JOURNAL OF THEORETICAL POLITICS

Article

Participation and punishment

Journal of Theoretical Politics
2016, Vol. 28(4) 537–551

DOI: 10.1177/0951629815586882
jtp.sagepub.com
SAGE

C. Daniel Myers
University of Minnesota, Minneapolis, MN, USA

Abstract
Studies have demonstrated that altruistic punishment can motivate pro-social behavior in settings that require group cooperation. This paper extends these findings to the voter participation game. Using a laboratory experiment we find that some subjects are willing to punish abstainers and that punishment increases participation. However, not all subjects respond to punishment equally. We find that punishment increases turnout more among subjects with a egoist social value orientation (SVO) than among subjects with a pro-social SVO. This suggests that increased opportunities for altruistic punishment do not just increase participation, but may also change the composition of the electorate.

Keywords
Political Participation, altruistic punishment, voter turnout, social value orientation

1. Introduction

Recent work suggests that altruistic punishment, the punishment of norm-violators by another when the punisher has nothing instrumental to gain (Boyd et al., 2003; Fehr and Gächter 2002), can account for some of the ability of human societies to overcome the free-rider problem in many situations (e.g. Fehr and Gächter, 2002). This study extends findings about altruistic punishment to the voter participation game (Palfrey and Rosenthal, 1983). In this game, two teams compete over a single office. Each player can decide whether to participate at a cost or abstain at no cost. The team with the most members participating gains the office, and each member of that team benefits from the office whether or not they voted. We modify the

Corresponding author:
C. Daniel Myers, University of Minnesota, 1414 Social Sciences, Minneapolis, MN 55455, USA.
Email: cdmyers@umn.edu

game by allowing players to see the actions of their fellow team members and to pay a cost to punish them.

Punishment opportunities increase participation in public goods games (Fehr and Gächter 2000; Fehr and Gächter 2002). However, the effects of punishment in the voter participation game are not as clear. In public goods games, contributing to the public good helps all members of the group and is efficiency-enhancing in that larger contributions increase the total amount earned by all players. In the voter participation game, participation helps members of the participant's team but harms players on the other team. Further, as will be seen when the game is described in detail, participation is efficiency-reducing in that participation is costly and does not increase the value of the resource for which the teams are competing. Thus it is unclear who the target for punishment will be; players who wish to punish those who deviate from the efficiency-maximizing result will punish participators, while those who wish to punish those who harm their team will punish abstainers. Drawing on work by Bornstein (1992: 2003) we characterize these two varieties of punishment as exhibiting "collective rationality" or "group-focused rationality." Past studies show that when mechanisms for coordination are available, some groups exhibit collective-focussed rationality and successfully coordinate on the efficient (no-participation) outcome (Bornstein, 1992).

If altruistic punishment increases participation in the voter participation game, it may help to explain the so-called "paradox of political participation" (Fiorina, 1990). Scholars have long noted that interactions within social networks can increase political participation (e.g. Huckfeldt and Sprague, 1995; Verba et al., 1995). However, evidence for the mechanism that underlies the relationship between social ties and political participation is frequently missing. Altruistic punishment is one mechanism through which interactions in social networks might increase turnout. This possibility is supported by studies showing that increasing the visibility of voters' decision to vote, and thus the opportunity for others to punish abstention, increases turnout (Funk, 2010; Gerber et al., 2008). Still, demonstrating that altruistic punishment, or the threat of it, causes the increase in turnout in these situations is difficult, since doing so requires observing the day-to-day interactions among friends and neighbors where altruistic punishment within social networks would occur. Here, a lab experiment can be useful in demonstrating the plausibility of this mechanism since it allows us to manipulate the possibility of punishment and observe interactions in a way that is difficult in the world at large.

Our study is closely related to recent work on group competition in the Tullock contest game (Abbink et al., 2010). This game differs in several important ways from the voter participation game studied in the present paper. For example, in the Tullock contest game contribution levels are continuous, while in the voter participation they are binary (vote/abstain). Perhaps most importantly, subjects in the study of Abbink et al. (2010) stay in the same groups for 20 periods, creating a possible instrumental motivation for punishment. As will be seen, subjects in our study are randomly rematched into new teams of four at the end of each period, reducing any instrumental motivation for punishment and thus rendering it purely

"altruistic." In this way, the present study can be thought of as a harder test of the ability of punishment to motivate participation in team competition settings.

We go beyond testing the effect on participation to examine how punishment effects the composition of the electorate. Some voters' voting decisions may be highly sensitive to the amount of social pressure, while others may vote regardless of whether or not they are subject to social pressure. Existing work shows that this is the case in the dictator game, where the possibility of punishment does not change the behavior of pro-social subjects, but increases the contributions of subjects with an egoist social value orientation (SVO) to the point that they are indistinguishable from pro-social subjects (Simpson and Willer, 2008). Following this, we hypothesize that pro-social subjects will participate regardless of the possibility of punishment, while egoist subjects will participate at low rates when punishment is not possible and high rates when punishment is possible. If this is the case, then increasing the opportunities for altruistic punishment may do more than increase the overall rate of turnout, it may also produce a less pro-social electorate.

2. A participation game with altruistic punishment

The model tested in this experiment begins with a standard voter participation game from Palfrey and Rosenthal (1983). In this game, there are two teams or parties, A and B. All members of team A favor candidate A; likewise all members of team B favor candidate B. The teams are of equal size with four members each. Each player decides simultaneously whether or not to participate. Voters who decide to participate pay a cost, c; those who do not participate pay no cost. After all voters have decided, the number of participants on each team is compared, and the team with the most participants wins. Each member of the winning team receives w, regardless of whether or not they individually participated. Members of the losing team receives nothing. In general, $w > c > 0$. In the case of a tie, players on both teams receive $\frac{w}{2}$. This point marks the end of the standard voter participation game, which we will henceforth refer to as the "participation stage." The strategic dilemma players face is fairly simple: players would like their team to win, but would prefer that the team win without their help. Importantly, because participation is costly and does not affect the value of winning, the outcome that maximizes social welfare in this game is for no one to participate, unlike in public goods games.

We add a "punishment stage" to the standard voter participation game. In this stage, each player is informed of the election outcome and the first stage decisions of each other member of her team. She can then punish any of her teammates. She punishes by selecting a teammate or teammates and imposing a cost p on that player or players. The punisher also incurs a cost p' for each teammate punished. Each player may punish as many of her teammates as she wants, paying p' each time. After the punishment stage, players receive their final payoff and the game ends. Each player's payoff is as follows:

$$Earnings_i = A_i w - B_i c - C_i p - D_i p' \quad (1)$$

where A_i is an indicator for whether the player i's team won, B_i and indicator for whether i participated, C_i a count of the number of teammates that i punished and D_i a count of the number of teammates who punished i. For the experiment discussed below, we set the values as follows: $w = 3$, $c = 1$, $p = 2$ and $p' = 1$.

Importantly, and in contrast to Abbink et al. (2010), this game is modeled as a one-shot interaction in that players do not interact with the same team multiple times. The lack of repeated interactions removes the instrumental reasons voters might have to punish such as creating a reputation for punishing abstainers in the hopes of winning more elections in the future.[1] The current design allows us to isolate the effects of altruistic punishment; in doing so it presents a particularly hard test for the claim that such punishment can drive participation. It also allows us to test whether such punishment can increase turnout in cases where voters do not anticipate repeated interactions, for example, when elections are sufficiently far apart that voters discount the importance of future contests.

It is straightforward to show that this game reduces to a single-stage voter participation game as described in Palfrey and Rosenthal (1983). Since punishment is costly and has no impact on payoffs except for this cost, each player can maximize her final payoff by never punishing. Since no player ever punishes in the punishment period, behavior in the participation period is not changed by the presence of the second period. The focus of this paper is not on predicting the level of turnout, as it is in other papers that investigate the voter participation game (Levine and Palfrey, 2007; Schram and Sonnemans, 1996). However, it is worth mentioning the equilibrium predictions for the game configured in this manner. The voter participation game has three Nash equilibria in which all players play the same pure or mixed strategy: a full participation equilibrium in pure strategies, a "high-turnout" equilibrium in mixed strategies, and a "low-turnout" equilibrium in mixed strategies. In this game, these equilibria are turnout rates of 1, 0.81, and 0.19, respectively. In general, point predictions by Nash equilibria do a poor job of predicting behavior in the voter participation game. However, quantal response equilibria do a much better job (Levine and Palfrey, 2007). The quantal response equilibria for two levels of "error" in player's actions that Goeree and Holt (2005) characterize the actions of experienced and inexperienced players in the voter participation game are 0.44 and 0.48, respectively.

2.1. Hypotheses

Equilibrium predictions notwithstanding, the extensive literature cited above suggests that individuals will engage in costly punishment. Thus, the first hypothesis predicts that experimental subjects will punish.

H1. Punishment hypothesis: subjects will punish.

However, as discussed in the introduction, the target of punishment in the voter participation game is less clear than in the public goods setting. The team competition setting creates a tension between punishing team members who harm the team's earnings and punishing team members who deviate from the socially

efficient outcome. In public goods games, these two motivations result in the same action: players who refuse to contribute to the public good harm the earnings of the team and also deviate from the efficient outcome of full contribution. In the voter participation game players who vote help their team. However, this participation does not increase the value of the prize for which the teams are competing. Thus, the efficient outcome is one in which no one participates and the winner is randomly determined; players who participate deviate from this efficient outcome.

If players exhibit what Bornstein (2003) calls "group-focused rationality", they will punish fellow group members who abstain. Such rationality causes subjects to punish those who harm their group, regardless of whether such behavior harms society as a whole. In this case, group-focused punishers would punish those who abstain from participating; this punishment of those who refuse to take a costly action bears the closest resemblance to the kind of punishment seen in other games.

H2a. Group-focused punishment hypothesis: players will punish teammates who harm the group by abstaining; participation in sessions with punishment will increase.

Nevertheless, there is some evidence that players are likely to use a different kind of thinking, what Bornstein (2003) calls "collective rationality." Such players punish those who deviate from the socially efficient outcome by participating. While this kind of punishment may seem unlikely to readers, Bornstein (1992) finds that some experimental subjects attempt to reach the no-participation outcome if communication is available as a coordination device. Punishment of players who participate might provide an alternative coordination device that can be used to reach a similar outcome.[2]

H2b. Collective-focused punishment hypothesis: players will punish teammates who participate; participation in sessions with punishment will decrease.

Note that these two hypotheses are not mutually exclusive. It is possible that some subjects will punish abstainers while others will punish participators. Both types could coexist, or one might predominate in some sessions and the other in others.

In addition, this paper tests whether different types of people respond differently to the possibility of punishment. Some people participate because of internal factors, such as the desire to help other members of their group; these should not be affected by the possibility of punishment. Other people are less other-regarding; these may vote only when the threat of punishment overwhelms their inclination to free ride on the efforts of others. To test this we examine the relationship between the responsiveness of subjects to punishment and their SVO, their set of "preferences for particular patterns of outcomes for the self and others" and measures the tendency "to enhance the outcomes of a dyad, group, or collective, even when such actions are quite costly to the self" (Van Lange, 1999). SVO measures classify persons into three categories: pro-social individuals, who maximize outcomes for their self and others; individualists, who maximize outcomes for themselves; and competitive individuals, who maximize their outcome relative to others' outcomes. Following much of the literature, we combine the later two categories into an "egoist" category (de Creamer and Van Lange, 2001; Simpson and Willer, 2008).

H3. Pro-social subjects will participate at the same rate regardless of the possibility of punishment; egoist subjects will participate at lower rates in when punishment is not possible and higher rates where punishment is possible.

3. Method

We conducted a computerized experiment using the z-Tree software package (Fischbacher, 2007). The study took place in the Princeton Laboratory for Experimental Social Science. Subjects were undergraduate students. Subjects received instructions on the format of the game and had to participate in two practice rounds and pass a comprehension quiz before the session began.[3] We conducted 17 sessions, each lasting approximately 40 minutes. Ten sessions were conducted in December and January of 2008–2009, and seven additional sessions were conducted in April of 2012.

Each session had eight subjects and consisted of two experimental treatments. In the punishment treatment, subjects were allowed to punish the other members of their group; in the control, or non-punishment treatment, subjects were merely informed of the actions that their fellow groups members took and were not able to punish. We varied which treatment came first so that in nine of the sessions the punishment treatment ran first and in eight the control treatment ran first.

Each treatment consisted of 20 rounds. At the beginning of each round, the subjects were randomly assigned to the Yellow Team or the Brown Team so that four subjects were on each team. At the start of each round, each subject received an endowment of nine tokens. In the first period of each round, each subject decided whether to chose X or Y, where X represents the choice to participate and Y the choice to abstain.[4] If the subject chose Y, she received one token. If she chose X, she received no tokens. Each subject received three tokens if more members of her team chose X than the other team, and zero tokens otherwise. In the case of a tie, all subjects received 1.5 tokens.

In the second period of each round, each subject was shown the decisions of the three other members of her team. The actions were displayed in a random order so that is was clear that no connection could be drawn between the actions of other subjects across rounds. In the punishment treatment, each subject could then decide to punish any of her fellow teammates. Each punishment cost the punisher one token and the punished two tokens.[5] At the end of the second period, each subject was told how many times she was punished and shown her total earnings from that round. Play then proceeded to the next round and subjects were randomly assigned to new teams of four.[6]

After 20 rounds, the treatment ended. After the subjects completed the punishment and the control treatments, they were asked to fill out a number of psychological batteries as well as a series of demographic questions and several open ended questions asking them about the strategies they employed. Subjects were then paid in private. Subjects earned an average of 10.5 tokens per round, with a minimum of 1, a maximum of 13 and a standard deviation of 1.6 (or an average of 31.5 cents per round with a minimum of 3 cents a maximum of 39 cents and a standard

deviation of 4.8 cents after conversion). Average cash earnings for the entire experiment were 22.66 dollars and ranged from 21 dollars to 24 dollars.

We measure subjects' SVO using the instrument in Van Lange (1999). In this instrument, subjects make nine allocations of "points" between themselves and an unidentified other. Each decision has three options: a pro-social option where both the allocator and the other get equal amounts, an individualistic option where the allocator gets slightly more than in the pro-social option but the other gets considerably less, and a competitive option where the allocator gets roughly the same amount as in the pro-social option, but the other gets almost nothing. An example of this choice is as follows:

(A) you get 520, other gets 520;
(B) you get 580, other gets 320;
(C) You get 520, other gets 120;

here (A) is the pro-social option, (B) the individualistic option, and (C) the competitive option. Following standard practice, subjects who made at least six out of nine allocations of the same type were classified as that type.[7] For example, a subject who chose six pro-social allocations and three individualistic allocations would be classified as pro-social. Out of 136 subjects, 67 were classified as pro-social, 59 as individualistic, and 3 as competitive. The remaining seven could not be classified because they did not consistently choose one type of allocation, and are not included in the following analyses which involve SVO.[8] As noted above, we classified subjects as either pro-social or egoist, with the egoist category covering individualistic and competitive subjects.[9]

4. Results

4.1. Punishment and its effects

Figure 1 graphs the average participation rate in each condition across all sessions. Across all rounds, subjects participated 60 percent of the time in the no-punishment treatment and 88 percent of the time in the punishment treatment. The dashed line indicates the average for treatments with punishment, the solid line those without punishment. Visual inspection suggests that the presence of the punishment period caused a large increase in participation. The participation rate in sessions with punishment starts higher and shows, if anything, an upward trend across rounds. In sessions without punishment, participation starts lower and declines slightly over the course of the session. A Wilcoxon signed rank test comparing the group-mean turnout rate in punishment treatments to non-punishment treatments finds that this difference is statistically significant ($p < 0.001$). Since each subject took part in one punishment treatment and one non-punishment treatment we can see the effect of punishment at the individual level as well. Out of 136 subjects, 101 participated more often in the punishment treatment than in the non-punishment treatment

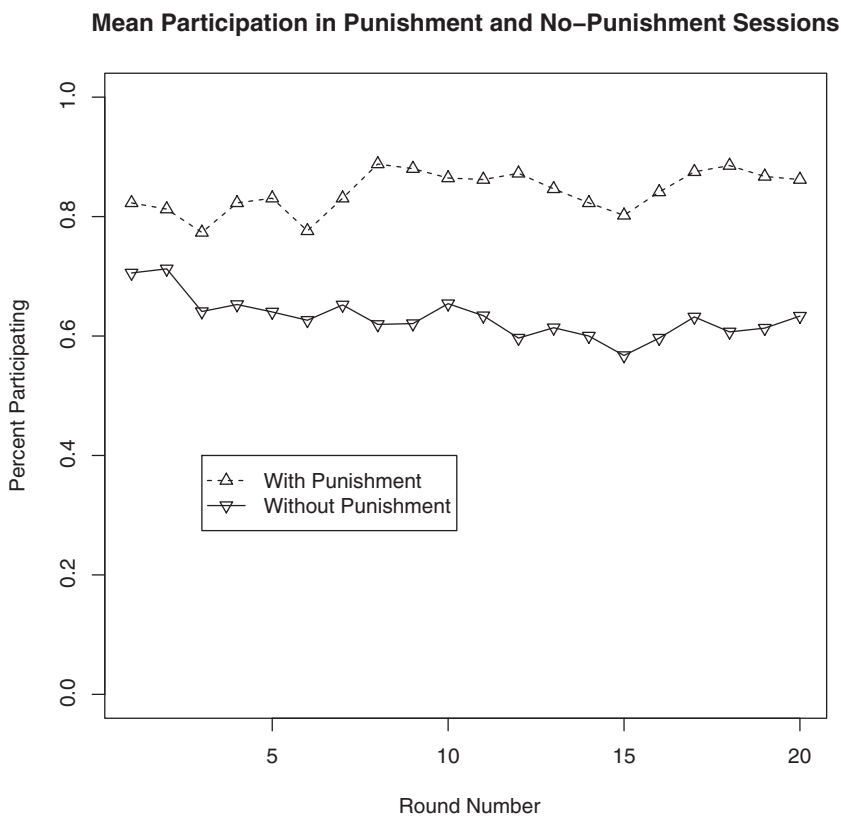

Figure 1. Turnout by session.

than in the non-punishment treatment, while only 16 participated at a lower rate and 19 at the same rate.

To test the effect of punishment more formally we following the example of models used in the analysis of public good games contributions (e.g. Croson, 2007) and use a random-effects logit model.[10] The dependent variable is individual *i*'s decision to participate or not to participate in period *t*. The independent variable of interest, *punish* is an indicator variable that equals one in the punishment treatment. To account for interdependence among observations in the same group we include random effects for each individual, each group, and each period number. As Table 1 shows, this model again finds a strong positive effect of punishment on participation.

4.2. Rates of punishment

Figure 1 shows that turnout was higher from the very beginning of sessions when punishment was possible. This suggests that subjects anticipated being punished if they free-rode. This expectation was correct: across all sessions abstainers were

Table 1. Modeling participation decisions (random effects logit).

	Coefficient (Standard error)
Intercept	0.734 (0.196)
Punish	2.027 (0.084)
N (decisions) = 5440	
N (groups) = 17	
N (periods) = 20	
N (individuals) = 129	

punished an average of 0.95 times per abstention. Thus while abstainers gained one token for abstaining, they lost an average of 1.9 tokens from being punished, a significant net loss. This clearly demonstrates that subjects were willing to engage in punishment despite its irrational character, validating *H1*, the punishment hypothesis.

Subjects primarily punished abstainers, but occasionally punished cooperators. Subjects who participated were punished an average of .06 punishments per participation action. The rates of punishment of abstainers and participants strongly supports the group focused punishment hypothesis (*H2a*). However, a few subjects punished in the manner suggested by *H2b*, the society focused punishment hypothesis. Ten subjects punished cooperators more often than they punished abstainers while 90 subjects punished abstainers more often (the remainder never punished).[11] While these subjects were in the clear minority, it is not inconceivable that such behavior might come to predominate in some situations, especially those with a greater opportunity for coordination between competing teams.

4.3. Who Participates?

The possibility of punishment increases turnout, but whose turnout does it increase? Consistent with *H3*, we find that egoist subjects are more responsive to punishment than pro-social subjects. In the no-punishment treatment pro-social subjects participated at a 66 percent rate compared with 55 percent for egoist subjects, a significant difference ($p = 0.029$). However, in the punishment treatment pro-social and egoist subjects participated at almost identical rates, 87 percent of the time for pro-social subjects and 88 percent of the time for egoist subjects. This difference also had a small but meaningful effect on the composition of the electorate in these two conditions. In the punishment treatment an average of 7 subjects participated, of whom 3.6 or 51 percent were pro-social. In the no-punishment treatment an average of 4.8 subjects participated, of whom 2.7 or around 57 percent were pro-social.[12]

We test these results formally with a random-effects logit model, as shown in Table 2. Model 1 replicates the results shown in Table 1. Model 2 adds in an indicator variable for SVO in which 1 indicates an egoist subject and an interaction between this variable and *Punish*, which equals 1 in the punishment treatment. As before, there is a strong positive main effect of punishment on participation. *SVO*

Table 2. Modeling participation decisions (random effects logit).

	Model 1	Model 2
Intercept	.734 (0.196)	1.089 (0.257)
Punish	2.027 (0.084)	1.685 (0.117)
SVO	—	−0.779 (0.326)
Punish * SVO	—	0.776 (0.173)
N (decisions) = 5440		
N (groups) = 17		
N (periods) = 20		
N (individuals) = 129		

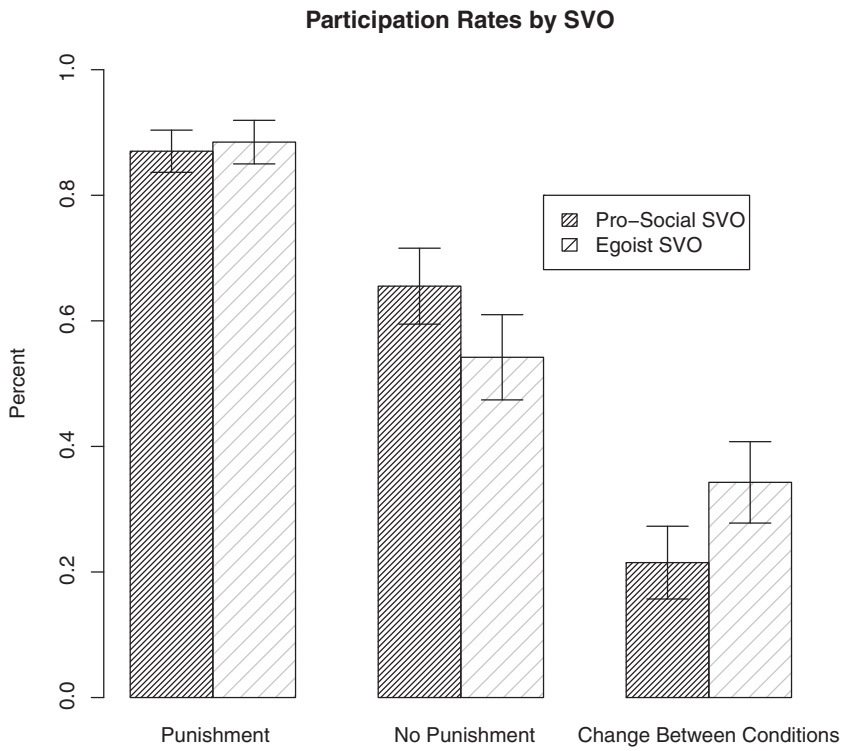

Figure 2. Mean turnout by SVO and experimental condition.

has a significant and negative effect on participation in the no-punishment treatment. However, the coefficient on the interaction term indicates that this negative effect is entirely wiped away when punishment is possible. Consistent with *H3*, egoist subjects participate less than pro-social subjects only when punishment is not

possible; when punishment is possible, egoists and pro-social subjects participate at the same rate.

5. Discussion

These results show that subjects are willing to engage in costly punishment in the voter participation game, even when such punishment is individually irrational and increased participation in future rounds will not increase the future earnings of the individual doing the punishing or the total earnings of all subjects. Punishment is overwhelmingly directed at players who abstain from participation and drastically increases the rate of participation, supporting *H1* and *H2a*, the group-focused punishment hypothesis. However, a small set of subjects punished cooperators as predicted by *H2b*, the society-focused punishment hypothesis. This raises the possibility that a group with a sufficient number of this type of punishers would develop a norm of zero percent turnout, enforced by the punishment of any subject who "defected" by participating.

This study found a link between SVO and the sensitivity of a subject's turnout to the possibility of punishment. As hypothesized, turnout among egoist subjects was significantly lower than among pro-social subjects when punishment was not possible. However, the introduction of punishment increased turnout by egoist subjects more than it increased turnout by pro-social subjects, such that turnout among subjects with pro-social and egoist SVOs was the same in punishment sessions. This finding supports the conditional relationship Simpson and Willer (2008) found between punishment and pro-social behavior: punishment forces egoist subjects to behave as though they were pro-social. One important difference is that while Simpson and Willer (2008) find that only egoist subjects are responsive to punishment, we find that both types of subjects respond to punishment, but that egoist subjects respond more.

6. Conclusion

This experiment shows the role that altruistic punishment can play in motivating political participation by extending existing findings about altruistic punishment to the voter participation game. It shows that punishment increases participation and changes the composition of the electorate by boosting participation among egoist subjects more than among pro-social subjects. The results offers evidence that punishment within social networks is one mechanism underlying the connection between social ties and voter participation. For example, voters in Gerber et al.'s (2008) study who turned out because they were told that their voting record would be revealed to their neighbors may have done so out of the fear of punishment. Moreover, this work suggests that the SVO of those who voted because of this threat was likely different from the SVO of those who would have voted in the absence of this threat. Future work should examine the form that such social punishment takes in more realistic settings, as well as replicate the finding regarding SVO.

Altruistic punishment can only function if the actions that potential voters take are observable; non-participators cannot be punished if no one knows that they did not participate. The role of altruistic punishment in motivating participation will thus depend on how visible the decision to participate is. Indeed, voting, a mostly private act, may be less affected by punishment than other, more public forms of participation, such as putting up a lawn sign or attending a fundraiser. One implication of these results is that policy changes that alter how visible the decision to vote is may have unintended consequences. Policy changes that reduce the visibility of the decision to vote, such as mail-in voting, are likely to reduce the social pressure to vote (c.f. Funk 2010). This effect may help to explain the failure of some studies to find a positive effect on turnout from vote-by-mail reforms (Kousser and Mullin, 2007). Moreover, even if such reforms have no effect on overall turnout, this study suggests that they may change the composition of the electorate by removing a major factor that drives egoist voters to the polls.

In addition to highlighting the importance of the visibility of participation, this experiment assumes that voters' social networks are politically homogeneous. Subjects in this experiment could only punish other subjects who supported the same candidate as them; if they could punish opponents then punishment might have a more mixed effect on turnout. This would be in line with research that finds that heterogeneous social networks reduce turnout (Mutz, 2002). The present results have greater relevance for societies where voters tend to have politically homogeneous networks than for societies where voters have politically heterogeneous networks; future work should extend this research to punishment in heterogeneous networks.

Acknowledgements

Thanks go to Jens Großer, Tali Mendelberg, Adam Meirowitz, Scott Ashworth, Ken Williams, Dustin Tingley, Stephanie Wang, Ben Lauderdale, Eric Dickson and participants at the 2008 NYU-CESS Experimental Political Science Conference for helpful feedback. All errors of fact or interpretation are, of course, mine.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Princeton University's Research Program in Political Economy and the Princeton Laboratory for Experimental Social Science.

Notes

1. Note, however, that in the experiment that follows subjects are randomly rematched into 2 groups of 4 across 20 rounds, allowing for the possibility that subjects would interact in the same team configurations more than once. Thus, instrumental motivations for punishing are greatly reduced, but not eliminated, as they are in the theoretical model.
2. Such punishment might also reflect anti-social punishment, or the "punishment of people who behave pro-socially" (Herrmann et al., 2008; cf. Cinyabuguma et al., 2006; Nikiforakis, 2008). Most public goods experiments find some punishment of high contributors; Herrmann et al. (2008) replicates this finding across 16 societies and finds that anti-social punishment is most common in societies with weak norms of social cooperation and weak adherence to the rule of law. Notably for this experiment, rates of anti-social punishment are lowest in the United States, where this study was conducted.
3. Complete instructions and screen shots are available in the online appendix.
4. As is standard in voter participation experiments, the instructions and on-screen display did not explicitly refer to voting, participating, or abstaining.
5. It was impossible for subjects to end the round with negative earnings; the nine tokens each subjects received at the beginning of each round is equal to the cost of punishing three other players and being punished three times.
6. This design follows most voter participation game experiments by using stranger matching, not absolute stranger matching as in most public goods experiments with punishment (Fehr and Gächter, 2000: 2002). Under strangers matching subjects are on different teams in each round, but play with the same set of subjects across all rounds. Under absolute stranger matching two subjects never interact with each other more than once. Absolute stranger matching is necessary in public goods games with punishment because encouraging another subject to contribute more by punishing them can only help the punisher in future rounds. In contrast, punishing a subject to encourage them to participate in future rounds of the voter participation game can either help or hurt the punisher, since in the next round the punished subject is equally likely to be on the opposing team as to be on the punisher's team.
7. While this measure of SVO is widely used, it is not incentivized, a factor that may create some suspicion in the minds of readers. Past studies have shown that SVO measured in this way is highly predictive of play in the dictator game, something readers might view as a preferred measure of altruism (Simpson and Willer, 2008). Strong relationships have also been noted between SVO, measured with an unincentivized instrument, and volunteering (McClintock and Allison, 1989), play in a common pool resource games (Roch and Samuelson, 1997), and play in an incentivized one-shot prisoner's dilemma (Van Lange, 1999). All suggest that despite not being incentivized, this technique produces a valid measure of the desired construct.
8. The substantive results do not change when we relax the classification scheme so that these subjects are classified as the type they choose most often.
9. A total of 80 subjects completed the SVO battery immediately after playing the voter-participation game, while 56 completed the battery several days before and immediately after playing the game. While SVO is supposed to be a stable psychological trait, it is possible that playing the voter-participation game affected the measurement of some subject's SVO. Comparing the pre- and post-measurements of SVO among the 56 who completed both show that only three subjects were categorized differently in the post-measurement than they were in the pre-measurement, with all three shifting from pro-social to egoist SVOs. To ensure that the results reported below are not the product of

mis-measuring SVO among the 80 subjects who completed only the post-measurement of SVO, we conducted a sensitivity analysis reported in the online appendix.

10. Thanks to an anonymous reviewer for suggesting this analytic strategy.
11. The comments of several subjects who punished cooperators provide some insight as to their motives. When describing their strategy, one wrote that they were trying to promote a different equilibrium where everyone chose X. Another justified her actions by saying "If the 8 of us were picking Y, then we would make 2.5 rather than 1.5 on average." This rationale closely tracks the logic of the society focused hypothesis.
12. Interestingly, this effect appears to be primarily the product of an interaction between gender and SVO. SVO had little effect on how women responded to punishment. Female egoists increased their participation from 62 percent without punishment to 88 percent with punishment, while pro-social females increased from 66 to 87 percent, respectively. However, men with different SVOs responded very differently: male egoists increased their participation from 46 to 89 percent in punishment sessions, while pro-social males increased from 64 to 87 percent. These findings require replication, as this study did not expect an effect of gender on responsiveness to punishment, much less an interaction of gender and SVO. However, they are an intriguing hint for future research.

References

Abbink K, Brandts J, Herrmann B and Orzen H (2010) Intergroup conflict and intra-group punishment in an experimental contest game. *American Economic Review* 100(1): 420–447.

Bornstein G (1992) The free-rider problem in intergroup conflicts over step level and continuous public goods. *Journal of Personality and Social Psychology* 62(4): 597–606.

Bornstein G (2003) Intergroup conflict: individual, group and collective interests. *Personality and Social Psychology Review* 7(2): 129–145.

Boyd R, Gintis H, Bowles S and Richerson PJ (2003) The evolution of altruistic punishment. *Proceedings of the National Academy of Science* 100(6): 3531–3535.

Cinyabuguma M, Page T and Putterman L (2006) Can second-order punishment deter perverse punishment? *Experimental Economics* 9(3): 265–279.

Croson RTA (2007) Theories of commitment, altruism and reciprocity: evidence from linear public goods games. *Economic Inquiry* 45(2): 199–216.

de Creamer D and Van Lange PAM (2001) Why prosocials exhibit greater cooperation than proselfs. *European Journal of Personality* 15(1): 5–18.

Fehr E and Gächter S (2000) Cooperation and punishment in public goods experiments. *The American Economic Review* 90(4): 980–994.

Fehr E and Gächter S (2002) Altruistic punishment in humans. *Nature* 415: 137–140.

Fehr E and Fischbacher U (2004) Third-party punishment and social norms. *Evolution and Human Behavior* 25: 63–87.

Fiorina MP (1990) Information and rationality in elections. In Ferejohn JA and Kuklinski JH (ed.) *Information and Democratic Processes*. Urbana, IL: University of Illinois Press.

Fischbacher U (2007) z-Tree: Zurich toolbox for readymade economic experiments. *Experimental Economics* 10(2): 171–178.

Funk P (2010) Social incentives and voter turnout: evidence from the Swiss mail ballot system. *Journal of the European Economic Association*, 8(5): 1077–1103.

Gerber AS, Green DP and Larimer CW (2008) Social pressure and voter turnout: evidence from a large-scale field experiment. *American Political Science Review* 102(1): 33–48.

Goeree JK and Holt CA (2005) An explination of anamalous behavior in models of political participation. *American Political Science Review* 99(2): 201–213.

Güth W, Schmittberger R and Schwarze B (1982) An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization* 3(4): 367–388.

Herrmann B, Thöni C and Gächter S (2008) Antisocial punishment across societies. *Science* 319: 1362–1367.

Huckfeldt R and Sprague J (1995) *Citizens, Politics and Social Communications: Information and Influence in an Election Campaign*. Cambridge: Cambridge University Press.

Kousser T and Mullin M (2007) Does voting by mail increase participation? Using matching to analyze a natural experiment. *Political Analysis* 15(4): 428–445.

Levine DK and Palfrey TR (2007) The paradox of voter participation? A laboratory study. *American Political Science Review* 101(1): 143–158.

McClintock CG and Allison ST (1989) Social value orientation and helping behavior. *Journal of Applied Social Psychology* 19(4): 353–362.

Mutz DC (2002) The consequence of cross-cutting networks for political participation. *American Journal of Political Science* 46(4): 838–855.

Nikiforakis N (2008) Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics* 92(1): 91–112.

Palfrey TR and Rosenthal H (1983) A strategic calculus of voting. *Public Choice* 41: 7–53.

Roch SG and Samuelson CD (1997) Effects of environmental uncertainty and social value orientation in resource dilemmas. *Organizational Behavior and Human Decision Processes* 70(3): 221–235.

Schram A and Sonnemans J (1996) Voter Turnout as a Participation Game: An Experimental Investigation. *International Journal of Game Theory* 25: 385–406.

Simpson B and Willer R (2008) Altruism and indirect reciprocity. *Social Psychology Quarterly* 71(1): 37–52.

Van Lange PAM (1999) The pursuit of joint outcomes and equality in outcomes. *Journal of Personality and Social Psychology* 77(2): 337–349.

Verba S, Schlozman KL and Brady H (1995) *Voice and Equality: Civic Voluntarism in American Politics*. Cambridge, MA: Harvard University Press.