

# Investing in institutions for cooperation

Alexander Smith<sup>1</sup>  · Xi Wen<sup>1</sup>

Received: 26 April 2016 / Revised: 6 February 2017 / Accepted: 14 February 2017 /  
Published online: 10 April 2017  
© Economic Science Association 2017

**Abstract** We examine the effects of endogenizing contribution productivity in a repeated public good game. In our experimental treatment, subjects collectively decide (by voting) how much to invest in augmenting the technology for producing the public good, and subsequently make individual voluntary contributions to provision. In the control, contribution productivity is exogenous. Contributions in the two treatments are similar.

**Keywords** Public good games · Endogenous institutions · Production technology · Contribution productivity

**JEL Classification** C71 · C91 · C92 · H41

## 1 Introduction

Incentives to free ride are pervasive in team production, creating a tension for individuals between serving the team by providing full effort and serving themselves by shirking. Standard voluntary contribution mechanism (VCM) public good games capture this tension very well. However, they do not reflect that in real-

---

**Electronic supplementary material** The online version of this article (doi:[10.1007/s40881-017-0033-2](https://doi.org/10.1007/s40881-017-0033-2)) contains supplementary material, which is available to authorized users.

---

✉ Alexander Smith  
adksmith@wpi.edu

Xi Wen  
xwen@alum.wpi.edu

<sup>1</sup> Department of Social Science and Policy Studies (SSPS), Worcester Polytechnic Institute (WPI), 100 Institute Road, Worcester, MA 01609, USA

world team production, teams often have some ability to change the productivity of effort. By investing in better workspaces, computers, equipment, tools, support staff, etc., the productivity of effort can be increased, potentially affecting the subsequent provision of effort. Furthermore, being involved in determining the productivity of effort might affect provision decisions.

To study any such effects, we conduct a public good game experiment in which contribution productivity is endogenous. Subjects collectively decide (by voting) how much to invest in augmenting the technology used for producing the public good, and subsequently make individual voluntary contributions to provision, as in standard games (Marwell and Ames 1979; Isaac et al. 1984). We compare contributions to a control treatment in which contribution productivity is on average the same, but imposed exogenously. Our aim is determining if having subjects choose contribution productivity reduces the free riding problem typical in the voluntary provision of public goods (Ledyard 1995; Chaudhuri 2011).

We find that while subjects in the experimental treatment invest considerable amounts in contribution productivity, their contributions to provision of the public good are not significantly different from in the control treatment. Our main contribution to the literature is thus a null result on how endogenizing contribution productivity affects contributions. We also contribute to the broader literature on how the endogenous selection of parameters and/or institutions affects behavior in cooperative games. Specifically, our results are consistent with those of Kingsley and Brown (2016), who fail to find that endogenous institutional choice promotes contributions, but we contrast Dal Bo et al. (2010) and Sutter et al. (2010), who report that pro-social behavior increases when subjects choose their institutions.

Previous papers on endogenous institutions in public good games focus mainly on the implementation and subsequent effects of punishment mechanisms (Gurerk et al. 2006; Tyran and Feld 2006; Ertan et al. 2009; Kosfeld et al. 2009; Putterman et al. 2011; Markussen et al. 2014). Related to this, Sutter et al. (2010) find that choosing whether to have reward or punishment mechanisms increases contributions compared to when the same institutions are imposed exogenously. Kingsley and Brown (2016) have a central authority that detects self-interested behavior only with some probability. One would think that endogenously choosing the detection probability might increase contributions, but Kingsley and Brown (2016) fail to find evidence of this. On the topic of choosing the benefits of cooperation, Dal Bo et al. (2010) report that when subjects vote to change the payoffs in a prisoner's dilemma game in a way that promotes cooperation, subjects are more cooperative than when the same payoffs are implemented exogenously.

Norton and Isaac (2010) also study the endogenous determination of the benefits of cooperation. Each group's manager chooses contribution productivity and others make voluntary contributions. Managers typically choose high productivity, and when the manager's decision history is provided to the others, contribution decay is often avoided. Under the right conditions, endogenous institutional choice, therefore, promotes contributions, which is different from what we find. Finally, in Isaac and Norton (2013), subjects vote over taxes that serve as preliminary contributions that can later be supplemented by individual contributions. Total provision is higher than in a baseline treatment that does not allow taxation.

Endogenous institutional choice, therefore, increases provision, once again contrasting our results.

## 2 The experiment

Contribution productivity is endogenous in the *experimental* treatment and exogenous in the *control*. In both treatments, subjects are assigned to groups of four that are fixed for ten rounds. They receive endowments of ten lab dollars (LD; later converted to USD at a rate of 1 LD = 0.1 USD) at the start of each round.

### 2.1 The experimental treatment

Each round has *investment* and *contribution* stages. In the investment stage, subjects vote over how many LD each subject will invest in contribution productivity (each subject invests the same amount). Specifically, each subject submits a number between 0 and 10 (inclusive; up to one digit after the decimal point allowed). Investment is the median of the four votes (mean of the two middle votes).<sup>1</sup> Note that it is not incentive compatible to reveal true preferences. Subjects have an incentive to exaggerate their preferences to move the outcome toward their true preferences. An alternative voting scheme, such as majority voting between two investment amounts, would make it weakly dominant to truthfully reveal. However, it would also restrict the potential outcomes. We opt to instead elicit a continuum of investment outcomes, so that when we later examine the relationship between contributions and investment, both are continuous variables.<sup>2</sup>

Each subject has the investment amount deducted from her endowment of ten LD and contribution productivity is:

$$M = \sqrt{0.1 \cdot I} \quad (1)$$

where  $I$  is investment and  $M$  stands for “multiplier” (the amount by which the sum of contributions is later multiplied to determine each subject’s return from the public good). The returns to investment are positive ( $M_I > 0$ ), but diminishing ( $M_{II} < 0$ ), which incorporates empirical realism, and moves the social optimum of the game away from the center of the action space, where it would be if the relationship between  $I$  and  $M$  was linear.<sup>3,4</sup>

<sup>1</sup> With four person groups, there is not a unique vote that is the median vote. However, four person groups are the most common in the literature (Ledyard 1995; Chaudhuri 2011), and we want for our only departure from the literature to be the addition of the investment stage.

<sup>2</sup> There exist voting rules that make it weakly dominant to truthfully reveal and elicit a continuum of investment outcomes. For example, investment could be the lower (or higher) of the two middle votes. However, this potentially creates downward (or upward) pressure on investment.

<sup>3</sup> The endogenous determination of  $M$  means that  $M$  could be less than 0.25, in which case the contribution stage is no longer a social dilemma. However, the concavity of  $M$  means that very low investment is required for this to happen. The lowest investment that occurred was  $I = 0.75 \rightarrow M = 0.27$ .

<sup>4</sup> The investment stage adds complexity to the conventional game. To prevent confusion, the instructions (see Appendix) include multiple examples of determining  $M$  from different sets of votes. In addition, we

After the voting, subjects learn  $M$ , and proceed to the contribution stage, which is like conventional games. However, instead of having their full endowments of ten LD, subjects have a budget of  $10 - I$  LD. In addition, contribution productivity is the  $M$  that was chosen in the investment stage. As such, the payoffs are:

$$\pi_s = 10 - I - c_s + M(I) \cdot \sum_{t=1}^4 c_t \quad (2)$$

where  $c_s$  is the contribution of subject  $s$ , whose four group members are indexed by  $t$ .

Having subjects complete the contribution stage with a budget of  $10 - I$  LD allows us to later analyze how they make unconstrained allocations between investing, contributing, and keeping money for themselves. An alternative design would give subjects separate budgets for investing and contributing, but this would restrict substitution between investment and contributions. In addition, specifying payoffs as linear in contributions is consistent with the bulk of the previous literature on VCM public good games (Ledyard 1995; Chaudhuri 2011). Again, we want for our primary departure from the previous literature to be the addition of the investment stage, and not changes to the VCM.

At the end of each round, subjects are told the sum of contributions to the group account and their payoffs from the round. At the end of the ten rounds, each subject's ten payoffs are converted to USD and added to a \$5 show-up fee.

### 2.1.1 Subgame perfect Nash equilibrium

Assuming individual wealth maximization, a subgame perfect Nash equilibrium to the game is for every round to have:

1. Votes  $(v_1, v_2, v_3, v_4)$  such that  $v_1 = v_2 = v_3 = 0 \leq v_4$ , and
2. Contributions  $c_1 = c_2 = c_3 = c_4 = 0$ .

The votes achieve  $I = 0$ , and since there are no contributions, the payoffs are ten LD for each subject in each round (subjects simply keep their endowments).<sup>5</sup>

### 2.1.2 Social optimum

The social optimum, in contrast, involves all subjects fully expending their endowments. The critical thing is finding the optimal balance between investment and contributions. The social optimum is for every round to have:

---

Footnote 4 continued

are an engineering school, with students who are generally very comfortable with equations. All students are required to do a year of calculus to fulfill general degree requirements.

<sup>5</sup> There are other Nash equilibria. They involve sets of votes such that a unilateral deviation in voting does not change the investment outcome. Such equilibria result in positive investment. However, it is never optimal to contribute anything.

1. A set of votes  $v_1 \leq v_2 \leq v_3 \leq v_4$  such that  $\frac{(v_2+v_3)}{2} = \frac{10}{3}$ , and
2. Contributions  $c_1 = c_2 = c_3 = c_4 = \frac{20}{3}$ .

The votes imply  $I = 3.33 \rightarrow M = 0.58$ , and when subjects subsequently contribute all that remains (6.67 LD each), their payoffs are 15.40 LD.

## 2.2 Control treatment

The control treatment is a conventional repeated public good game with parameters chosen to on average replicate the incentives of the contribution stage in the experimental treatment. This creates a baseline for identifying the effect on contributions of having subjects choose  $M$ . Using the data from the experimental treatment, we calculate the average  $M$  and associated remaining budget ( $10 - I$ ) in each round. Subjects in the control treatment play a game in which the  $M$  in each round is the average  $M$  from the same round in the experimental treatment, and the endowments in each round are the average remaining budgets (after investment) from the same round in the experimental treatment.

An alternative design would match each group in the control to a group in the experimental treatment, and have each group in the control face the sequence of  $M$ s chosen by its experimental treatment counterpart. The problem with this type of matched pairs design is that without being able to match groups based on similarity of observable characteristics (which is impractical given the need to conduct many sessions one at a time), variations in behavior due to group characteristics would create noise that would bias the comparison of the treatments toward failing to reject the null hypothesis of no difference.

To create parallelism between the treatments, the control treatment also has two stages: the *information* and *contribution* stages. In the information stage, subjects are told their endowments and  $M$  in the round. The contribution stage then proceeds as in the experimental treatment. In this game, the Nash equilibrium is contributing nothing, and the social optimum is contributing everything.

## 3 Results

The experiment was programmed and conducted with the experiment software z-Tree (Fischbacher 2007). Student subjects were recruited from introductory economics classes at our school. A total of 136 participated. There were 88 subjects in the experimental treatment (11 sessions of eight) and 48 in the control (six sessions of eight). Experimental and control sessions lasted about 45 and 35 min. Average total earnings were \$15.31 and \$15.58 (including the \$5 show-up fee). Summary statistics are given in Table 1.

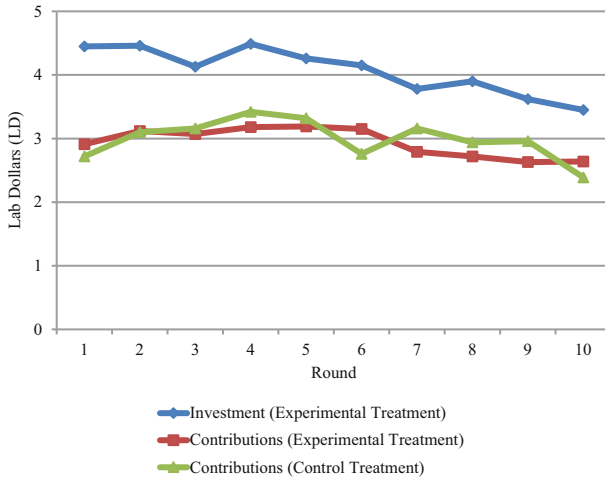
**Table 1** Summary statistics

	Mean	Std dev	Median	Min	Max	<i>n</i>
Experimental treatment						
vote	4.24	2.43	4.00	0.00	10.00	880
investment	4.07	1.41	4.13	0.75	9.00	220
<i>M</i>	0.63	0.12	0.64	0.27	0.95	220
contribution	2.94	1.85	3.00	0.00	9.05	880
payoff	10.31	2.22	10.39	2.37	16.85	880
Top 11 group level outcomes (by average cumulative earnings)						
ave. vote	4.08	0.78	4.21	2.92	5.14	11
ave. invest.	3.93	0.96	3.95	2.46	5.20	11
ave. <i>M</i>	0.62	0.08	0.63	0.48	0.72	11
ave. cont.	3.55	0.66	3.57	2.79	5.19	11
ave. cum. earn.	119.50	8.49	112.79	102.74	132.08	11
Bottom 11 group level outcomes (by average cumulative earnings)						
ave. vote	4.40	0.68	4.54	3.48	5.33	11
ave. invest.	4.20	0.77	4.32	3.00	5.40	11
ave. <i>M</i>	0.64	0.06	0.65	0.53	0.73	11
ave. cont.	2.33	0.35	2.28	1.71	2.93	11
ave. cum. earn.	94.20	6.19	94.37	82.56	101.30	11
Control treatment						
contribution	2.99	1.49	3.00	0.00	6.51	480
payoff	10.58	1.73	10.55	4.24	15.09	480
Top 6 group level outcomes (by average cumulative earnings)						
ave. cont.	3.48	0.12	3.50	3.32	3.62	6
ave. cum. earn.	113.20	1.73	113.27	110.83	115.20	6
Bottom 6 group level outcomes (by average cumulative earnings)						
ave. cont.	2.51	0.47	2.67	1.70	3.01	6
ave. cum. earn.	98.49	7.09	100.84	86.34	105.92	6

### 3.1 The trends

In Fig. 1, we plot the trends of average investment and contributions in the experimental treatment, and average contributions in the control.<sup>6</sup> In the experimental treatment, investment begins considerably higher than the socially optimal 3.33 LD, but declines toward 3.33 LD as the rounds proceed. Using a Wilcoxon signed rank test, and treating each group as an independent observation, investment is significantly different from 3.33 LD in each of the first seven rounds ( $p \leq 0.02$ ;  $n = 22$  for each test), but not significantly different from 3.33 LD in each of the final three rounds ( $p \geq 0.19$ ;  $n = 22$  for each test). Contributions, on the other hand, are

<sup>6</sup> We omit the trend of average votes in the experimental treatment because it is very similar to the trend of average investment.



**Fig. 1** Trends of investment and contributions

**Table 2** Random effects regressions of key variables on a linear time trend

dep var	(1) vote	(2) invest	(3) cont	(4) cont	(5) payoff	(6) payoff
round	-0.08 (0.03)***	-0.11 (0.03)***	-0.05 (0.03)*	-0.04 (0.03)	-0.05 (0.04)	-0.05 (0.04)
constant	4.68 (0.21)***	4.68 (0.24)***	3.24 (0.17)***	3.21 (0.16)***	10.56 (0.30)***	10.85 (0.27)***
treatment	Exp	Exp	Exp	Cont	Exp	Cont
subjects	88	-	88	48	88	48
groups	-	22	-	-	-	-
rounds	1-10	1-10	1-10	1-10	1-10	1-10
<i>n</i>	880	220	880	480	880	480
<i>R</i> <sup>2</sup>	0.01	0.05	0.01	0.01	0.00	0.01

Standard errors adjusted for clustering at the group level are reported in parentheses

\*\*\*  $p < 0.01$

\*\*  $p < 0.05$

\*  $p < 0.1$

consistently well short of the socially optimal 6.67 LD, and appear to be quite similar between treatments.

We present statistical evidence on the significance of the trends in Table 2. While the tendencies for votes and investment in the experimental treatment to decline over time are highly significant, the evidence on the decline of contributions in the two treatments is at most weakly significant. Payoffs are also fairly stable over time.

Now treating group level averages as the independent units of observation, we have 22 observations from the experimental treatment and 12 from the control. The difference in average contributions between the treatments ( $2.99 - 2.94 = 0.05$  LD) is not significant (Wilcoxon ranksum  $p = 0.77$ ), failing to provide evidence that having subjects endogenously determine contribution productivity promotes cooperation. A power analysis indicates that to reject the null hypothesis of no difference at the 5% level of significance 80% of the time, a difference of 0.69 LD is required. Increasing the sample size of the control treatment to 22 independent groups would only reduce the required difference to 0.61 LD. Furthermore, for our observed difference of 0.05 LD to be significant at the 5% level 80% of the time, we would need a sample of 3190 groups in each treatment.<sup>7</sup>

### 3.2 Differences between groups

In addition to the basic summary statistics that we present in Table 1, we also report statistics on group level heterogeneity. Specifically, for each treatment, we partition the full set of groups based on group level average cumulative earnings. For the top and bottom halves of the groups, we report summary statistics on a variety of variables, where the individual values of the variables are the group level averages. While votes, investment and  $M$ s in the experimental treatment all tend to be slightly higher for the bottom half of groups than for the top half, none of the differences are statistically significant (Wilcoxon ranksum  $p \geq 0.41$  in all three cases). However, for both treatments, the differences in average contributions between the two halves are highly significant (Wilcoxon  $p < 0.01$  in both cases).<sup>8</sup> The nonparametric analysis, therefore, suggests that it is not differences in voting and investment (and thus  $M$ ) between groups that underlie differences in cumulative earnings. Rather, cumulative earnings appear to be most closely related to contributions.

We continue our examination of differences between groups by regressing average payoffs at the group level (in each round) on  $M$  and average contributions (see Table 3).  $M$  has a negative effect on average payoffs in the experimental treatment, but the effect of  $M$  is not significant in the control. In contrast, the effect of average contributions is highly significant in both treatments, emphasizing the important role of contributions in determining payoffs.

### 3.3 Analysis of individual votes and contributions

We now examine the correlates of individual votes and contributions (see Table 4). We present the results of fixed effects models in light of consistent evidence (in the models that we present in Table 4) that the corresponding random effects models are biased (see the Hausman  $p$  values at the bottom of the table). Regressing votes and

<sup>7</sup> The analysis of the difference in average payoffs between treatments is very similar (results available upon request).

<sup>8</sup> The differences in average cumulative earnings, of course, are also highly significant (Wilcoxon  $p < 0.01$  in both cases).



**Table 3** Random effects regressions of average payoffs at the group level

dep variable	(1) payoff	(2) payoff
<i>M</i>	-2.45 (1.13)**	-0.97 (0.61)
ave. cont.	1.23 (0.16)***	1.51 (0.02)***
constant	8.21 (0.40)***	6.67 (0.41)***
treatment	Exp	Cont
groups	22	12
rounds	1–10	1–10
<i>n</i>	220	120
<i>R</i> <sup>2</sup>	0.79	0.99

Standard errors adjusted for clustering at the group level are reported in parentheses

\*\*\*  $p < 0.01$

\*\*  $p < 0.05$

\*  $p < 0.1$

**Table 4** Fixed effect regressions of individual votes and contributions

dep variable	(1) vote	(2) cont	(3) cont	(4) cont	(5) cont
vote <sub>-1</sub>	0.09 (0.06)	–	–	–	–
vote deviation <sub>-1</sub>	-0.08 (0.07)	–	–	–	–
vote	–	0.01 (0.06)	–	–	–
vote deviation	–	–	0.02 (0.06)	–	–
vote deviation <sup>+</sup>	–	–	–	0.02 (0.08)	–
vote deviation <sup>-</sup>	–	–	–	0.02 (0.09)	–
contribution <sub>-1</sub>	0.14 (0.06)**	0.07 (0.06)	0.07 (0.06)	0.07 (0.06)	0.18 (0.10)*
cont. deviation <sub>-1</sub>	-0.02 (0.02)	-0.03 (0.03)	-0.03 (0.03)	-0.03 (0.03)	-0.07 (0.03)**
<i>M</i>	–	-1.08 (1.27)	-0.94 (1.24)	-0.94 (1.29)	0.98 (2.77)
round	-0.07 (0.03)*	-0.08 (0.03)**	-0.08 (0.03)**	-0.08 (0.03)**	-0.07 (0.03)**
constant	3.78 (0.47)***	3.71 (0.84)***	3.66 (0.85)***	3.66 (0.85)***	1.83 (1.71)
treatment	Exp	Exp	Exp	Exp	Cont
subjects	88	88	88	88	48
rounds	2–10	2–10	2–10	2–10	2–10
<i>n</i>	792	792	792	792	432
<i>R</i> <sup>2</sup>	0.05	0.10	0.11	0.11	0.20
Hausman <i>p</i>	<0.01	<0.01	<0.01	<0.01	<0.01

Standard errors adjusted for clustering at the group level are reported in parentheses

\*\*\*  $p < 0.01$

\*\*  $p < 0.05$

\*  $p < 0.1$

contributions on variables capturing previous votes and contributions introduce multiple potential sources of endogeneity, including omitted variable biases and simultaneity. Fixed effects control for the effects of any time invariant omitted

variables (such as unobservable individual factors) that might simultaneously cause votes and contributions.<sup>9</sup>

In model (1) (experimental treatment), contributions in the previous round have a positive effect on votes, suggesting that there is some “momentum” or “inertia” associated with pro-social behavior such that when one makes a high contribution, she then has an increased inclination to vote high in the next round, perhaps as a means of incentivizing others to contribute (due to a higher  $M$ ). In model (2), the round is the only significant correlate of contributions. In model (3), to include how the subject’s vote deviated from the median, we have to drop votes from the model, since vote deviations are a linear combination of votes and  $M$ ; vote deviations do not have a significant effect.

Model (4) has two vote deviation terms: one for the amount by which the subject’s vote exceeded the median (takes on a value of zero if the subject’s vote was less than the median), and another for the amount by which the subject’s vote was less than the median (takes on a value of zero if the subject’s vote was higher than the median). Neither vote deviation variable is significant, suggesting that voting differently from other group members is not associated with higher or lower contributions. The most important result from models (2)–(4), however, is that the effect of  $M$  is never significant, failing to provide evidence that a higher  $M$  creates sufficient incentive for subjects to contribute more.<sup>10</sup>

In model (5)(control treatment), contributions are positively associated with contributions in the previous round and negatively associated with the subject’s deviation from the average contribution of others in the previous round. The latter effect is consistent with previous literature (Ashley et al. 2010; Smith 2015). The effect of  $M$  is once again not significant.

## 4 Discussion

Our two main results are that: (1) compared to the social optimum, people over-invest and under-contribute, and (2) endogenously determining contribution productivity does not have a significant effect on contributions. Both findings merit further discussion. We designed our experiment to measure the levels of investment and contributions when contribution productivity is endogenously determined by subjects. Definitively determining *why* the levels are what they are is beyond the scope of the study.

A possible explanation is that the early round over-investment may reflect an attempt to encourage others to cooperate. Alternatively, it could be the result of the incentives of our voting mechanism to exaggerate preferences over  $I$ . In either case,

<sup>9</sup> Smith (2013) discusses using fixed effects to help correct the bias associated with estimating autoregressive models using least squares. See Smith (2012) for a paper on the importance of correctly modeling the autoregressive nature of behavior in public good games.

<sup>10</sup> At least in absolute terms. We do find that subjects contribute higher proportions of their remaining money when they have higher  $M$ s, consistent with Isaac and Walker (1988) finding that subjects choose higher contribution percentages when they have higher MPCRs. These regressions are available upon request.

it seems plausible that as subjects learn that a higher  $M$  does not lead to higher contributions, they reduce their investment until in the final rounds, it converges to the social optimum. The under-contributions are similar to that in previous experiments in which subjects start off by contributing about 50% (Ledyard 1995; Chaudhuri 2011). A potential reason for why contributions do not unravel over time in our experiment is that it is a reaction to the falling  $M$  in both treatments (and increasing amounts of money, or “remaining budgets,” from which to make contributions).

Our null result on the effect of endogenizing contribution productivity, which contrasts previous literature on endogenous institutions (Dal Bo et al. 2010; Norton and Isaac 2010; Sutter et al. 2010; Isaac and Norton 2013), is potentially attributable to a variety of factors. First, votes and contributions are chosen from continuous action spaces. Sutter et al. (2010), who have subjects choose one of three institutions, speak directly on the importance of the size of the action space, explaining that they “wanted to keep the design simple so that subjects fully understood the available institutions in the endogenous treatments. Any more complicated reward or punishment technology would have made the choice task of the participants more difficult” (Sutter et al. 2010, p. 1544).

Dal Bo et al. (2010) prisoner’s dilemma of course has a binary action space, and rates of cooperation are known to be higher with smaller action sets (Gangadharan and Nikiforakis 2009), creating the possibility of observing different treatment effects for no other reason than two games having action sets of different sizes. Dal Bo et al. (2010) experiment is also, to some extent, a different kind of game from ours because when subjects elect to change the payoffs in the prisoner’s dilemma game, the game transforms into a coordination game, with high- and low-payoff Nash equilibria, making whether the game is a coordination game potentially relevant as well.

Finally, our voting mechanism could be part of the explanation behind our null result. Since some subjects, under certain circumstances, have an incentive to exaggerate their preferences, our voting mechanism does not provide as strong a signaling opportunity as it could. However, it is not clear that signaling opportunities are better in other environments. In Dal Bo et al. (2010), subjects vote in groups of four over changing prisoner’s dilemma payoffs. They learn something about the distribution of votes from the implemented outcome (chosen by majority), but they are never informed of the distribution of votes, nor whether the computer had to break a tie (which happens a lot). In Sutter et al. (2010), unanimity is required for an institution to be implemented. Each of four subjects in the group has to “accept” a particular institution for it to be implemented, and while this indicates something about each subject’s preference for a particular institution, it says nothing about any individual’s preference ordering over the three institutions.

In our experiment, a different decision rule, such as majority voting between two investment levels, or defining the investment level as the lower (or higher) of the two middle votes, might lead to different investment outcomes, and different contribution behavior. We can only speculate about how a different voting mechanism might affect our findings. Our null result should thus be taken with caution. What is clear is that we provide contrasting results to previous studies

reporting that choosing institutions improves outcomes (Dal Bo et al. 2010; Sutter et al. 2010), suggesting that more research is required to determine the conditions under which the “democracy premium” occurs.

**Acknowledgements** Financial support from Worcester Polytechnic Institute (WPI) is gratefully acknowledged. We thank seminar participants from Temple University, the College of the Holy Cross, and the University of Massachusetts at Lowell. We thank conference participants from the 2013 North American ESAs in Santa Cruz, the 2014 International ESAs in Honolulu, and the 2014 North American ESAs in Fort Lauderdale. We are also grateful to the anonymous referees and Co-Editor Nikos Nikiforakis for excellent, and very helpful, comments and suggestions.

## References

- Ashley, R., Ball, S., & Eckel, C. (2010). Motives for giving: A reanalysis of two classic public goods experiments. *Southern Economic Journal*, 77(1), 15–26.
- Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics*, 14(1), 47–83.
- Dal Bo, P., Foster, A., & Putterman, L. (2010). Institutions and behavior: Experimental evidence on the effects of democracy. *American Economic Review*, 100(5), 2205–2229.
- Ertan, A., Page, T., & Putterman, L. (2009). Who to punish? Individual decisions and majority rule in mitigating the free rider problem. *European Economic Review*, 53(5), 495–511.
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171–178.
- Gangadharan, L., & Nikiforakis, N. (2009). Does the size of the action set matter for cooperation? *Economics Letters*, 104(3), 115–117.
- Gurerk, O., Irlenbusch, B., & Rockenbach, B. (2006). The competitive advantage of sanctioning institutions. *Science*, 312(5770), 108–111.
- Isaac, R. M., & Norton, D. (2013). Endogenous institutions and the possibility of reverse crowding out. *Public Choice*, 156(1–2), 253–284.
- Isaac, R. M., & Walker, J. (1988). Group size effects in public goods provision: The voluntary contributions mechanism. *Quarterly Journal of Economics*, 101(1), 179–199.
- Isaac, R. M., Walker, J., & Thomas, S. (1984). Divergent evidence on free riding: An experimental examination of possible explanations. *Public Choice*, 43(2), 113–149.
- Kingsley, D., & Brown, T. (2016). Endogenous and costly institutional deterrence in a public good experiment. *Journal of Behavioral and Experimental Economics*, 62(1), 33–41.
- Kosfeld, M., Okada, A., & Riedl, A. (2009). Institution formation in public goods games. *American Economic Review*, 99(4), 1335–1355.
- Ledyard, J. (1995). Public goods: A survey of experimental research. In J. Kagel & A. Roth (Eds.), *Handbook of Experimental Economics*. Princeton: Princeton University Press.
- Markussen, T., Putterman, L., & Tyran, J.-R. (2014). Self-organization for collective action: An experimental study of voting on formal, informal and no sanction regimes. *Review of Economic Studies*, 81(1), 301–324.
- Marwell, G., & Ames, R. (1979). Experiments on the provision of public goods 1: Resources, interest, group size, and the free-rider problem. *American Journal of Sociology*, 84(6), 1335–1360.
- Norton, D., & Isaac, R. M. (2010). Endogenous production technology in a public goods enterprise. In R. M. Isaac & D. Norton (Eds.), *Research in Experimental Economics. Charity with Choice* (13th ed.). Boston: Emerald.
- Putterman, L., Tyran, J.-R., & Kamei, K. (2011). Public goods and voting on formal sanction schemes. *Journal of Public Economics*, 95(9–10), 1213–1222.
- Smith, A. (2012). Comment on social preferences, beliefs, and the dynamics of free riding in public good experiments. *Economics Bulletin*, 32(1), 923–931.
- Smith, A. (2013). Estimating the causal effect of beliefs on contributions in repeated public good games. *Experimental Economics*, 16(3), 414–425.

- Smith, A. (2015). Contribution heterogeneity and the dynamics of contributions in repeated public good games. *Journal of Behavioral and Experimental Economics*, 58(1), 149–157.
- Sutter, M., Haigner, S., & Kocher, M. (2010). Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations. *Review of Economic Studies*, 77(4), 1540–1566.
- Tyran, J.-R., & Feld, L. (2006). Achieving compliance when legal sanctions are non-deterrent. *Scandinavian Journal of Economics*, 108(1), 135–156.