referendum. On average, subjects expect more people to comply with *Rule:Don't* (panel c) than with *Rule:Give* (panel b). The bottom graphs (cumulative densities)

---

[21]Note that our dictator game version differs in many respects from standard implementations of the game, namely by having ex-ante choices with role uncertainty, binary decisions, risky prospects with a small probability to win a high price, and by having an online participant sample. Still, the observation that 61% of subjects chose to give tickets does not deviate much from previous findings 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.
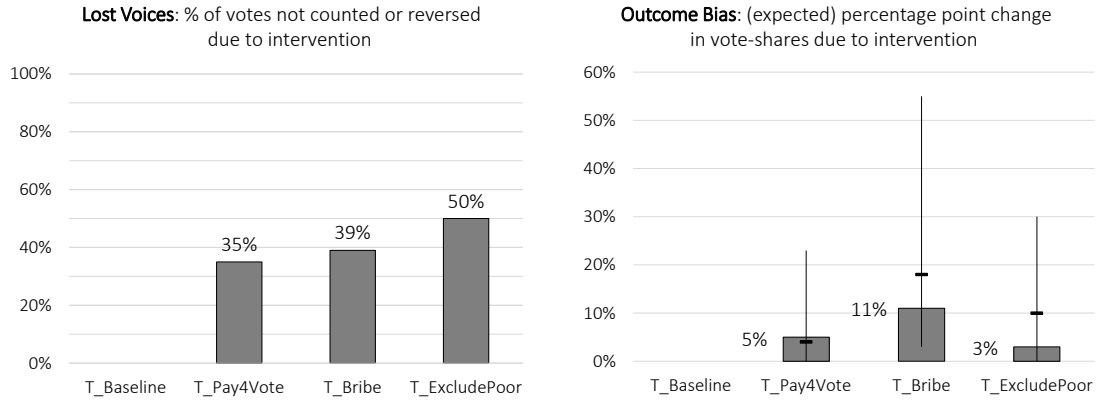
Figure 5: Left panel: Share of votes not counted or reversed in each treatment. Right panel: Outcome bias (absolute difference in the share of votes for *Rule:Give* before and after the intervention) in percentage points. The figure shows the actual outcome bias (as bars) as well as the distribution of subjects' beliefs about the outcome bias (median and 10th-90th percentile).
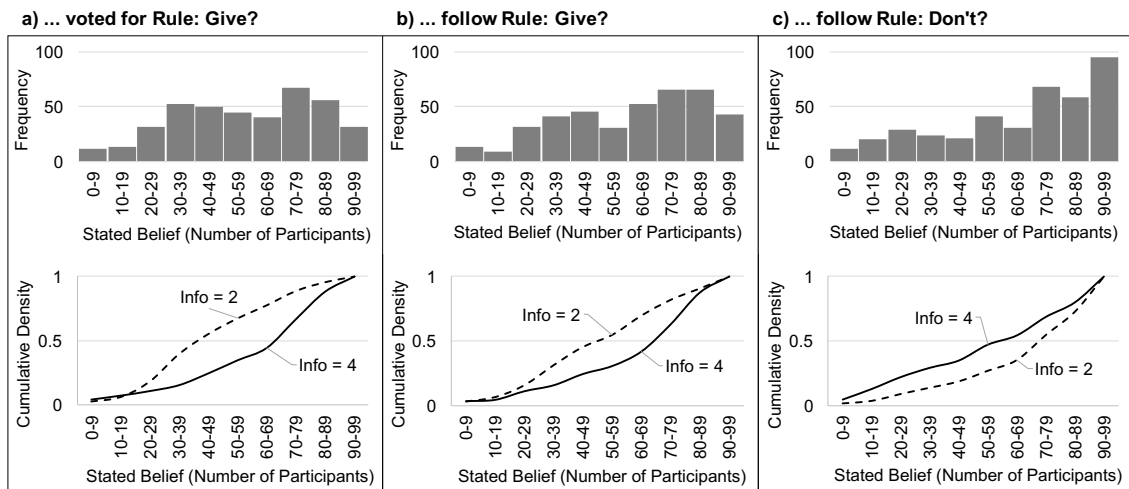
Figure 6: Beliefs about the choices of other participants (data from all treatments pooled, N=400). Top: Frequency of beliefs by answer bracket. Bottom: Cumulative density of answers among subjects having received *info*= 2 and *info*= 4, respectively.

show that our information treatment was successful in shifting beliefs regarding the number of *Givers* in their group: among subjects who received the information that four out of five subjects in an earlier study chose to give (*info*= 4), beliefs about the number of participants voting for *Rule:Give* (panel a) and following *Rule:Give* (panel b) are consistently higher than among those subjects who received the information that only two out of five subjects chose to give (*info*= 2). They also believe that less people choose to follow *Rule:Don't* (panel c).

3.1 Rule Compliance and Treatment Effects

Figure 7 delivers a first impression of the levels of rule-compliance with and without malpractice. The figure shows results separately for subjects who chose to *not* give in round 1 (*Non-Givers*, panel a) and those who chose to give in round 1 (*Givers*, panel b). Bar charts at the top of the figure depict compliance rates in the baseline treatment (*T_Baseline*). Here, we observe very high compliance rates: Almost every subject (98% of *Non-Givers* and 93% of *Givers*) follows the rule that prescribes the action that she preferred in round 1. More importantly, a significant fraction of subjects also follows the opposite rule: 65% of *Non-Givers* decide to follow rule *Rule:Give* and 53% of *Givers* decide to follow *Rule:Don't*. These numbers confirm a basic prediction of our model, namely that a democratically elected rule is voluntarily followed by more than just the original proponents of the action. As a consequence, overall giving rates in the baseline treatment react strongly to rules. The share of subjects who give increases from 57% in round 1 to 81% under *Rule:Give* and drops to only 28% under *Rule:Don't*.

Result 1 (Rule-Compliance without Malpractice). *In the absence of electoral malpractice, democratically elected rules have strong influence on voluntary behavior: Conditional on Rule:Give (Rule:Don't) being elected, 81% (72%) of subjects in T_Baseline voluntarily comply. 54% of subjects in T_Baseline are "rule-followers" who comply with either rule given its election.*
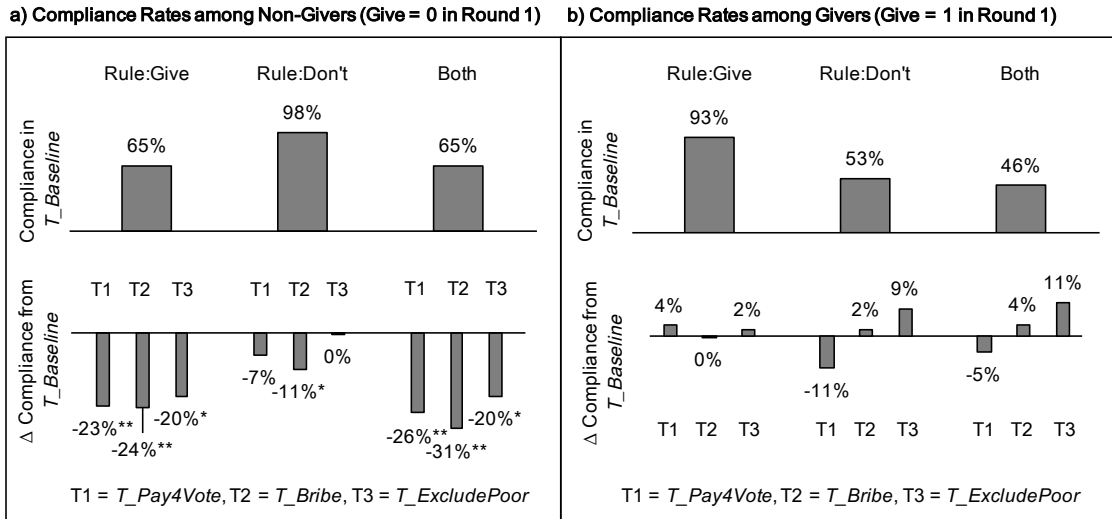
Figure 7: Compliance rates among a) *Non-Givers* (left panel) and b) *Givers* (right panel). *Both* = 1 if $Give_i | Rule:Give = 1$ and $Give_i | Rule:Don't = 0$. Top: Compliance rates in *T_Baseline*. Bottom: Percentage point change in compliance rates (Δ Compliance) from *T_Baseline*: T1 = *T_Pay4Vote*, T2 = *T_Bribe*, T3 = *T_ExcludePoor*. Stars denote significance level of one-sided Fisher-exact tests (H_1: Δ Compliance ¡ 0): $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

18

The bottom graphs in Figure 7 show percentage point differences between compliance rates in *T_Baseline* and compliance rates in each of the treatments involving electoral manipulation. We immediately see strong and significant treatment effects among subjects whose individual choice in round 1 was to *not* give (*Non-Givers*, panel a): Of them, roughly 20-25 percent less can be convinced to follow *Rule:Give* if this rule is elected in the presence of a voting fee (*T_Pay4Vote*), monetary offers to vote differently (*T_Bribe*), or without the participation of low-income voters (*T_ExcludePoor*). The share of *Non-Givers* who can be identified as rule-compliers—those who voluntary comply with either rule, if elected—drops from 65% without malpractice to only 34–45%. These responses are in line with our prediction that the manipulation of election processes lowers the utility to follow elected rules and thus diminishes voluntary rule-compliance. Maybe surprisingly, we find no evidence for such treatment effects being present among *Givers* (panel b): It seems that compliance with *Rule:Don't*—the rule we were expecting to see a deterioration in compliance among subjects who indicated a preference to give in round 1—is not affected by concerns about electoral manipulation.

To yield a deeper understanding of treatment differences and in order to calculate population average treatment effects, we classify subjects by

$$Type_i = Give_i \text{ (Round 1) } \times Vote_i \in \{(0,0),(0,1),(1,0),(1,1)\}$$

and estimate effects of electoral malpractice for each type separately using OLS regressions. We present results from this approach in Table 2:[22]

Panel a) reports the number of subjects of each type in the experimental population. Panel b) reports baseline compliance rates (the share of compliant subjects in *T_Baseline*) conditional on *Rule:Give* being elected (left-hand side) and conditional on *Rule:Don't* being elected (right-hand side). Panel c) reports treatment effects: It shows estimates of the change in compliance rates when going from *T_Baseline* to a treatment with electoral malpractice. Here, we first report separate treatment effects for each of the three malpractice treatments (*T_Pay4Vote*, *T_Bribe*, and *T_ExcludePoor*). In the lowermost section of panel c) we then report a "generalized" malpractice effect by pooling these data.

White cells in Table 2 panel c) show how malpractice affects the compliance

---

[22]We discussed the necessity to control for $Give_i$ (Round 1) $\in \{0,1\}$ and $Vote_i \in \{0,1\}$ in the identification section 2.3. In Table 2 we also control for possible effects of exogenous information $info_i \in \{2,4\}$. Controlling for $info_i$ avoids sampling bias when running estimations on the smaller samples defined by types: Figure 6 shows that $info_i$ influences beliefs about the share of *Givers* in the population. Via this belief channel, the information treatment might influence compliance decisions. Although this is not a cause of concern in large samples—given that $info_i$ is individual randomly drawn from a uniform distribution—, deviations from uniformity in smaller samples might bias the estimates of treatment effects.

(a) Population by Type:

	By Vote$_i$	By Give$_i$ (Round 1)		
		0	1	all
All Treatm.	0	92	17	109
	1	63	228	291
	all	155	245	400

(b) Compliance Rates in the Baseline:

	Share of subjects complying with					
		Rule: Give			Rule: Don't		
	By Vote$_i$	By Give$_i$ (Round 1)			By Give$_i$ (Round 1)		
		0	1	all	0	1	all
T_Baseline	0	.57	.50	.56	.96	.63	.89
	1	.80	1	.95	1	.51	.63
	all	.65	.93	.81	.98	.53	.72

(c) Treatment Effects on Compliance Rates (vs. Baseline):

		Rule: Give			Rule: Don't		
	By Vote$_i$	By Give$_i$ (Round 1)			By Give$_i$ (Round 1)		
		0	1	all	0	1	all
T_Pay4Vote	0	-.15	.59	-.04	-.05	-.63	-.15*
		(.14)	(.)	(.13)	(.07)	(.)	(.08)
	1	-.35**	-.04	-.11**	-.09	-.07	-.08
		(.16)	(.03)	(.04)	(.08)	(.10)	(.08)
	all	-.24**	.01	-.09*	-.06	-.11	-.09
		(.11)	(.04)	(.05)	(.05)	(.09)	(.06)
T_Bribe	0	.00	-.02	.00	-.09	.28	-.03
		(.15)	(.)	(.14)	(.08)	(.)	(.08)
	1	-.57***	-.04	-.16***	-.16*	.04	-.01
		(.18)	(.03)	(.05)	(.09)	(.09)	(.08)
	all	-.23*	-.04	-.11**	-.12**	.05	-.01
		(.12)	(.04)	(.05)	(.06)	(.09)	(.06)
T_ExcludePoor	0	-.16	.17	-.11	-.01	.27	.03
		(.14)	(.)	(.13)	(.07)	(.)	(.07)
	1	-.33*	-.02	-.09*	.00	.10	.08
		(.18)	(.03)	(.05)	(.09)	(.10)	(.08)
	all	-.23**	.00	-.09*	-.01	.12	.07
		(.11)	(.04)	(.05)	(.05)	(.09)	(.06)
Pooled	0	-.12	.21	-.07	-.04	.02	-.03
		(.11)	(.)	(.10)	(.06)	(.)	(.06)
	1	-.41***	-.03	-.11***	-.08	.02	.00
		(.14)	(.03)	(.04)	(.07)	(.08)	(.07)
	all	-.23***	-.02	-.10***	-.06	.02	-.01
		(.09)	(.03)	(.04)	(.04)	(.08)	(.05)

Standard errors in parentheses. $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

Table 2: Number of subjects (a), baseline compliance rates (b) and treatment effects (c) by $Type_i = Give_i$ (Round 1) $\times Vote_i$ as well as average treatment effects for the entire population. White cells in (c) show coefficients and standard errors from OLS regressions of binary treatment variables on the compliance of types to *Rule:Give* ($Give_i|Rule:Give = 1$) and *Rule:Don't* ($Give_i|Rule:Don't = 0$), respectively, controlling for $info_i$. Grey cells show estimates of average treatment effects when types are weighted by population shares according to table (a).

of each type. For instance, the first four cells in the top-left corner of panel c) report the effects of implementing a voting fee on compliance with *Rule:Give* (*T_Pay4Vote*): Compliance drops by 15 percentage points among *Non-Givers* who voted for *Rule:Don't*, by 35 percentage points ($p < 0.05$) among *Non-Givers* who voted for *Rule:Give* and by 4 percentage points among *Givers* who voted for *Rule:Give*. Only among the $n = 3$ *Givers* in *T_Pay4Vote* who voted for *Rule:Don't*

we measure a positive (and clearly, insignificant) effect.[23] To arrive at population average treatment effects, which are reported in the grey cells of the same panel, we weight types by their share in the experimental population. For example, we calculate the population average treatment effect of bribing voters (*T_Bribe, Rule:Give*) as $(92/400) \cdot (.00) + (63/400) \cdot (-.57) + (17/400) \cdot (-.02) + (228/400) \cdot (-.04) = -.11^{**}$. Standard errors for weighted averages are calculated using the Delta method.[24]

Overall, Table 2 reinforces the impression from Figure 7: Electoral malpractice significantly affects compliance with rules promoting redistribution (*Rule:Give*), but seems to have little impact on compliance with rules opposing it (*Rule:Don't*). Treatment differences for *Rule:Don't* are small and (mostly) insignificant across all types. When pooling malpractice treatments (panel c, lowermost section), the population average treatment effect on compliance with *Rule:Don't* is estimated to be basically zero (-0.01, $p = 0.87$). In contrast, apart from type $(Give_i, Vote_i) = (1,0)$—who only constitute 4% of the population—all types consistently show (weakly) lower compliance with *Rule:Give* if the vote aggregation process is manipulated in one way or the other. Compliance of subjects who did not give in round 1 but indicated a preference for *Rule:Give*—that is, compliance of type $(Give_i, Vote_i) = (0,1)$—is most volatile to whether the group selects this rule by democratic means: Among these participants, the share of subjects who follow *Rule:Give* drops by 35 percentage points in *T_Pay4Vote*, 57 percentage points in *T_Bribe* and 33 percentage points in *T_ExcludePoor*. Across all subjects who did not give in round 1, treatment effects closely match the effects displayed in Figure 7 (-24, -23, and -23 percentage points, respectively). Weighting these types in the total population we estimate that all three forms of electoral malpractice significantly reduce the overall share of individuals complying with *Rule:Give* by roughly 10 percentage points ($p < 0.1$, $p < 0.05$). Note that all three treatments show very similar effects on compliance rates, both on the type- and the aggregate level. Pooling the data (panel c, lowermost section), treatment effects for *Rule:Give* are significant at the 1 percent level.

Our analysis suggests that what is losing out under malpractice is the (non-coercive) power of a democratic vote to change individual behavior. A different way to look at the results is to make this loss in power explicit. Figure 8 shows the average difference between an individual's choice to give conditional on *Rule:Give* (*Rule:Don't*) being elected (round 2) and her choice before the referendum (round 1)—that is, the average of

---

[23]We do not report standard errors or significance levels for *Givers* who vote for *Rule:Don't* due to the tiny sample sizes. For the same reason we do not attempt to interpret their behavior.

[24]For example, the standard error for the average treatment effect we just calculated can be determined from $\sqrt{(92/400)^2 \cdot (.15)^2 + (63/400)^2 \cdot (.18)^2 + (17/400)^2 \cdot (.37)^2 + (228/400)^2 \cdot (.03)^2} = .05$
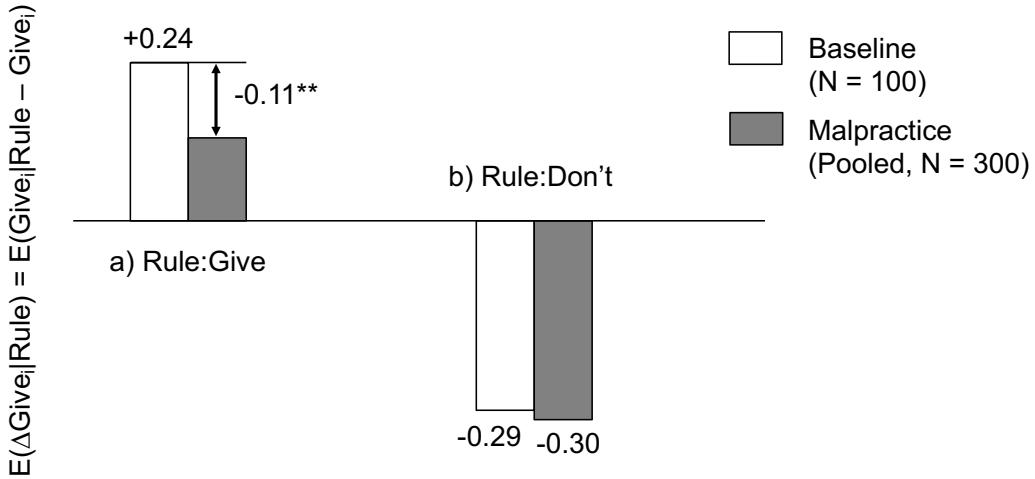
Figure 8: Power of the democratic vote to change individual behavior. Left-hand side (panel a): Average of $\Delta_i(Give|Rule:Give) := Give_i|Rule:Give - Give_i$. Right-hand side (panel b): Average of $\Delta_i(Give|Rule:Don't) := Give_i|Rule:Don't - Give_i$. Stars denote significance level of the coefficient on a binary treatment variable for malpractice ($= 1$ if individual i is in treatment *T_Pay4Vote*, *T_Bribe* or *T_ExcludePoor*) in a univariate OLS regression on $\Delta_i Give|Rule:Give$ (=Difference-in-Differences estimator). $^{**}p < 0.05$

$\Delta_i Give|Rule:Give := Give_i|Rule:Give - Give_i$ (Round 1) (on the left-hand side), and the average of $\Delta_i Give|Rule:Don't := Give_i|Rule:Don't - Give_i$ (Round 1) (on the right-hand side), respectively. If the democratic vote has power, one would expect *Rule:Give* to increase giving rates ($E(\Delta_i Give|Rule:Give) > 0$) and, conversely, *Rule:Don't* to decrease giving rates ($E(\Delta_i Give|Rule:Give) < 0$). This is also what we observe in the data. Consistent with our previous analysis, manipulations of the electoral process do not affect the power of *Rule:Don't*. *Rule:Give*, on the other hand, looses roughly half of its power to positively affect behavior. We summarize our findings regarding treatment effects below.

Result 2 (Main Result: Treatment Effects). *The manipulation of electoral processes significantly lowers voluntary compliance with Rule:Give. Of subjects who did not give before the election, on average 23 percent less ($p < 0.01$) can be convinced to follow Rule:Give in the presence of a voting fee (T_Pay4Vote), monetary offers to vote differently (T_Bribe), or without the participation of low-income voters (T_ExcludePoor). This translates into a 10 percentage points reduction of the compliance rate in the total population ($p < 0.01$) and is equivalent to the rule loosing roughly half of its non-coercive power to change individual behavior. We find no evidence of electoral manipulation affecting compliance with Rule:Don't.*

4 Understanding Treatment Effects

What drives the strong adverse treatment effect on voluntary compliance with *Rule:Give*? Why is compliance with *Rule:Don't* not affected by manipulations of the electoral process? In this section, we will try to better understand the psychological determinants of rule compliance by analyzing the role of beliefs in driving behavior. In addition, we will exploit variance in the individual effects of the treatment interventions as well as information we obtained from the questionnaire about subject characteristics to account for individual heterogeneity and thus, better understand the behavioral pattern.

4.1 Beliefs about the Behavior of Other Subjects

We observe that rules have strong influence on voluntary behavior (see, for example, Figure 7). Do people follow rules because they want to follow others? Can this explain the treatment effects? Visually comparing the distribution of individual beliefs about the behavior of other participants in treatment *T_Baseline* with the respective distributions in treatments *T_Pay4Vote*, *T_Bribe* and *T_ExcludePoor*, we do not observe systematic differences.[25]

Confirming this are the results of two-sample Kolmogorov-Smirnov tests which can also not reject equality of these distributions. This makes beliefs about others an unlikely candidate to explain treatment differences. 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 explain the overall pattern of choices observed in the experiment.

Table 3 presents the results of an instrumental variable approach to estimating the role of others in guiding behavior under *Rule:Give* (panel a) and *Rule:Don't* (panel b). The main variable of interest in this analysis is $E_i(Comply_{-i}|Rule)$, which is the share of the 99 other participants whom individual i believes to be complying with *Rule:Give* or *Rule:Don't*, respectively.[26] Because $E_i(Comply_{-i}|Rule)$ might be endogenous in a regression on $Give_i|Rule$, we instrument it with the binary variable $1.[info_i = 4]$. As Figure 6 shows, $info_i$ on average has a strong effect on $E_i(Comply_{-i}|Rule)$. Because it is exogenously randomized, it is a valid instrument.

---

[25]Figure 6 plots the distribution of these beliefs when pooling all four treatments. Beliefs in each individual treatment follow very much the same distribution.

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

Table 3 is structured as follows. Columns (1) present results of an OLS regression of $1.[info_i = 4]$, a dummy for malpractice,[27] and type controls $Give_i \times Vote_i$ on $E_i(Comply_{-i}|Rule:Give)$ (panel a) and $E_i(Comply_{-i}|Rule:Don't)$ (panel b), respectively. The small and insignificant coefficients on malpractice are in line with the Kolomogorov-Smirnov tests indicating that treatments did *not* systematically alter beliefs about the behavior of other subjects. At the same time, the large and highly significant coefficients on $1.[info_i = 4]$ confirm the observation from Figure 6: Going from $info_i = 2$ to $info_i = 4$ increases (decreases) an individual's belief about the share of participants complying with *Rule:Give* (*Rule:Don't*) on average by 13 percentage points ($p < 0.01$). Variable $info_i$ is thus a powerful instrument to assess the causal effect of beliefs about the behavior of others on choices under both rules. Columns (2) report results of an OLS regression using the same explanatory variables on compliance with *Rule:Give* (panel a) and *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. However, due to possible endogeneity, this correlation does not imply causality. For this reason, in columns (3), we use an IV (2SLS) estimator. Using $1.[info_i = 4]$ as an instrument for $E_i(Comply_{-i}|Rule)$, we find strong evidence that beliefs about the behavior of others causally explain compliance with *Rule:Don't* (panel b). Specifically, a 1 percentage point increase in the expected share of others who comply is estimated to increase the probability of individual i to also comply and not give by 0.87 percentage points ($p < 0.01$). Accounting for this effect, no other explanatory variable is significant at the 5 percent level. Maybe surprisingly, we find no evidence that compliance with *Rule:Give* (panel a) is driven by similar motivations: $E_i(Comply_{-i})$ is insignificant for compliance with *Rule:Give* at any reasonable confidence level. Most importantly, irrespective of whether we control for beliefs about the behavior of others directly (column 2) or via instrument $info_i$ (column 3), malpractice is identified to have virtually the same effect on rule-compliance as before, that is, reducing compliance with *Rule:Give* by approximately 10 percentage points in the total population while having no significant effect on compliance with *Rule:Don't*. These results imply that the drop in voluntary compliance with *Rule:Give* which we observe in the presence of electoral manipulation (*T_Pay4Vote*, *T_Bribe* or *T_ExcludePoor*) is not mediated by mean-variance shifts of beliefs about the behavior of others. On this hand, our results speak against a signaling theory of legitimacy. Rather, manipulations of electoral processes seem to directly impact the intrinsic motivation of individuals to follow *Rule:Give*. The

---

[27]*Malpractice = 1 if individual i is a subject in treatment T_Pay4Vote, T_Bribe or T_ExcludePoor.*

(a) Rule:Give

| | $E_i(Comply_{-i})$ | $Comply_i = 1$ | | | |
	(1) OLS	(2) OLS	(3) IV	(4) OLS	(5) OLS
$info_i = 4$	.13*** (.02)			-.04 (.04)	-.04 (.04)
$E_i(Comply_{-i})$		.46*** (.07)	-.32 (.30)		
$Malpractice = 1$	-.02 (.03)	-.10*** (.04)	-.11** (.05)	-.11*** (.04)	-.10** (.04)
$Female_i$					.08* (.04)
$Risk\_Seeking_i$					.02* (.01)
$Betrayal\_Aversion_i$					.00 (.01)
Constant	Yes	Yes	Yes	Yes	Yes
Control for $Type_i$	Yes	Yes	Yes	Yes	Yes
Add. Controls					Yes
Observations	400	400	400	400	375

(b) Rule:Don't

| | $E_i(Comply_{-i})$ | $Comply_i = 1$ | | | |
	(1) OLS	(2) OLS	(3) IV	(4) OLS	(5) OLS
$info_i = 4$	-.13*** (.03)			-.11*** (.04)	-.09** (.04)
$E_i(Comply_{-i})$		.51*** (.07)	.87*** (.33)		
$Malpractice = 1$	.01 (.03)	-.02 (.05)	-.02 (.05)	-.01 (.05)	-.05 (.05)
$Female_i$					.00 (.05)
$Risk\_Seeking_i$					.00 (.01)
$Betrayal\_Aversion_i$					.04*** (.01)
Constant	Yes	Yes	Yes	Yes	Yes
Control for $Type_i$	Yes	Yes	Yes	Yes	Yes
Add. Controls					Yes
Observations	400	400	400	400	375

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

Table 3: 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_Bribe or $T\_ExcludePoor$. IV regressions are 2SLS with $E_i(Comply_{-i})$ being instrumented by $1.[info_i = 4]$.

analysis of *Rule:Don't* shows, on the other hand, that concerns regarding the process of rule selection may not necessarily be the prime drivers of compliance with *any* type of rule. Here, in stark comparison to *Rule:Give*, a strategic motivation to follow the behavior others is the dominant explanation. Given that beliefs about the behavior of other subjects do not vary significantly between treatments, this observation goes some way in explaining why malpractice does not significantly affect the share of subjects following *Rule:Don't*.

Columns (4) and (5) of Table 3 underline the robustness of our findings by presenting variations on the same scheme. Columns (4) present results 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 dependency between individual behavior and beliefs about the share of pro-social agents in the population—which are shifted by $info_i \in \{2, 4\}$—instead of specifically controlling for strategic complementarity in compliance. Columns (5) extend this analysis by including an extensive battery of individual characteristics and questionnaire answers as controls.[28] In both cases, our findings—in particular, regarding the effects of electoral manipulation (reflected in the coefficient on *Malpractice*) and the role of others in guiding behavior (now reflected in the coefficient on $info_i$)—are unchanged. We summarize our results below.

Result 3 (Beliefs about the Behavior of Other Subjects)**.** *Beliefs about the behavior of other subjects causally explain voluntary compliance with Rule:Don't: A 1 percentage point increase in $E_i(Comply_{-i})$ increases the probability of the average subject to also comply with Rule:Don't by 0.87 percentage points ($p < 0.01$). We find no evidence of beliefs about others causally affecting voluntary compliance with Rule:Give. In particular, variance in the beliefs about other subjects cannot explain the observed adverse effects of electoral malpractice (T_Pay4Vote, T_Bribe, T_ExcludePoor) on compliance rates: Treatment effects are likely to be driven by a loss in the intrinsic motivation of individuals to follow the rule.*

---

[28]$Risk\_Seeking_i$ is questionnaire-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)?". $Betrayal\_Aversion_i$ is questionnaire-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)?". Control for $Type_i$ includes $Give_i$ (Round 1), $Vote_i$, and $Give_i$ (Round 1) $\times Vote_i$. Additional controls in (5) are: $Western_i$, $Student_i$, $UGrad_i$, number of mistakes in control questions, factor variables measuring political and social values in questionnaire, as well as *Big Five* personality test measures on 7-point Likert scales. All controls not shown in the table are estimated to have small, insignificant effects ($p > 0.1$).

4.2 Individual Disenfranchisement and Beliefs about the Outcome Bias

While treatments *T_Pay4Vote*, *T_Bribe* and *T_ExcludePoor* differ in the particular form of electoral malpractice, they have in common that due to the intervention (a) many individuals lose their voice in the decision making process and (b) many individuals believe that the outcome of the referendum is biased compared to a fair majority vote (see Figure 5). Could it be that these two effects—being *personally* disenfranchised in the election and having doubts about the referendum's *overall* representativeness—are driving the loss in intrinsic motivation to follow *Rule:Give*?

Let

$$
Lost\_Voice_i = \begin{cases} 1 & \text{if } i \text{ is in } T\_Pay4Vote \text{ and } Accept\_Pay_i = 0 \\ 1 & \text{if } i \text{ is in } T\_Bribe \text{ and } Accept\_Bribe_i = 1 \\ 1 & \text{if } i \text{ is in } T\_ExcludePoor \text{ and } Income_i < 40\text{K} \\ 0 & \text{otherwise.} \end{cases}
$$

Also, let $E_i[Outcome\_Bias]$ be the belief of individual i about the absolute size of the outcome bias.[29] As shown in Figure 5, there is substantial heterogeneity between subjects regarding these two variables *within* each treatment. In Table 4 we test whether this variance captures the variance in compliance with *Rule:Give* that we observe between treatments.

The table presents results from OLS regressions of treatment dummies and controls on $Give_i|Rule:Give$, to which we successively add $Lost\_Voice_i$ and $E_i[Outcome\_Bias]$ as additional explanatory variables. Column (1) repeats our main finding that all three forms of malpractice (*T_Pay4Vote*, *T_Bribe*, and *T_ExcludePoor*) significantly reduce compliance with *Rule:Give*. Column (2) adds $Lost\_Voice_i$ as an explanatory variable, column (3) adds $E_i[Outcome\_Bias]$ as an explanatory variable, and column (4) adds both. Table 4 suggests that, indeed, (a) the experience of having one's voice not being counted in the referendum and (b)

---

[29] *Outcome_Bias* is defined as the absolute difference between the share of votes for *Rule:Give* when counting the original votes of all 100 subjects (before the intervention) and the share of votes for *Rule:Give* that are finally counted in the referendum (after the intervention). The belief about the size of this bias is calculated from elicited beliefs with the following formula:

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

	$Comply_i \| \textbf{\textit{Rule:Give}} = 1$			
	(1) OLS	(2) OLS	(3) OLS	(4) OLS
$Lost\_Voice_i = 1$		-.11** (.04)		-.10** (.04)
$E_i[Outcome\_Bias]$			-.34*** (.12)	-.33*** (.12)
$T\_Pay4Vote$	-.11** (.05)	-.07 (.05)	-.08* (.05)	-.05 (.05)
T_Bribe	-.12** (.05)	-.08 (.05)	-.04 (.06)	.00 (.06)
$T\_ExcludePoor$	-.09* (.05)	-.04 (.05)	-.06 (.05)	-.01 (.05)
Constant	Yes	Yes	Yes	Yes
Add. Controls	Yes	Yes	Yes	Yes
Observations	400	400	400	400

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

Table 4: Explaining treatment variance in *Rule:Give* by variance in $Lost\_Voice_i$ ($= 1$ if individual i's original vote is not counted in the referendum) and $E_i[Outcome\_Bias] \in [0, 1]$ (individual i's subjective belief about absolute size of the outcome bias). OLS estimates. Regression includes constant and the following controls: $Give_i$ (Round 1), $Vote_i$, $Give_i$ (Round 1) $\times Vote_i$ and $info_i$.

doubts about the overall representativeness of the election may be the underlying cause for the loss in intrinsic motivation: Including either of the two in the regression leads to a strong reduction in the size and significance of treatment effects. Including both in the regression basically wipes out the treatment effects observed for *T_Bribe* and *T_ExcludePoor*. Only a small but insignificant effect remains for *T_Pay4Vote*.

Result 4 (Individual Disenfranchisement and Beliefs about the Outcome Bias). *Variance in $Lost\_Voice_i$ and $E_i[Outcome\_Bias]$ explains the variance between treatments: The experience of personally being disenfranchised in the election and having doubts about the referendum's overall representativeness may be underlying the loss in intrinsic motivation to follow Rule:Give that is observed in treatments T_Pay4Vote, T_Bribe, and T_ExcludePoor.*

4.3 Experience and Valuation of Democracy

Table 5 shows treatment effects separately for (1) subjects of western and non-western nationality, (2) subjects who state a high importance of living in a democratic country and those who do not, (3) subjects who claim to always participate in elections and those who do not, and (4/5) subjects who indicate a low justifiabil-

ity for bribes and lobbying activities in the political sphere and those who do not. Information on nationality is provided to us by the survey platform (prolific.ac). Data for the separation in Columns (2) to (5) comes from our questionnaire.

Table 5 suggests that our treatments may have affected a psychological domain that is associated with judgements of real world institutions: Significant treatment effects are found only among individuals who are likely to live in established democracies (column 1), who value democratic institutions (columns 2-3) and who strongly condemn violations of democratic principles (columns 3-4). Column (4) provides maybe the strongest support for this claim: Those who indicate a very high sensitivity to bribery in the real world also react very sensitively to electoral malpractice in our experiment. Those who find the acceptance of bribes in the course of one's duties at least sometimes acceptable, on the other hand, show only small and insignificant responses.

Result 5 (Experience and valuation of democracy)**.** *The adverse effect of malpractice on compliance with Rule:Give is strong and significant only (1) among subjects who have a Western nationality, (2) among subjects who self-identify to value democratic institutions highly and (3) among subjects who indicate a low justifiability for bribes and (political) lobbying in the real world.*

5 Conclusion

We have presented the results of an online experiment that allows us to causally estimate how the introduction of a voting fee, monetary incentives to change voting behavior or the exclusion of poor voters from the ballot affect compliance with elected rules of behavior in a dictator game. Our results show that such attempts at manipulating a democratic voting process can have strong and significant adverse effects on the willingness of people to follow rules promoting redistribution (*Rule:Give*). We conclude that electoral malpractices, which are prevalent in many countries around the world, may undermine the positive effects of democracy on behavior that earlier research in public economics has established (see, for example, Frey, 1997; Tyran and Feld, 2006; Ertan, Page and Putterman, 2009; Sutter, Haigner and Kocher, 2010; Dal Bó, Foster and Putterman, 2010). Additional to this main result, our experiment provides insights into the psychological patterns underlying treatment effects and compliance behavior. We show that in our experiment, the adverse effects of vote buying and partial disenfranchisement on compliance cannot be explained by variance in beliefs about other participants' behavior. Rather, subjects seem to react intrinsically to violations of inclusiveness and unbiasedness in democratic elections. This connects to earlier literature in psychology and behav-

Table 5: Treatment effect heterogeneity by nationality and questionnaire responses to the following questions: (2) "How important is it for you to live in a country that is governed democratically?" (High = 10, Low = 1-9); (3) "When elections take place, do you vote always, usually, or never?" (Yes = always, No = other); (4,5) "Please indicate to what extent you think the following actions can be justified: (4) Accepting a bribe in the course of one's duties. (5) Lobbying politicians to influence legislation." (Low = can never be justified, High = other). OLS estimates. Regression includes constant and the following controls: $Give_i$ (Round 1), $Vote_i$, $Give_i$ (Round 1) $\times Vote_i$ and $info_i$.

$Comply_i | \textbf{Rule:} \textbf{Give} = 1$

| | Demographics | | Questionnaire Data | | | | | | | |
| | (1) Western Nationality | | (2) Importance of Democracy | | (3) Always participates in Elections | | (4) Justifiability of Bribes | | (5) Justifiability of Lobbying | |
	Yes	No	High	Low	Yes	No	Low	High	Low	High
$T\_Pay4Vote$	-.14**	-.08	-.13*	-.10	-.14**	-.06	-.16**	-.06	-.17**	-.06
	(.06)	(.08)	(.07)	(.07)	(.07)	(.08)	(.07)	(.08)	(.08)	(.07)
T_Bribe	-.12*	-.12	-.18**	-.06	-.15**	-.10	-.22***	-.02	-.18**	-.09
	(.06)	(.09)	(.08)	(.07)	(.07)	(.07)	(.07)	(.08)	(.08)	(.07)
$T\_ExcludePoor$	-.11*	-.01	-.08	-.10	-.15**	-.05	-.13*	-.06	-.12	-.07
	(.06)	(.09)	(.08)	(.07)	(.07)	(.07)	(.07)	(.08)	(.08)	(.07)
Constant	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Add. Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	272	128	183	201	190	194	214	170	205	179

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

ioral economics which suggests that procedural aspects of decision making affect preferences directly (Tyler, 1990; Frey, Benz and Stutzer, 2004; Cappelen et al., 2013; Bartling, Fehr and Herz, 2014, among others). Interestingly, we find no evidence for our treatments affecting the willingness of people to comply with rules *opposing* redistribution: Compliance with *Rule:Don't* is high both in the presence and absence of electoral malpractice. Moreover, in stark contrast to behavior under *Rule:Give*, beliefs about the behavior of others are in this case a very strong causal determinant of compliance. It seems that rules demanding subjects to behave egoistically—maybe because such rules are less prevalent in the real world and thus, subjects are less familiar with such demands—trigger psychological responses that make the wish to follow others weigh stronger than concerns regarding the procedure of rule selection. It remains to be shown by future research whether this observation is robust and generalizable.

We consider our results to be of interest to several neighboring fields of literature. The observation that a majority of subjects in our experiment voted for the rule that is in line with their previous action yields insights into the relationship of private giving decisions and preferences over related social rules as discussed, for example, by Corneo and Grüner (2000, 2002). By showing that democratically elected, non-binding rules can impact people's propensity to act in a pro-social way we add insight to how norms in giving behavior (e.g. Krupka and Weber, 2013), inequality acceptance (e.g. Almås et al., 2010) and defaults for donations (e.g. Altmann et al., 2014) may be shifted and mediated in society. A generalization of our main result would suggest that people are less likely to follow pro-social rules (for example, to be honest) when these rules are advocated by a corrupt authority (in our case a flawed election). This provides one possible explanation for the observation made in earlier experiments (see, for example, Gächter and Schulz, 2016) that the level of corruption in a society is correlated with measures of individual intrinsic honesty: Living in societies with high levels of corruption might undermine the trust in institutions per se and thus, lead people to behave dishonestly even in unrelated experimental situations. Whether electoral manipulation is indeed associated with such a ripple effect is an exciting question for future research. Finally, our finding that behavior under *Rule:Don't* is strongly driven by a wish to follow the behavior of others, while behavior under *Rule:Give* is largely immune to such "peer effects" resonates with previous research on the contagion of pro-social and anti-social behaviors by Offerman (2002), Croson and Shang (2008), Thöni and Gächter (2015) and Dimant (2017). Because pro-social behaviors are difficult to induce by peer-pressure, these studies have drawn the conclusion that an individual's own moral code of behavior is the main driving force behind pro-social choices. Our results show that group interactions *can* increase pro-social behavior, albeit not by appealing to the behavior

of others but by the democratic election of a pro-social code of conduct.

Of course, this essay can only be a first step towards understanding the effects of electoral malpractice on behavior under democratically elected institutions. More research is needed to draw definitive conclusions. We chose to study rule compliance in the domain of redistribution for its important role in economic research and policy. However, we see our study primarily as making a claim about compliance to behavioral rules in general. 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) is an important task for future research.

References

Akerlof, Robert. 2016. "Anger and Enforcement." *Journal of Economic Behavior and Organization*, 126: 110–124.

Akerlof, Robert. 2017. "The Importance of Legitimacy." *The World Bank Economic Review*, 30: 157–165.

Almås, Ingvild, Alexander W. Cappelen, and Bertil Tungodden. 2017. "Cutthroat Capitalism Versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians?" HCEO Working Paper 2017-003.

Almås, Ingvild, Alexander W Cappelen, Erik Ø Sørensen, and Bertil Tungodden. 2010. "Fairness and the Development of Inequality Acceptance." *Science*, 328(5982): 1176–1178.

Altmann, Steffen, Falk Armin, Paul Heidhues, and Rajshri Jayaraman. 2014. "Defaults and Donations: Evidence from a Field Experiment." CESifo Working Paper No. 5118.

Andreoni, James. 1989. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy*, 97(6): 1447–1458.

Andreoni, James. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *The Economic Journal*, 100(401): 464–477.

Bardhan, Pranab. 2000. "Irrigation and Cooperation: An Empirical Analysis of 48 Irrigation Communitiesin South India." *Economic Development and Cultural Change*, 48(4): 847–865.

Bartling, Björn, Ernst Fehr, and Holger Herz. 2014. "The Intrinsic Value of Decision Rights." *Econometrica*, 82(6): 2005–2039.

Basu, Kaushik. 2015. "The Republic of Beliefs: A New Approach to 'Law and Economics'." World Bank Policy Research Paper No. 7259.

Bénabou, Roland, and Jean Tirole. 2012. "Laws and Norms." *IZA Discussion Paper No. 6290.*

Berman, Eli, Michael Callen, Clark Gibson, and James D. Long. 2014. "Election Fairness and Government Legitimacy in Afghanistan." NBER Working Paper No. 19949.

Bernheim, B. Douglas. 1994. "A Theory of Conformity." *Journal of Political Economy*, 102(5): 8.

Black, Sandra E., and Lisa 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.

Bolton, Gary E., and Axel Ockenfels. 2000. "ERC: A Theory of Equity, Reciprocity, and Competition." *The American Economic Review*, 90(1): 166–193.

Bonin, John P., Derek C. Jones, and Louis Putterman. 1993. "Theoretical and Empirical Studies of Producer Cooperatives: Will Ever the Twain Meet?" *Journal of Economic Literature*, 31(3): 1290–1320.

Brusco, Valeria, Marcelo Nazareno, and Susan Carol Stokes. 2004. "Vote Buying in Argentina." *Latin American Research Review*, 39(2): 66–88.

Callen, Michael, and James D. Long. 2015. "Institutional Corruption and Election Fraud: Evidence from a Field Experiment in Afghanistan." *American Economic Review*, 105(1): 354–381.

Cappelen, Alexander W., Astri Drange Hole, Erik Ø. Sørensen, and Bertil Tungodden. 2007. "The Pluralism of Fairness Ideals: An Experimental Approach." *The American Economic Review*, 97(3): 818–827.

Cappelen, Alexander W., James Konow, Erik Ø. Sørensen, and Bertil Tungodden. 2013. "Just Luck: An Experimental Study of Risk-Taking and Fairness." *American Economic Review*, 103(4): 1398–1413.

Corneo, Giacomo, and Hans Peter Grüner. 2000. "Social Limits to Redistribution." *The American Economic Review*, 90(5): 1491–1507.

Corneo, Giacomo, and Hans Peter Grüner. 2002. "Individual Preferences for Political Redistribution." *Journal of Public Economics*, 83(1): 83–107.

Croson, Rachel, and Jen Shang. 2008. "The Impact of Downward Social Information on Contribution Decisions." *Experimental Economics*, 11(3): 221–233.

Dal Bó, Pedro. 2014. "Experimental Evidence on the Workings of Democratic Institutions." In *Institutions, Property Rights, and Economic Growth: The Legacy of Douglass North.* , ed. Sebastian Gallani and Itai Sened. New York:Cambridge University Press.

Dal Bó, Pedro, Andrew Foster, and Louis Putterman. 2010. "Institutions and Behavior: Experimental Evidence on the Effects of Democracy." *American Economic Review*, 100: 2205–2229.

De Alth, Shelley. 2009. "ID at the Polls: Assessing the Impact of Recent State Voter ID Laws on Voter Turnout." *Harvard Law and Policy Review*, 3(1): 185–202.

Dickson, Eric S, Sanford C Gordon, and Gregory A Huber. 2015. "Institutional Sources of Legitimate Authority: An Experimental Investigation." *American Journal of Political Science*, 59(1): 109–127.

Dimant, Eugen. 2017. "On Peer Effects: Contagion of Pro- and Anti-Social Behavior in Charitable Giving and the Role of Social Identity." mimeo.

Engel, Christoph. 2011. "Dictator Games: A Meta Study." *Experimental Economics*, 14(4): 583–610.

Ertan, Arhan, Talbot Page, and Louis Putterman. 2009. "Who to Punish? Individual Decisions and MajorityRule in Mitigating the Free Rider Problem." *European Economic Review*, 53(5): 495–511.

Fehr, Ernst, and Klaus M Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics*, 817–868.

Fehr, Ernst, and Simon Gächter. 2000. "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives*, 14(3): 159–181.

Fischbacher, Urs, and Franziska Föllmi-Heusi. 2013. "Lies in Disguise—An Experimental Study on Cheating." *Journal of the European Economic Association*, 11: 525–547.

Frey, Bruno S. 1997. "A Constitution For Knaves Crowds Out Civic Virtues." *The Economic Journal*, 107(443): 1043–1053.

Frey, Bruno S., Matthias Benz, and Alois Stutzer. 2004. "Introducing Procedural Utility: Not Only What, But Also How Matters." *Journal of Institutional and Theoretical Economics*, 160: 377–401.

Gächter, Simon, and Jonathan F. Schulz. 2016. "Intrinsic Honesty and the Prevalence of Rule Violations Across Societies." *Nature*, 531: 496–499.

Gonzalez-Ocantos, Ezequiel, Chad Kiewiet De Jonge, Carlos Meléndez, Javier Osorio, and David W Nickerson. 2012. "Vote Buying and Social Desirability Bias: Experimental Evidence from Nicaragua." *American Journal of Political Science*, 56(1): 202–217.

Gosling, Samuel D, Peter J Rentfrow, and William B Swann. 2003. "A Very Brief Measure of the Big-Five Personality Domains." *Journal of Research in Personality,* 37(6): 504–528.

Krupka, Erin L., and Roberto 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.

Manza, Jeff, and Christopher Uggen. 2008. *Locked Out: Felon Disenfranchisement and American Democracy.* Oxford University Press.

Norris, Pippa. 2014. *Why Electoral Integrity Matters.* Cambridge University Press.

Offerman, Theo. 2002. "Hurting Hurts More Than Helping Helps." *European Economic Review,* 46: 1423–1437.

Saito, Kota. 2013. "Social Preferences under Risk: Equality of Opportunity versus Equality of Outcome." *American Economic Review,* 103(7): 3084–3101.

Schlag, Karl H., and James Tremewan. 2016. "Simple Belief Elicitation." *mimeo.*

Sutter, Matthias, Stefan Haigner, and Martin 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.

Thöni, Christian, and Simon Gächter. 2015. "Peer Effects and Social Preferences in Voluntary Cooperation: A Theoretical and Experimental Analysis." *Journal of Economic Psychology,* 48: 72–88.

Tyler, Tom. 1990. *Why People Obey Rules.* Yale University Press.

Tyler, Tom R. 2006. "Psychological Perspectives on Legitimacy and Legitimation." *Annual Review of Psychology,* 57: 375–400.

Tyran, Jean-Robert, and Lars P. Feld. 2006. "Achieving Compliance when Legal Sanctions are Non-Deterrent." *The Scandinavian Journal of Economics,* 108(1): 135–156.

Weber, Max. 1978. *Economy and Society.* Berkeley: University of California Press.

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.

Zwick, Thomas. 2004. "Employee participation and productivity." *Labour Economics*, 11(6): 715–740.

Appendix

Theoretical Predictions for Voting Behavior

We extend our theory in Section 2 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 $U_i[Rule]$ denote i's expected utility given $Rule \in \{Rule{:}Give, Rule{:}Don't\}$. When voting, individual i takes into account how her own giving behavior will be affected by the rule as 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

$$\Delta F[\bar{u}^D] = F[+\bar{u}^D] - F[-\bar{u}^D]$$

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 $\Delta u_i(Give) \geq +\bar{u}^B$, individual i will choose $Give_i | Rule = 1$ irrespective of the rule. Individual i will then vote for *Rule:Give* ($Vote_i = 1$) if and only if

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

$$0.5 \cdot [u_i(1) + \bar{u}^B] + 0.5 \cdot \Delta F[\bar{u}^B] \cdot \Delta u_i(Receive) \geq 0.5 \cdot u_i(1)$$

$$\Leftrightarrow \bar{u}^B \geq -\Delta F(\bar{u}^B) \cdot \Delta u(Receive).$$

2. *Unconditional Non-Givers:* If $\Delta u_i(Give) < -\bar{u}^B$, individual i will choose $Give_i | Rule = 0$ irrespective of the rule. Individual i will then vote for *Rule:Give* ($Vote_i = 1$) if and only if

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

$$0.5 \cdot u_i(0) + 0.5 \cdot \Delta F[\bar{u}^B] \cdot \Delta u_i(Receive) \geq 0.5 \cdot [u_i(0) + \bar{u}^B]$$

$$\Leftrightarrow -\bar{u}^B \geq -\Delta F(\bar{u}^B) \cdot \Delta u(Receive).$$

3. *Rule-Followers:* If $-\bar{u}^B \leq \Delta u_i(Give) < +\bar{u}^B$, individual i will choose $Give_i(Rule) = 1$ under *Rule:Give* and $Give_i(Rule) = 0$ under *Rule:Don't*. Individual i will then vote for *Rule:Give* ($Vote_i = 1$) if and only if

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

$$0.5 \cdot [u_i(1) + \bar{u}^B] + 0.5 \cdot \Delta F[\bar{u}^B] \cdot \Delta u_i(Receive) \geq 0.5 \cdot [u_i(0) + \bar{u}^D]$$

$$\Leftrightarrow \Delta u_i(Give) \geq -\Delta F(\bar{u}^B) \cdot \Delta u(Receive)$$

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. If $\Delta F[\bar{u}^B]$ is close to zero, *Non-Givers* also vote according to their "natural" preferences, that is, $Vote_i = 0$. This case is illustrated in Figure 9, panel a). Increasing $\Delta F[\bar{u}^B]$ shifts voting preferences of non-givers in favor of *Rule:Give*. This first affects "moderate" *Non-Givers* who indeed would choose to give under the pro-social rule, i.e., those individuals who satisfy $-\bar{u}^B \leq \Delta u_i(Give) < 0$), see Figure 9, panel b). Only once $\Delta F[\bar{u}^B] \geq -\Delta\bar{u}^B/(\Delta u(Receive)$, also unconditional non-givers (and thus, all individuals) vote for *Rule:Give*, see Figure 9, panel c).
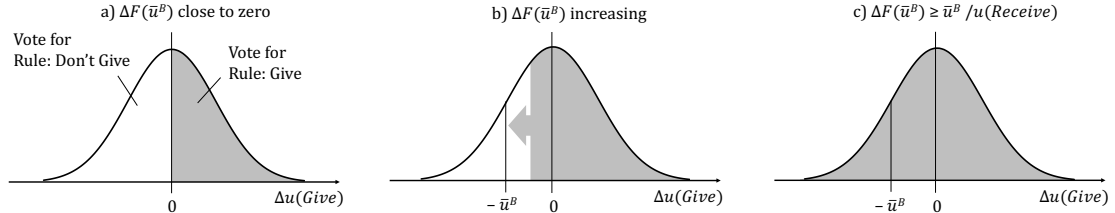
Figure 9: Theory: Share of Population voting for Rule: Give

Additional Data

	T_Baseline		T_Pay4Vote		T_Bribe		T_ExclPoor		Pooled	
Observations	100		100		100		100		400	
of which										
$Give_i$	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$	$=0$	$=1$
Observations	43	57	43	57	29	71	40	60	155	245
$Info_i = 4$	.51	.42	.63	.39	.52	.49	.45	.53	.53	.46
$Vote_i = 1$	.35	.86	.47	.97	.45	.96	.38	.93	.41	.93
$Don't\_Pay_i = 1$			.53	.21						
$Accept\_Bribe_i = 1$					.76	.24				
$Excl\_Poor_i = 1$							.48	.52		
$Give_i\|Rule{:}Give = 1$	.65	.93	.42	.97	.41	.93	.45	.95	.49	.94
$Give_i\|Rule{:}Don't = 0$	.98	.53	.91	.42	.86	.55	.98	.62	.94	.53
$Rule\_Complier_i$	.65	.46	.40	.40	.35	.49	.45	.47	.47	.46

Table 6: Summary of experimental data. $Dont\_Pay_i = 1$ if subject did not pay to make her vote count. $Accept\_Bribe_i = 1$ if subject accepted to change her vote against payment. $ExclPoor_i = 1$ if subject's vote was not counted because her stated household income is below 40.000 GBP. $Rule\_Complier_i = 1$ if subject complies with both rules, i.e., $Give_i(Rule{:}Give) = 1$ and $Give_i(Rule{:}Don't) = 0$.

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 <u>trust</u>...
...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 <u>justified</u>:
(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?

<div align="center">Never   Rarely   Usually   Almost always   Always</div>

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)

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 10: 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 11: Screenshot: General Instructions

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, <u>one</u> 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 12: Screenshot: Instructions about the Lottery

Distribution of lottery tickets

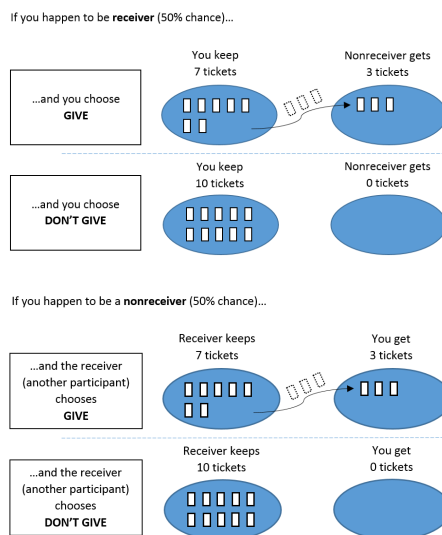In both rounds 1 and 2, the lottery tickets are distributed in two steps.

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.

100 Participants

Computer picks

50 Receivers   **50 Nonreceivers**

Each receiver gets: Each nonreceiver gets:
10 Lottery Tickets   0 Lottery Tickets
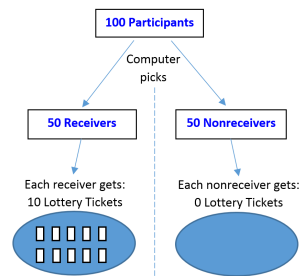
Step 2: Participants decide whether they want to share tickets with nonreceivers:

- 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)...

...and you choose **GIVE**   You keep 7 tickets   Nonreceiver gets 3 tickets

...and you choose **DON'T GIVE**   You keep 10 tickets   Nonreceiver gets 0 tickets

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

...and the receiver (another participant) chooses **GIVE**   Receiver keeps 7 tickets   You get 3 tickets

...and the receiver (another participant) chooses **DON'T GIVE**   Receiver keeps 10 tickets   You get 0 tickets

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 13: Screenshot: Instructions about the Distribution of Lottery Tickets

46

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 14: 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 15: 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 16: 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 17: 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 18: Screenshot: $Accept\_Bribe \in \{0, 1\}$ (Round 2, T_Bribe)

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 19: 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. |
| ☑ 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:

| ☑ These are my final answers. Please proceed. |

Figure 20: Screenshot: Beliefs about Others (Round 2, *T_Baseline*)