ECONtribute
Discussion Paper No. 068

Social Norms and Elections: How Elected Rules Can Make Behavior (In)Appropriate

Arno Apfelstaedt   Jana Freundt   Christoph Oslislo

February 2021   www.econtribute.de

UNIVERSITÄT BONN   Universität zu Köln

Social norms and elections:
How elected rules can make behavior (in)appropriate[*]

Arno Apffelstaedt   Jana Freundt   Christoph Oslislo

This Version: 27th February 2021

Abstract

Can elections change people's ideas about what is ethically right and what is wrong? A number of recent observations suggest that social norms can change rapidly as a result of election outcomes. We explore this conjecture using a controlled online experiment. In our experiment, participants rate the social appropriateness of sharing income with poorer individuals. We compare situations in which a rule has been elected that asks people to share or not to share, respectively, with situations in which no rule has been elected. In the absence of an election, sharing is widely considered socially appropriate, while not sharing is considered socially inappropriate. We show that elections can change this social norm: They shift the modal appropriateness perception of actions and, depending on the elected rule, increase their dispersion, i.e. erode previously existing consensus. As a result, actions previously judged socially inappropriate (not sharing) can become socially appropriate. This power prevails, albeit in weaker form, even if the election is subject to controversial practices such as vote buying or voter disenfranchisement. Drawing on behavioral data from another experiment, we demonstrate that election-induced norm shifts predict behavior change.

Keywords: social norms, elections, prosocial behavior, rule compliance
JEL Codes: D02, D91, C91

1 Introduction

Can elections shift people's ideas of what is ethically right and what is wrong? A number of recent observations suggest that social norms, often considered persistent long-standing social constructs, can change rapidly as a result of election outcomes. In 2016, shortly after the United Kingdom voted for Brexit, the country experienced a sharp rise in hate crime, which many observers attribute to a Brexit-induced increase in the social acceptability of xenophobic views and actions. As a result of the referendum, "anti-immigrant and anti-foreigner rhetoric had become 'normalised' ", making Britain effectively "a more racist country", the United Nations claim.[1] Similar claims were made after the election of Donald Trump as president of the United States that same year.[2] There are also examples where elections appear to have lead to more tolerant norms: Baskaran and Hessami (2018) and Kedia and Pareek (2020) observe that electoral successes of female candidates in Germany and the US, respectively, seem to have improved norms regarding the treatment of women and reduced bias against them in elections as well as in the workplace. Jung and Tavits (2021) document an increase in the social acceptability of abortion following a pro-abortion vote in the Irish referendum of 2018.

Social norms—defined as "shared understandings about actions that are obligatory, permitted, or forbidden" (Ostrom, 2000, p.144)—govern many parts of our everyday lives, ranging from economic and political decisions to cultural practices and are thus an important element of any social group. If the general conjecture that elections can influence norms is true, it most likely holds for other domains as well, be it norms regarding same-sex relationships (Aksoy et al., 2020), compliance with Covid-19 regulations (Galbiati et al., 2020), or everyday behavior such as alcohol consumption (Lane and Nosenzo, 2019). However, a fundamental challenge is that elections are typically not exogenous to the society in which social norms evolve and thus, observed differences in social norms may as well be attributed to other (potentially unobservable) factors that correlate with the election outcome, which makes it difficult to establish causality.

In this paper, we provide causal evidence on the effect of elections on social norms using a simple experiment. Our experiment abstracts from the complexity of a real world election and introduces a majority vote on a simple rule of behavior. The elected rule targets a rather general domain in which social norms have been shown to be highly important for behavior, namely prosociality. A major contribution of our study is that we provide *direct* clean evidence of whether social norms can change as a result of elections, thereby complementing previous studies that cannot directly identify social norms and disentangle their impact from, for example, personal moral views or changes in

---

[1]See Brown, D. and Coates, S., "UK 'more racist after Brexit'", *The Times*, May 12, 2018. Available at: `https://www.thetimes.co.uk/article/uk-more-racist-after-brexit-qb7hd7xl7` (accessed February 18, 2021).

[2]Many popular press articles reported an increase in racially motivated violence and sexism after Trump's election, the cause of which was attributed to a change in social norms. The election of a person who is "openly hostile to women [...] normalizes abusive behavior and gives implicit permission for others to perpetuate it", the Huffington Post (Nov 16, 2016), for instance, writes. See Jeltsen, M., "Trump's Election Raises Fears Of Increased Violence Against Women", *The Huffington Post*, Nov 16, 2016. Available at: `https://www.huffpost.com/entry/trump-women-rights-violence-fears_n_582a0f63e4b02d21bbc9f186` (accessed February 18, 2021).

preferences.

Our experimental setting is as follows. For each treatment, 100 international subjects are recruited via the online platform Prolific. Among the subjects of each treatment, we distribute experimental income unequally, making 50% of the subjects "rich" and 50% "poor". We then ask subjects to rate, on a scale from *very socially inappropriate* to *very socially appropriate*, two opposing actions, Give and Don't Give, where "Give" means that a rich subject shares her income with a randomly selected poor subject and "Don't Give" means that the subject does not share her income. Following the elicitation method proposed by Krupka and Weber (2013), subjects are incentivized to provide a rating that is identical with the most common rating in their session, making the elicited social approval a direct elicitation of social norms or "normative expectations".

In our main treatment (T_StdMajority), we elicit social approval ratings conditional on a simple majority vote among participants having elected a behavioral rule. The elected rule tells subjects either that "everybody should choose *Give*" (Rule:Give) or that "everybody should choose *Don't Give*" (Rule:Don't). To identify the effect of this election on social norms, we compare the ratings elicited in T_StdMajority to the ratings in a benchmark treatment T_NoRule in which subjects rate actions Give and Don't Give in the absence of an election. We find that the election indeed has a strong impact on social norms. Most impressively, and much in line with the anecdotal evidence reported above, we show that majority-elected rules can cause actions previously judged socially inappropriate (Don't Give) to become socially appropriate. This is the case, specifically, if the elected rule asks subjects to not give (Rule:Don't). We also find a statistically significant effect on social norms of Rule:Give being elected into power, but the size of the effect is much smaller.

Having established this main result, we answer two additional questions. First, we ask whether norm shifts require "free and fair" majority votes. To answer this question we conduct three additional treatments that describe a situation in which votes have been manipulated by asking voters to pay a fee, by offering them money to change their vote or by excluding poor voters. After a manipulated election, elected rules still shift social norms about prosocial actions, but their power to do so is weakened. Finally, we ask the question of whether the norm shifts we observe in our experiment can predict changes in behavior. Drawing on behavioral data from a previous experiment in Apffelstaedt and Freundt (2020), we show that social norms indeed predict giving decisions under each elected rule in a different sample of participants.

In a separate section, we discuss our results and relate them to possible channels of how elections can change normative expectations. As a first channel, we discuss the possibility that elections carry informative value about the underlying preferences in society. As a second channel, we discuss the possibility that showing compliance with the elected rule (irrespective of its content) may be a source of social approval. Our discussion leads to an important question: Do elections always lead to more agreement on social norms? In our experiment, in addition to observing shifts in modal appropriateness, we observe that elections can lead ratings to become more dispersed. In other words, elections can erode previously existing norm consensus. Both the election-induced mean

shift and the erosion of consensus could be responsible for affecting people's moral behavior.

Related Literature.   In this paper, we focus on so-called *injunctive norms* or normative expectations in a population.[3] Building on work by Cialdini and Trost (1998) and Ostrom (2000), Krupka and Weber (2013) define injunctive social norms as collective perceptions or judgments regarding the appropriateness of actions. This requires that norms are "jointly recognized, or collectively perceived, by members of a population" (Krupka and Weber, 2013, pp.498-499).

In this context, our results contribute to four different literatures. First, our work complements a handful of previous papers that examine factors in the institutional environment that can cause long-standing social norms to change. For instance, recent research has shown that social norms can change quickly as a result of policy interventions. One example are studies of so-called "norm-nudges" (for overviews see Bicchieri and Dimant, 2019; Hauser et al., 2018). In particular, our work adds to a recent literature that examines how public decision-making processes (such as elections, initiatives, or referenda) can lead to very rapid changes in social norms, which in turn can lead to changes in behavior. Jung and Tavits (2021) argue, based on results of a panel survey, that the outcome of the 2018 Irish abortion referendum changed Irish citizens' perceptions of the social norm regarding abortion. Baskaran and Hessami (2018) and Kedia and Pareek (2020) show that elections have the potential to affect norms regarding gender as well as gender-related behavioral outcomes. Using an experimental setup, Bursztyn et al. (2020a) show that Donald Trump's victory in the 2016 Presidential election increased individuals' willingness to publicly express xenophobic views as well as accept related expressions by others.[4] In a similar vein, Albornoz et al. (2020) argue that the increase in hate crime following the Brexit referendum should be attributed to a change in social norms: They show that hate crime spiked especially in regions in which the outcome of the election came as a surprise and thus, can be explained by an update of beliefs about whether xenophobic views are extreme or mainstream (Albornoz et al., 2020). In the existing studies, the effect of elections on social norms is inferred indirectly from observed behavior or revealed preferences. Our experimental design instead allows us to *directly* elicit social norms and thus examine the immediate response of norms to elections. Using a well-established norm elicitation method (Krupka and Weber, 2013), we highlight the role of shifts in the social appropriateness of actions in bringing about behavioral change. To our knowledge, our paper is the first to directly measure changes in social norms in response to the election of a behavioral rule.

Second, our paper contributes to a growing experimental literature on the effect of social norms on behavior. This literature assumes that most individuals tend to learn and follow social norms, leading, for instance, to a willingness to constrain selfish behavior (Ostrom, 2000, p.143,149). Following this conjecture, a number of recent experiments show that many people do indeed have

---

[3]We do not consider descriptive norms or empirical expectations, i.e. expectations about what others actually do. For a discussion of both concepts see also Bicchieri (2010, 2016); Bicchieri and Dimant (2019). Several studies have shown that both, injunctive and descriptive norms, can influence behavior (e.g., Bicchieri and Chavez (2010); Bicchieri et al. (2020a); Bursztyn et al. (2020b); Cialdini et al. (1990); Krupka and Weber (2009)).

[4]For related results see Crandall et al. (2018) and Huang and Low (2017).

an intrinsic preference to conform to what is collectively perceived as socially appropriate, and that norm conformity can explain behavior in a variety of social contexts (Kimbrough and Vostroknutov, 2016; Krupka et al., 2017; Gaechter et al., 2017). Krupka and Weber (2013) find that social norms vary with different framings in dictator games and argue that this variation can provide a plausible explanation for observed differences in behavior. By providing information about the moral views of others, Bursztyn et al. (2020b) manage to directly manipulate perceived social norms and show that this in turn changes behavior. Importantly, these studies suggest that there is a fairly stable preference for following social norms across different settings. Our results are consistent with these findings: In our experiment, we observe systematic shifts in individual perceptions of social norms and present evidence that—by assuming a stable preference for following social norms exists—these shifts are well suited to explain behavioral changes after elections. Note that we are primarily interested in *changes* in perceived social norms in a society, which is conceptually distinct from studies that examine changes in individual adherence to existing *stable* norms (e.g. Bicchieri et al., 2020b).

Third, our work adds to the literature on determinants of prosocial behavior and the analysis of the broader set of motives that shape people's social conduct. Norms and social pressure have been found to be important driving factors of altruistic behavior by attaching honor to good deeds and shame to selfish behavior (Bénabou and Tirole, 2006). A number of experimental studies has since then investigated the crucial role of social norms for prosocial behavior in various contexts (Krupka and Weber, 2009; Gaechter et al., 2012; Kimbrough and Vostroknutov, 2016). Our results confirm the view that prosocial behavior can be highly context-dependent and that a shift in social norms can lead to large shifts in prosocial outcomes. We add to the literature by showing that norms regarding prosocial behavior can be influenced by elections.

Fourth, we link to a recent literature that examines how rules and laws can change attitudes and social norms and thus influence behavior beyond the imposition of explicit sanctions. Using survey data, Galbiati et al. (2020) show that social distancing rules during the Covid-19 pandemic causally affected social norms regarding social interactions. Several other studies empirically investigate how laws shape attitudes on morally controversial issues, e.g., by studying the efficacy of anti-discrimination laws (Aksoy et al., 2020; Barron and Hebl, 2010, 2013). However, these studies are usually not able to disentangle precise channels to explain where changes in behavior or personal opinions stem from. An exception is Lane and Nosenzo (2019), who provide direct evidence that the legal status of an action causally affects its normative appropriateness. We investigate how *democratically elected* rules, in particular, affect social norms and thus behavior.

The paper proceeds as follows. In the next section, we explain the experimental setup in detail. In section 3 we present our results. In section 4, we discuss our findings with a focus on possible mechanisms and ways in which elections can influence and change norms. Section 5 concludes.

2 The Experiment

We design an experiment to analyze the influence of majority-elected rules on social norms. In our experiment, subjects have to rate two actions, Give and Don't Give, on a scale from very socially inappropriate to very socially appropriate. We are interested in how ratings are influenced by the majority election of a behavioral rule which either demands that "*everyone should choose GIVE*" (Rule:Give) or that "*everyone should choose DON'T GIVE*" (Rule:Don't). To be able to interpret the elicited ratings as a social norm, we use the elicitation method suggested by Krupka and Weber (2013): Subjects are told that "*by socially appropriate, we mean behavior that most people agree is the 'correct' or 'ethical' thing to do*" and are incentivized to provide a rating that is identical with how most of the other subjects in their treatment evaluate the action.

We run four election treatments and one benchmark treatment (see Table 1). In the election treatments, subjects are asked to rate actions Give and Don't Give under the assumption that Rule:Give or Rule:Don't has been elected into power. Between treatments, we vary the specifics of *how* the rule is elected. In our main treatment (T_StdMajority), we ask subjects to assume that the rule is elected by a simple majority vote, counting the votes of all 100 subjects who take part in the experimental session. The other three election treatments (T_Pay4Vote, T_MoneyOffer, and T_ExcludePoor) allow us to investigate whether social norms only shift in response to "free and fair" elections. Here, we consider majority votes that are subject to controversial practices which likely result in a subset of votes either being manipulated or not counted in the election. Finally, treatment T_NoRule is a benchmark treatment which provides us with social appropriateness ratings (i.e., social norms) of actions Give and Don't Give in the absence of an election. All treatments are between-subjects. We expect behavioral rules implemented by a referendum to have predictive power for the perceived social appropriateness of the respective behavior after the election. Specifically, we predict that the referendum will shift upward the social approval of actions that comply with the elected rule and will shift downward the social approval of actions opposed to the elected rule. Intuitively, we also expect that referenda that can be perceived as being less democratic will have a reduced (moral) power to shift social norms.

Treatment	Election?	Description	n
T_NoRule	No	Rating of actions in the absence of an election (benchmark rating)	100
T_StdMajority	Yes	Rules elected by standard majority vote (all votes counted)	100
T_Pay4Vote	Yes	Voters have to pay £0.20 to make vote count	100
T_MoneyOffer	Yes	Voters are offered £0.20 to vote for the opposite rule	100
T_ExcludePoor	Yes	Voters with household income < GBP 40K excluded from ballot	100

Table 1: Overview of Treatments

Our main predictions guiding the design and analysis of our experiment are as follows:

Prediction 1. *Majority-elected rules shift social norms. Relative to T_NoRule, the election of Rule:Give (Rule:Don't) will shift upward (downward) the social approval rating of action Give and will shift downward (upward) the social approval rating of action Don't Give.*

Prediction 2. *"Free and fair" elections have the strongest power to change norms. Relative to T_NoRule, the effect of elected rules on social approval ratings will be largest in T_StdMajority.*

2.1 Design

The experimental design builds on Apffelstaedt and Freundt (2020). In Apffelstaedt and Freundt (2020), we investigate the extent to which rules enacted through (more or less) democratic elections affect people's behavior. For the present paper, we adapted the design of the experiment to be able to investigate whether elections shift social norms.

Give or Don't Give. Social norms in our experiment concern choices in a simple behavioral paradigm. In a group of 100 subjects, income is distributed unequally. Before learning whether one is rich or poor, each subject has to decide privately whether to *Give* or *Don't Give*, where "Give" means that, conditional on being rich, the person shares her income with another (poorer) subject and "Don't Give" means that the subject does not share her income. We operationalize this paradigm using a lottery: In each treatment, we raffle a cash prize of £100 among the 100 participating subjects. At the beginning of the experiment, subjects learn that lottery tickets for the raffle will be distributed unequally: While 50 subjects receive 10 lottery tickets each, the remaining 50 subjects receive no (zero) lottery tickets. Actions Give and Don't Give are then introduced as follows: "*If you happen to be a receiver of lottery tickets, 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?*"

Social Norms. We elicit social norms using the coordination game method suggested by Krupka and Weber (2013). Following their definition, a social norm is an empirically measurable collective judgment that assigns to each action a degree of appropriateness or inappropriateness. For each of the two possible actions, Give and Don't Give, we ask subjects to evaluate the "social appropriateness" of the action on a 6-point scale. The scale allows subjects to evaluate the action negatively as "very socially inappropriate", "socially inappropriate", or "somewhat socially inappropriate", or positively as "somewhat socially appropriate", "socially appropriate", or "very socially appropriate".[5] Two measures are taken to ensure that the elicited rating reveals a social norm—that is, a coordinated belief about what is wrong and what is right: First, we tell subjects that "*by socially appropriate, we mean behavior that most people agree is the 'correct' or 'ethical' thing to*

---

[5]We chose to omit a "neutral" category from the scale for the same reason as Krupka and Weber (2013), namely that such a midpoint would represent an artificial focal point that is separate from the focal point arising from the social norm, and thus, could interfere with the coordination incentive.

do". Second, subjects are incentivized to provide a rating that is identical with how most of the other subjects evaluate the action. Specifically, we pay the subject a bonus payment of £2.00 if, for a randomly selected rating, the subject's rating matches the modal rating among the 99 other subjects in her treatment.[6]

Elections: Rule:Give vs. Rule:Don't. Our prime interest is in understanding how social norms regarding Give and Don't Give change in response to a behavioral rule being elected in a (more or less) democratic manner. In our four election treatments, we therefore ask subjects to evaluate actions Give and Don't Give assuming that a "*code of conduct*" has been elected by participants demanding either that "*everyone should choose GIVE*" (Rule:Give) or, contrarily, that "*everyone should choose DON'T GIVE*" (Rule:Don't). Subjects are informed that the final rule is non-binding: "*Once a rule has been set, each individual can decide privately and anonymously whether he/she wants to follow the rule or not.*" Treatments differ as to the exact manner in which the rule is presumed to have been voted into power.

Treatments. There are five between-subjects treatments, each with 100 subjects (see Table 1). In a benchmark treatment (T_NoRule), we ask subjects to evaluate actions Give and Don't Give without a behavioral rule being in place. In this treatment, we neither talk about an election nor about a rule, but simply let subject evaluate each action in a neutral setting. In our four election treatments, subjects evaluate actions Give and Don't Give conditional on Rule:Give or Rule:Don't being elected. In T_StdMajority, the rule is assumed to have been elected with a "free and fair" majority vote: Subjects are told that "*all 100 individuals who take part in the lottery are asked to vote for the rule they prefer to have implemented as the code of conduct. The rule that receives more votes in total will be implemented as the code of conduct.*" In T_Pay4Vote, subjects are presented with the same statement, to which we add, however, that "*only the votes of individuals who pay £0.20 will be counted*". In T_MoneyOffer, we instead add the information that before the votes are finally counted, "*all individuals are offered an additional payment of £0.20 to vote for the rule that is opposite to the one they originally intended to vote for*". Finally, in T_ExcludePoor, we add the information that "*only the votes of individuals with a household income above £40,000 are counted*". Participants are not informed about the share of voters who decide to pay the fee in T_Pay4Vote, about the share who accept the bonus payment in T_MoneyOffer, or about the share whose votes are excluded due to their household income in T_ExcludePoor.

---

[6]The definition and measurement of social norms suggested by Krupka and Weber (2013) differs from previous binary conceptions of social norms by allowing for actions to vary in the degree to which they are regarded as socially (in)appropriate. This is especially important for our purpose of studying changes in these perceptions dependent on the election procedure. We speculate that focal points in the coordination game might be created via "cultural values" such as prosociality, obedience to societal rules or the importance of democratic principles.

2.2 Procedures

The exact setup of the experiment and the wording of instructions can be found in appendix A.2. Below, we summarize the procedures of the election treatments and of the benchmark treatment.

Procedure of Election Treatments. Each session consist of two tasks in the same fixed order. In task 1, subjects are informed about the raffle and the fact that lottery tickets are distributed unequally among the 100 participants of the treatment. Subjects are then asked to privately decide whether—in the case of being awarded 10 tickets—they want to Give or Don't Give 3 of their tickets to a randomly selected nonreceiver. There is no mention of an election (nor of a rule) at this stage of the experiment. Subjects' decisions in this part of the experiment determine the final allocation of lottery tickets and thus each participant's chances of winning the £100 cash prize. The winner of the raffle is drawn after the experiment. Subjects are not informed about the choices of other participants in their treatment. In part 2 of the experiment, subjects are confronted with a "*hypothetical choice situation*", in which they are instructed to assume that a "*code of conduct*" has been voted into power that either demands that "*everyone should choose GIVE*" (Rule:Give) or, contrarily, that "*everyone should choose DON'T GIVE*" (Rule:Don't). The details of the voting procedure depend on the treatment (see above). We then ask them to rate the social appropriateness of actions Give and Don't Give conditional on Rule:Give or Rule:Don't being elected. Subjects can earn a bonus payment of £2.00 for matching the modal rating among the 99 other participants in the treatment. This concludes the experiment.

Procedure of Benchmark Treatment (NoRule). This treatment consists of one task. At the beginning of the session, subjects are informed about the raffle and the fact that lottery tickets are distributed unequally among the 100 participants of the treatment. There is no decision to be made: The initial allocation of lottery tickets by the computer determines each participant's final chances of winning the £100 cash prize. Subjects are then confronted with a "hypothetical choice situation", in which they are instructed to assume that each participant would be asked to privately decide whether—in the case of being awarded 10 tickets—they want to Give or Don't Give 3 of their tickets to a randomly selected nonreceiver. There is no mention of an election (nor of a rule). Instead of making the choice themselves, we ask them to rate the social appropriateness of actions Give and Don't Give. Subjects can earn a bonus payment of £2.00 for matching the modal rating among the 99 other participants in the treatment. This concludes the experiment.

2.3 Implementation

The experiment was programmed using LimeSurvey and conducted on the online survey platform Prolific using a randomly drawn sample of international participants. Prolific automatically provides us with basic (self-declared) demographic information about individual subjects. Additional

to this basic information, which includes gender, age, and student status, we required that participants had filled in information about their nationality and country of residence. The four election treatments were conducted over a period of two weeks in September 2018. On average, subjects spent about 15 minutes to go through the experiment. In addition to the chance to win a cash prize of £100 and a possible bonus payment of £2.00 in the social norm task, subjects received a base payment of £1.60 for completing the experiment. The benchmark treatment (NoRule) was conducted as a separate control treatment in November 2020. Since this experiment took only 10 minutes to complete, we reduced the base payment to £1.10. Data collection for each of the five treatments was preset to stop when the number of subjects reached 100.

The entire sample of 500 participants has a mean age of 28.73 years (SD 9.59), 46.60 percent of participants are female, and 38.87 percent are students. The largest share of participants have a British nationality (38.08 percent), followed by 11.62 percent US Americans.[7] The share of "Non-Western" subjects is 21.84 percent.[8]

3 Results

3.1 Do elections shift norms? StdMajority vs. NoRule

To what extent can elections change social norms? We begin our analysis by comparing social approval ratings after a "free and fair" majority election of Rule:Give or Rule:Don't (T_StdMajority) with those elicited in the absence of an election (T_NoRule).

Figure 1 displays the mean and median of social approval ratings across treatments T_NoRule and T_StdMajority. The left-hand side of the figure shows how subjects rate action Give in the absence of an election (T_NoRule), when Rule:Give is elected by majority vote (T_StdMajority) and when Rule:Don't is elected by majority vote (T_StdMajority). The right-hand side of the figure (grey bars) shows how subjects rate action Don't Give under the same conditions. Following the convention introduced by Krupka and Weber (2013), we have converted subjects' responses into numerical scores. A rating of "very socially inappropriate" received a score of -1, "socially inappropriate" a score of -2/3, "somewhat socially inappropriate" a score of -1/3, "somewhat socially appropriate" a score of 1/3, "socially appropriate" a score of 2/3, and "very social appropriate" a score of 1.

Recall our main prediction (Prediction 1): We predicted that majority-elected rules will shift upward the social approval of actions that comply with the rule and will shift downward the social approval of actions opposed to the rule. Specifically, we predicted that the election of Rule:Give (Rule:Don't) will shift upward (downward) the social approval of action Give and will

---

[7]Not every participant had filled out all questions about demographics. Of 500 subjects, only 494 subjects filled in information on their student status and only 499 filled provided their nationality.

[8]Western = 1 if Nationality = United Kingdom (190 participants), United States (58), Austria (4), Australia (7), Belgium(6), Canada (19), Denmark (4), Finland (5), Germany (11), Greece (12), Ireland (4), Italy (30), Netherlands (8), Norway (1), New Zealand (1), Portugal (25), Sweden (4), Switzerland (1).
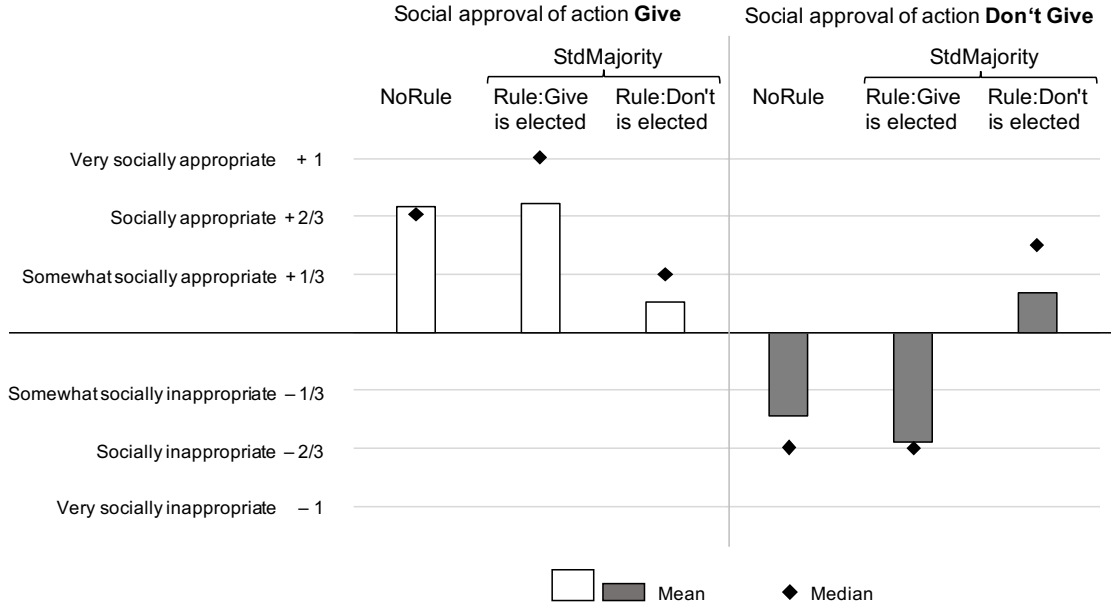
Figure 1: Elicited social approval (mean and median) of actions Give (left panel) and Don't Give (right panel) in the absence of an election (T_NoRule) and after a standard majority election of Rule:Give or Rule:Don't (T_StdMajority).

shift downward (upward) the social approval of action Don't Give. This is exactly what we find. Consider first action Give (left-hand side of Figure 1): In the absence of an election (NoRule), the action is rated as "socially appropriate" (mean: .72, median: .67). The election of Rule:Give shifts the rating of action Give moderately upward toward "very socially appropriate" (mean: .74, median: 1.00), whereas the election of Rule:Don't leads to a strong downward shift to at or below "somewhat socially appropriate" (mean: .17, median: .33). Although the mean and median ratings of action Give remain positive throughout, close to 40% of subjects rate action Give *negatively* under Rule:Don't, which is an increase of 35 and 32 percentage points, respectively, compared to NoRule and Rule:Give. We find a similar, flipped version of this pattern for action Don't Give (right-hand panel of Figure 1): In the absence of an election (NoRule), action Don't Give is rated as moderately socially *in*appropriate (mean: -.48, median: -.67). The election of Rule:Give pushes this rating slightly further into the negative (mean: -.63, median: -.67), whereas the election of Rule:Don't produces a strong effect in the opposite direction. Under Rule:Don't, mean and median social approval of action Don't Give are *positive* (mean: .23, median: .50), *and even higher* than that of the opposite action Give. Two thirds (66%) of subjects rate action Don't Give positively under Rule:Don't, an increase of 54 and 58 percentage points, respectively, compared to NoRule and Rule:Give. This finding shows that a majority election can cause actions previously judged socially inappropriate (Don't Give) to become socially appropriate.

Detailed information on the distributions of approval ratings can be found in Table 2. Similar shifts to those observed in the mean and median are also observed in the modal rating of actions

	T_StdMajority				T_NoRule	
	Rule:Give		**Rule:Don't**		**NoRule**	
Rating	Give	Don't Give	Give	Don't Give	Give	Don't Give
− − −	5%	36%	8%	10%	0%	15%
− −	1%	41%	13%	11%	2%	41%
−	0%	15%	17%	13%	1%	32%
+	5%	4%	24%	16%	11%	8%
+ +	34%	1%	20%	32%	49%	4%
+ + +	55%	3%	18%	18%	37%	0%
Mean	.74	−.63	.17	.23	.72	−.48
Median	1.00	−.67	.33	.50	.67	−.67
Rule vs. NoRule	−2.24**	3.68***	6.34***	−7.00***	(Rank-sum test (z))	
Rule:Don't vs. Rule:Give			6.41***	−7.48***	(Signed rank test (z))	

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$; all two-tailed.

Ratings are: "very socially inappropriate" (− − −), "socially inappropriate" (− −),"somewhat socially inappropriate" (−), "somewhat socially appropriate" (+), "socially appropriate" (+ +), "very socially appropriate" (+ + +); modal ratings are shaded. For means and medians, responses are converted into numerical scores −1 (− − −), −2/3 (− −), −1/3 (−), +1/3 (+), +2/3 (+ +), +1 (+ + +).

Table 2: Elicited social approval (full distribution and non-parametric tests) of actions *Give* and *Don't Give* in the absence of an election (T_NoRule) and and after a standard majority election of Rule:Give or Rule:Don't (T_StdMajority).

Give and Don't Give (shaded values in Table 2). Non-parametric rank-sum and signed rank tests reported at the bottom of the table verify that social norms are significantly altered by the outcome of the majority election: Compared to NoRule, Rule:Give significantly shifts upward the social approval of action Give ($z = −2.24$, $p = .025$) and significantly shifts downward the social approval of action Don't Give ($z = 3.68$, $p < .001$). Analogously, Rule:Don't significantly shifts downward the social approval of action Give ($z = 6.34$, $p < .001$) and significantly shifts upward the social approval of action Don't Give ($z = −7.00$, $p < .001$).[9] We summarize our findings below:

Result 1. *Majority-elected rules (T_StdMajority) shift social norms. The election of Rule:Give makes action Give (Don't Give) more (less) socially appropriate. The election of Rule:Don't makes action Don't Give (Give) more (less) socially appropriate. Majority-elected rules can cause actions previously judged socially inappropriate (Don't Give) to become socially appropriate.*

3.2 Do norm shifts require "free and fair" elections?

Having established that an inclusive and unbiased majority election has the power to shift social norms, we now ask whether this power is sensitive to controversial practices that can make the

---

[9]A notable additional observation (which we did not predict) is that social approval ratings become more dispersed following the election of Rule:Don't (see Table 2). That is, compared to the case of Rule:Give or NoRule, people seem to agree less on which action constitutes "the right thing to do" if Rule:Don't has been elected into power. One possible interpretation of this finding is that elections can sometimes lead to a fragmentation rather than to a consolidation of normative expectations. We will elaborate on this additional finding and the particular aspects of our setting that may be driving it in section 4.

election appear less democratic: Are norms affected less if voting is costly (T_Pay4Vote), voters are bribed (T_MoneyOffer), or parts of the electorate are excluded from the ballot (T_ExcludePoor)?

Figure 2 displays mean and median ratings of actions Give (left panel) and Don't Give (right panel) following the election of Rule:Give or Rule:Don't, respectively, across the four election procedures T_StdMajority, T_Pay4Vote, T_MoneyOffer, and T_ExcludePoor. In this figure, the benchmark average rating for the case where there exists no rule (Give: .72, Don't Give: −.48) is represented by a dashed line. Complementing the figure, in Table 3 we present OLS estimates of the effect of elected rules on mean approval ratings by treatment. Column (1) shows estimates for the effect of rules on the mean social approval of action Give. Column (3) shows estimates for the effect of rules on the mean social approval of action Don't Give. Here, the benchmark rating of NoRule serves as the constant. Columns (2) and (4) replicate and add individual-specific controls for gender, age, student status and country of origin.[10]
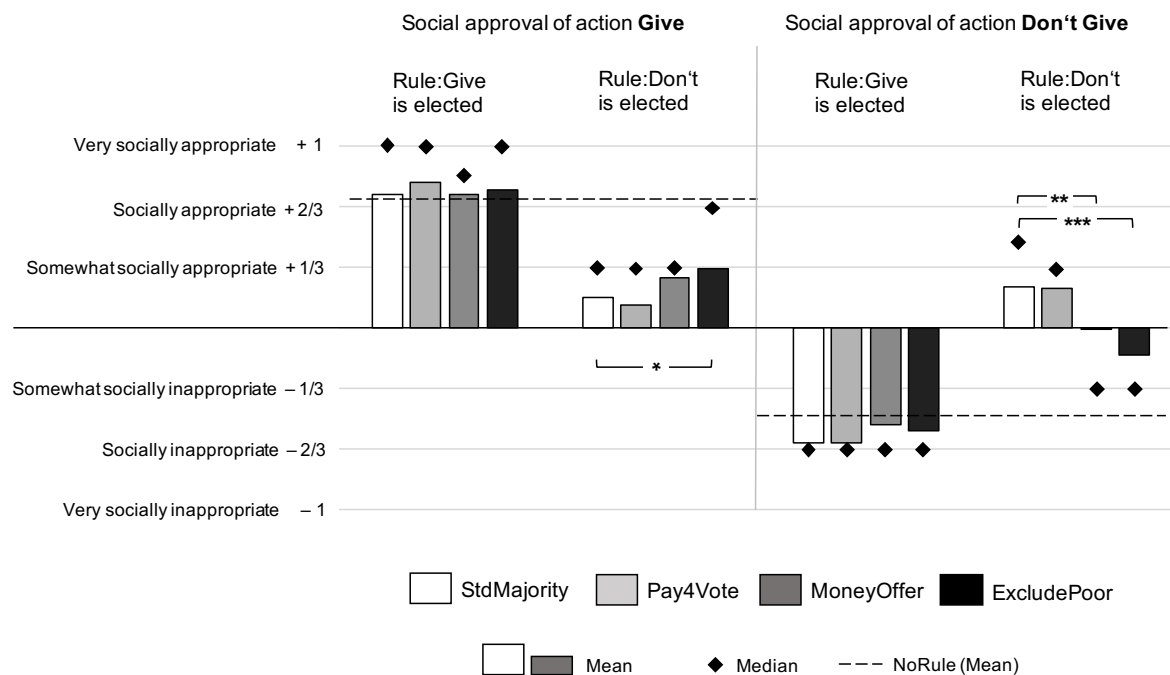
Figure 2: Elicited social approval (mean and median) of actions *Give* (left panel) and *Don't Give* (right panel) across different election procedures. Dashed line denotes mean social approval in the absence of an election (T_NoRule; Give: .72, Don't Give: -.48). Stars denote significant results of Rank-sum tests comparing the social approval after a standard majority election (T_StdMajority) with the social approval after a non-standard majority election (T_Pay4Vote, T_MoneyOffer, T_ExcludePoor): $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Analyzing Figure 2 and Table 3, we first see that *all* four election procedures have the power to change norms: Relative to the baseline rating of NoRule, the election of Rule:Give shifts the social

---

[10] The entire distribution of ratings for each of the four treatments T_StdMajority, T_Pay4Vote, T_MoneyOffer, and T_ExcludePoor, including non-parametric Rank-sum test vs. NoRule and StdMajority can be found in Table A.1 in the Appendix.

Dep. Var.	Mean social approval			
	Give		Don't Give	
	(1)	(2)	(3)	(4)
Election of **Rule:Give**				
StdMajority	.02	.03	-.16**	-.12*
	(.056)	(.059)	(.061)	(.064)
Pay4Vote	.09*	.09*	-.16**	-.12*
	(.045)	(.051)	(.064)	(.067)
MoneyOffer	.02	.02	-.06	-.02
	(.053)	(.056)	(.066)	(.067)
ExcludePoor	.04	.04	-.09	-.04
	(.049)	(.051)	(.068)	(.069)
Election of **Rule:Don't**				
StdMajority	-.55***	-.54***	.71***	.75***
	(.073)	(.075)	(.079)	(.082)
Pay4Vote	-.59***	-.59***	.70***	.74***
	(.071)	(.076)	(.076)	(.078)
MoneyOffer	-.44***	-.44***	.47***	.51***
	(.066)	(.071)	(.075)	(.077)
ExcludePoor	-.39***	-.39***	.33***	.38***
	(.070)	(.073)	(.076)	(.075)
Constant	.72	.63	-.48	-.30
	(.031)	(.080)	(.041)	(.087)
Mean rating NoRule	.72		-.48	
Controls	No	Yes	No	Yes
Observations	900	900	900	900
(Subjects)	(500)	(500)	(500)	(500)
R^2	.219	.224	.270	.279

Robust standard errors (clustered at subject level) in parentheses:
$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.
Controls in columns (2) and (4) are: female (1/0), age (continuous), student (1/0), and Western (1/0).

Table 3: Marginal effects of elected rules (Rule:Give and Rule:Don't) on mean social approval of actions Give and Don't Give across different election procedures: OLS Regressions. Mean ratings in the absence of an election (T_NoRule; Give: .72, Don't Give: -.48) serve as the benchmark.

approval of action Give (Don't Give) slightly upward (slightly downward) relative no NoRule, while the election of Rule:Don't shifts ratings strongly into the opposite direction.[11] Qualitatively, this is true for all four treatments, i.e, regardless of whether voting is costly (T_Pay4Vote), voters are bribed (T_MoneyOffer), or parts of the electorate are excluded from the ballot (T_ExcludePoor). As Table 3 shows, the effect of Rule:Don't on mean approval ratings is always highly significant, while the effect of Rule:Give is only sometimes weakly so. On average, the election of Rule:Give shifts the mean approval rating of action Give upward by .05 points and the mean approval rating of action Don't Give downward by $-.12$ points, which is about one-tenth and one-third, respectively, of a discrete step in the approval rating (where a discrete step means, e.g., going from "weakly

_____

[11]Note from Figure 2 that the median is typically (but not always) shifted by a larger amount than the mean.

socially appropriate" to "socially appropriate"). In comparison, the election of Rule:Don't leads to average shifts of $-.49$ points and $.55$ points, respectively, which converts to between one and two discrete steps on the rating scale.

A second observation we make is that not all of the elections shift norms to the same extent as the inclusive and unbiased majority election T_StdMajority. While T_Pay4Vote has virtually the same power as T_StdMajority, T_MoneyOffer and T_ExludePoor perform significantly worse in shifting social norms. Throughout Table 3, T_MoneyOffer and T_ExludePoor show systematically smaller coefficients than T_Majority. In Figure 2, stars denote statistically significant differences to T_StdMajority according to non-parametric Rank-sum tests (for detailed test results, see Table A.1 in the Appendix). We see that shifts of social approval ratings are significantly smaller following the election of Rule:Don't when this rule comes into force with an election in which voters received bribes (T_MoneyOffer) or in which voters with a low household income were excluded from the ballot (T_ExludePoor). The results are particularly pronounced and meaningful for the social approval ratings of action Don't Give (right panel in Figure 2): While in T_StdMajority and in T_Pay4Vote, the election of Rule:Don't pushes the mean and median social approval rating of action Don't Give into the positive, both mean and median rating remain negative following the election of Rule:Don't in T_MoneyOffer and T_ExcludePoor. Through these two treatments, then, one distinct aspect of the power of a free and fair election is unveiled: Elections that are perceived to be manipulated, while still powerful, can no longer cause an action previously judged socially inappropriate to become socially appropriate. We conclude:

Result 2. *Elected rules can shift social norms, but bribing voters (T_MoneyOffer) or excluding parts of the electorate (T_ExcludePoor) weaken this ability. While T_Pay4Vote has virtually the same power as T_StdMajority, elected rules shift social approval ratings significantly less in T_MoneyOffer and T_ExludePoor.*

3.3 Do election-induced norm shifts predict behavior change?

Literature on social norms typically argues that people have an intrinsic preference to conform to what is collectively perceived as socially appropriate and, in a variety of social contexts, refrain from maximizing material profits in order to comply with social norms (Elster, 1989; Bicchieri, 2006; López-Pérez, 2008; Kimbrough and Vostroknutov, 2016; Krupka et al., 2017; Gaechter et al., 2017). If elections can change norms, can these changes predict how people adapt their behavior to the election outcome? To answer this question, we draw on data from another experiment in which we elicited actual choices (Give or Don't Give) following the election instead of social approval ratings. The experiment, which was conducted in spring 2017 with a separate group of subjects on the same online platform as our rating experiment, forms the core of another paper (Apffelstaedt and Freundt, 2020). In that paper, we analyze the effects of different voting procedures on people's willingness to comply with elected rules. For a detailed description of the experiment and its results, see Apffelstaedt and Freundt (2020). In the following, we refer to this experiment as the "behavior
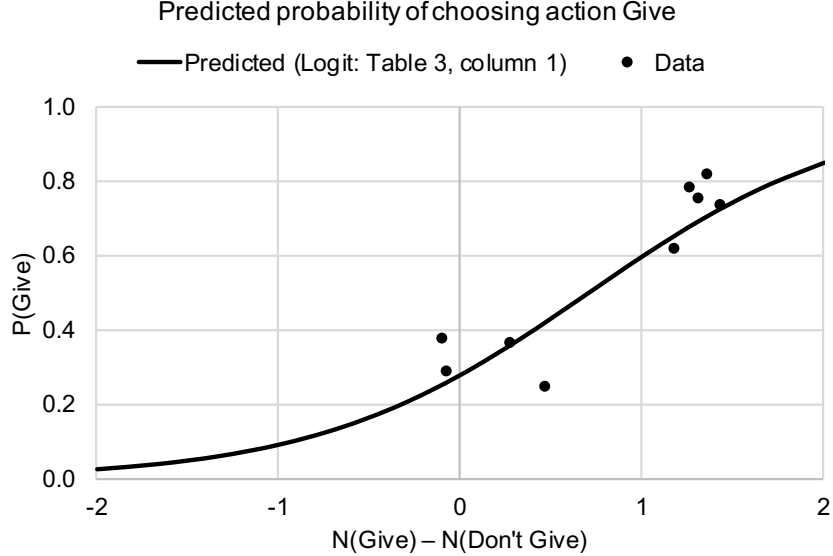
experiment".[12]

Figure 3 presents data from the actions taken in the behavior experiment: Black bars denote the share of subjects choosing action Give across treatments and in different choice situations (NoRule, Rule:Give, Rule:Don't). In the absence of an election (NoRule), 61.25% of subjects (n=400) choose to Give. The election of Rule:Give increases the share of Givers on average by 15.5 percentage points ($t = 25.11$, $p < 0.001$) to 76.75%, while the election of Rule:Don't lowers that share on average by 30 percentage points ($t = 36.29$, $p < 0.001$) to 31.25%. Thus, behavior responds significantly to rules that have been elected in a (more or less) democratic manner.
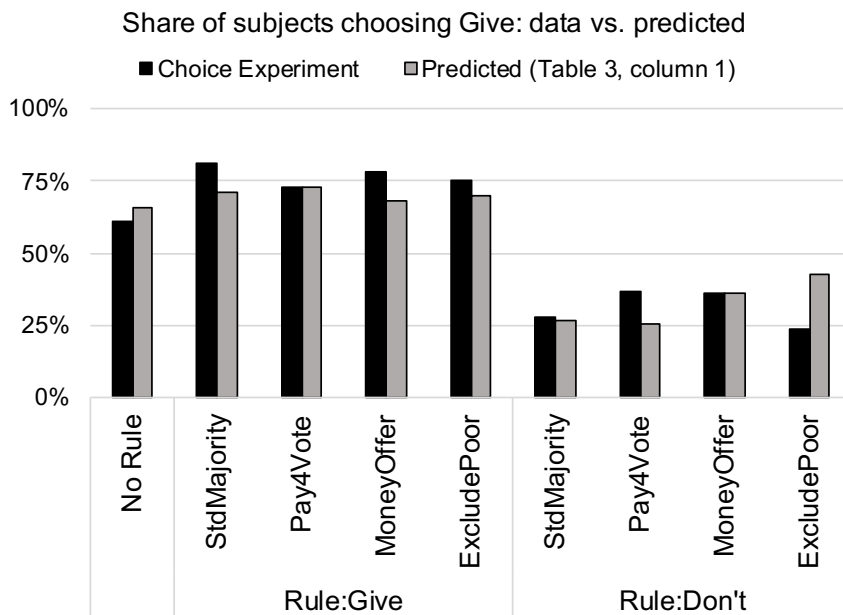
Figure 3: Share of subjects choosing action Give: data vs. predicted. Data from behavior experiment (Apffelstaedt and Freundt, 2020). Predictions using norms elicited through rating experiment (this paper), utility parameters according to specification Table 4, column (1).

How much of the change in behavior can be *directly* predicted by changes in social norms? Note that we use a setup without strategic incentives to follow a social norm. Rather, we assume that behavioral changes are driven by changes in the moral appeal associated with each action. Let us denote by $N(\text{Give}) \in [-1, 1]$ the elicited mean social approval of action Give in a given situation and by $N(\text{Don't Give}) \in [-1, 1]$ the elicited mean social approval of action Don't Give in the same

---

[12]The design of the behavior experiment is very close the rating experiment conducted for this paper. In fact, the decision situation subjects phase in the behavior experiment corresponds exactly to the hypothetical situation described to subjects in the rating experiment. Subjects first go through a choice phase (the same as task 1 in the rating experiment), in which they are asked to Give or Don't Give in the absence of a rule. Then, in part 2, subjects are first asked to vote for Rule:Give or Rule:Don't. After voting, each subject is then asked to decide between action Give or Don't Give conditional on Rule:Give or Rule:Don't being elected. Subjects know that their decision whether or not to give determines the number of lottery tickets they hold at the end of the experiment and thus their chances to win £100. Treatments are exactly the same as in our rating experiment (T_StdMajority, T_Pay4Vote, T_MoneyOffer, T_ExcludePoor).

Predicted probability of choosing action Give

Figure 4: Predicted probability of choosing action Give as a function of $N(\text{Give}) - N(\text{Don't Give})$. Predictions using norms elicited through rating experiment (this paper) and actions from behavior experiment (Apffelstaedt and Freundt, 2020). Conditional Logit prediction with utility parameters according to specification Table 4, column (1).

situation. We seek to understand how much the propensity to choose action Give over Don't Give in the behavior experiment depends on the difference in social approval, $N(\text{Give}) - N(\text{Don't Give})$, elicited in the rating experiment. For this, assume that the utility from taking action Give takes the form

$$U_{\text{Give}} = const + \gamma \cdot [N(\text{Give}) - N(\text{Don't Give})], \tag{1}$$

and normalize the utility from taking action Don't Give to zero (i.e., $U_{\text{Don't Give}} = 0$).[13] In this simple utility framework, γ measures the weight that individuals attach to norms: A positive weight γ implies a utility gain from following that action (Give or Don't Give) which yields a higher social approval. The constant (*const*) captures the average utility individuals derive from choosing action Give over Don't Give that is independent of norms. Following the procedure in Krupka and Weber (2013), we combine the data from the behavior experiment with the data from the rating experiment to estimate the parameters of the utility function using conditional Logit. The results of this estimation are found in Table 4: In column (1), we estimate γ by fitting the utility function to the share of Givers in the behavior experiment using as only explanatory variable the elicited difference in social approval, $N(\text{Give}) - N(\text{Don't Give})$. We find a large, positive and highly significant estimate, $\gamma = 1.347$ ($p < .001$). This estimate tells us that, on average, the

---

[13]Because subjects can only choose between two actions, Give and Don't Give, only differences in utility matter for decisions. The normalization of $U_{\text{Don't Give}} = 0$ is thus without loss of generality.

relative utility from taking action Give strongly increases with the difference in social approval between actions Give and Don't Give. Vice versa, if that difference in ratings becomes smaller or even turns negative as, for instance, when Rule:Don't is elected, the propensity to choose action Don't Give will become larger. Columns (2)-(4) show that the estimate of γ is robust to including demographic controls and does not vary significantly if we estimate it separately by treatment (column 3).

Dep. Var.	Utility according to Eq. (1)			
	(1)	(2)	(3)	(4)
Appropriateness rating (γ)	1.347***	1.371***	1.403***	1.444***
	(.103)	(.106)	(.143)	(.148)
Appropriateness rating X				
Pay4Vote			-.194	-.233
			(.192)	(.198)
MoneyOffer			.248	.203
			(.216)	(.221)
ExcludePoor			-.154	-.140
			(.199)	(.199)
Constant (*const*)	-.944	-1.656	-.970	-1.671
	(.125)	(.393)	(.130)	(.393)
Controls	No	Yes	No	Yes
Observations	1200	1182	1200	1182
(Subjects)	(400)	(394)	(400)	(394)
Log-likelihood	-746.8	-725.9	-742.7	-722.4

Robust standard errors (clustered at subject level) in parentheses.
$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.
Controls in columns (2) and (4) are: female (1/0), age (continuous), student (1/0), and Western (1/0).

Table 4: Conditional logit estimation of choice determinants (utility according to Eq. (1)). Choice data from behavior experiment (Apffelstaedt and Freundt, 2020). Includes mean appropriateness ratings N(Give) and N(Don't Give) from rating experiment (this paper) as explanatory variable.

To get a better sense of the estimated relationship between norms and behavior, Figure 4 plots the predicted probability of choosing action Give according to the model specification in Table 4, column (1): When there is no difference between the social approval of actions Give and Don't Give, N(Give) $- N$(Don't Give) $= 0$, P(Give) is predicted at 28%. That is, in the absence of clear guidance by a social norm, our model predicts that the majority of subjects will choose selfishly. From this position, increasing the social approval of action Give (or, equivalently, decreasing the social approval of action Don't Give) will lead on average to a 10 percentage point increase in the probability of taking action Give for every step on the rating scale (e.g., going from "weakly socially appropriate" to "socially appropriate"). This is a strong relationship: In the case of a standard majority election of Rule:Give, for instance, for which we elicit a difference in social approval of N(Give) $- N$(Don't Give) $= 1.37$, the probability of action Give is predicted at 71%—an increase of more than 40 percentage points relative to the case without normative guidance.

Finally, in Figure 3, we plot—next to the original data from the choice experiment—the share

of Givers predicted by the model specification in Table 4, column (1), for each of our experimental settings. The figure demonstrates that, overall, the simple model of norm-dependent utility specified in Eq. (1) performs well in reproducing behavioral changes across different choice situations, i.e., going from NoRule to Rule:Give to Rule:Don't. Behavioral differences between treatments *within* a given rule (within Rule:Give/Rule:Don't) are not predicted as consistently as differences between NoRule, Rule:Give and Rule:Don't. In line with the analysis in Apffelstaedt and Freundt (2020), we suggest that these differences should be attributed to intrinsic preferences regarding the rule-selection procedure rather than social norms.[14] We summarize our results on the relationship of norms and behavior as follows:

Result 3. *Election-induced norm shifts predict behavior change. Using choice data from Apffelstaedt and Freundt (2020), we predict a one-step increase in the mean approval rating of an action (e.g., going from "weakly socially appropriate" to "socially appropriate") to increase the probability of taking that action by on average 10 percentage points.*

4   Discussion

Our results show that majority-elected rules can shift social norms. Under certain circumstances, elections can cause actions previously judged socially inappropriate to become socially appropriate. This power prevails, albeit in weaker form, even if the election is subject to controversial practices such as vote buying or voter disenfranchisement. Moreover, the observed changes in norms seem to perform well in predicting how behavior will change as a result of the election.

A notable observation is that the election of Rule:Don't consistently has a much stronger effects on social norms and behavior relative to the case of NoRule than the election of Rule:Give. The shift in ratings following the election of Rule:Don't is about five times larger than the shift observed following the election of Rule:Give (see Table 3). The respective shift in behavior is about twice as large (see Figure 3). Why does the election of Rule:Don't have such a more profound relative impact on social norms than the election of Rule:Give? Below, we discuss two mechanisms of election-induced norm shifts that can explain the observed pattern. First, we discuss the possibility, formalized e.g. in Bursztyn et al. (2020a), that elections can provide information about the distribution of preferences in the population. Intuitively, if the election of Rule:Don't is more 'surprising' with regard to those preferences than the election of Rule:Give, then social norms should react stronger to the election of the former than to the election of the latter. As an alternative explanation, we discuss the possibility that the introduction of an elected rule may make social norms more complex because, once elected, behavior must also be evaluated in terms of whether it is socially appropriate to follow the rule. We argue that this added complexity may offer an explanation for an interesting, but so far little discussed, effect of elections on norms: If the rule

---

[14]For a detailed analysis of treatment effects on rule compliance (i.e. on behavior conditional on Rule:Give or Rule:Don't being elected) and possible mechanisms we refer the interested reader to Apffelstaedt and Freundt (2020).

elected contradicts a prior norm (Rule:Don't vs. NoRule), the election may lead to an erosion of social consensus about social norms. As a result, people may no longer agree on what constitutes ethically "right" behavior.

Channel 1: The informative value of elections. An intuitive explanation for the observed shifts in social appropriateness ratings is that elections cause subjects to update their beliefs about the distribution of moral views in society. To see why, note that learning that Rule:Give (Rule:Don't) has won the election implies that more (respectively, less) than 50 percent of people want a prosocial code of conduct to guide behavior. Depending on the outcome of the election, as well as prior beliefs about voting shares, this information may lead people to update their beliefs about which action is considered socially appropriate. Such updating processes play a particularly important role for social norms, which are *coordinated* beliefs about the social appropriateness of actions: The election result serves both as an informative signal about what people personally believe is appropriate and as a coordination device that helps people better agree on a common set of actions to approve or disapprove of.

How can such updating processes explain the observed differences in treatment effect size of Rule:Don't vs. Rule:Give? Ratings in T_NoRule show that the vast majority of subjects ex-ante expect others to rate the prosocial action (Give) as socially appropriate (97%) and the selfish action (Don't Give) as socially inappropriate (88%). If subjects expect voting behavior to be consistent with what people perceive as appropriate, most subjects should expect Rule:Give to be elected with a high degree of certainty. If this is the case, the actual election of Rule:Give merely confirms subjects' prior expectations, implying that social approval ratings should change only modestly. In contrast, the election of Rule:Don't would be perceived as a more surprising outcome because it comes unexpectedly given subjects' priors. As a result, social approval ratings should shift considerably, with action Don't Give winning significant approval, and action Give losing approval. This is exactly what we observe in our experiment.

Channel 2: Prosociality vs. rule compliance. There is an alternative explanation for the observed asymmetry in treatment effects. This explanation is based on the idea that the election of a behavioral rule modifies the set of relevant dimensions that must be considered when assessing the social appropriateness of actions. To get an intuition for this idea, note that in the absence of an election, the appropriateness of actions Give and Don't Give can arguably be evaluated on a one-dimensional scale—specifically, whether the action being evaluated is prosocial or not. Accordingly, in T_NoRule, we observe that the vast majority of subjects (97% and 88%, respectively) rate the prosocial action (Give) as socially appropriate and the selfish action (Don't Give) as socially inappropriate.

The election of a behavioral rule adds one further dimension to the task of evaluating the social appropriateness of actions: *How socially appropriate is it to follow the rule?* Both the prosociality of an action as well its compliance with the elected rule may now influence the final rating provided

by a subject. Consider, first, the task of evaluating actions conditional on Rule:Give being elected. Since following an elected rule might generally be considered socially appropriate, both dimensions now point in the same direction: Action Give is socially appropriate *both* from the viewpoint of prosociality as well as from the viewpoint of rule compliance. Analogously, action Don't Give is socially inappropriate according to both dimensions. Because these ratings by and large agree with how the actions are judged if prosociality is the only dimension, we should see little movement in social approval ratings between NoRule and Rule:Give. Our actual findings are consistent with this conjecture (see Figure 2). Now consider, instead, the case of Rule:Don't being elected. Here, of course, the two dimensions disagree. Depending on how strongly a subject values the aspect of rule compliance (and also how strongly the subject thinks that other people do), her ratings of actions Give and Don't Give may now change considerably relative to the case of NoRule. More precisely, the higher the subject's weight on the dimension of rule compliance, the more strongly action Don't Give wins social approval, and action Give loses social approval. Assuming that a substantial share of people consider compliance with elected rules to be an important dimension of social appropriateness, this can explain the stronger shift in appropriateness ratings in the case of Rule:Don't compared to Rule:Give.

Notice that this channel can also explain differences in the extent to which social norms are affected between T_StdMajority, T_MoneyOffer, and T_ExcludePoor. While we can only speculate about this, it seems intuitive that when evaluating the social appropriateness of actions, the more democratic the election of the rule, the greater the weight that should be given to rule compliance. Following this argumentation, controversial practices such as vote buying and the disenfranchisement of poor voters should decrease the power of elections to change social norms. In line with our results, we would expect this effect to be more pronounced in the case of Rule:Don't since it is only here where the weight attached to rule compliance can shift social approval ratings relative to NoRule.

Elections and norm consensus. In the results so far, we have focused on modal, mean, and median ratings to analyze effects of elections on social norms. Another noteworthy finding which we have only referred to in a footnote (see footnote 9) is an observation that may, prima facie, appear counterintuitive: We find that elections do not only have the potential to cause a shift in modal, median and mean social appropriateness ratings but also an increase in the variance of the distribution of individual ratings (Table 2). We interpret this finding as a decrease in *norm consensus*, i.e. in the degree to which members of a society agree on which action constitutes "the right thing to do". How can we explain the fact that social norms may become *less* clear even though elections are thought to provide *better* information about the moral preferences in society? In our data, this finding is again particularly prevalent in the case of Rule:Don't.

We can make sense of this finding by acknowledging that people may be heterogeneous when it comes to weighting the social appropriateness of following elected rules against the social appropriateness of taking prosocial actions. In the case of Rule:Don't, these two dimensions disagree regard-

ing the evaluation of actions Give (high on prosociality, low on rule compliance) and Don't Give (high on rule compliance, low on prosociality). A subject who weights rule compliance strongly (and believes other people to do so as well), will rate action Give socially inappropriate and action Don't Give socially appropriate. At the same time, however, a respondent who strongly weights prosociality over compliance will give the opposite rating. As people become unsure or begin to disagree about what the social norm is, rating decisions diverge as a consequence of the election of Rule:Don't. Ultimately, either the shift in modal, median, and mean social appropriateness ratings that we observe under Rule:Don't *or* the observed erosion of norm consensus could be responsible for why people's behavior is affected by the election. Psychological research underscores the importance of perceived social consensus or "norm clarity" in shaping one's opinions and the ability of norms to guide behavior (Lewandowsky et al., 2019; Zitek and Hebl, 2007). Thus, to fully understand how elections influence norms and how these norms in turn shape behavior, we believe it is an important task of future research to shed light on the role of norm consensus as well as the role of individual uncertainty about social norms in driving behavioral responses.

5   Conclusion

We have investigated how elected rules can affect what is perceived as socially appropriate behavior. Participants in our online experiment rate the social appropriateness of sharing versus not sharing experimental income with other participants. We find that majority-elected rules that ask people to share or not to share, respectively, can change social norms: They shift the modal appropriateness perception of actions and, depending on the elected rule, increase their dispersion, i.e. erode previously existing norm consensus. As a result, elections can cause actions previously judged socially inappropriate (not sharing) to become socially appropriate. Comparing different voting procedures, we show that the power of elections to shift social norms prevails, albeit in weaker form, if the election is subject to controversial practices such as vote buying or voter disenfranchisement. Using behavioral data from Apffelstaedt and Freundt (2020), we show that the norm shifts we observe are able to predict changes in behavior that result from the election of rules.

We hope that our paper will stimulate future research on the importance of democratic procedures in general, and elections in particular, for the formation and dissolution of social norms in a society. We especially hope for further (laboratory) experiments that attempt to disentangle the exact psychological mechanisms behind such norm changes. Our data show that elected rules not only shift modal appropriateness ratings of behavior, but can also alter their distribution and lead to an erosion of a previously existing norm consensus. We consider this observation to be of great interest and importance for understanding the dynamics of norm changes in response to democratic procedures in the form of deliberative and information aggregation processes, especially elections and referenda. We hope that future experimental studies will help to further understand the mechanisms behind this phenomenon.

In addition, the role of social norms in behavior change and in willingness to comply with newly implemented behavioral rules plays an important role in public policy. More evidence is needed to understand their effects, to assess their robustness, and to determine in which contexts social norms can predict behavior. Particularly in the context of the current Covid-19 pandemic, voluntary compliance with prosocial behavioral rules has become an important public health issue, and information on how to promote such compliance, for example by targeting social norms, can be of great importance for implementing efficient public health policies.

References

Aksoy, C.G, Carpenter, C.S, De Haas, R, Tran, K.D, 2020. Do laws shape attitudes? Evidence from same-sex relationship recognition policies in Europe. European Economic Review 124, 103399.

Albornoz, F, Bradley, J, Sonderegger, S, 2020. The Brexit Referendum and the Rise in Hate Crime; Conforming to the New Norm. Working Paper. This Version: July 31, 2020.

Apffelstaedt, A, Freundt, J, 2020. Corrupted Votes and Rule Compliance. Working Paper. This Version: February 4, 2020.

Barron, L.G, Hebl, M, 2010. Reducing "Acceptable" Stigmatization Through Legislation. Social Issues and Policy Review 4, 1–30.

Barron, L.G, Hebl, M, 2013. The force of law: The effects of sexual orientation antidiscrimination legislation on interpersonal discrimination in employment. Psychology, Public Policy, and Law 19, 191–205.

Baskaran, T, Hessami, Z, 2018. Does the election of a female leader clear the way for more women in politics? American Economic Journal: Economic Policy 10, 95–121.

Bénabou, R, Tirole, J, 2006. Incentives and prosocial behavior. American Economic Review 96, 1652–1678.

Bicchieri, C, 2006. The Grammar of Society, Cambridge: Cambridge University Press.

Bicchieri, C, 2010. Norms, preferences, and conditional behavior. Politics, Philosophy & Economics 9, 297–313.

Bicchieri, C, 2016. Norms in the wild: How to diagnose, measure, and change social norms, Oxford: Oxford University Press.

Bicchieri, C, Chavez, A, 2010. Behaving as expected: Public information and fairness norms. Journal of Behavioral Decision Making 23, 161–178.

Bicchieri, C, Dimant, E, 2019. Nudging with care: The risks and benefits of social information. Public Choice 2019, 1–22.

Bicchieri, C, Dimant, E, Sonderegger, S, 2020a. It's Not A Lie if You Believe the Norm Does Not Apply: Conditional Norm-Following with Strategic Beliefs. Working Paper. This version: October 9, 2020.

Bicchieri, C, Dimant, E, Gächter, S, 2020b. Observability, social proximity, and the erosion of norm compliance. Working Paper. This version: March 24, 2020.

Bursztyn, L, Egorov, G, Fiorin, S, 2020a. From extreme to mainstream: The erosion of social norms. American Economic Review 110, 3522–3548.

Bursztyn, L, González, A.L, Yanagizawa-Drott, D, 2020b. Misperceived social norms: Women working outside the home in saudi arabia. American Economic Review 110, 2997–3029.

Cialdini, R.B, Reno, R.R, Kallgren, C.A, 1990. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. Journal of Personality and Social Psychology 58, 1015.

Cialdini, R.B, Trost, M.R, 1998. Social influence: Social norms, conformity and compliance. In: Gilbert, D.T, Fiske, S.T, Lindzey, G (Eds.), The Handbook of Social Psychology, New York. McGraw-Hill, 151–192.

Crandall, C.S, Miller, J.M, White, M.H, 2018. Changing norms following the 2016 US presidential election: The Trump effect on prejudice. Social Psychological and Personality Science 9, 186–192.

Elster, J, 1989. Social norms and economic theory. Journal of economic perspectives 3, 99–117.

Gaechter, S, Gerhards, L, Nosenzo, D, 2017. The importance of peers for compliance with norms of fair sharing. European Economic Review 97, 72–86.

Gaechter, S, Nosenzo, D, Sefton, M, 2012. Peer effects in pro-social behavior: Social norms or social preferences? Journal of the European Economic Association 11.

Galbiati, R, Henry, E, Jacquemet, N, Lobeck, M, 2020. How Laws Affect the Perception of Norms: Empirical Evidence from the Lockdown. Working Paper. This version: July 2020.

Hauser, O.P, Gino, F, Norton, M.I, 2018. Budging beliefs, nudging behaviour. Mind & Society 17, 15–26.

Huang, J, Low, C, 2017. Trumping norms: Lab evidence on aggressive communication before and after the 2016 US presidential election. American Economic Review 107, 120–124.

Jung, J.h, Tavits, M, 2021. Do referendum results change norm perceptions and personal opinions? Electoral Studies 71, 102307.

Kedia, S, Pareek, A, 2020. Women in Politics: The Effect on Board Diversity. Working Paper. This version: December 2020.

Kimbrough, E, Vostroknutov, A, 2016. Norms make preferences social. Journal of the European Economic Association .

Krupka, E, Weber, R.A, 2009. The focusing and informational effects of norms on pro–social behavior. Journal of Economic Psychology 30, 307–320.

Krupka, E.L, Leider, S, Jiang, M, 2017. A meeting of the minds: Informal agreements and social norms. Management Science 63, 1708–1729.

Krupka, E.L, Weber, R.A, 2013. Identifying Social Norms Using Simple Coordination games: Why Does Dictator Game Sharing Vary? Journal of the European Economic Association 11, 495–524.

Lane, T, Nosenzo, D, 2019. Law and Norms: Empirical Evidence. Working Paper. This version: September 9, 2019.

Lewandowsky, S, Cook, J, Fay, N, Gignac, G.E, 2019. Science by social media: Attitudes towards climate change are mediated by perceived social consensus. Memory and Cognition 47, 1445–1456.

López-Pérez, R, 2008. Aversion to norm-breaking: A model. Games and Economic behavior 64, 237–267.

Ostrom, E, 2000. Collective action and the evolution of social norms. Journal of Economic Perspectives 14, 137–158.

Zitek, E.M, Hebl, M.R, 2007. The role of social norm clarity in the influenced expression of prejudice over time. Journal of Experimental Social Psychology 43, 867–876.

Appendix

A.1 Additional Analyses

Panel (a): Action **Give**	**Rule:Give**				**Rule:Don't**				**NoRule**
Rating	StdMaj	P4Vote	MOffer	ExPoor	StdMaj	P4Vote	MOffer	ExPoor	NoRule
− − −	5%	1%	3%	2%	8%	9%	6%	10%	0%
− −	1%	1%	2%	1%	13%	11%	4%	3%	2%
−	0%	1%	0%	1%	17%	21%	21%	13%	1%
+	5%	4%	10%	6%	24%	23%	25%	19%	11%
+ +	34%	36%	31%	40%	20%	23%	28%	37%	49%
+ + +	55%	57%	54%	50%	18%	13%	16%	18%	37%
Mean	.74	.80	.74	.76	.17	.13	.27	.33	.72
Median	1.00	1.00	1.00	.83	.33	.33	.33	.67	.67
vs. NoRule	−2.24**	−2.88***	−1.82*	−1.80*	6.34***	7.08***	5.99***	4.94***	(Rank-sum test (z))
vs. StdMaj		−0.52	0.31	0.51		0.49	−1.06	−1.80*	(Rank-sum test (z))

Panel (b): Action **Don't Give**	**Rule:Give**				**Rule:Don't**				**NoRule**
Rating	StdMaj	P4Vote	MOffer	ExPoor	StdMaj	P4Vote	MOffer	ExPoor	NoRule
− − −	36%	43%	35%	37%	10%	5%	11%	12%	15%
− −	41%	32%	25%	35%	11%	14%	15%	30%	41%
−	15%	16%	27%	15%	13%	17%	25%	17%	32%
+	4%	2%	7%	4%	16%	20%	20%	19%	8%
+ +	1%	4%	3%	5%	32%	26%	22%	16%	4%
+ + +	3%	3%	3%	4%	18%	18%	7%	6%	0%
Mean	−.63	−.63	−.53	−.57	.23	.22	−.01	−.15	−.48
Median	−.67	−.67	−.67	−.67	.50	.33	−.33	−.33	−.67
vs. NoRule	3.68**	3.87***	1.70*	2.92***	−7.00***	−7.36***	−5.30***	−3.27***	(Rank-sum test (z))
vs. StdMaj		0.51	−1.45	−.42		0.24	2.73**	4.00***	(Rank-sum test (z))

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$; all two-tailed.

Ratings are: "very socially inappropriate" $(- - -)$, "socially inappropriate" $(- -)$, "somewhat socially inappropriate" $(-)$, "somewhat socially appropriate" $(+)$, "socially appropriate" $(+ +)$, "very socially appropriate" $(+ + +)$; modal ratings are shaded. For means and medians, responses are converted into numerical scores -1 $(- - -)$, $-2/3$ $(- -)$, $-1/3$ $(-)$, $+1/3$ $(+)$, $+2/3$ $(+ +)$, $+1$ $(+ + +)$.

Table A.1: Elicited social approval of actions *Give* (panel a) and *Don't Give* (panel b) across all treatments.

A.2 Experimental Instructions

Welcome and Consent Form

This study is hosted by the University of Hamburg [Fribourg/Cologne].

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.

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.

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 GBP 1.60 for completing the study.

- One participant will receive an extra cash prize of GBP 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 GBP 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.co 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.

Instructions about the Lottery

- 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 GBP 100. The final distribution of lottery tickets depends partly on luck and partly on the choices you and other participants make.

- You will begin with task 1 on the next screen.

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

☐ I have read the information and want to proceed.

Instructions about the Distribution of Lottery Tickets

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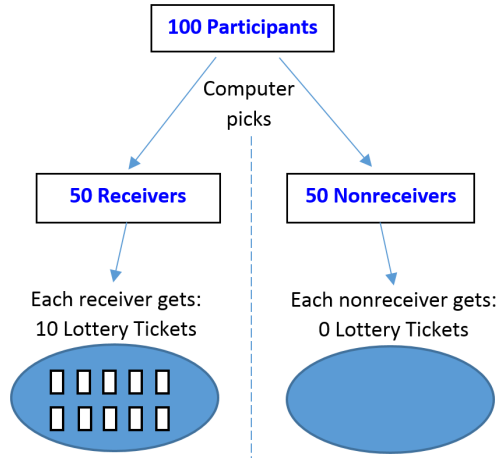
Step 2: Participants decide whether they want to share tickets with nonreceivers:

- 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 a participant decides to
GIVE

Participant keeps: Nonreceiver gets:
7 tickets   3 tickets

If a participant decides to
DON'T GIVE

Participant keeps: Nonreceiver gets:
10 tickets   0 tickets
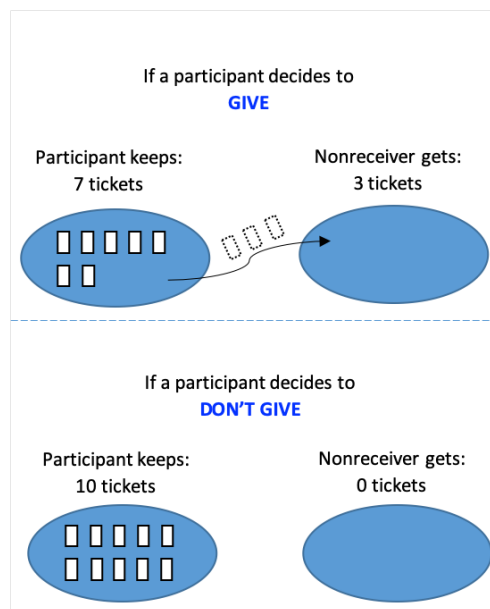
When taking the decision whether to GIVE or DON'T GIVE, you will not know whether you have been selected to be a receiver or a nonreceiver. Nor will anybody else. You will receive a message with this information after all participants have finished the study.

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

30

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.

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.

To show that you have read and understood the instructions, please answer the following control questions:

If you happen to be a receiver and choose GIVE, you will hold: (numbers appear in random order)

○ 3 tickets

○ 10 tickets

○ 7 tickets (correct)

○ 0 tickets

If you happen to be a receiver and choose DON'T GIVE, you will hold:

○ 7 tickets

○ 10 tickets (correct)

○ 0 tickets

○ 3 tickets

If you happen to be a nonreceiver and the other participant chooses GIVE, you will hold:

○ 0 tickets

○ 10 tickets

○ 3 tickets (correct)

○ 7 tickets

If you happen to be a nonreceiver and the other participant chooses DON'T GIVE, you will hold:

○ 0 tickets (correct)

○ 3 tickets

○ 7 tickets

○ 10 tickets

Instructions for The Giving Decision (Task 1)

Task 1
Your Choice: Give or Don't Give
If you happen to be a receiver, 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.

Instructions for the Evaluation Task

Task 2
Evaluate choices in a similar situation
On the following screen, you will read the description of a hypothetical choice situation that is very similar to the choice situation you just faced: 100 individuals take part in a lottery that has the exact same structure as the lottery you just took part in. Similar to your choice, each of these individuals has to decide whether to GIVE or DON'T GIVE 3 out of 10 lottery tickets to a nonreceiver.

In this new situation, however, you will NOT be asked to choose yourself. Instead, you will be asked to EVALUATE the different choices available to the other individuals. For each of the possible actions, you will have to decide whether taking that action would be

- "socially appropriate" and "consistent with moral or proper social behavior", or

- "socially inappropriate" and "inconsistent with moral or proper social behavior."

By socially appropriate, we mean behavior that most people agree is the "correct" or "ethical" thing to do.

All of the 99 other participants of today's study will evaluate the same choices in the same hypothetical situation. We will compare your evaluation with the evaluation of the 99 other participants. **If your evaluation is the same as the evaluation most frequently given by the other 99 participants, then you will receive an additional payment of GBP 2.00!**

Note: You and the other 99 participants will evaluate several choices. For the extra payment of GBP 2.00 we will select one of these choices at random. If you evaluate this choice the same way as most of the other 99 participants do then you will receive an additional payment of GBP 2.00. Each of your evaluations has the same chance to be selected for your payment. That is, you maximize your chances to earn GBP 2.00 by trying to always match the most common evaluation in your group.

Note: Your evaluation in Task 2 does NOT influence your chances to win the lottery! The lottery tickets for your group have been distributed in Task 1.

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

To show that you have read and understood the instructions, please answer the following control questions:

	True	False
You will receive GBP 2.00 if your evaluation is the same as the evaluation provided by most of the 99 other participants.	○	○
By socially appropriate, we mean behavior that most people agree is the "correct" or "ethical" thing to do.	○	○

Description of the Hypothetical Situation

Consider the following hypothetical situation: 100 other individuals take part in a lottery that has exactly the same structure as the lottery you took part in a few minutes ago. However, in this new lottery, **before anyone of the 100 individuals decides 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 ($\rightarrow$ RULE: GIVE) or whether everyone should choose DON'T GIVE ($\rightarrow$ RULE: DON'T GIVE). Only one of the two rules will be implemented.

In the choice situation you have to evaluate, the rule for the code of conduct will be determined in the following way:

Treatment T_StandardMajority:

- All 100 individuals who take part in the lottery are asked to vote for the rule (RULE: GIVE or RULE: DON'T GIVE) they prefer to have implemented as the code of conduct. The rule that receives more votes in total will be implemented as the code of conduct.

Treatment T_Pay4Vote:

- All 100 individuals who take part in the lottery are asked to vote for the rule (RULE: GIVE or RULE: DON'T GIVE) they prefer to have implemented as the code of conduct. The rule that receives more votes in total will be implemented as the code of conduct.

- However, only the votes of those participants who pay GBP 0.20 to make their vote count will be counted in the election.

Treatment T_MoneyOffer:

- All 100 individuals who take part in the lottery are asked to vote for the rule (RULE: GIVE or RULE: DON'T GIVE) they prefer to have implemented as the code of conduct. The rule that receives more votes in total will be implemented as the code of conduct.

- However, before the final votes are counted, 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.

Treatment T_ExcludePoor:

- All 100 individuals who take part in the lottery are asked to vote for the rule (RULE: GIVE or RULE: DON'T GIVE) they prefer to have implemented as the code of conduct. The rule that receives more votes in total will be implemented as the code of conduct.

- However, only the votes of participants with a household income above GBP 40,000 will be counted in the election.

Once a rule has been set, each individual can decide privately and anonymously whether he/she wants to follow the rule or not.

This is the situation you will be asked to evaluate on the next screen. Please make sure to remember it will. In particular, **on the next screen, you will be asked whether it is socially appropriate to follow the rule under the circumstances it has come into force.**

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

To show that you have read and understood the instructions, please answer the following control questions: In the hypothetical choice situation, the code of conduct will equal...
Please choose only one of the following: [$T\_StandardMajority$, items appear in random order]

○ the rule that the majority of individuals voted for.

○ the rule that was selected using a coin-flip by the computer.

○ the rule that was selected by the researcher.

In the three malpractice treatments, the additional (correct) items were, respectively:

○ the rule that the majority of those individuals who pay £0.20 voted for. [$T\_Pay4Vote$]

○ the rule that the majority of individuals finally voted for—after being offered GBP 0.20 to change their vote. [$T\_MoneyOffer$]

○ the rule that the majority of individuals with annual household income above GBP 40,000 voted for. [$T\_ExcludePoor$]

The Evaluation Decision

Please evaluate: Choice in the presence of a code of conduct
Here is a reminder of how the rule for the code of conduct is determined: [Example $T\_Baseline$]

- All 100 individuals who take part in the lottery are asked to vote for the rule (RULE: GIVE or RULE: DON'T GIVE) they prefer to have implemented as the code of conduct. The rule that receives more votes in total will be implemented as the code of conduct.

For each of the two possible rules (RULE: GIVE and RULE: DON'T GIVE), please indicate below how socially appropriate you believe it is to follow the rule and how socially appropriate you believe it is to not follow the rule. Remember that you will earn money (GBP 2.00) if your evaluation is identical with the most common evaluation given by the other 99 participants of this study.

RULE: GIVE. Below you see the choices available for each individual if RULE: GIVE is implemented as the code of conduct. Please indicate how socially appropriate you believe each choice to be.

	Very socially inappropriate	Socially inappropriate	Somewhat socially inappropriate	Somewhat socially appropriate	Socially appropriate	Very socially appropriate
Follow the rule and GIVE.	○	○	○	○	○	○
Don't follow the rule and DON'T GIVE.	○	○	○	○	○	○

RULE: DON'T GIVE. Below you see the choices available for each individual if RULE: DON'T GIVE is implemented as the code of conduct. Please indicate how socially appropriate you believe each choice to be.

	Very socially inappropriate	Socially inappropriate	Somewhat socially inappropriate	Somewhat socially appropriate	Socially appropriate	Very socially appropriate
Follow the rule and DON'T GIVE.	○	○	○	○	○	○
Don't follow the rule and GIVE.	○	○	○	○	○	○

Feedback Screen and End of the Study

Thank you very much for your participation. Your evaluations have been saved.

Did you feel that it makes a large difference whether to choose GIVE or DON'T GIVE or were you very much indifferent between choosing any of the two?

○ Was very much indifferent/did not care.

○ Felt that it makes a large difference.

Below, you have the opportunity to leave a more general personal comment. We greatly appreciate your feedback.

Please make sure to click "Submit" and then the link on the next page in order to prove that you have completed the study! Thank you!

Example Screenshots: T_StandardMajority

Task 2

Evaluate choices in a similar situation

On the following screen, you will read the description of a <u>hypothetical</u> choice situation that is very similar to the choice situation you just faced: 100 individuals take part in a lottery that has the exact same structure as the lottery you just took part in. Similar to your choice, each of these individuals has to decide whether to GIVE or DON'T GIVE 3 out of 10 lottery tickets to a nonreceiver.

In this new situation, however, you will NOT be asked to choose yourself. Instead, you will be asked to EVALUATE the different choices available to the other individuals. For each of the possible actions, you will have to decide whether taking that action would be

- "socially appropriate" and "consistent with moral or proper social behavior", or
- "socially *in*appropriate" and "*in*consistent with moral or proper social behavior."

By socially appropriate, we mean behavior that most people agree is the "correct" or "ethical" thing to do.

All of the 99 other participants of today's study will evaluate the <u>same</u> choices in the <u>same</u> hypothetical situation. We will compare your evaluation with the evaluation of the 99 other participants. **If your evaluation is the same as the evaluation most frequently given by the other 99 participants, then you will receive an additional payment of £2.00!**

<u>Note</u>: You and the other 99 participants will evaluate several choices. For the extra payment of £2.00 we will select one of these choices at random. If you evaluate this choice the same way as most of the other 99 participants do then you will receive an additional payment of £2.00. Each of your evaluations has the same chance to be selected for your payment. That is, you maximize your chances to earn £2.00 by trying to always match the the most common evaluation in your group.

<u>Note</u>: Your evaluation in Task 2 does NOT influence your chances to win the lottery! The lottery tickets for your group have been distributed in Task 1.

Figure A.1: Screenshot: Description of the evaluation decisions 1 (*T_Baseline*)

Consider the following <u>hypothetical</u> situation:

100 other individuals take part in a lottery that has exactly the same structure as the lottery you took part in a few minutes ago. However, in this new lottery, **before anyone of the 100 individuals decides 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.

In the choice situation you have to evaluate, the rule for the code of conduct will be determined in the following way:

- All 100 individuals who take part in the lottery are asked to vote for the rule (RULE: GIVE or RULE: DON'T GIVE) they prefer to have implemented as the code of conduct. The rule that receives more votes in total will be implemented as the code of conduct.

Once a rule has been set, each individual can decide privately and anonymously whether he/she wants to follow the rule or not.

This is the situation you will be asked to evaluate on the next screen. Please make sure to remember it well. In particular, **on the next screen, you will be asked whether it is socially appropriate to follow the rule <u>under the circumstances it has come into force</u>.**

Figure A.2: Screenshot: Description of the evaluation decisions 2 (*T_Baseline*)

Please evaluate: Choice in the presence of a code of conduct

Here is a reminder of how the rule for the code of conduct is determined:

> - All 100 individuals are asked to vote for the rule (RULE: GIVE or RULE: DON'T GIVE) they prefer to have implemented as the code of conduct. The rule that receives more votes in total will be implemented as the code of conduct.

For each of the two possible rules (RULE: GIVE and RULE: DON'T GIVE), please indicate below how socially appropriate you believe it is to follow the rule and how socially appropriate you believe it is to *not* follow the rule. Remember that you will earn money (£2.00) if your evaluation is identical with the most common evaluation given by the other 99 participants of this study.

RULE: GIVE. Below you see the choices available for each individual if RULE: GIVE is implemented as the code of conduct. Please indicate how socially appropriate you believe each choice to be.

	Very socially inappropriate	Socially inappropriate	Somewhat socially inappropriate	Somewhat socially appropriate	Socially appropriate	Very socially appropriate
Follow the rule and GIVE.	○	○	○	○	○	○
Don't follow the rule and DON'T GIVE.	○	○	○	○	○	○

RULE: DON'T GIVE. Below you see the choices available for each individual if RULE: DON'T GIVE is implemented as the code of conduct. Please indicate how socially appropriate you believe each choice to be.

Figure A.3: Screenshot: Evaluation decisions (*T_Baseline*)

Example Screenshots: T_NoRule

Instructions: Evaluate Choices in a Similar Situation

For the following questions we ask you to please consider the following <u>hypothetical</u> situation:

100 individuals in another group take part in a lottery that has <u>the exact same structure</u> as the lottery you are taking part in. That is, of the 100 participants, the computer selects 50 as receivers (who each get 10 lottery tickets) and 50 as nonreceivers (who get 0 lottery tickets):
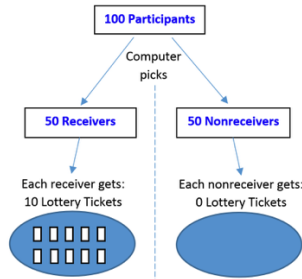
However, <u>in this new hypothetical situation</u>, the distribution of lottery tickets does not stop here.

Before the winning ticket is drawn, all participants get the chance to decide whether they want to share tickets with nonreceivers:

- Before participants learn whether they are a receiver or nonreceiver, each participant decides whether he/she wants to give away 3 of the 10 lottery tickets to a nonreceiver.
- This decision (**GIVE** or **DON'T GIVE**) has the following consequences:

Figure A.4: Screenshot: Description of the Evaluations 1

38

Figure A.5: Screenshot: Description of the Evaluations 2

Your Evaluation

Your answers to the control questions were correct.

Below, we ask you to evaluate actions GIVE and DON'T GIVE in the hypothetical choice situation explained on the previous screen.

For each of those two actions, we ask you to decide whether you think that taking that action would be

- "socially appropriate" and "consistent with moral or proper social behavior", or
- "socially inappropriate" and "inconsistent with moral or proper social behavior".

By socially appropriate, we mean behavior that most people agree is the "correct" or "ethical" thing to do.

With your evaluation you can earn an additional £2.00 to your base payment!

Whether you receive the additional £2.00 is determined as follows:

- All of the 99 other participants of today's study evaluate the same actions in the same hypothetical situation.
- We will compare your evaluation with the evaluation of the 99 other participants. If your evaluation is the same as the evaluation most frequently given by the other 99 participants, then you will receive an additional payment of £2.00.*

*For the extra payment of £2.00 we will select one of the two evaluations (GIVE or DON'T GIVE) at random. Both of your evaluations have the same chance to be selected for the extra payment. You maximize your chances to earn an additional £2.00 by trying to always match the most common evaluation among the other 99 participants. Neither your base payment of £1.10, nor your chances to win the cash prize of £100.00 in the lottery is affected by whether you receive the extra payment.

Figure A.6: Screenshot: Evaluations of Giving Decisions 1

Evaluation 1. Please indicate how socially appropriate you believe action **GIVE** to be:

	Very socially inappropriate	Socially inappropriate	Somewhat socially inappropriate	Somewhat socially appropriate	Socially appropriate	Very socially appropriate
Action **GIVE** is	○	○	○	○	○	○

Evaluation 2. Please indicate how socially appropriate you believe action **DON'T GIVE** to be:

	Very socially inappropriate	Socially inappropriate	Somewhat socially inappropriate	Somewhat socially appropriate	Socially appropriate	Very socially appropriate
Action **DON'T GIVE** is	○	○	○	○	○	○

Figure A.7: Screenshot: Evaluations of Giving Decisions 2