Democracy and Resilient Pro-Social Behavioral Change: An Experimental Study

Kenju Kamei[1]

First version: November, 2010
This version: September, 2011

Abstract

This paper shows experimentally that when subjects democratically choose a mild sanction rule that encourages cooperation, they behave more cooperatively not only in the democratic setting, but also in a second environment in which the policy was adopted without consideration of members' preferences, relative to subjects in an environment in which the policies are always non-democratically implemented. In this paper, the democratic effects are identified while controlling for selection effects. The finding implies that the impact of democracy on behavior, and specifically, the treatment effect of a policy implemented through democratic institutions, may be underestimated if the effect on external environments is ignored.

JEL classification: C91, C92, D70, D72, D78, H41

Keywords: cooperation, democracy, experiment, public goods, social dilemma, treatment effect

[1]I am grateful to Pedro Dal Bó for invaluable guidance, discussions, encouragement, and making NSF grant SES-0720753 available for this experiment. I am also grateful to Louis Putterman for continuous advice, discussions and constant support. I would also like to thank Brian Knight for insightful suggestions and discussions, and Kenneth Chay, Andrew Foster, Jordi Blanes-i-Vidal, Stefano Dellavigna, Raj Chetty and seminar audiences at Brown University for their helpful suggestions. I thank Urs Fischbacher for sharing his z-tree file used in Falk, Fischbacher and Gächter (forthcoming) with me.

E-mail address: kenju.kamei@gmail.com. Deloitte Tax LLP. Tel.: 401-499-5491.

1 Introduction

Does using a democratic procedure to choose policies influence their effectiveness? If the process used does influence a policy's effectiveness, could participating in democratic choice spill over and affect individuals' behaviors in non-democratic settings? These are the questions that this paper addresses.

Empirical research and experiments suggest that the process by which a policy is implemented may substantially change its outcome by affecting civic behaviors.[2] However, the effects of democratic decision processes on individual behavior may go beyond the direct effect. This is because democratic decision processes in one environment may affect behavior in other environments, including even non-democratic ones (those where democratic decisions are not the norm). Previous experimental research has not considered the effects of democratic processes on behavior beyond the immediate environment.

This paper provides the first experimental evidence on the spill-over effects of democratic participation in decision-making and draws attention to its importance. I use the framework of a simple linear voluntary contribution game (public goods game) and a non-deterrent or mild sanction policy.[3] A deterrent or severe sanction policy, if implemented in social dilemmas, forces people to cooperate at the socially optimal level, whereas the mild policy simply mitigates the incentive to defect; thus, with a mild policy, the dilemmas remain after the policy has been implemented (Tyran and Feld, 2006). When I measure the effects of the policy, I control thoroughly for selection effects (the effects of a selection bias through democratic processes). Controlling for selection effects is important because demo-

[2]Empirical studies include Frey (1997) for tax laws, Bardhan (2000) for irrigation rules, and Ichniowski, Shaw and Prennushi (1997) and Ichniowski and Shaw (2003) for human resource management practices. Experimental evidence comes from Bortolami and Mittone (2009) for the probability of being audited and the type of sanction and reward, Dal Bó, Foster and Putterman (2010) for a policy that changes the material payoff of a game from a prisoner's dilemma to a coordination game, Tyran and Feld (2006) for a non-deterrent law, and Sutter *et al.* (2010) for an informal sanctioning opportunity. For a summary of experimental evidence, see Dal Bó (2010). While these studies show the superiority of democratic institutions, the existence of democratic effects appears to depend on policies. See Rauchdobler *et al.* (2010), Sausgruber and Tyran (2007), and Sutter and Weck-Hannemann (2004). Other experiments involving endogenous institution formation focus less on separating the effects of the institution from that of the implementation process (Botelho *et al.* 2005, Ertan *et al.* 2009, Gürerk *et al.* 2006, Kamei *et al.* 2011, Kosfeld *et al.* 2008, Markussen *et al.* 2011, Nese and Sbriglia 2009, Putterman *et al.* forthcoming, Rockenbach and Wolff 2010, Walker *et al.* 2000).

[3]There are many real examples of mild policies, ranging from global policies such as environmental policy to various ordinances such as prohibition of littering.

cratic processes are channels through which subjects can choose institutional conditions; in other words, subjects' voting decisions and subsequent actions in a game may be correlated with unobservable factors.[4] To control for selection effects, every individual makes a voting decision and then the computer randomly assigns them democratic or undemocratic institutional conditions as in Dal Bó, Foster and Putterman (2010). I use the term "direct effect of democracy" for the effect of democratic processes on a social dilemma that is not captured by the effect of an exogenously imposed policy and by the selection effects. I use the term "spill-over effect of democracy" for the effect of a person's exposure to democracy in one setting on her behavior in another setting. I refer to these effects together as the "effects of democracy."

In addition to evidence on the spill-over effect, experimental evidence on the direct effect of democracy in a public goods experiment in this paper also significantly contributes to the literature. This study is the first to thoroughly control for selection effects in order to identify the direct effect of democracy in a voluntary contribution dilemma.[5]

A key feature of the experimental design is that each subject belongs to two groups with distinct partners and simultaneously plays two modified public goods games. For some subjects, a mild sanction policy is implemented democratically in one group, and by the computer in the other group. For other subjects, the policy is implemented undemocratically by the computer in both their assigned groups. The former are treatment subjects, and the latter are control subjects. My experiment indicates that treatment subjects who support the policy contribute to the *democratic* group significantly more than control subjects with the same preference. Surprisingly, the former subjects contribute to their *non-democratic* group significantly more than the latter subjects do. The magnitude of the spill-over effect of democracy is almost the same as the direct effect of democracy. This substantial behavioral change that extends beyond the democratic group is striking, considering that the policy

[4]See Dal Bó, Foster and Putterman (2010).

[5]Policies used in public-goods experiments are, unlike Dal Bó, Foster and Putterman (2010), usually weak in that irrespective of whether or not they are implemented, standard theory predicts the same individual behavior. A well-known example of evidence based on numerous previous experiments is that some individuals exhibit a preference for cooperation. Even though standard theory predicts that full defection is a unique Nash Equilibrium in a dilemma game, some subjects cooperate positively in an experiment (Ledyard 1995).

implemented only mitigates the incentive to defect, and that the social dilemma remains.

Since democracy has a spill-over effect along with a direct effect, people exposed to democracy in social dilemmas become more cooperative on many civic issues in a society. This experimental evidence has implications for collective action dilemmas, which might bridge an extensive literature on democracy and that on the emergence of cooperation norms in a society or organization. Dilemmas exist at many levels in a society: from concerns about the environment or crime at the national level to concerns about air pollution at the community level. Policies at various levels have been implemented to resolve these dilemmas. The result of this paper implies that the more dense the network of democratic processes, the more likely citizens are to become public spirited and cooperative on civic issues.[6] Resilient pro-social behavioral changes beyond one issue might facilitate social interactions with other citizens with similarly high public spirits, through which civic behavior would be strengthened and cooperation turned into a social norm.[7,8] Enhanced civic culture may result in better functioning institutions and increased economic performance in a society (e.g., Knack and Keefer 1997, Tabellini 2005, Tabellini 2007).

The finding in this paper also has broad implications for the methods used to measure the treatment effect of a policy. Since the impact of democracy on behavior may be substantial and may not be restricted to one democratic setting, the treatment effect of the policy might be underestimated if the effect on external settings is ignored when a policy is implemented

[6]In the theory of the firm, it is well-known that intangible intrinsic motivation may significantly affect an agent's performance. On one hand, an agent's intrinsic motivation may be crowded out by the introduction of a monetary incentive (Gneezy and Rustichini 2000a, Frey and Oberholzer-Gee 1997, Gneezy and Rustichini 2000b). On the other hand, it may be enhanced by exposure to democratic decision-making. Intrinsic motivation which is crowded in or crowded out in one context may influence the agent's behavior beyond that context (Frey 1993, Osterloh and Frey 2000, Frey and Jegen 2001). However, there is no clear empirical and experimental evidence of the additional effect of motivation crowding-in (positive motivation spill-over effect). Researchers in political science have proposed that civic participation and democratic decision-making may have a substantial indirect impact on those involved; however, this research suffers from econometric problems such as a self-selection bias. One indirect effect is of the citizen's initiative — which is one form of direct democracy in the United States — on an increase in voter turnout (Schlozman and Yohai 2008, Smith 2001, Tolbert *et al.* 2001, Tolbert *et al.* 2003, Tolbert *et al.* 2009). Another possible effect is that participating in a public problem-solving activity or in collective decision-making change a citizen's belief about how he or she should behave and increase the likelihood that the citizen will engage in future public activities (Haney, Borgida and Farr 2002).

[7]See Gächter and Thöni (2005), Page *et al.* (2005) and Gunnthorsdottir *et al.* (2007) for related experiments on this claim.

[8]Previous research in political economy has proposed that democracy might help foster civic attitudes such as trust (Muller and Seligson 1994, Tabellini 2005) and social capital (Guiso *et al.* 2008) in a society.

through democratic institutions.

The rest of the paper is organized as follows: Section 2 provides details of the experimental design, theoretical predictions, and the identification strategy. Section 3 reports and discusses results from the experiment. Section 4 concludes. An Appendix providing supporting material, including instructions, is available on-line.

2 Experimental Design

I use a public goods game framework. The instructions for all experimental treatments are neutrally framed (See Appendix A for instructions). The experiment was programmed in z-tree (Fischbacher, 2007). Communication between subjects was not permitted during a session.

The experiment has two periods. The key vehicle for analyzing the effects of democracy comprehensively, including the additional effect on the public good that are within the exogenous process (the spill-over effect of democracy), is that each subject faces two public good provision problems.[9] At the onset of the experiment, each subject is randomly assigned to two groups of two individuals. The other individuals in the two groups are different from one another. To prevent repeated encounters, regardless of his or her decision in Period 1, after that period is over, each individual is assigned to two new groups with two different individuals.[10] In Period 2, as in Period 1, the other individuals in the two groups are different from each other. The smallest group size, set in this experiment, makes it possible to obtain as many independent observations as possible.[11]

This paper uses an inventive device, linking triplets of subjects into "triads" of the type

[9] In the literature on public goods experiments, Fellner and Lünser (2008) and Falk, Fischbacher and Gächter (forthcoming) use a setting in which each subject faces two public good provision problems. In Fellner and Lünser (2008), each subject belongs to a local group and a global group. Each global group includes two mutually exclusive local groups. In Falk, Fischbacher and Gächter (forthcoming), each individual belongs to two groups, and each group has three members. In both experiments, subjects make a sequence of simple voluntary contribution decisions in the standard linear public goods game.

[10] At the beginning of the experiment, each individual is given clear instructions about changing partners and is told the total number of periods in the main part of the experiment.

[11] In a one-shot environment in which no communication is allowed, small group size may make cooperation difficult (Goeree *et al.* 2000). However, in standard public goods environments with finite periods, the impact of group size on contribution decisions is small (Ledyard 1995, Zelmer 2003).

shown in Figure 1; during the experiment, three groups are linked with each other. This device effectively generates treatment and control subjects. As explained in this section, such a structure assures that each subject exposed to the endogenous introduction of a policy also belongs to a group in which the policy is exogenously imposed, and that some other subjects participate only in groups in which the policy is exogenously imposed. The former are treatment subjects and the latter are control subjects in this experiment (see Section 2.2). The triad structure is unknown to subjects; they only know that they belong to two groups, each with a different partner (See Appendix A.1 for instructions).

In Section 2.1, I summarize a series of decisions subjects make in the experiment, along with standard theory predictions. In Section 2.2, I discuss the identification strategy used to measure the effects of democracy.

2.1 Experimental Procedure

In Period 1, each subject is given an endowment of 20 points (experimental currency unit where 6 points = US$1) in each of their two assigned groups; in each group subjects decide how to allocate their 20 points between a private account and a group account. For their choices, only integers between 0 and 20 are allowed. The marginal per capita return (MPCR) from the public account equals 0.6 in each group. Each subject's total earnings in Period 1 are the sum of their earnings in each group. On one hand, since the MPCR equals 0.6, contributing nothing to the public account is a strictly dominant strategy for each individual. On the other hand, since 0.6 times 2 is 1.2, which is greater than 1, if two individuals in a group contribute 20 points to the public account, the total material payoff achieves its maximum. Thus, the unique Nash Equilibrium (NE) in each group is full free-riding (contributing 0 points to the public account), whereas the social optimum is full cooperation (contributing 20 points to the public account). Period 1 plays a role in familiarizing subjects with the public goods game. In Period 1, each three groups forms a triad (Figure 1). The triad device does not operate in the first period, but makes it possible to cluster subjects by interaction unit when I analyze the data. At the end of Period 1, subjects are aware of their counterparts' allocations to the public account and of their own earnings.

In Period 2, a mild sanction policy may be implemented in each group. The MPCR equals 0.6 in this period as well. When the policy is in place, each point that a subject allocates to his or her private account results in a fine of 0.3 points. The earnings in each group are as follows: for a contribution x_i to the public account, individual i obtains:

$$\begin{aligned} Earnings &= (20 - x_i) \cdot (1 - 0.3 \cdot I_{\{implement\}}) + 0.6 \cdot \sum_{j \in \{1,2\}} x_j, \\ \text{where } I_{\{implement\}} &= 1 \text{ if the policy is implemented; } = 0 \text{ otherwise.} \end{aligned}$$

As in Period 1, since each individual belongs to two groups, their total earnings in Period 2 are the sum of their earnings in each group. Since $1 - 0.3 > 0.6$, regardless of whether or not the policy is in use, contributing nothing to the public account is still the strictly dominant strategy. As a result, this is the unique NE for a subject who is maximizing material payoffs, while contributing the full endowment to the public account is the social optimum.

Whether the sanction policy is in place depends on individual voting decisions and on the computer's random choices. Period 2 begins with the voting and implementation stage, followed by the above contribution stage.

At the onset of Period 2, each individual votes on whether or not to use a mild sanction policy in *each* of their two assigned groups. I refer to subjects as follows: "Yes-Yes voters" are those who voted for the policy in both their assigned groups, "No-No voters" are those who voted against it in both their assigned groups, and "Yes-No voters" are those who voted for the policy in one group and against it in the other. After every subject makes decisions, the triad device comes into play: in each triad the computer randomly assigns the endogenous condition to one group and the exogenous condition to the remaining two groups. As a result, two-thirds of the subjects are exposed to the endogenous process in one group and one-third are never exposed to it.[12] Under the endogenous condition, the computer counts the votes and the majority determines the implementation of policy in that group. When one individual votes in favor of the policy and the other votes against it, the computer randomly breaks the tie. Under the exogenous condition, the computer does not

[12]Subjects are not told the exact process; they are only told that their votes are randomly counted for each of their two assigned groups.

consider the votes and randomly decides whether or not to use the policy.[13] Since subjects make a voting decision even in the exogenous condition, I am aware of their preferences toward the policy, irrespective of implementation processes; thus, I can fully control for selection effects.

I use the following notation for the analysis: "Endo" means that majority rule determined whether to use the policy, and "Exo" means that the computer randomly determined whether to use the policy. "Imp" means that the policy was implemented, and "Not" means that the policy was not implemented in Period 2. The seven pairs of possible implementation outcomes for each subject in their two assigned groups are: "Endo Imp, Exo Imp," "Endo Imp, Exo Not," "Endo Not, Exo Imp," "Endo Not, Exo Not," "Exo Imp, Exo Imp," "Exo Imp, Exo Not," and "Exo Not, Exo Not." Note that since I set a high probability for the policy to be exogenously implemented and for the policy to be implemented when the vote share in favor of the policy is 50 percent in the endogenous condition, the outcome "Exo Imp" is much more frequent than the outcome "Exo Not;" the outcome "Endo Imp" is much more frequent than the outcome "Endo Not" when the vote share is 50 percent (See footnote 13).

Under the endogenous condition, a subject has partial information about the distribution of individual votes in a group from the vote outcome. For example, suppose a subject in a democratic group votes not to use the policy, but the policy is implemented in that group. The subject immediately knows that his or her counterpart voted to use the policy.[14] To also make that information available in the exogenous condition, each individual is informed of whether "at most" one person, or "at least" one person, voted for the policy in each

[13] Under the exogenous condition, or when the vote share is 50 percent under the endogenous condition, the computer will implement the policy with a probability of 90 percent and will not implement the policy with a probability of 10 percent as in Dal Bó *et al.* (2010). The unequal probabilities set in the experiment make it possible to produce as many as treatment and control subjects. Subjects are not told the exact probabilities; they are only told that the computer decides randomly.

[14] If we hypothesize that a subject makes contribution decisions and develops beliefs about partners' future cooperation based on partial information, the difference in information between the two conditions becomes a problem when identifying the effects of democracy. To eliminate a competing hypothesis, I cancel out the difference in the information.

group, regardless of whether the votes are counted.[15,16] I refer to the former as "information *At-Most*," and the latter as "information *At-Least*." In summary, every individual is aware of the following after the voting decision:

(1) whether "at most" one or "at least" one person in each group voted to implement the policy;

(2) whether votes or the computer decided on the implementation of the policy;

(3) whether the policy will be implemented in each of the individual's two assigned groups.

Some subjects are simultaneously assigned to a democratic group and a non-democratic group; these two groups are different from each other even if the same policy is implemented in each group because the implementation process is different in each group. In addition to the announcement of outcomes, in the contribution stage, subjects have reminders of what determined the scheme in each of their two assigned groups and how the vote went in each group on their computer screens (See Appendix Figure B.1 for examples of reminders on the screen).

As described above, a rational, self-interested subject will contribute nothing in Period 2, regardless of the outcome of voting and implementation. In other words, the sanction policy, if implemented, will cause him or her to pay a maximum fine of 6 points. Since the group size is two, each subject's vote is pivotal in a group. Thus, under common knowledge that the probability with which their votes will be considered is positive, as well as that the probability with which the computer will implement the policy when the endogenous condition is assigned to a group but the vote share is 50 percent is also positive, standard theory predicts that each subject will vote against the use of the policy in both assigned groups in this stage (Subgame Perfect Nash Equilibrium; Appendix Figure B.2).

Subjects may vote to implement a policy and then voluntarily contribute to the group when the policy is implemented. This is because, first of all, a mild sanction policy reduces the temptation to defect, possibly making mutual cooperation more attainable if subjects

[15] If the vote share is 50 percent, the computer chooses whether to report at least one voted or at most one voted for the policy with a probability of 50 percent each. Subjects are not told the exact probability, and are only told that the computer decides randomly.

[16] By including information feedback, the *only* difference between a democratic group and a non-democratic group is the implementation processes of the policy.

are not rationally self-interested. Second, voting for the mild sanction policy could be considered a credible way of signaling an intention to cooperate, since each subject has partial information on the distribution of votes by the information feedback function, irrespective of the condition (endogenous or exogenous). Of course, subjects may also benefit from the mild sanction policy if it induces their partners, but not themselves, to contribute more.

Although this experimental design is simple enough for subjects to clearly understand the procedures of the experiment, a few comprehension questions are included with the instructions.

All subjects take the Cognitive Reflection Test (CRT), consisting of three questions (Frederick 2005), after Period 2.[17] Subjects with higher CRT scores are less likely to make an error. In addition, Frederick (2005) shows that subjects with higher CRT scores are less likely to favor a gamble involving losses. Thus, among subjects who simultaneously belong to democratic and non-democratic groups, those with higher CRT scores might be more likely to differentiate their contribution decisions between these two groups. I include this test to examine whether the spill-over effect of democracy that subjects exhibit differs by CRT score.

They subsequently participate in a Beauty Contest Game (See Nagel 1995 and Dal Bó, Foster and Putterman 2010).[18] I use responses to this game to explore whether strategically sophisticated subjects exhibit the spill-over effect of democracy.

Finally, each subject answers exit questions about his or her gender, SAT scores (Math, Verbal), and other personal characteristics. Every participant receives payment immediately after the experiment.

2.2 Identification Strategy

How should we measure the direct effect of democracy? If the spill-over effect of democracy does not exist, one way is to use the set of subjects who implement the mild sanction policy

[17] The CRT, which is included in Appendix A.2, measures "the ability to suppress an erroneous answer that springs impulsively" or "how precise the cognitive process executed quickly with little conscious deliberation is" (Frederick, 2005). A subject earns seven points for each correct response to the three questions.

[18] Each subject chooses an integer between 0 and 100. The subject with the closest number to two thirds of the average of all numbers in a session earns 120 points.

endogenously in a group and find it exogenously imposed in the other ("Endo Imp, Exo Imp"), and to compare contribution decisions between their democratic and non-democratic groups. If the direct effect of democracy exists, these subjects will make significantly different decisions between the two groups because they experience the democratic process in only one group. Since we are comparing the same individual's contribution decisions, we need not control for selection effects. Note that in this case, testing of the effects of democracy coincides with testing the direct effect of democracy.

I split subjects into two subsets: Yes-Yes voters and No-No voters. Based on the findings of a previous study (Dal Bó *et al.* 2010), I hypothesize that only Yes-Yes voters will exhibit the direct effect of democracy, and formulate the following identification strategy.

Strategy 1 *The Direct Effect of Democracy* $= \bar{C}^{Endo}_{\text{“}Endo,Exo,\text{”}YY} - \bar{C}^{Exo}_{\text{“}Endo,Exo,\text{”}YY}$

$C^{Endo}_{\text{“}Endo,Exo,\text{”}s}$ denotes the set of Period 2 contributions by voters of type $s \in \{YY, NN\}$ in the *democratic* group under "Endo Imp, Exo Imp" (treatment condition). Here, YY indicates a Yes-Yes voter and NN indicates a No-No voter. Similarly, $C^{Exo}_{\text{“}Endo,Exo,\text{”}s}$ denotes the set of Period 2 contributions by voters of type s in the *non-democratic* group under "Endo Imp, Exo Imp." $C^{Exo}_{\text{“}Exo,Exo,\text{”}s}$ denotes the set of Period 2 average contributions by voters of type s under "Exo Imp, Exo Imp" (control condition). Furthermore, $\bar{C}^{Endo}_{\text{“}Endo,Exo,\text{”}s}$, $\bar{C}^{Exo}_{\text{“}Endo,Exo,\text{”}s}$ and $\bar{C}^{Exo}_{\text{“}Exo,Exo,\text{”}s}$ denote the average of $C^{Endo}_{\text{“}Endo,Exo,\text{”}s}$, $C^{Exo}_{\text{“}Endo,Exo,\text{”}s}$ and $C^{Exo}_{\text{“}Exo,Exo,\text{”}s}$, respectively. When I employ Strategy 1, I use individual-level Wilcoxon Signed Ranks tests to test for the existence of the direct effect.

If the spill-over effect of democracy exists and is not negligible, using Strategy 1 will produce a result that is indecisive or uninterpretable: we will not only underestimate the direct effect of democracy, but will also fail to measure the spill-over effect of democracy. Thus, if the spill-over effect of democracy exists, we need to use two categories of Yes-Yes voters: Yes-Yes voters who are exposed to democracy (those who are in the treatment condition: "Endo Imp, Exo Imp"), and Yes-Yes voters who are never exposed to democracy (those who are in the control condition: "Exo Imp, Exo Imp"). The use of these two Yes-Yes voters makes it possible to measure the two kinds of effects of democracy simultaneously (Figure 2). I denote the set of Yes-Yes voters in the treatment condition ("Endo Imp, Exo Imp") as

"Endo Imp-yes, Exo Imp-yes," and the set of Yes-Yes voters in the control condition ("Exo Imp, Exo Imp") as "Exo Imp-yes, Exo Imp-yes." The motivation for this experiment is the conjecture that democratic decision-making in one group affects an individual's behavioral response to a policy not only in that group, but also in his or her other group; the following identification strategy is formulated and used:

Strategy 2 *(1) The Direct Effect of Democracy* $= \bar{C}^{Endo}_{\text{"}Endo,Exo,\text{"}YY} - \bar{C}^{Exo}_{\text{"}Exo,Exo,\text{"}YY}$

(2) The Spill-Over Effect of Democracy $= \bar{C}^{Exo}_{\text{"}Endo,Exo,\text{"}YY} - \bar{C}^{Exo}_{\text{"}Exo,Exo,\text{"}YY}$

I use individual-level Mann-Whitney tests to test the effects of democracy under this strategy.

Although group composition does change across periods, contribution decisions in Period 2 may be correlated between members in each Period 1 triad. As a robustness check, I conduct Somers' D tests with standard errors adjusted for clusters in Period 1 triad ID.[19]

My experiment has implications for which identification strategies should be used to measure the effects of democracy. Although Strategy 2 is the appropriate identification strategy for this study, I also examine the consequences of employing Strategy 1.

3 Results

The experiment for this study was conducted at a computer lab at Brown University between April and May, 2010. Three hundred undergraduate students participated in 15 sessions in the experiment.[20] No subject participated in more than one session. The sessions lasted around an hour on average, and the average earnings were $21.38 (including a $5 participation fee) with a standard deviation of $5.04.

The average contribution in Period 1 was 5.96 with a standard deviation of 6.06; a contribution of 30 percent of the endowment in this one-shot experiment is near the lower end

[19] As for Somers' D tests, see Newson (2002).

[20] Participants were recruited from the general student population at Brown University, using the Brown University Social Science Experiment Lab (BUSSEL) on-line recruitment tool. Of the 300 subjects, 46 listed economics as one of their concentrations, which was similar to the proportion of economics majors in the university's undergraduate population. Other participants came from fields ranging from the humanities to the sciences. Each session had 21, 18 or 15 participants. All subjects were first-time participants in public goods game experiments.

of the range which previous literature suggests is to be expected, given the lower efficiency of the group optimum (0.6 times 2 = 1.2).[21]

3.1 Vote Decisions and Implementation Outcomes

Table 1 presents the distribution of voter type and the sample mean of each characteristic variable by voter type. In total, 127 subjects (42.3%) were Yes-Yes voters, 143 subjects (47.7%) were No-No voters, and the remaining 30 subjects (10.0%) were Yes-No voters. I do not include Yes-No voters in my analysis because the sample size is too small. The average CRT score, the average SAT score (Math), the average SAT score (Verbal), and the proportion of economics concentrations of the Yes-Yes voters are all insignificantly higher than those of the No-No voters.

Of the 200 subjects who belonged to democratic and non-democratic groups in the voting stage, 120 faced the treatment condition ("Endo Imp, Exo Imp"). Of these 120 subjects, 67 were Yes-Yes voters and 37 were No-No voters. Of the 100 subjects who belonged to two non-democratic groups, 76 faced the control condition ("Exo Imp, Exo Imp"); of these 76 subjects, 31were Yes-Yes voters and 37 were No-No voters. The ratio of Yes-Yes voters under the treatment condition to Yes-Yes voters under the control condition was approximately 2:1 as intended; we compare behavior without any bias since Yes-Yes voters were *randomly* assigned to either the treatment condition or the control condition.[22] Table 2 reports the number of subjects in each of the seven voting and implementation stage outcomes by individual voting decision.

[21] See Ledyard (1995). Similarly, a result of meta-analyses conducted by Zelmer (2003) predicts that the ratio of the average contribution to the total endowment will be around 40 percent in my experimental setting, and 30 percent falls in the confidence interval. Botelho *et al.* (2005) used a MPCR of 0.6 and a group size of two in a standard public goods game; the average contribution in Period 1 was around 25 percent of the endowment in a perfect stranger treatment. Dal Bó and Dal Bó (2010) used a MPCR of 0.7 and a group size of two in a standard public goods game; the average contribution in Period 1 was around 35 percent of the endowment in a stranger treatment.

[22] While it is true that the power of statistical analyses decreases when there are fewer subjects in one category, I detect both the direct effect of democracy and the spill-over effect of democracy as shown in this section. In this section, I conduct various robustness checks for this finding.

3.2 Test of the Existence of the Direct and Spill-Over Effects of Democracy

In this section, I measure the effects of democracy using Strategies 1 and 2.

3.2.1 Individual-Level Treatment Effects

Let me begin by measuring the effects of democracy using Strategy 1. In the experiment, $\bar{C}^{Endo}_{\text{“}Endo,Exo,\text{”}YY} = 16.6$ and $\bar{C}^{Exo}_{\text{“}Endo,Exo,\text{”}YY} = 15.7$; the difference in the average contribution is less than 1 point and is not statistically significant (p-value = .16, 2-tailed). Thus, the direct effect of democracy is not detected. Does this imply that the direct effect of democracy does not exist under my experimental conditions? The answer is no. A test using Strategy 2 provides a different, interesting result.

Under the control condition ("Exo Imp, Exo Imp"), $\bar{C}^{Exo}_{\text{“}Exo,Exo,\text{”}YY} = 12.1$. Thus, the difference between $\bar{C}^{Endo}_{\text{“}Endo,Exo,\text{”}YY}$ and $\bar{C}^{Exo}_{\text{“}Exo,Exo,\text{”}YY}$ is more than 4 points, and is significant at the 1% level. This result indicates that even when the policy does not change the set of equilibria for a rational, self-interested subject in a public goods game, Yes-Yes voters exhibit the substantial direct effect of democracy if they are able to implement the policy by their votes.

Surprisingly, the difference between $\bar{C}^{Exo}_{\text{“}Endo,Exo,\text{”}YY}$ and $\bar{C}^{Exo}_{\text{“}Exo,Exo,\text{”}YY}$ is a little less than 4 points, and is significant at the 5% level. This means that Yes-Yes voters under the treatment condition ("Endo Imp, Exo Imp") also exhibit a similar magnitude of the democratic effect on contribution decisions in the *other* non-democratic group.[23]

The clear contrast between results, depending on how we measure the effects of democracy, provides substantial implications not only for economists working on the topic, but also for those using evidence on the causal consequences of the democratic process for their research: the difference between the treatment effects of an endogenous and exogenous policy may be substantially bigger than previously thought. If we were not aware of the potential

[23] A reader may pose the possibility that some treatment subjects were confused about what determined the scheme in each of their two assigned groups in the contribution stage. However, this is very unlikely since the voting process proceeded separately in each of their two assigned group; in the contribution stage subjects received reminders about implementation outcomes, including what determined the scheme in each of their two assigned groups, on their computer screens (See Section 2.1 and Appendix Figure B.1).

importance of the spill-over effect of democracy, even if the effects of democracy were by far more than the direct effect, we might end up measuring a negligible amount which was much less than even the true direct effect of democracy, as the effects of democracy; we might decide not to consider the considerable additional effect that was not captured by the direct effect of democracy for our empirical work or theoretical work.

I now consider No-No voters under the treatment and control conditions. On the one hand, $\bar{C}^{Endo}_{\text{“}Endo,Exo,\text{”}NN} = 7.35$ and $\bar{C}^{Exo}_{\text{“}Endo,Exo,\text{”}NN} = 7.65$; this difference in the average contribution is not significant at the 10% level. On the other hand, $\bar{C}^{Exo}_{\text{“}Exo,Exo,\text{”}NN} = 5.84$; there is no statistical difference in the average contribution of the two sets of No-No voters in Period 2 at the 10% level. Thus, I fail to show that No-No voters exhibit the effects of democracy. The result that only Yes-Yes voters contribute more in a democratic group in which the social dilemma *remains* is similar to a result in Dal Bó, Foster and Putterman (2010) in which a prisoner's dilemma game *is changed* to a coordination game. The result that Yes-Yes voters exposed to the endogenous process also contribute more in their non-democratic groups – that is, the impact spills over to their behavior in a second domain – is a new finding.[24]

3.2.2 Period 1 Average Contribution and the Effects of Democracy

One might regard the results in Section 3.2.1 as an accident, given the possibility that Yes-Yes voters who were intrinsically more cooperative might have been more frequently assigned to the treatment condition ("Endo Imp, Exo Imp") by chance. Period 1 contribution decisions have been thought of as indicators of the willingness to cooperate. How different were the distributions of Period 1 average contributions in "Endo Imp-yes, Exo Imp-yes" and in "Exo Imp-yes, Exo Imp-yes"?

The average contributions of Yes-Yes voters in Period 1 under the treatment and control conditions are 9.12 and 5.00, respectively; the difference in the average contribution

[24]Note that Yes-Yes voters earned less under the treatment condition than the control condition. The average earnings of Yes-Yes voters under the treatment condition are 20.22 in their democratic group, and 17.01 in their non-democratic group; the difference in average earnings is significant according to a Wilcoxon Signed Ranks test (p-value = .0005, 2-tailed). The average earnings of Yes-Yes voters under the control condition are 20.28. According to a Mann-Whitney test, the difference in average earnings between the non-democratic groups under the treatment and control conditions is significant (p-value = .0031, 2-tailed).

is statistically significant (p-value = .012, 2-tailed). The histogram of Period 1 average contributions for these two sets of Yes-Yes voters shows how different the distributions are between them (Figure 3). In my samples, the fraction of observations with Period 1 average contributions less than 4.0 is much smaller in the treatment condition than in the control condition, whereas the fraction of observations with Period 1 average contributions more than 12.5 is much larger in the treatment condition than in the control condition.[25]

Given this unbalanced Period 1 behavior between the treatment and control conditions, it is important to study whether the results in Section 3.2.1 are robust for subjects with similar contribution decisions in Period 1.[26] As a robustness check, I measure individual-level treatment effects using the [4, 12.5] range in this section.[27] In addition to this range, I choose a more robust range, considering the nature of finite samples in experiments: I narrow the [4, 12.5] range into the intersection between the [4, 12.5] range and the IQR (Inter-Quartile Range) of all 300 subjects' Period 1 average contributions. This analysis provides a test for the effects of democracy on "near" median subjects, but we should note that using this narrower range also has a disadvantage because it decreases the number of available subjects. Since the lower quartile of all 300 subjects' Period 1 average contributions is 0.0 and the upper quartile is 9.5, the IQR is [0.5, 9.5] (See Appendix Figure B.3 for the histogram and empirical distribution of all 300 subjects). In both the [4, 12.5] and [4, 9.5] ranges, neither the average contributions nor the empirical distributions of contribution decisions in Period 1 are significantly different between treatment and control subjects at the 10% level (Figure 4).

As a randomization check, I test the null hypothesis that the mean is the same for Yes-Yes voters under the treatment and control conditions in each of the [4, 12.5] and [4, 9.5] ranges for every other individual characteristic and the subject's Period 1 partner's Period 1 average contribution. If we employ a conventional 5% threshold for statistical significance in each sample pool, we fail to reject the null hypothesis for all variables (Panel (II), Appendix Table B.4).

[25] There were no observations in Period 1 in the [0.5, 3.5] range in "Endo Imp-yes, Exo Imp-yes." There were only 2 observations in that period in the [13.0, 20.0] ranges in "Exo Imp-yes, Exo Imp-yes."

[26] Besides this, the proportions of female subjects are significantly different between the treatment and control subjects (Panel (I), Appendix Table B.4).

[27] The use of this range is supported by Figure 3, but see also footnote 28.

I confirm the findings revealed in Section 3.2.1 even if I use these two sample pools. That means that whereas Strategy 1 rejects the existence of the direct effect of democracy at the 10% level, Strategy 2 finds both a direct effect of democracy and a spill-over effect of democracy at the 5 or 1% level (Table 3).[28]

For No-No voters, neither the average contributions nor the empirical distributions of contributions are significantly different between "Endo Imp, Exo Imp" and "Exo Imp, Exo Imp." Therefore, there is no need to conduct a robustness check using a conservative range as I did for Yes-Yes voters.

3.2.3 Empirical Distributions of Contribution Decisions

An empirical distribution of the contribution decisions is more informative than their mean. In this section, I compare the decisions between Yes-Yes voters under the treatment condition ("Endo Imp, Exo Imp") and under the control condition ("Exo Imp, Exo Imp"). Figure 4 presents the empirical distributions of contributions for these two sets of Yes-Yes voters whose Period 1 average contributions are in the [4, 12.5] range (a conservative range).

In Period 2, when the policy was implemented, the distribution of contribution decisions all shifted down (to the right) regardless of whether a group was under the treatment or control conditions and whether the process was endogenous or exogenous.[29] However, the empirical distributions of $C^{Endo}_{\text{“}Endo,Exo,\text{”}YY}$ and of $C^{Exo}_{\text{“}Endo,Exo,\text{”}YY}$ shifted significantly more than the distribution of $C^{Exo}_{\text{“}Exo,Exo,\text{”}YY}$. It is striking that in Period 2 both the medians of $C^{Endo}_{\text{“}Endo,Exo,\text{”}YY}$ and $C^{Exo}_{\text{“}Endo,Exo,\text{”}YY}$ are 20, and that more than 80 percent of contribution decisions fall in the range [15, 20]. In contrast, under the control condition in the same period, the median of $C^{Exo}_{\text{“}Exo,Exo,\text{”}YY}$ is 13.5 and only 42.9 percent of average contributions

[28] As Figure 3 shows, there are relatively many observations for which Period 1 average contributions are 0 both in "Endo Imp-yes, Exo Imp-yes" and in "Exo Imp-yes, Exo Imp-yes." Fischbacher *et al.* (2001) categorize these types of subjects as "free riders." An analysis using observations for which Period 1 average contributions are in the $\{0\}\cup[4,\ 12.5]$ range yields qualitatively similar results (See Appendix Table B.5, Figure B.6 and related notes).

If I use only observations for which Period 1 average contributions are 0, $\bar{C}^{Endo}_{\text{“}Endo,Exo,\text{”}YY} = 11.7$, $\bar{C}^{Exo}_{\text{“}Endo,Exo,\text{”}YY} = 9.72$ and $\bar{C}^{Exo}_{\text{“}Exo,Exo,\text{”}YY} = 6.00$. These results are consistent with the idea that democratic decision-making in one setting amplifies the behavioral response to a policy not only in that setting, but also in other settings. However, these results are not decisive since the differences are not statistically significant.

[29] In "Exo Imp-yes, Exo Imp-yes," there are two contribution decisions per subject; thus, the average of the two decisions is used as the data for each subject.

fall in the range [15, 20]. Appendix Figure B.4 and Appendix Figure B.5 report empirical distributions for all Yes-Yes voters and Yes-Yes voters whose Period 1 average contributions are in the [4, 9.5] range, respectively; an analysis using these sets of Yes-Yes voters yields qualitatively similar results, confirming the existence of not only the direct effect but also the spill-over effect of democracy.[30]

3.2.4 Does Information Feedback Affect Cooperation Decisions?

In the experimental design, subject receives the information *At-Least* or the information *At-Most*, regardless of whether they belong to a democratic or non-democratic group, Figure 5 reports average contributions of Yes-Yes voters under the treatment condition ("Endo Imp, Exo Imp") and under the control condition ("Exo Imp, Exo Imp") by information feedback.

Yes-Yes Voter:

If a Yes-Yes voter receives the information *At-Least* in an "Imp" condition, the feedback does not give the voter certain information about how his or her counterpart voted, since the event occurs even when the counterpart votes not to use the policy. In contrast, if a Yes-Yes voter receives the information *At-Most* in the "Imp" condition, the voter learns that his or her counterpart voted against the use of the policy. The idea that a Yes-Yes voter's contribution decisions are influenced by a counterpart's voting decision can be seen under the control condition ("Exo Imp, Exo Imp"): $\bar{C}^{Exo}_{\text{“}Exo,Exo,\text{”}YY}$ is higher when Yes-Yes voters are not aware of their counterpart's vote than when they are aware that their counterpart voted not to use the policy, although the difference is not statistically significant.

Regardless of the difference in the information feedback, in all categories, $\bar{C}^{Endo}_{\text{“}Endo,Exo,\text{”}YY}$ and $\bar{C}^{Exo}_{\text{“}Endo,Exo,\text{”}YY}$ are higher than $\bar{C}^{Exo}_{\text{“}Exo,Exo,\text{”}YY}$. Specifically, when subjects are aware that their counterpart voted not to use the policy, the size of the spill-over effect of democracy

[30] Also see Figure B.6 for Yes-Yes voters whose Period 1 average contributions are in the $\{0\}\cup[4, 12.5]$ range and who are under the treatment and control conditions; an analysis using these two sets of Yes-Yes voters yields qualitatively similar results.

For Period 1 average contributions, the empirical distributions of all Yes-Yes voters are significantly different between the treatment and control conditions; however, the empirical distributions of Yes-Yes voters are not significantly different between the two conditions for those whose Period 1 average contributions are in the [4, 9.5] range, in the [4, 12.5] range, and in the $\{0\}\cup[4, 12.5]$ range.

is more than 6 points irrespective of the range of Period 1 average contributions I use; the effect is significant at the 5% level (Figure 5.ii). When subjects do not receive certain information on their counterpart's voting preference, the size of the spill-over effect of democracy is more than 3.5 points but is not significant at the 10% level for the set of all Yes-Yes voters and Yes-Yes voters whose Period 1 average contributions are in the [4, 12.5] range; the size is more than 5 points and is significant when using observations with Period 1 contributions in the [4, 9.5] range (Figure 5.i). Combined with the result that the difference between $\bar{C}^{Endo}_{\text{“}Endo,Exo,\text{”}YY}$ and $\bar{C}^{Exo}_{\text{“}Endo,Exo,\text{”}YY}$ is smaller than that between $\bar{C}^{Exo}_{\text{“}Endo,Exo,\text{”}YY}$ and $\bar{C}^{Exo}_{\text{“}Exo,Exo,\text{”}YY}$, these results suggest that the spill-over effect of democracy, besides the direct effect of democracy, is substantial regardless of which information they received.

The impact of information about a counterpart's voting decision is weak: in each implementation condition, Wilcoxon Signed Ranks/Mann-Whitney tests fail to reject the null hypothesis that average contributions are the same between the two information conditions: the information *At-Least* and the information *At-Most*. This result also suggests that contribution decisions are not sensitive to information about the distribution of votes, and that my results concerning the substantial size of the effects of democracy can be robust to it.[31]

No-No Voter:

If a No-No voter is faced with the "Endo Imp" condition, or if a No-No voter receives the information *At-Least*, the voter learns that his or her counterpart voted to use the policy. If a No-No voter receives the information *At-Most* in the "Exo Imp" condition, this feedback gives the voter no additional information.

In most categories, I do not detect the effects of democracy. But, when No-No voters

[31] In the observations with Period 1 average contributions in the [4, 12.5] range, Yes-Yes voters under the treatment condition received the information *At-Least* in 18 out of 27 democratic groups (66.7%) and in 17 out of 27 non-democratic groups (62.9%); the corresponding number was 20 out of 28 non-democratic groups (71.4%) in "Exo Imp-yes, Exo Imp-yes." In the [4, 9.5] range, the former Yes-Yes voters received that information in 13 out of 19 democratic groups (68.4%) and in 10 out of 19 non-democratic groups (52.6%), and the latter Yes-Yes voters received the same information in 14 out of 20 non-democratic groups (62.5%). The *At-Least* Information was more likely to be received in the non-democratic group under the control rather than the treatment condition. This indicates that the information feedback incorporated in this experiment functioned well in eliminating informational difference across conditions; this implies that results in Sections 3.2.2 and 3.2.3 are not due to the information effect. The fractions of the information *At-Least* received are slightly lower than the theoretical probabilities (See Appendix B).

under the treatment condition learned only in their democratic group that their counterpart voted to use the policy, they contributed weakly significantly more in both their democratic and non-democratic groups, relative to No-No voters who received the same information under the control condition. Although results are not decisive, this suggests that even No-No voters might exhibit both the direct effect of democracy and the spill-over effect of democracy.

3.2.5 Cognitive Reflection and Strategic Sophistication

Why does a subject exhibit the spill-over effect of democracy in addition to the direct effect of democracy?

I examine whether the effects of democracy are related to cognitive reflection and the strategic sophistication of a subject. I find that Yes-Yes voters exhibited both the direct and spill-over effects of democracy, regardless of CRT scores and strategic sophistication under the treatment condition.[32] This result suggests that findings in the previous sections are not due to confusion or errors that subjects might have made (Figure 6).

4 Conclusions

This paper provides evidence that democratic collective decision-making may generate a spill-over effect on behavior beyond the group. Those who experienced an endogenous change in the payoff structure in one group contributed to the group significantly more than those who experienced the same change only exogenously. Furthermore, when those who had experienced the endogenous change were placed in other groups in which the change was imposed exogenously, they contributed significantly more than subjects who had never experienced the endogenous change, even though the other members in the two groups were

[32] I define more-sophisticated subjects as those whose responses to the Beauty Contest Game are less than the median of all 300 subjects' responses. It is noteworthy that whereas more-sophisticated treatment subjects significantly differentiated contribution decisions between their democratic and non-democratic groups, the difference between $\bar{C}^{Exo}_{\text{“}Endo,Exo,\text{”}YY}$ and $\bar{C}^{Exo}_{\text{“}Exo,Exo,\text{”}YY}$ is still around 4 points; although the difference is not significant due to high standard errors (diminished sample size) the impact is a little bit larger than when using all subjects' data (See Section 3.2.1). In the [4, 12.5] and [4, 9.5] ranges, the difference between $\bar{C}^{Exo}_{\text{“}Endo,Exo,\text{”}YY}$ and $\bar{C}^{Exo}_{\text{“}Exo,Exo,\text{”}YY}$ by more-sophisticated subjects is statistically significant.

different from one another. I provide these results while not only controlling thoroughly for selection effects but also cancelling out a comparative advantage of the endogenous process over the exogenous process regarding the information effects.

Resilient pro-social behavioral changes play an important role in civically-engaged societies: norms created within a group can be generalized to interactions outside the group (Putnam 1995 and Larsen *et al.* 2004). Suppose that those who are civically motivated by a particular issue also changed their behaviors in other domains. We could expect that a more intense social system of democratic participation, a greater likelihood that those who become active would meet with like-minded citizens, and cooperation norms, as well as trust, would emerge; eventually, the border of a group may disappear. In this experiment, strikingly, increased cooperation in non-democratic groups by those who previously participated in democratic processes in other groups was almost the same as that in the democratic groups by the same subjects.

This result suggests that when a policy is implemented through a democratic process, the positive effects of an endogenously implemented policy are not restricted to the relevant social dilemma, but may emerge in another social dilemma through resilient behavioral changes triggered by the democratic process. In a firm with more participatory management, workers may have a higher perceived self-determination and intrinsic motivation; this may make coordinating employees within and across sections easier, making the firm more productive. If we do not recognize the crucial spill-over effects of an endogenous process, we may make incorrect decisions about policy implementation. Furthermore, it is noteworthy that even though introducing some rule does not change the substance of the problem in a material sense, democratic processes themselves may resolve not only the problem targeted by such rule, but also various other problems.

References

Anderson, Theodore Wilbur. 1962. "On the Distribution of the Two-Sample Cramer–von Mises Criterion." *The Annals of Mathematical Statistics*, 33(3): 1148-59.

Bardhan, Pranab. 2000. "Irrigation and Cooperation: An Empirical Analysis of 48 Irrigation

Communities in South India." *Economic Development and Cultural Change*, 48(4): 847-65.

Bortolami, Francesca and Luigi Mittone. 2009. "Does Participation in a Collective Decision Affect the Levels of Contributions Provided?" CEEL Working Paper 2-2009.

Botelho, Anabela, Glenn W. Harrison, Lígia M. Costa Pinto and Elisabet E. Rutström. 2005. "Social Norms and Social Choice." Working Paper.

Dal Bó, Pedro. 2010. "Experimental Evidence on the Workings of Democratic Institutions." Working paper, available at http://cniss.wustl.edu/files/cniss/dalbopaper.pdf.

Dal Bó, Ernesto and Pedro Dal Bó. 2010. ""Do the Right Thing:" The Effects of Moral Suasion on Cooperation." Working Paper.

Dal Bó, Pedro, Andrew Foster and Louis Putterman. 2010. "Institutions and Behavior: Experimental Evidence on the Effects of Democracy." *American Economic Review*, 100(5): 2205-29.

Ertan, Arhan, Talbot Page and Louis Putterman. 2009. "Can Endogenously Chosen Institutions Mitigate the Free-Rider Problem and Reduce Perverse Punishment?" *European Economic Review*, 53(5): 495-511.

Fellner, Gerlinde and Gabriele K. Lünser. 2008. "Cooperation in Local and Global Groups." Working paper, Vienna University of Economics & B.A. Department of Economics.

Falk, Armin, Urs Fischbacher and Simon Gächter. forthcoming. "Living in Two Neighborhoods – Social Interaction Effects in the Lab." *Economic Inquiry*.

Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics*, 10(2): 171-8.

Fischbacher, Urs, Simon Gächter and Ernst Fehr. 2001. "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." *Economics Letters*, 71(3): 397-404.

Frederick, Shane. 2005. "Cognitive Reflection and Decision-Making." *Journal of Economic Perspectives*, 19(4): 25-42.

Frey, Bruno S. 1993. "Motivation as a limit to pricing." *Journal of Economic Psychology*, 14(4): 635-64.

Frey, Bruno S. 1997. "A Constitution for Knaves Crowds Out Civic Virtues." *Economic Journal*, 107(443): 1043-53.

Frey, Bruno S. and Reto Jegen. 2001. "Motivation Crowding Theory." *Journal of Economic Survey*, 15(5): 589-611.

Frey, Bruno S. and Felix Oberholzer-Gee. 1997. "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out." *American Economic Review*, 87(4): 746-55.

Gächter, Simon and Christian Thöni. 2005. "Social Learning and Voluntary Cooperation Among Like-Minded People." *Journal of the European Economic Association*, 3(2/3) Papers and Proceedings: 303-14.

Goeree, Jacob K., Charles A. Holt and Susan K. Laury. 2000. "Private Costs and Public Benefits: Unraveling the Effects of Altruism and Noisy Behavior." *Journal of Public Economics*, 83(2): 257-78.

Gneezy, Uri and Aldo Rustichini. 2000a. "Pay Enough or Don't Pay At All." *Quarterly Journal of Economics*, 115(2): 791-810.

Gneezy, Uri and Aldo Rustichini. 2000b. "A Fine is a Price." *Journal of Legal Studies*, 29(1): 1-17.

Gürerk, Őzgür, Bernd Irlenbusch and Bettina Rockenbach. 2006. "The Competitive Advantage of Sanctioning Institutions." *Science*, 312(5770): 108-10.

Guiso, Luigi, Paola Sapienza and Luigi Zingales. 2008. "Long Term Persistence." Working Paper No. 20, Initiative on Global Markets, Graduate School of Business, The University of Chicago.

Gunnthorsdottir, Anna, Daniel Houser and Kevin McCabe. 2007. "Disposition, history and contributions in public goods experiments." *Journal of Economic Behavior and Organization*, 62(2): 304-15.

Haney, Beth, Eugene Borgida and James Farr. 2002. "Citizenship and civic engagement in public problem-solving." In M. X. Delli-Carpini, L. Huddy, and R.Y. Shapiro (eds.), *Research in Micropolitics: New Directions in Political Psychology*: 225-52, Greenwich, CT: JAI Press.

Ichniowski, Casey and Kathryn Shaw. 2003. "Beyond Incentive Pay: Insiders' Estimates of the Value of Complementary Human Resource Management Practices." *Journal of Economic Perspectives*, 17(1): 155-80.

Ichniowski, Casey, Kathryn Shaw and Giovanna Prennushi. 1997. "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines." *American Economic Review*, 87(3): 291-313.

Kamei, Kenju, Louis Puttermana and Jean-Robert Tyran. 2011. "State *or* Nature? Formal vs. Informal Sanctioning in the Voluntary Provision of Public Goods." Department of Economics Working Paper 2011-3, Brown University.

Knack, Stephen and Philip Keefer. 1997. "Does Social Capital Have an Economy Payoff? A Cross-Country Investigation." *Quarterly Journal of Economics*, 112(4): 1251-88.

Kosfeld, Michael, Akira Okada, and Arno Riedl. 2008. "Institution Formation in Public Goods Games." *American Economic Review*, 99(4): 1335-55.

Larsen, Larissa, Sharon L. Harlan, Bob Bolin, Edward J. Hackett, Diane Hope, Andrew Kirby, Amy Nelson, Tom R. Rex, and Shaphard Wolf. 2004. "Bonding and Bridging: Understanding the Relationship between Social Capital and Civic Action." *Journal of Planning Education and Research*, 24(1): 64-77.

Ledyard, John O. 1995. "Public Goods: A Survey of Experimental Research." in John H. Kagel and Alvin E. Roth, eds., *Handbook of experimental economics*. Princeton: Princeton UniversityPress, pp. 111–94.

Markussen, Thomas, Louis Putterman and Jean-Robert Tyran. 2011. "Self-Organization for Collective Action: An Experimental Study of Voting on Formal, Informal, and No Sanction Regimes." Department of Economics Working Paper 2011-4, Brown University.

Muller, Edward N. and Mitschell A. Seligson. 1994. "Civic Culture and Democracy: The Question of Causal Relationships." *American Political Science Review*, 88(3): 635-52.

Nagel, Rosemarie. 1995. "Unraveling in Guessing Games: An Experimental Study." *American Economic Review*, 85(5): 1313-26.

Nese, Annamaria and Patrizia Sbriglia. 2009. "Individuals' Voting Choice and Cooperation in Repeated Social Dilemma Games." Labsi Working Paper.

Newson, Roger. 2002. "Parameters behind "nonparametric" statistics: Kendall's tau, Somers' D and median differences." *The Stata Journal* 2(1): 45–64.

Osterloh, Margit and Bruno S. Frey. 2000. "Knowledge Transfer, and Organizational Forms." *Organization Science*, 11(5): 538-50.

Page, Talbot, Louis Putterman, and Bulent Unel. 2005. "Voluntary Association in Public Goods Experiments: Reciprocity, Mimicry and Efficiency." *Economic Journal*, 115(506): 1032–53.

Putnam, Robert. 1995. "Tuning In, Tuning Out: The Strange Disappearance of Social Capital in America." *Political Science and Politics*, 28(4): 664-83.

Puttermana, Louis, Jean-Robert Tyran and Kenju Kamei. forthcoming. "Public Goods and Voting on Formal Sanction Schemes: An Experiment." *Journal of Public Economics*.

Rauchdobler, Julian, Rupert Sausgruber and Jean-Robert Tyran. 2010. "Voting on Thresholds for Public Goods: Experimental Evidence." *Public Finance Analysis*, 66(1): 34-64.

Rockenbach, Bettina and Irenaeus Wolff. 2009. "Institution design in social dilemmas: How to design if you must?" working paper.

Sausgruber, Rupert and Jean-Robert Tyran. 2007. "Pure redistribution and the provision of public goods." *Economics Letters*, 95(3): 334-8.

Schlozman, Daniel and Ian Yohai. 2008. "How Initiatives Don't Always Make Citizens: Ballot Initiatives in the American States, 1978-2004." *Political Behavior*, 30(4): 469-89.

Scholz, Fritz and Michael Stephens. 1987. "K-sample Anderson-Darling Tests." *Journal of the American Statistical Association* 82(399): 918–24.

Smith, Mark. 2001. "The Contingent Effects of Ballot Initiatives and Candidate Races on Turnout." *American Journal of Political Science*, 45(3): 700-6.

Sutter, Matthias, Stefan Haigner and Martin Kocher. 2010. "Choosing the Carrot or the Stick? – Endogenous Institutional Choice in Social Dilemma Situations." *Review of Economic Studies*, 77(4): 1540-66.

Sutter, Matthias and Hannelore Weck-Hannemann. 2004. "An experimental test of the public-goods crowding-out hypothesis when taxation is endogenous." *Finanzarchiv*, 60(1): 94-110.

Tabellini, Guido. 2005. "Institutions and Culture: Economic Development in the Regions of Europe." CESIFO Working Paper 1492.

Tabellini, Guido. 2007. "Institutions and Culture." CEPR Discussion Papers 6589, C.E.P.R. Discussion Papers.

Tolbert, Caroline J., John A. Grummel, and Daniel A. Smith. 2001. "The Effects of Ballot Initiatives on Voter Turnout in the United States." *American Politics Research*, 29(6): 625-48.

Tolbert, Caroline J., Ramona S. McNeal and Daniel A. Smith. 2003. "Enhancing Civic Engagement: The Effect of Direct Democracy on Political Participation and Knowledge." *State Politics and Policy Quarterly*, 3(1): 23-41.

Tolbert, Caroline J., Daniel Bowen and Todd Donovan. 2009. "Initiative Campaigns: Direct Democracy and Voter Mobilization." *American Politics Research*, 37(1): 155-92.

Tyran, Jean-Robert and Lars P. Feld. 2006. "Achieving Compliance when Legal Sanctions are Non-deterrent." *Scandinavian Journal of Economics*, 108(1): 135-56.

Walker, James M., Roy Gardner, Andrew Herr, Elinor Ostrom. 2000. "Collective Choice in the Commons: Experimental Results on Proposed Allocation Rules and Votes." *Economic Journal*, 110(460): 212-34.

Zelmer, Jennifer. 2003. "Linear Public Goods Experiments: A Meta-Analysis." *Experimental Economics*, 6(3): 299-310.

Table 1: Summary Statistics: the Characteristics of Subjects

(a) Distribution of Votes

	(i) All data	(ii) Voter Type		
		Yes-Yes Voter	No-No Voter	Yes-No Voter
Number of subjects	300	127	143	30

(b) Sample Mean of Individual Characteristics

	(i) All data	(ii) Voter Type		
		Yes-Yes Voter	No-No Voter	Yes-No Voter
(1) Average CRT Score	1.40	1.57	1.36	.83
(2) Average beauty contest game response	36.5	35.1	36.7	41.7
(3) Average SAT (Math) score	731.8	744.1	735.1	663.8
(4) Average SAT (Verbal) Score	716.4	726.8	713.8	685.4
(5) Average number of college or university semesters completed	3.09	3.25	2.92	3.23
(6) Average # of economics courses taken	1.53	1.57	1.64	0.80
(7) Average # of political science courses taken	0.80	0.78	0.82	0.80
(8) The proportion of economics concentrators	15.3%	18.1%	14.0%	10.0%
(9) Average earnings in experiment	$21.38 (S.D. = $5.04)			

Table 2: Individual Voting Decisions and Implementation Outcomes

		Voting decisions in individuals' two assigned groups			
	Total	'yes, yes'	'yes, no'	'no, yes'	'no, no'
"Endo Imp, Exo Imp"	120	67	10	6	37
"Endo Imp, Exo Not"	20	15	1	0	4
"Endo Not, Exo Imp"	54	2	0	3	49
"Endo Not, Exo Not"	6	0	0	1	5
"Exo Imp, Exo Imp"	76	31	8	0	37
"Exo Imp, Exo Not"	22	10	1	0	11
"Exo Not, Exo Not"	2	2	0	0	0
Total	300	127	20	10	143

Notes: "Yes" and "no" indicate that the individual voted for and against, respectively, the use of the sanction policy. Numbers in each cell are the number of subjects who were faced with "X, Y" when their voting decisions in the X group and in the Y group had been a and b, where $a, b \in \{\text{yes, no}\}$, respectively, which I denote by 'a, b.'

Table 3: Average Contributions and Non-Parametric Tests for Yes-Yes voters under the treatment condition ("Endo Imp, Exo Imp") vs. Yes-Yes voters under the control condition ("Exo Imp, Exo Imp")

(a) Average Contributions

	(1) All Data	Yes-Yes voters Period 1 Average Contribution (2) [4,12.5]	(3) [4, 9.5] #1	No-No voters (1) All Data
Treatment condition				
Number of subjects	67	27	19	37
Average contribution in Period 1	9.12	8.00	6.82	3.55
Average contribution in Period 2				
(i) Democratic group	16.55	17.30	18.00	7.35
(ii) Non-democratic group	15.73	17.33	17.89	7.65
Control condition				
Number of subjects	31	14	10	37
Average contribution in Period 1	5.00	7.96	6.60	2.66
Average contribution in Period 2				
(iii) Non-democratic group	12.05	12.96	11.30	5.84

(b) Mann-Whitney/Wilcoxon Signed Ranks Tests

	(1) All Data	Yes-Yes voters Period 1 Average Contribution (2) [4,12.5]	(3) [4, 9.5]	No-No voters (1) All Data
Average Contribution in Period 2 #2				
Wilcoxon Signed Ranks Test (i) vs. (ii)	.16	1.00	.68	.60
Mann-Whitney Test (i) vs. (iii) [Direct effect of democracy]	.0051***	.0234**	.0048***	.45
Mann-Whitney Test (ii) vs. (iii) [Spill-over effect of democracy]	.018**	.018**	.0044***	.41
Average Contribution in Period 1				
Mann-Whitney Test (i) vs. (ii) (iii)	.012**	.835	.64	.21

Notes: Panel (b) reports p-value for two-sided tests. Average contributions in more detailed categories along with the non-parametric test results are found in Appendix Table B.3. Also, average contributions and distribution of votes in other vote outcomes are found in the same table. #1 Intersection between the [4, 12.5] range and Inter-Quartile Range of all 300 subjects' Period 1 average contributions makes the choice of the range more robust to outliers. #2 Somers' D tests where standard errors are adjusted for clusters in Period 1 triad ID, instead of Mann-Whitney tests and Wilcoxon Signed Ranks test, yield qualitatively similar results; results are omitted to conserve space. ** and *** indicate significance at the .05 level and the .01 level, respectively.

Figure 1: Triad Device

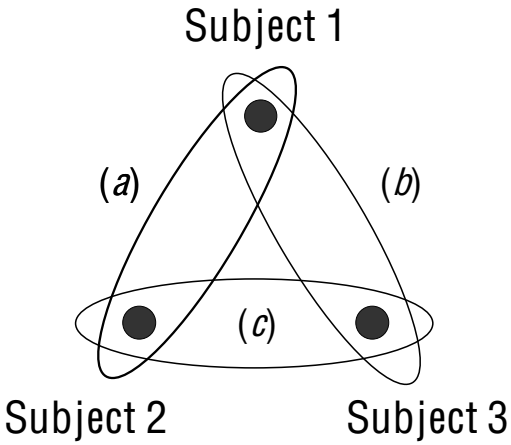
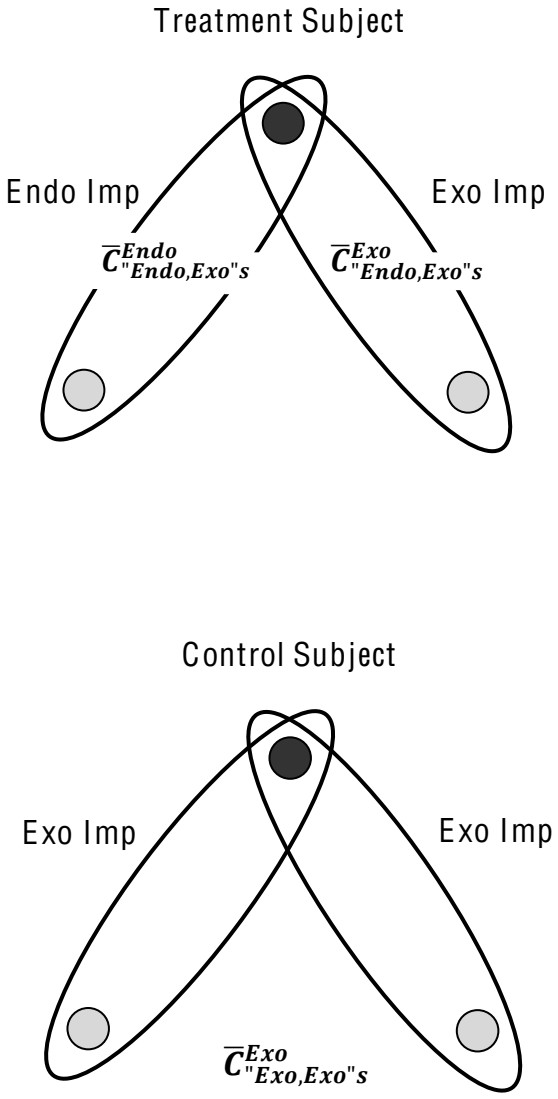
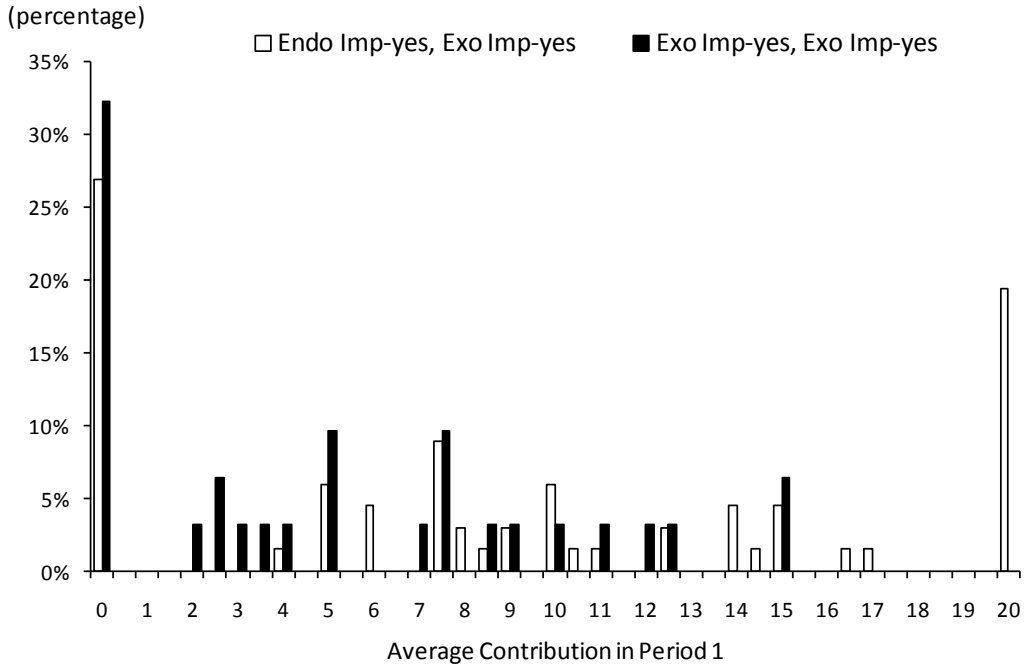


Figure 2: Treatment Subject and Control Subject



Notes: $\bar{C}^{Endo}_{"Endo,Exo"s}$ is the average contribution of voters of type s in their democratic group under the treatment condition ("Endo Imp, Exo Imp") in Period 2. $\bar{C}^{Exo}_{"Endo,Exo"s}$ is the average contribution of voters of type s in their non-democratic group under the same treatment condition in Period 2. $\bar{C}^{Exo}_{"Exo,Exo"s}$ is the average contribution of voters of type s in their non-democratic groups under the control condition ("Exo Imp, Exo Imp") in Period 2. Here, $s \in \{YY, NN\}$, where YY and NN indicate Yes-Yes voters and No-No voters respectively.

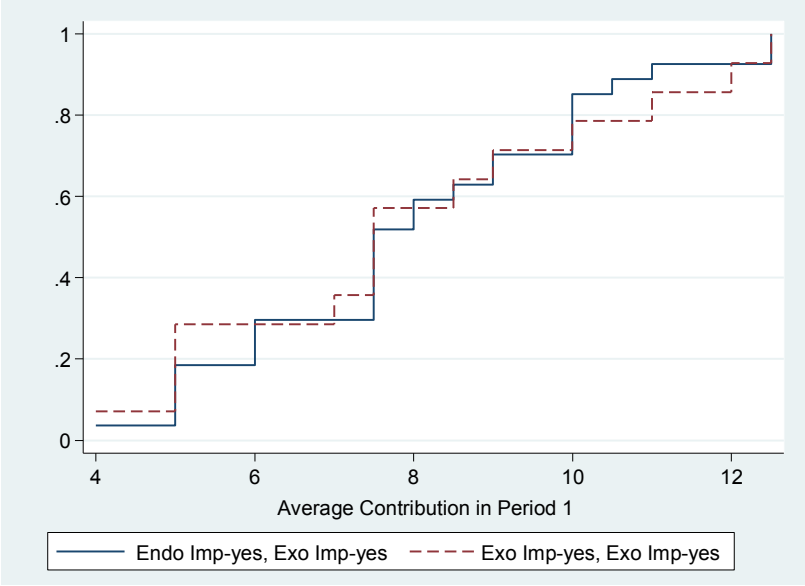
Figure 3: Period 1 Average Contributions by Two Sample Pools: "Endo Imp-yes, Exo Imp-yes" and "Exo Imp-yes, Exo Imp-yes"



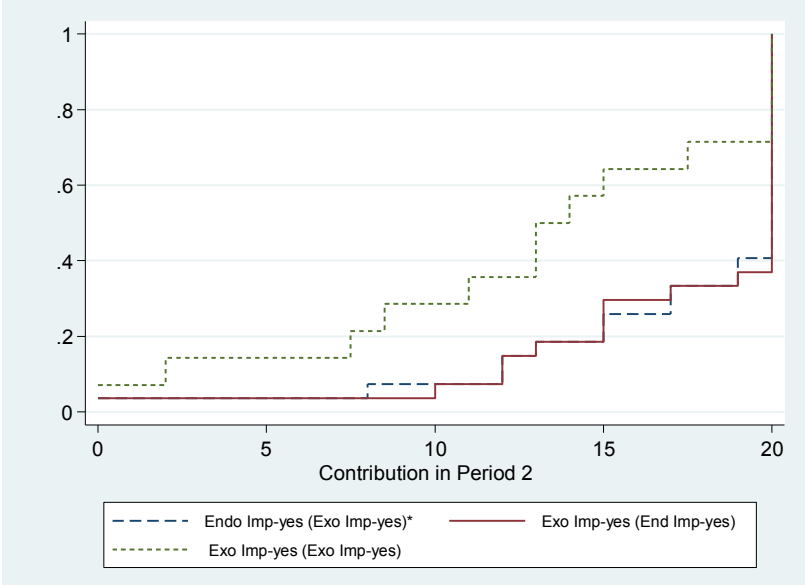
Notes: The difference in the average contribution between "Endo Imp-yes, Exo Imp-yes" and "Exo Imp-yes, Exo Imp-yes" is significant according to a Mann-Whitney test (p-value = .012, 2-tailed). The difference in empirical distributions between "Endo Imp-yes, Exo Imp-yes" and "Exo Imp-yes, Exo Imp-yes" is significant (p-value = .011, 2-tailed, according to a Cramér-von-Mises test; p-value = .006, 2-tailed, according to an Anderson-Darling test). See Anderson (1962) for Cramér-von-Mises tests. Also, see Scholz and Stephens (1987) for Anderson-Darling tests.

Figure 4: Empirical Distribution Before and After the Policy Was Implemented

(a) *Before* the policy was implemented



(b) *After* the policy was implemented

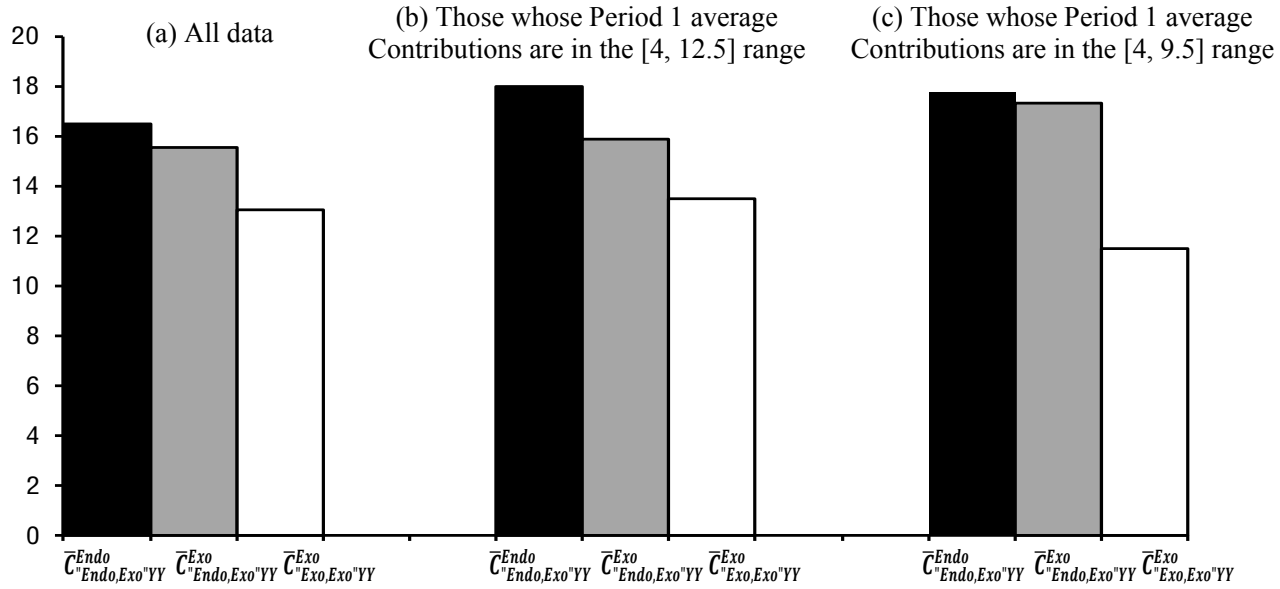


Notes: 1. Yes-Yes voters who are under the treatment condition ("Endo Imp, Exo Imp") or the control condition ("Exo Imp, Exo Imp") and whose Period 1 average contributions are in the [4,12.5] range only. In the control condition, there are two contribution decisions per subject, and thus the average of the two decisions is used as the data for each subject. See Appendix Figure B.4 for empirical distributions of all observations under the treatment condition or the control condition and Appendix Figure B.5 for those of Yes-Yes voters whose Period 1 average contributions are in the [4, 9.5] range. * "X Imp-yes (Y Imp-yes)," where X, Y ∈{Endo, Exo}, indicates the set of contributions in the X group by Yes-Yes voters who are in "X Imp-yes, Y Imp-yes."

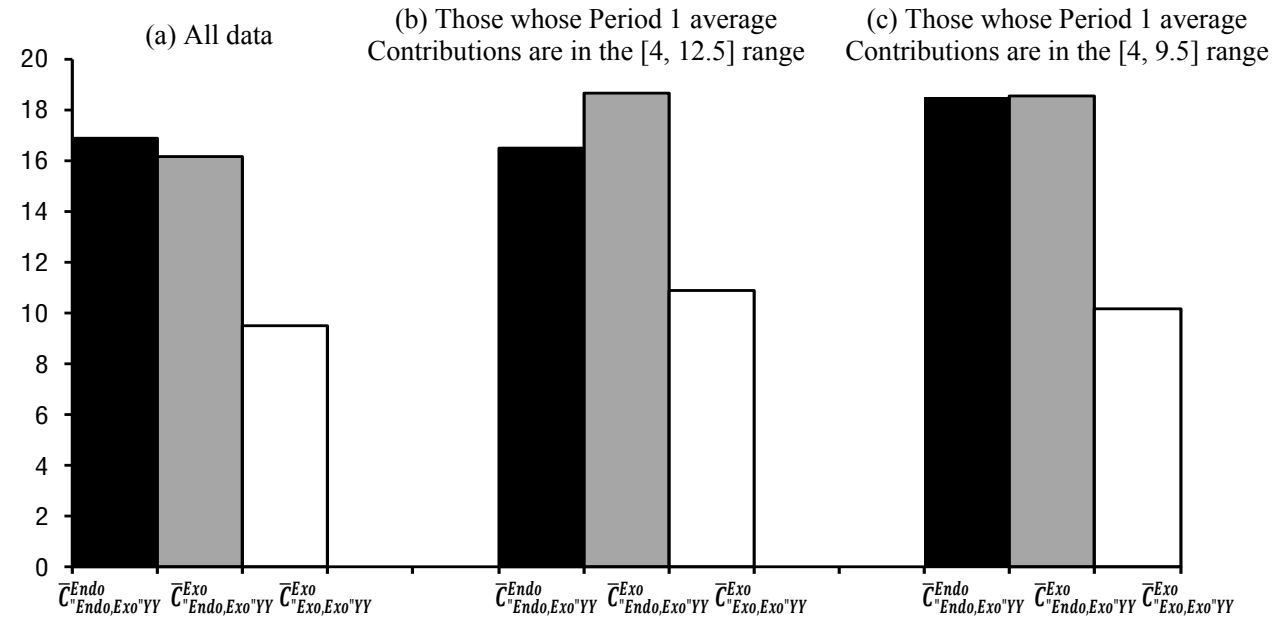
2. There is no statistical difference in the empirical distribution of Period 1 average contributions between these two implementation conditions (p-value = .979, 2-tailed, according to a Cramér-von-Mises test; p-value = .653, 2-tailed, according to an Anderson-Darling test).

3. The difference in the distribution between $C^{Endo}_{"Endo,Exo"YY}$ and $C^{Exo}_{"Exo,Exo,"YY}$ is weakly significant according to a Crémer-von Mises test (p-value = .0839, 2-tailed) but significant according to an Anderson Darling test (p-value = .0121, 2-tailed); the difference in the distribution between $C^{Exo}_{"Endo,Exo,"YY}$ and $C^{Exo}_{"Exo,Exo,"YY}$ is also weakly significant according to a Crémer-von Mises test (p-value = .0946, 2-tailed) but significant according to an Anderson Darling test (p-value = .00697, 2-tailed). In contrast, the difference in the distribution between $C^{Endo}_{"Endo,Exo,"YY}$ and $C^{Exo}_{"Endo,Exo,"YY}$ is not significant according to a Crémer-von Mises test (p-value = 1.00, 2-tailed) as well as an Anderson Darling test (p-value = 0.704, 2-tailed).

Figure 5: Contribution Decisions by the Information Feedback for the Treatment Subjects ("Endo Imp-yes, Exo Imp-yes") and the Control Subjects ("Exo Imp-yes, Exo Imp-yes")



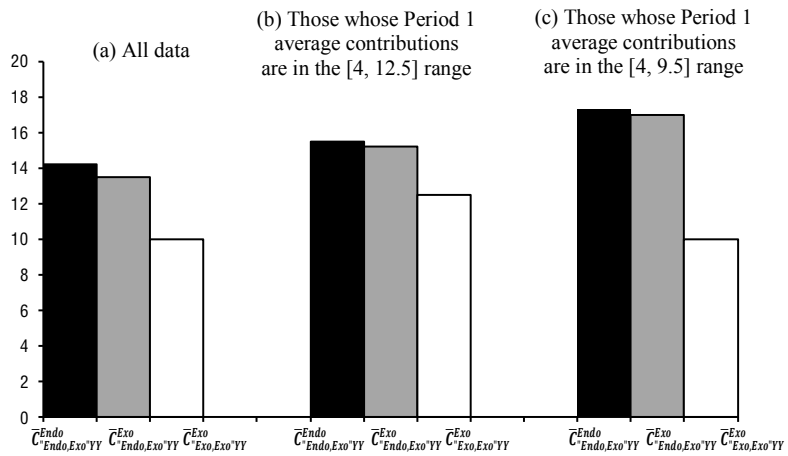
(i) When Yes-Yes voters *did not* learn that their counterpart voted not to use the policy (Yes-Yes voters receive the information *At-Least*)



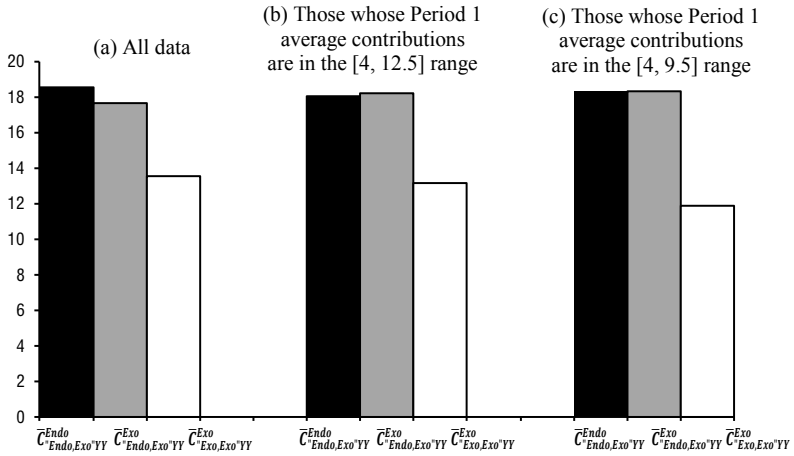
(ii) When Yes-Yes voters *did* learn that their counterpart voted not to use the policy (Yes-Yes voters receive the information *At-Most*)

Notes: Each subject received one kind of information for each of their two assigned groups. $\bar{C}^{Endo}_{"Endo,Exo"YY}$, $\bar{C}^{Exo}_{"Endo,Exo"YY}$ and $\bar{C}^{Exo}_{"Exo,Exo"YY}$ in Figure 5 are the means of all contribution decisions under the information *At-Least* or the information *At-Most*. Appendix Table B.6 reports average contributions and non-parametric test results by the information feedback and the implementation outcome.

Figure 6: Contribution Decisions by the CRT score and by Strategic Sophistication for the Treatment Subjects ("Endo Imp-yes, Exo Imp-yes") and the Control Subjects ("Exo Imp-yes, Exo Imp-yes")

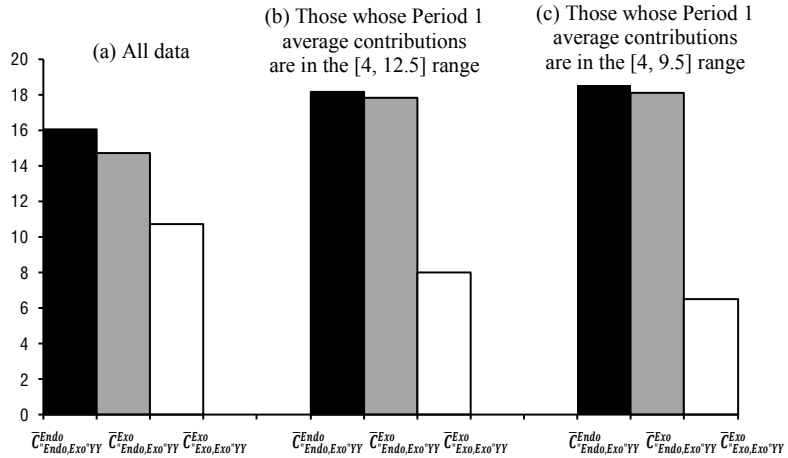


(i) High CRT score (= 2, 3)

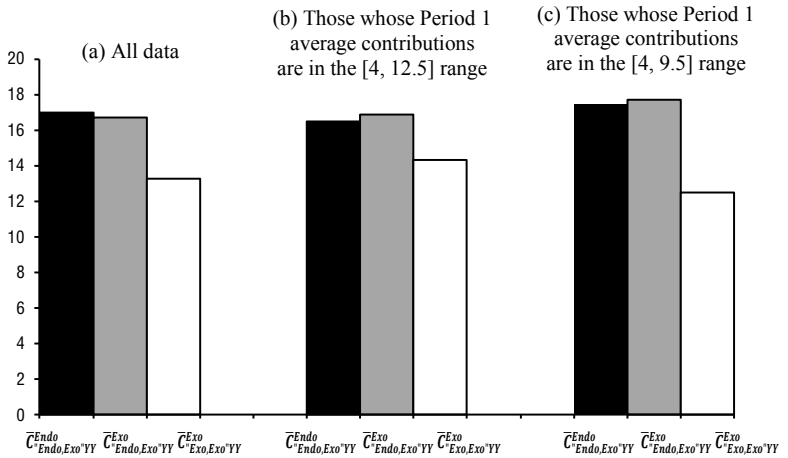


(ii) Low CRT score (= 0, 1)

(1) CRT score



(i) Low guess response (< 35)



(ii) High guess response (≥ 35)

(2) Strategic Sophistication

Note: The average of all 300 subjects' CRT scores is 1.40. The median of all 300 subjects' Beauty Contest Game responses is 35. Detailed results on statistical tests are found in notes and Tables B7-B10 in Appendix B.