

Online Appendix for: Corrupted Votes and Rule Compliance

Arno Apffelstaedt and Jana Freundt

B.1 Randomization on observables

	StdMajority	Pay4Vote	MoneyOffer	ExcludePoor	All
Female	0.38 (100)	0.42 (100)	0.42 (100)	0.45 (100)	0.42 (400)
Age	30.22 (97)	29.27 (98)	33.06 (99)	31.66 (100)	31.06 (394)
Western	0.70 (100)	0.64 (100)	0.66 (100)	0.78 (100)	0.69 (400)

Number of observations in parentheses.

Table B.1: Share of female participants, mean age and share of participants from western countries in each treatment and in total.

B.2 Treatment effects: Type-level and type-weighted estimations

In the main text, we estimate treatment effects on rule compliance using regressions that control for type fixed effects, where $Type_i = (Give_i|NoRule) \times Vote_i$. While this approach effectively controls for unbalanced type distributions across treatments, this appendix presents the results of an alternative approach to estimating treatment effects, which explicitly accounts for the possibility of treatment effects varying across types. As we will demonstrate, results remain unchanged. The approach relies on a weights-based identification strategy proposed by Dal Bó, Foster and Kamei (2019). It entails estimating individual treatment effects for each type and then computing the average treatment effect by assigning weights to each type's treatment effect based on its prevalence in the entire population.

It is worth noting that Dal Bó, Foster and Kamei (2019) introduce this method primarily to address issues related to endogenous selection into treatment when a subject's type is not consistently observable across all conditions. These issues are not present in our setting since the type of each subject $Type_i = (Give_i|NoRule) \times Vote_i$ is observed in all

treatments. Instead, we employ their weights-based approach to account for potential heterogeneous treatment effects across types, offering insights into how various types respond to malpractice.¹

The detailed results of the type-weighting approach are presented in Table B.2. Panel (a) shows the number of subjects per type across all four treatments: There are

- 92 subjects of type $((Give_i|NoRule), Vote_i) = (0, Rule:Don't)$,
- 17 subjects of type $((Give_i|NoRule), Vote_i) = (1, Rule:Don't)$,
- 63 subjects of type $((Give_i|NoRule), Vote_i) = (0, Rule:Give)$, and
- 228 subjects of type $((Give_i|NoRule), Vote_i) = (1, Rule:Give)$.

To calculate type-weighted average treatment effects, we first obtain baseline compliance rates and treatment effects for each of the four types separately. We then weigh types according to the relative frequency with which they occur in our sample. In Panel (b) of Table B.2, we present the proportion of compliant subjects per type in the baseline treatment StdMajority. In Panel (c) of Table B.2 we report treatment effects: White cells display the individual treatment effects for each type, estimated through simple (univariate) OLS regressions while conditioning on the type. Using these estimates, we compute population average treatment effects (reported in gray cells) by weighing types according to their prevalence in the population. For example, to compute the average treatment effect of offering money to voters (MoneyOffer) on compliance with Rule:Give, we use the formula:

$$(92/400)(.01) + (63/400)(-.57) + (17/400)(-.17) + (228/400)(-.04) = -.12.$$

Standard errors for weighted averages are calculated using the Delta method. For example, the standard error for the average treatment effect we just calculated can be determined from

$$\sqrt{(92/400)^2(.16)^2 + (63/400)^2(.18)^2 + (17/400)^2(.36)^2 + (228/400)^2(.03)^2} = .05.$$

The results obtained using the type-weighting approach closely align with the estimates derived from regressions with type fixed effects, as presented in the main text. To ensure better comparability, we reproduce Table 3 and Figures 3 and 4 below using the results from our type-weighting approach. Table B.3 replicates Table 3. Figures B.1 and B.2 replicate Figures 3 and 4. As both the table and figures illustrate, compliance rates and treatment effects, at both the population and subgroup levels, are nearly identical across the two estimation methods.

¹With types being perfectly observed, there are several methods available to accommodate type-dependent treatment effects. The type-weighted approach we employ is essentially equivalent to using a matching estimator with exact matching on (discrete) type covariates. Identically sized effects (usually with higher levels of significance) are found with similar methods such as inverse probability weighting or regression adjustment.

(a) All treatments: n by $Type_i$ (b) *StdMajority*: Share of n complying with...

$Vote_i$	$Give_i NoRule$		Σ	...Rule:Give			...Rule:Don't		
	0	1		0	1	w.avg.	0	1	w.avg.
<i>Rule:Don't</i>	92	17	109	0.57	0.50	0.56	0.96	0.62	0.91
<i>Rule:Give</i>	63	228	291	0.80	1.00	0.96	1.00	0.51	0.62
Σ	155	245	400	0.66	0.97	0.85	0.98	0.52	0.70

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

	$Vote_i$...Rule:Give			...Rule:Don't		
		0	1	w.avg.	0	1	w.avg.
<i>Pay4Vote</i>	<i>Rule:Don't</i>	-0.18 (0.14)	0.50 (0.42)	-0.07 (0.14)	-0.05 (0.07)	-0.63 (0.30)	-0.14 (0.08)
	<i>Rule:Give</i>	-0.35 (0.16)	-0.04 (0.03)	-0.10 (0.04)	-0.10 (0.08)	-0.07 (0.10)	-0.08 (0.08)
	w.avg.	-0.25 (0.11)	0.00 (0.04)	-0.10 (0.05)	-0.07 (0.05)	-0.11 (0.09)	-0.10 (0.06)
<i>MoneyOffer</i>	<i>Rule:Don't</i>	-0.01 (0.16)	-0.17 (0.36)	-0.03 (0.14)	-0.09 (0.08)	0.37 (0.26)	-0.02 (0.08)
	<i>Rule:Give</i>	-0.57 (0.18)	-0.04 (0.03)	-0.16 (0.05)	-0.15 (0.09)	0.02 (0.09)	-0.02 (0.08)
	w.avg.	-0.24 (0.12)	-0.05 (0.04)	-0.12 (0.05)	-0.12 (0.06)	0.04 (0.09)	-0.02 (0.06)
<i>ExcludePoor</i>	<i>Rule:Don't</i>	-0.13 (0.14)	-0.00 (0.33)	-0.11 (0.13)	0.00 (0.07)	0.37 (0.23)	0.06 (0.07)
	<i>Rule:Give</i>	-0.33 (0.17)	-0.02 (0.03)	-0.09 (0.04)	0.00 (0.09)	0.08 (0.10)	0.06 (0.08)
	w.avg.	-0.21 (0.11)	-0.02 (0.04)	-0.09 (0.05)	0.00 (0.05)	0.10 (0.09)	0.05 (0.06)
<i>Pooled</i>	<i>Rule:Don't</i>	-.12 (0.11)	.06 (0.26)	-0.09 (0.10)	-0.04 (0.06)	0.15 (0.23)	-0.01 (0.06)
	<i>Rule:Give</i>	-0.40 (0.14)	-0.03 (0.03)	-0.11 (0.04)	-0.08 (0.07)	0.01 (0.08)	-0.01 (0.07)
	w.avg.	-0.23 (0.09)	-0.03 (0.03)	-0.11 (0.04)	-0.06 (0.04)	0.02 (0.08)	-0.01 (0.05)

Standard errors in parentheses.

Table B.2: Number of subjects (a), baseline compliance rates (b) and treatment effects (c) by $Type_i = (Give_i|NoRule) \times Vote_i$. White cells in panel (b) show compliance rates (share of subjects for whom $Comply_i|Rule = 1$) for each $Type_i$ in *StdMajority*. White cells in panel (c) show treatment effects for each $Type_i$ from OLS regressions. Gray cells show weighted-average baseline compliance rates (panel b) and weighted-average treatment effects (panel c). Weighted standard errors calculated assuming normally distributed standard errors (Delta method).

Dep.Var.	a) Compliance with Rule:Give					b) Compliance with Rule:Don't				
	Subgroups					Subgroups				
	All Subjects	$Give_i NoRule$		$Vote_i$		All Subjects	$Give_i NoRule$		$Vote_i$	
	0	1	<i>Rule:Don't</i>	<i>Rule:Give</i>		0	1	<i>Rule:Don't</i>	<i>Rule:Give</i>	
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Pay4Vote	-0.10 (0.04)	-0.25 (0.11)	0.00 (0.04)	-0.07 (0.14)	-0.10 (0.04)	-0.10 (0.06)	-0.07 (0.05)	-0.11 (0.09)	-0.14 (0.08)	-0.08 (0.08)
MoneyOffer	-0.12 (0.05)	-0.24 (0.12)	-0.05 (0.04)	-0.03 (0.14)	-0.16 (0.05)	-0.02 (0.06)	-0.12 (0.06)	0.04 (0.09)	-0.02 (0.08)	-0.02 (0.08)
ExcludePoor	-0.09 (0.05)	-0.21 (0.11)	-0.02 (0.04)	-0.11 (0.13)	-0.09 (0.04)	0.06 (0.06)	-0.00 (0.05)	0.10 (0.09)	0.05 (0.07)	0.06 (0.08)
Constant (Compliance in StdMajority)	0.85 (0.03)	0.66 (0.08)	0.97 (0.03)	0.56 (0.09)	0.96 (0.03)	0.70 (0.04)	0.98 (0.04)	0.52 (0.07)	0.91 (0.05)	0.62 (0.06)
Malpractice (Pooled)	-0.11 (0.04)	-0.23 (0.09)	-0.03 (0.03)	-0.09 (0.10)	-0.11 (0.04)	-0.01 (0.05)	-0.06 (0.04)	0.02 (0.08)	-0.01 (0.06)	-0.01 (0.07)
Observations	400	155	245	109	291	400	155	245	109	291

Standard errors in parentheses.

Table B.3: Treatment effects of Pay4Vote, MoneyOffer and ExcludePoor on rule compliance: Type-weighted estimates. The coefficient on Malpractice (shaded in grey) represents the pooled treatment effect estimated from conducting the same regression using a common indicator variable $Malpractice = 1$ if $Treatment = Pay4Vote, MoneyOffer$ or $ExcludePoor$. Full results and details of the type-weighted analysis can be found in Table B.2. Replication of Table 3.

Baseline compliance rates (share of subjects complying with the elected rule after a standard majority vote)

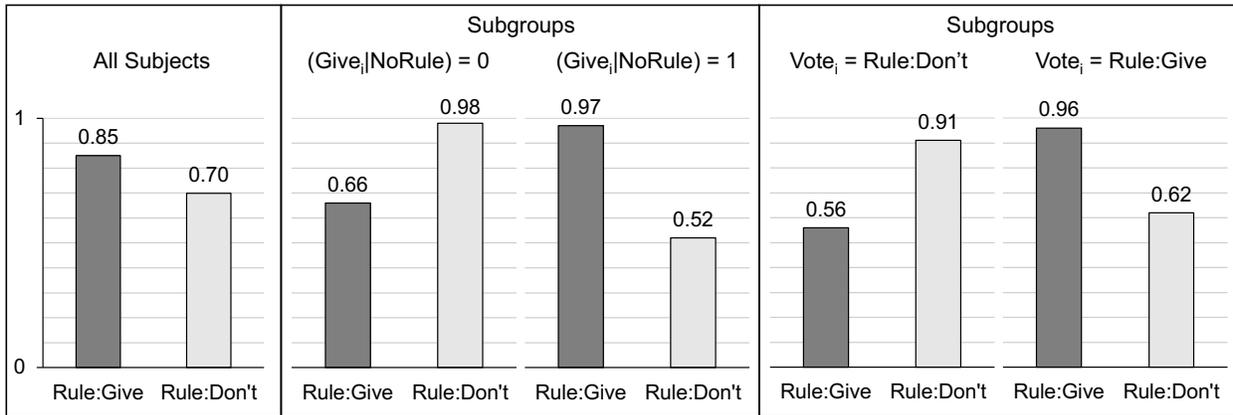


Figure B.1: Share of subjects complying with majority-elected rules in treatment StdMajority. Graphs show type-weighted shares. For details see Table B.2, Panel (b). Replication of Figure 3.

Effect of interventions Pay4Vote (P), MoneyOffer (M) and ExcludePoor (E) on rule compliance (percentage point change from baseline compliance rates)

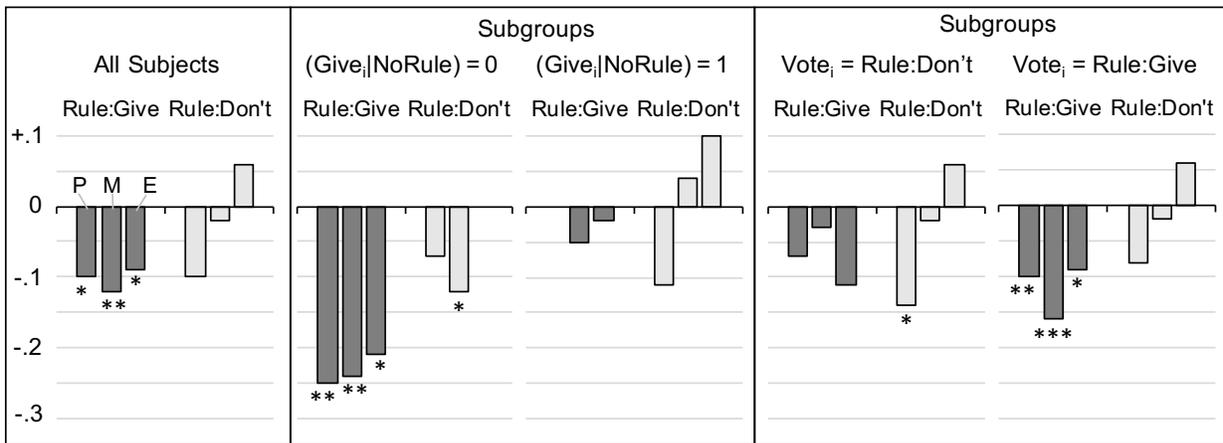


Figure B.2: Treatment effects of interventions Pay4Vote (P), MoneyOffer (M) and ExcludePoor (E) on rule compliance. Graphs show type-weighted estimates of treatment effects as reported in Table B.2 (panel c) and Table B.3. Replication of Figure 4. Stars denote statistically significant differences to compliance rates in StdMajority: * $p < .1$, ** $p < .05$, *** $p < .01$.

B.3 Replication of Table 4 with Individual Treatment Effects

Dep. Var.	a) Compliance with Rule:Give					b) Compliance with Rule:Don't				
	$E_i(Comply_{-i})$	$Comply_i = 1$		$Comply_i = 1$		$E_i(Comply_{-i})$	$Comply_i = 1$		$Comply_i = 1$	
	OLS	OLS	2SLS	OLS	OLS	OLS	OLS	2SLS	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
$info_i = 4$	0.13 (0.02)			-0.04 (0.04)	-0.05 (0.04)	-0.12 (0.03)			-0.11 (0.04)	-0.09 (0.04)
$E_i(Comply_{-i})$		0.46 (0.07)	-0.32 (0.31)				0.51 (0.08)	0.89 (0.33)		
Pay4Vote	-0.04 (0.03)	-0.09 (0.05)	-0.12 (0.06)	-0.11 (0.05)	-0.11 (0.05)	0.02 (0.04)	-0.09 (0.06)	-0.09 (0.06)	-0.08 (0.06)	-0.12 (0.06)
MoneyOffer	-0.03 (0.03)	-0.11 (0.05)	-0.13 (0.06)	-0.12 (0.05)	-0.11 (0.05)	-0.02 (0.04)	-0.00 (0.06)	0.01 (0.06)	-0.01 (0.06)	-0.04 (0.06)
ExcludePoor	0.00 (0.03)	-0.10 (0.05)	-0.09 (0.05)	-0.09 (0.05)	-0.07 (0.05)	0.04 (0.04)	0.04 (0.06)	0.03 (0.06)	0.06 (0.06)	0.04 (0.06)
Constant	0.49 (0.03)	0.31 (0.06)	0.74 (0.18)	0.58 (0.05)	0.25 (0.17)	0.78 (0.04)	0.57 (0.08)	0.29 (0.24)	0.99 (0.06)	0.52 (0.21)
Type fixed eff.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Add. Controls					Yes					Yes
Observations	400	400	400	400	369	400	400	400	400	369

Standard errors in parentheses.

Table B.4: The role of others in guiding behavior. $E_i(Comply_{-i})$ is individual i 's belief about the share of other participants complying with the rule. IV regressions are 2SLS with $E_i(Comply_{-i})$ being instrumented by $1.[info_i = 4]$. Additional controls in (5) are: $Female_i$, Age_i , $Western_i$, $Student_i$, $UGrad_i$, $Risk_Seeking_i$, $Trust_i$, number of mistakes in control questions, factor variables measuring political and social values in questionnaire, as well as Big Five personality measures. $Female_i$ and $Risk_Seeking_i$ (answer on 11-point scale to "Are you a person who is generally willing to take risks (10) or do you try to avoid taking risks (0)?") are (weakly) significant for compliance with Rule:Give (0.08 ($p = 0.05$) and 0.02 ($p = 0.06$), respectively). $Trust_i$ (answer on 11-point scale to "Do you think that most people would try to take advantage of you if they got the chance (10), or would they try to be fair (0)?") is highly significant for compliance with Rule:Don't (0.04, $p = 0.001$). All other demographic and questionnaire controls are insignificant.

B.4 Calculation of subjective beliefs about *Outcome_Bias*

We define *Outcome_Bias* as the absolute value of the difference between the share of votes for Rule:Give before and after the intervention. We calculate subjective beliefs about the expected outcome bias ($E_i(\textit{Outcome_Bias})$) from elicited beliefs about the share of subjects paying for their vote (Pay4Vote), the share of subjects accepting the bribe (MoneyOffer), or the voting behavior among “poor” and “rich” subjects (ExcludePoor).

In the Pay4Vote treatment, *Outcome_Bias* is equal to the absolute value of the difference between (1) the share of subjects who paid for the vote and voted for Rule:Give times the overall share of subjects who voted for Rule:Give, divided by the overall share of subjects who paid for the vote, and (2) the share of subjects who voted for Rule:Give. Expressed in form of beliefs directly elicited from subjects in the Pay4Vote treatment, the belief of subject i regarding *Outcome_Bias* is computed as $E_i(\textit{Outcome_Bias}|\textit{Treatment} = \textit{Pay4Vote}) :=$

$$\left| \frac{E_i(\textit{Pay}_j = 1|\textit{Vote}_j = \textit{Rule:Give}) \cdot E_i(\textit{Vote}_j = \textit{Rule:Give})}{\left[E_i(\textit{Pay}_j = 1|\textit{Vote}_j = \textit{Rule:Give}) \cdot E_i(\textit{Vote}_j = \textit{Rule:Give}) + E_i(\textit{Pay}_j = 1|\textit{Vote}_j = \textit{Rule:Don't}) \cdot (1 - E_i(\textit{Vote}_j = \textit{Rule:Give})) \right]} - E_i(\textit{Vote}_j = \textit{Rule:Give}) \right|$$

where $E_i(\textit{Pay}_j = 1|\textit{Vote}_j = \textit{Rule:Give})$ ($E_i(\textit{Pay}_j = 1|\textit{Vote}_j = \textit{Rule:Don't})$) is the subject’s belief regarding the share of individuals who pay for the vote, given that they voted for Rule:Give (Rule:Don’t), and $E_i(\textit{Vote}_j = \textit{Rule:Give})$ is the subject’s belief regarding the total share of individuals voting for Rule:Give.

In the MoneyOffer treatment, *Outcome_Bias* is equal to the absolute value of the difference between (1) the share of subjects who accepted the bribe and voted for Rule:Don’t times the share of subjects who voted for Rule:Don’t plus the share of subjects who did not accept the bribe and voted for Rule:Give times the share of subject who voted for Rule:Give, and (2) the share of subjects who voted for Rule:Give. Expressed in form of beliefs directly elicited from subjects in the MoneyOffer treatment, the belief of subject i regarding *Outcome_Bias* is computed as $E_i(\textit{Outcome_Bias}|\textit{Treatment} = \textit{MoneyOffer}) :=$

$$\left| E_i(\textit{Accept}_j = 1|\textit{Vote}_j = \textit{Rule:Don't}) \cdot (1 - E_i(\textit{Vote}_j = \textit{Rule:Give})) + (1 - E_i(\textit{Accept}_j = 1|\textit{Vote}_j = \textit{Rule:Give})) \cdot E_i(\textit{Vote}_j = \textit{Rule:Give}) - E_i(\textit{Vote}_j = \textit{Rule:Give}) \right|$$

where $E_i(\textit{Accept}_j = 1|\textit{Vote}_j = \textit{Rule:Give})$ ($E_i(\textit{Accept}_j = 1|\textit{Vote}_j = \textit{Rule:Don't})$) is the subject’s belief regarding the share of individuals who accept the bribe, given that they voted for Rule:Give (Rule:Don’t), and $E_i(\textit{Vote}_j = \textit{Rule:Give})$ is the subject’s belief regarding the total share of individuals voting for Rule:Give.

In the ExcludePoor treatment, *Outcome_Bias* is equal to the absolute value of the difference between (1) the share of subjects with income larger than £40,000 who voted for Rule:Give, and (2) the share of subjects who voted for Rule:Give. Expressed in form of beliefs directly elicited from subjects in the ExcludePoor treatment, the belief of subject i regarding *Outcome_Bias* is computed as $E_i(\textit{Outcome_Bias}|\textit{Treatment} = \textit{ExcludePoor}) :=$

$$\left| E_i(\textit{Vote}_j = \textit{Rule:Give}|\textit{Income}_j > \textit{£40,000}) - E_i(\textit{Vote}_j = \textit{Rule:Give}) \right|$$

where $E_i(\text{Vote}_j = \text{Rule:Give} | \text{Income}_j > \pounds 40,000)$ is the subject’s belief regarding the share of individuals with income larger than $\pounds 40,000$ who vote for Rule:Give, and $E_i(\text{Vote}_j = \text{Rule:Give})$ is the subject’s belief regarding the total share of individuals voting for Rule:Give.

By definition, $E_i(\text{Outcome_Bias} | \text{Treatment} = \text{StdMajority}) := 0$ for all subjects in the baseline treatment StdMajority.

B.5 Theoretical framework

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

$$(\text{Give}_i | \text{NoRule}) = 1 \Leftrightarrow \Delta u_i \geq 0.$$

A positive Δu_i may reflect social preferences of individual i such as inequality aversion or “warm glow” utility.² Let Δu_i be distributed in the population with cumulative density function $F[\cdot]$. The share of Givers in the population is then given by $1 - F[0]$ as illustrated in Figure B.3, panel a), below.

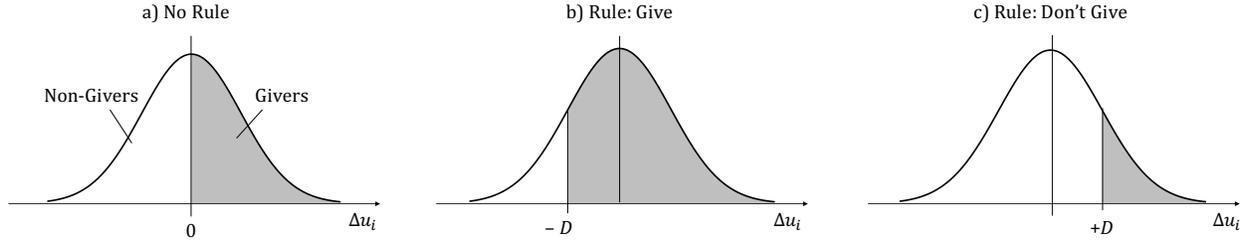


Figure B.3: Theory: Illustration of population shares choosing to give ($\text{Give}_i = 1$) and not to give ($\text{Give}_i = 0$) when there exists no code of conduct (panel a) and when there exists a code of conduct that came into force with a standard majority vote (panels b and c).

Consider next the situation with a code of conduct, either Rule:Give or Rule:Don’t. If the code has come into force with a standard majority vote (StdMajority) we assume that it adds fixed utility $D \geq 0$ to the action that is prescribed by the code. This constant can be interpreted as an emotional utility some people derive from following a rule elected by the majority. It follows that

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

²Typical examples in standard settings are Fehr and Schmidt (1999), Bolton and Ockenfels (2000) and Andreoni (1989, 1990). Inequity aversion over chances to win a prize has been modeled by, for example, Saito (2013). Experimental evidence showing how prosocial behavior extends to choices over risky payoffs can be found in Brock, Lange and Ozbay (2013) and Freundt and Lange (2017), among others.

Compared to the case without a rule, the share of subjects choosing to give increases or decreases, see Figure B.3, panel (b) and (c), respectively. Note, importantly, that rules only affect the behavior of those individuals who in the absence of a rule would have chosen the opposite action. While Rule:Give may convince a Non-Giver to give, it will leave the behavior of a Giver ($\Delta u_i \geq 0$) unaffected. Similarly, Rule:Don't may induce some Givers to stop giving, but will not affect the choice of Non-Givers ($\Delta u_i < 0$). We assume that electoral malpractice (in our experiment Pay4Vote, MoneyOffer, ExcludePoor) alters the value some people derive from obeying the elected rule. Instead of generating utility D , rule compliance is now associated with a lower utility $D - M$. Constant $M \geq 0$ measures the loss in utility induced by malpractice. As a result, individual i 's propensity to comply with the elected rule is reduced. In particular,

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

First and foremost, we thus expect that malpractice leads people to revert back to their individually preferred behavior: As M increases, a lower share of Non-Givers will follow Rule:Give, see Figure B.4, panel b). Similarly, a lower share of Givers will be willing to follow Rule:Don't (Figure B.4, panel c)). As M becomes sufficiently large such that $D - M$ turns negative, people may even turn against rules that match their individual giving preferences. For example, it is theoretically possible that giving under Rule:Give will deteriorate below rates observed in the absence of a rule, although such a strong reaction might be unlikely to be observed in the experiment.

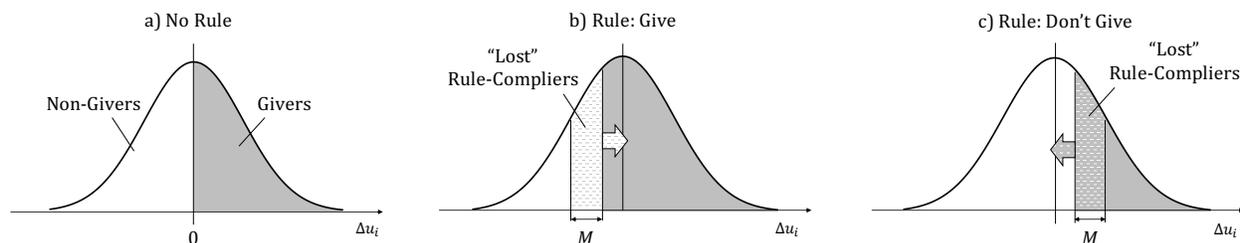


Figure B.4: Theory: Illustration of population shares choosing to give ($Give_i = 1$) and not to give ($Give_i = 0$) when there exists no code of conduct (panel a) and when there exists a code of conduct that came into force with malpractice during the election (panels b and c).

Voting behavior. We can extend the above theory to yield predictions about voting behavior. Note that in all treatments, subjects vote before interventions take place that may undermine the democratic election. Voting decisions are therefore unbiased by the exposure to a particular treatment. We assume that each subject votes *sincerely* in the sense that she chooses to vote for the outcome that yields her a higher expected utility. Let $U_i[\textit{Rule}]$ denote i 's expected utility given $\textit{Rule} \in \{\textit{Rule:Give}, \textit{Rule:Don't}\}$. When voting, individual i takes into account how her own giving behavior will be affected by the rule as 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 $\Delta u(\text{Receive}) > 0$ denote the difference in utility between receiving three tickets from another subject and not receiving any tickets. Because the average subject in the population is more likely to give under Rule:Give than under Rule:Don't, the conditional probability that i will receive three tickets from another subject increases by

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

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

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

$$\begin{aligned} U_i[\text{Rule:Give} | (\text{Give}_i | \text{Rule}) = 1] &\geq U_i[\text{Rule:Don't} | (\text{Give}_i | \text{Rule}) = 1] \\ 0.5 \cdot [u_i(\text{Give}_i = 1) + D] + 0.5 \cdot \Delta F[D] \cdot \Delta u_i(\text{Receive}) &\geq 0.5 \cdot u_i(\text{Give}_i = 1) \\ \Leftrightarrow \underbrace{\Delta F(D)}_{>0} &\geq \underbrace{-\frac{D}{\Delta u(\text{Receive})}}_{<0}. \end{aligned}$$

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

$$\begin{aligned} U_i[\text{Rule:Give} | (\text{Give}_i | \text{Rule}) = 0] &\geq U_i[\text{Rule:Don't} | (\text{Give}_i | \text{Rule}) = 0] \\ 0.5 \cdot u_i(\text{Give}_i = 0) + 0.5 \cdot \Delta F[D] \cdot \Delta u_i(\text{Receive}) &\geq 0.5 \cdot [u_i(\text{Give}_i = 0) + D] \\ \Leftrightarrow -D &\geq -\Delta F(D) \cdot \Delta u(\text{Receive}) \\ \Leftrightarrow \Delta F(D) &\geq \frac{D}{\Delta u(\text{Receive})} \end{aligned}$$

and otherwise will vote for Rule:Don't.

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

$$\begin{aligned} U_i[\text{Rule:Give} | (\text{Give}_i | \text{Rule}) = 1] &\geq U_i[\text{Rule:Don't} | (\text{Give}_i | \text{Rule}) = 0] \\ 0.5 \cdot [u_i(\text{Give}_i = 1) + D] + 0.5 \cdot \Delta F[D] \cdot \Delta u_i(\text{Receive}) &\geq 0.5 \cdot [u_i(\text{Give}_i = 0) + D] \\ \Leftrightarrow \Delta u_i &\geq -\Delta F(D) \cdot \Delta u(\text{Receive}) \\ \Leftrightarrow \Delta F(D) &\geq -\frac{\Delta u_i}{\Delta u(\text{Receive})}, \end{aligned}$$

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

We can see that there is a monotonic relation between $\Delta u_i(\text{Give})$ and the tendency to vote for Rule:Give. Givers always vote for Rule:Give. This is true for both, unconditional givers and rule-followers. Non-Givers, on the other hand, only vote for Rule:Give if they expect that rules have sufficiently large effect on the giving behavior of others. Otherwise, they vote for Rule:Don't. If $\Delta F[D]$ is close to zero, all Non-Givers vote for Rule:Don't. This case is illustrated in Figure B.5, panel a). Increasing $\Delta F[D]$ shifts voting preferences of non-givers in favor of Rule:Give. This first affects rule-following Non-Givers who indeed would choose to give under the pro-social rule, i.e., those individuals who satisfy $-D \leq \Delta u_i(\text{Give}) < 0$, see Figure B.5, panel (b). Only once $\Delta F(D) \geq \frac{D}{\Delta u(\text{Receive})}$, also unconditional non-givers (and thus, all individuals) vote for Rule:Give, see Figure B.5, panel c).

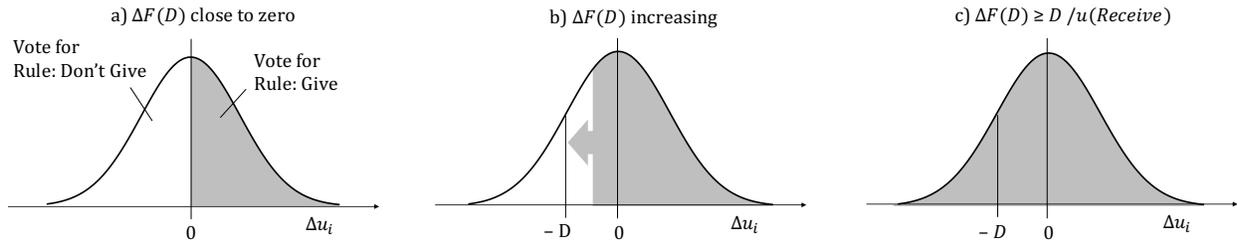


Figure B.5: Theory: Share of Population voting for Rule:Give

B.6 Questionnaire

Questionnaire: Politics

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

1. In political matters, people talk of “the left” and “the right”. On a scale from 0 to 10, where would you place your views, generally speaking?

(Scale: 0 = Left, 10 = Right)

2. On a scale from 0 to 10, how important is it for you to live in a country that is governed democratically?

(Scale: 0 = not at all important, 10 = extremely important)

3. How democratic do you think your country is overall?

(Scale: 0 = not at all democratic, 10 = completely democratic)

4. How important is it for you to personally express your voice when it comes to political decision making?

(Scale: 0 = not at all important, 10 = extremely important)

5. It is important that you pay attention to this study. Please tick number 7 to show that you pay attention. The scale below does not play a role.

(Scale: 0 = not at all important, 10 = very important)

6. On a scale from 0 to 10, where 0 means “no trust at all” and 10 means “very much trust”, how much do you personally trust...

...politicians?

...large corporations?

...the results of elections?

7. Please indicate for each of the following actions to what extent you think that action can be justified:

(Scale: 0= can never be justified, 10= can always be justified)

- Violating the instructions of one’s superiors (for example at work or school).
- Accepting a bribe in the course of one’s duties.
- Cheating on taxes if one has the chance.
- Influencing the actions of people by giving them money.
- Lobbying politicians to influence legislation.

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

0:
“The rich have an obligation to subsidize the poor. If necessary, they have to be forced to do so.”

10:
“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:
“For a society to be fair, the incomes of all people should be equal.”

10:
“There is nothing unfair in having more money than somebody else, no matter how large the difference.”

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

Never Rarely Usually Almost always Always

Questionnaire: General questions

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

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

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

2. How much do you agree with the following statement: “Money brings out the worst in people.”?

(Scale: 0 = Do not agree at all, 10 = Agree completely)

3. Do you think that most people would try to take advantage of you if they got the chance, or would they try to be fair?

(Scale: 0 = All people would try to be fair, 10 = All people would try to take advantage of you)

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

Would be willing to pay at most... (dropdown menu with answer choices from 0 GBP to 20

GBP in steps of 1)

5. Here are a number of personality traits that may or may not apply to you. Please indicate to what extent you agree or disagree that these personality traits apply to you.

Note: You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other.

I see myself as...

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

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

B.7 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 B.6: 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 B.7: Screenshot: General Instructions (Pay4Vote)

The Lottery

There are two rounds in this lottery:

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

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

I have read the instructions carefully and want to proceed.

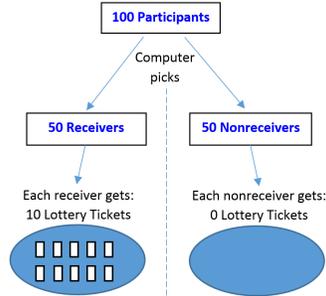
Figure B.8: Screenshot: Instructions about the Lottery

Distribution of lottery tickets

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

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

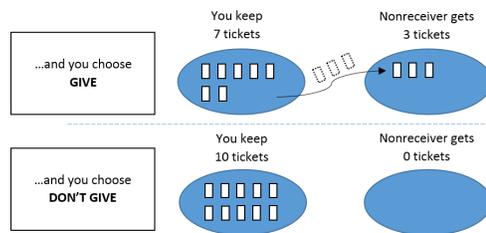
- The computer randomly selects 50 out of 100 participants to be "Receivers". Each receiver gets 10 lottery tickets from the computer.
- The other 50 participants are "Nonreceivers". Nonreceivers get no tickets from the computer.
- No participant learns whether he/she has been chosen to be a receiver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



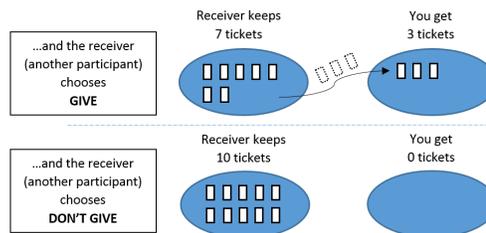
Step 2: Participants decide whether they want to share tickets with nonreceivers:

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

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



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



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

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

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

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

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

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

Figure B.9: Screenshot: Instructions about the Distribution of Lottery Tickets

Round 1

Your Choice: Give or Don't Give

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

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

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

Remind me of the way lottery tickets are distributed.

Lottery tickets are distributed in two steps:

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

Step 2: Each participant decides privately whether he/she wants to GIVE or DONT 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 B.10: Screenshot: Choice $Give_i | NoRule \in \{0, 1\}$ (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 B.11: Screenshot: Randomized social information $info_i \in \{2, 4\}$ (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 DONT 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 DONT GIVE, a code of conduct will be set.

- The code of conduct says whether everyone should choose GIVE (\Rightarrow **RULE: GIVE**) or whether everyone should choose DONT GIVE (\Rightarrow **RULE: DONT 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: DONT GIVE) he/she prefers to have implemented as the code of conduct for all participants. Please select a rule below.

Vote for RULE: GIVE

Vote for RULE: DONT GIVE

Once you have made your decision, please tick below:

This is my final answer. Please proceed.

Figure B.12: 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 subtracted 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 B.13: Screenshot: $Accept_Pay4Vote \in \{0, 1\}$ (Round 2, 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 B.14: Screenshot: $Accept_MoneyOffer \in \{0, 1\}$ (Round 2, MoneyOffer)

Round 2

Your choice: Follow the rule or not

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

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

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

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

Figure B.15: Screenshot: Information about intervention Exclude_Poor (Round 2)

Round 2

Your choice: Follow the rule or not

Your vote for the code of conduct has been counted.

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

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

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

Follow the rule and GIVE.

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

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

Follow the rule and DON'T GIVE.

Don't follow the rule and GIVE.

Once you have made your decision, please tick below:

This is my final answer. Please proceed.

Figure B.16: Screenshot: $Give_i | Rule \in \{0, 1\}$ (Round 2, StdMajority)

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
	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>						

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
	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>					

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
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>					

Once you have made your decisions, please tick below:

These are my final answers. Please proceed.

Figure B.17: Screenshot: Beliefs about Others (Round 2, StdMajority)

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

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

	% 0-9	% 10-19	% 20-29	% 30-39	% 40-49	% 50-59	% 60-69	% 70-79	% 80-89	% 90-100
	<input type="radio"/>									

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

	% 0-9	% 10-19	% 20-29	% 30-39	% 40-49	% 50-59	% 60-69	% 70-79	% 80-89	% 90-100
	<input type="radio"/>									

Once you have made your decisions, please tick below:

Figure B.18: Screenshot: Beliefs about Intervention (Round 2, Pay4Vote)

References (Online Appendix)

- 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.
- Bolton, Gary E., and Axel Ockenfels.** 2000. “ERC: A Theory of Equity, Reciprocity, and Competition.” *American Economic Review*, 90(1): 166–193.
- Brock, J. Michelle, Andreas Lange, and Erkut Y. Ozbay.** 2013. “Dictating the Risk -Experimental Evidence on Giving in Risky Environments.” *American Economic Review*, 103(1): 415–437.
- Dal Bó, Pedro, Andrew Foster, and Kenju Kamei.** 2019. “The Democracy Effect: a weights-based identification strategy.” NBER Working Paper No. 25724.
- Fehr, Ernst, and Klaus M Schmidt.** 1999. “A Theory of Fairness, Competition, and Cooperation.” *Quarterly Journal of Economics*, 817–868.
- Freundt, Jana, and Andreas Lange.** 2017. “On the Determinants of Giving under Risk.” *Journal of Economic Behavior & Organization*, 142: 24–31.
- Saito, Kota.** 2013. “Social Preferences under Risk: Equality of Opportunity versus Equality of Outcome.” *American Economic Review*, 103(7): 3084–3101.