Political Psychology

Political Psychology, Vol. 31, No. 4, 2010
doi: 10.1111/j.1467-9221.2010.00768.x

The Behavioral Logic of Collective Action: Partisans Cooperate and Punish More Than Nonpartisans

Oleg Smirnov
Stony Brook University

Christopher T. Dawes
University of California, San Diego

James H. Fowler
University of California, San Diego

Tim Johnson
Stanford University

Richard McElreath
University of California, Davis

Laboratory experiments indicate that many people willingly contribute to public goods and punish free riders at a personal cost. We hypothesize that these individuals, called strong reciprocators*, allow political parties to overcome collective action problems, thereby allowing those organizations to compete for scarce resources and to produce public goods for like-minded individuals. Using a series of laboratory games, we examine whether partisans contribute to public goods and punish free riders at a greater rate than nonpartisans. The results show that partisans are more likely than nonpartisans to contribute to public goods and to engage in costly punishment. Given the broad theoretical literature on altruistic punishment and group selection as well as our own formal evolutionary model, we hypothesize that it is being a partisan that makes an individual more likely to be a strong reciprocator and not vice versa.*

KEY WORDS: Partisanship, Public goods game, Collective action, Altruistic punishment, Strong reciprocity

Published by Wiley Periodicals, Inc., 350 Main Street, Malden, MA 02148, USA, 9600 Garsington Road, Oxford, OX4 2DQ, and PO Box 378 Carlton South, 3053 Victoria Australia

> *Political parties are collections of individuals, so that virtually everything they do involves collective action, and they provide public goods for their members, since much of what they do affects many, if not all, partisans*.—Aldrich (1995, 31)

More than four decades has passed since Mancur Olson's *The Logic of Collective Action* (1965) focused attention on partisan activity as the exemplary collective action problem. Yet today political scientists still ponder: why do individuals engage in costly partisan activities when they could obtain more benefit by free riding on others' partisan efforts? Slow progress on this question, however, is less a sign of stagnant scholarship than it is evidence of the richness of Olson's puzzle. Indeed, if (1) individuals act according to rational self-interest and (2) the benefits of partisan activity can be consumed by those who do not provide them, then—as definitively stated by Olson—"the average person will not be willing to make a significant sacrifice for the party he favors, since a victory for his party provides a collective good" (1965, 164).

Individuals do engage in partisan activity, however, and some even choose to identify strongly with one political party. Thus, a central focus for political science is explaining why individuals engage in such behavior. Olson proposed that party activity will only result if the party is organized for some *other* purpose (Olson, 1965). For instance, we might attend a party meeting if it allows us to exchange beneficial automotive repair tips; but, if the meeting does not offer external benefits that compensate the costs of attendance, we will not attend. Similarly, others propose that individuals will engage in costly partisan activities if institutions alter the incentive structure of party participation (Aldrich, 1995, 21–22). That is, if registering as a partisan—a costly activity—allows one to vote in a party primary (a collective action problem in itself), then individuals will do so if the benefits of voting outweigh the costs of registering. Such propositions, in sum, serve as contemporary political science's explanation of why partisan activity exists: political parties provide sufficient compensation for self-interested individuals to incur the costs of engaging in partisan efforts.

That solution, however, appears dubious. For one, the costs and benefits of partisan activity are poorly specified, thus making it difficult to determine whether a personal benefit or an exogenous institution actually solves the collective action problem. Second, there exist instances in which partisan activities cost much more than the personal benefits they produce—or so most partisan volunteers would agree after spending four hours sealing envelopes for a same-party candidate from another district. For these reasons, explanations of partisan activity that focus on the profitability of partisanship appear incomplete.

With that inadequacy noted, the question remains: Why do people engage in partisan activities? One possibility is that people do not maintain self-interested preferences and, thus, they ignore the collective action problem inherent in partisan efforts. For the past decade, scholars in anthropology and economics have

found increasing evidence that humans hold such preferences. Specifically, humans appear to follow the directives of *strong reciprocity* (Bowles & Gintis, 2002; Camerer & Fehr, 2006). Strong reciprocity consists of two behavioral predispositions: (1) a willingness to contribute to public goods and (2) an enthusiasm for sanctioning those who fail to contribute to public goods (Bowles & Gintis, 2002, 425).

In this study, we examine whether individuals who identify themselves as partisans (and thus, assumedly, engage in more partisan activity than nonpartisans) abide by the behavioral patterns of strong reciprocity more frequently than nonpartisans. We place laboratory participants in a series of experimental games in which they earn real money based on their choices. These games allow us to measure subjects' willingness to contribute to a public good and, as well, their willingness to sanction those who do not contribute. Our results show that partisans, as opposed to nonpartisans, are more likely to be strong reciprocators. Specifically, partisans contribute to public goods at higher rates than nonpartisans, and they use punishment to enforce cooperation more often than do nonpartisans. Moreover, our experimental design allows us to establish that partisan punishment is based upon norm enforcement motives (punishing the defectors) rather than egalitarian motives (punishing the rich; Dawes et al., 2007)—this distinction is impossible to make in a classical public goods game with punishment since the defectors and the rich are the same people in the classical setting (Fowler, Johnson, & Smirnov, 2005).

In addition to offering evidence that contributes to an understanding of why political parties exist despite collective action problems, the results reported here also contribute to the literature on partisanship. Not only do we show that partisans have distinctive behavioral patterns in social dilemmas outside the domain of voting and elections, but we also offer evidence that yields insight into the decline of partisanship (Abramson & Aldrich, 1982) and the tendency of partisans to view politics through an ingroup/outgroup lens (Greene, 1999).

Before presenting these findings, however, we first present a general theory of partisans as strong reciprocators. With our theory presented, we outline laboratory procedures designed to test our theory, and we examine the relationship between partisanship and behavior in these laboratory games. With our results, we present a number of robustness checks and then conclude with a discussion of how our findings reshape the way political science should contemplate both collective action and partisanship.

Partisanship and Strong Reciprocity

Political scientists have traditionally defined partisanship as a loyalty to one political party that shapes an individual's electoral behavior in a lasting manner (Campbell, Converse, Miller, & Stokes, 1960). As such, partisanship serves as a set of individual principles guiding broad evaluation of political candidates and events (Key, 1966). Converse (1969) explains the origin of individual partisanship as a

combination of intergenerational transmission from parents (especially, one's father) and social learning. Partisanship is acquired in one's youth and remains stable, if not fixed, over one's lifetime. Moreover, strength of partisanship is increasing with age: Young adults are weak partisans while the elderly exhibit stronger partisanship (Converse, 1976). Empirical evidence supports these claims by showing that both intergenerational transmission and developmental learning contribute to an individual's partisanship (Cassel, 1993).

It is important to note that partisanship is often equated with party identification, which Miller (1991) points out is not technically correct. Party identification is a relatively narrow concept that describes an individual's self-identification with a certain political party; the concept thus omits general behavioral attributes of partisans. Contrary to partisanship, party identification may in fact be short-lived given individual social and political experience (Fiorina, 1981). As Franklin and Jackson (1983) put it, party identifications are "a person's accumulated evaluations from previous elections and are dependent upon the events and the actions of political leaders during these elections and during subsequent terms in office" (p. 968). Party identification must converge with core values, beliefs, and preferences over issues (Franklin, 1984). We agree with Miller (1991), Converse and Pierce (1987), as well as the myriad other scholars who argue that partisanship is a broader construct with multiple facets. In addition to developmental learning and intergenerational transmission, partisanship also emerges as a complex transition to general consistency in one's behavior (loyalty to a party), political preferences, and expectations (Brader & Tucker, 2001). Given past scholars' sound recognition that partisanship is a more complex phenomenon than party identification, we use a measure of partisanship that does not distinguish between partisans who identify with different parties.

Parting with past work, however, we want to examine differences between partisans and nonpartisans that extend beyond their electoral behavior. Previous scholarship offers evidence that warrants such an investigation. For instance, it is already well known that partisans are more likely to participate in politics (Verba et al., 1995), but recent laboratory experiments suggest that a willingness to donate money to anonymous recipients in the dictator game greatly increases the partisan motivation to vote (Fowler, 2006). Partisans also appear to use special cognitive information-processing mechanisms when making voting decisions (Lodge & Hamill, 1986); could such unique cognitive strategies lead partisans to act in a peculiar manner in other social situations? In addition, previous work fails to find statistically significant differences between partisans and nonpartisans when it comes to issue voting (Gant & Luttbeg, 1987); such results suggest that the instrumental value of partisanship is not just brand labelling or assistance in deciding which political candidate to choose.

Instead of solely guiding vote choice, partisanship may also signal (either intentionally or unintentionally) an individual's willingness to provide and maintain public goods. That is, since political parties are in the business of providing

public goods (Aldrich, 1995), then we should expect partisans to exhibit a more pronounced tendency than nonpartisans to cooperate and punish in nonparty settings. For instance, one situation in which differences between partisans and nonpartisans might manifest is the *public goods game*, which is an extension of the *n*-person Prisoner's Dilemma (see Ledyard, 1995, for a review of the public goods game literature). In the game, individuals possess personal resources and belong to a group that can pool resources in order to increase social welfare. However, contributing personal resources to the common pool also decreases the contributors' payoffs (by the amount of their contribution), regardless of the behavior of other group members. Thus, since there is always an incentive to free ride, no strategically minded, self-interested individual will make a contribution to the common pool. As a result, no one contributes to the common pool and the group is forced to recognize that everyone would have received greater wealth had everyone contributed.

Cooperation, however, is possible. Recent laboratory work shows that when experimenters modify the public goods game to give participants the option of costly decentralized punishment, then contribution to the public good increases (Fehr & Gächter, 2002; Fehr & Fischbacher, 2004). In the game with punishment, any player can decrease the payoff of any other member(s) in the group at a personal cost. If targeted at a noncontributor, such punishment becomes altruistic since the individual pays a cost to punish, while the benefit (modifying the behaviour of noncontributors so that they contribute in future rounds) is distributed to other people.

Fehr and Gächter (2002) show that experimental participants punish defectors frequently and that this punishment promotes cooperation even when reputation building is not possible (no two players interact twice in the game). Field studies also support this research. Various forms of costly self-enforcement of cooperative behavior are customary in communities around the world (Henrich et al., 2006), and it is common to punish those who free ride on others' personally costly efforts to use natural resources like fisheries, water, grazing lands, forests, and wildlife (see Ostrom, 1990, and Smirnov, 2007, for more specific references and examples). Given this theoretical and empirical evidence, it is now well established that "communities often are capable of enforcing norms because a considerable fraction of members are willing to engage in the costly punishment of shirkers without a reasonable expectation of being personally repaid for their efforts" (Bowles & Gintis, 2002 p. 425).

Individuals who provide and maintain public goods engage in strong reciprocity—that is, they return others' behavior in kind even when they do not expect benefits from such actions (Gintis, 2000; Sethi & Somanathan, 1996). Scholars have begun to uncover the mechanisms underlying strong reciprocity. Laboratory experiments suggest that emotions, such as anger, may explain why individuals engage in costly acts of punishment (Fehr & Gächter, 2002). Another emotion explaining altruistic punishment may be spite: individuals punish

wealthier individuals—who happen to be defectors in public goods games (Fowler et al., 2005)—at high rates (Dawes et al., 2007; Johnson, Dawes, Fowler, McElreath, & Smirnov, 2007). Whatever the emotional/cognitive basis of punishment, physiological evidence obtained in a neuroscientific study of economic behavior showed that a crucial part of the brain's reward circuit—the caudate nucleus—is activated when individuals punish defectors (de Quervain et al., 2004). Such evidence suggests that strong reciprocity is a stable and common characteristic of human populations. Nevertheless, researchers should exercise caution when using emotions to explain strong reciprocity. For example, the caudate nucleus is part of a reward processing system, and it may not be connected with anger. Strong reciprocators may anticipate punishment and get rewards processed through the striatum (we are thankful to an anonymous reviewer for bringing this to our attention).

Despite the fact that many people feel anger and spite that compels them to punish defectors, the extent to which individuals engage in strong reciprocity varies considerably from one individual to another. Not all participants in past experiments show emotions or spend personal resources to punish free riders. Due to this heterogeneity, we hypothesize that people who engage in collective political efforts are more likely than those who do not to experience the emotions mentioned above and, therefore, to act as strong reciprocators. In other words, we expect partisans to cooperate and punish free riders more than nonpartisans.

Correlation or a Causal Relationship?

An important and challenging question is whether the relationship between strong reciprocity and partisanship is causal or not. For instance, strong reciprocators may become partisans due to their greater willingness to engage in behaviors that foster collective action, or they may acquire their behavioral disposition via involvement in activities that require them to produce and maintain public goods. Or, perhaps, strong reciprocity and partisanship amplify each other in a synergistic exchange. Any of these possible causal trajectories appear reasonable and future longitudinal studies may shed light on which is accurate.

At the same time, various theoretical analyses suggest an unambiguous causal relationship between partisanship and strong reciprocity, if partisan activities can be seen as a form of intergroup competition. On the one hand, strong reciprocity evolves conditional upon the existence of competition among groups such that individuals who belong to more successful groups obtain greater individual payoffs (Boyd, Gintis, Bowles, & Richerson, 2003; Sober & Wilson, 1998). On the other hand, participation in intergroup competition—which is a proxy for partisanship—does not necessarily depend on the existence of strong reciprocity.

For illustrative purposes, consider the following game theoretic model. A population of n players consists of cooperators (C), defectors (D), and

reciprocators (R-defined as cooperators who punish defectors at a cost to themselves). Additionally, assume that each individual can belong to one of the two competing groups, or "parties." Some individuals may not belong to any of the groups. The model, therefore, has nine behavioral types: cooperators who belong to group 1 (C-1), group 2 (C-2), or no group at all (C-N); defectors in group 1 (D-1), group 2 (D-2), and nonaffiliated defectors (D-N). Similarly, there are three kinds of reciprocators: R-1, R-2, and R-N.

All players are randomly matched with each other, giving us a 9×9 payoff matrix. Those players who belong to either group 1 or group 2 also pay a cost of intergroup competition. The group that has a greater number of cooperators and reciprocators wins the competition and obtains an exogenous benefit, which is then equally divided among all members of the winning group.

Evolutionary selection is based upon the discrete version of replicator dynamics, which is more appropriate for finite populations than the continuous version. In this case, individuals engaging in a low utility behavior are likely to be replaced by those engaging in high-utility behaviors (or, equivalently, those who receive low utility are likely to change their behavior to imitate the higher utility individuals). Random variation is added to the selection mechanism to avoid premature convergence.

Thus, the payoff to a player $i = 1, 2, \ldots, n$ who belongs to a group $j = 1, 2, \mathrm{N}$ can be expressed as follows:

$$U_{i,j} = (\alpha+\beta)b - c\alpha_i - x\beta_i(1-\alpha-\beta) - z\gamma_i\beta + \pi_i\left(w\left(\Pr\left[\alpha_j+\beta_j > \alpha_{\sim j}+\beta_{\sim j}\right]\right) - y\right) \tag{1}$$

where α_i, β_i, γ_i, and π_i are all *dummy variables* describing whether the individual is respectively a cooperator, strong reciprocator, or defector and whether or not the player belongs to one of the two competing groups; α and β are *proportions* of cooperators and reciprocators in the population, with the proportion of defectors being equal to $1 - \alpha - \beta$; α_j and β_j are respectively the *total number* of cooperators and reciprocators who belong to group j; finally, b and c are the benefit and cost of individual cooperation, x is the cost of punishment for a reciprocator, z is the cost of punishment for a defector, y is the cost of intergroup competition, and w is the benefit of winning the competition.

Given the complexity of the evolutionary game theoretic model above, we study it computationally (see the online appendix available at http://ms.cc.sunysb.edu/~osmirnov/—for the R code of the computational model). Figures 1 and 2 below show typical evolutionary dynamics for two cases: low and high cost of intergroup competition ($y = 1$ and $y = 10$; other parameters can be found in the R code in the online appendix).

When the cost of intergroup competition is *low*, there is selection for "partisan" types, i.e., individuals belonging to one of the two groups. Partisan defectors cannot sustain their numbers, always losing to partisan cooperators or

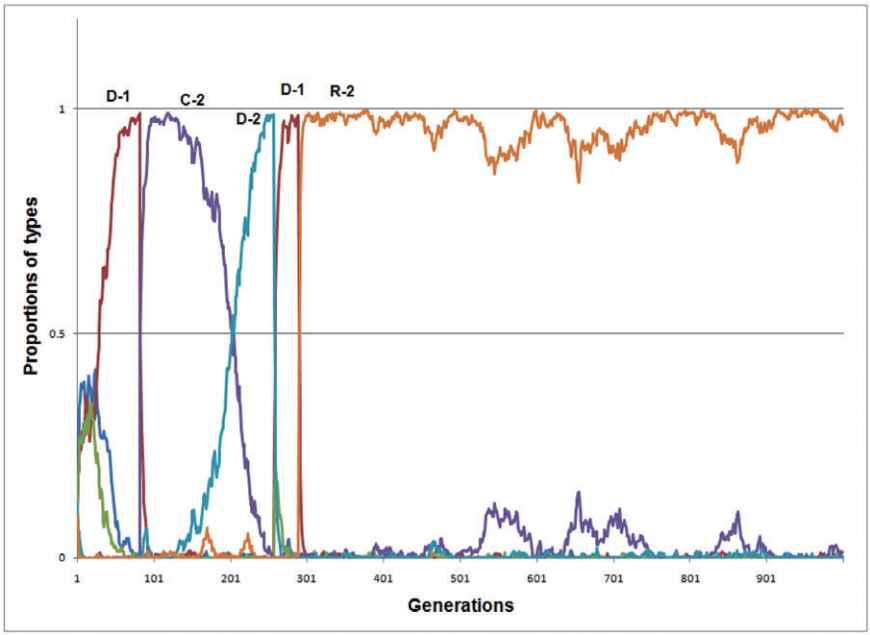

Figure 1. Low cost of intergroup competition.

reciprocators from *another* group. Partisan cooperators also cannot sustain their numbers, always losing to defectors from the *same* group. On the contrary, partisan reciprocators are much more robust as a behavioral type. This type cannot "lose" to any other type. It may only be vulnerable—in the long run—to nonpunishing cooperators from the *same* group if there is large random variation in the behavioral types. In this case, the population may randomly "drift" from reciprocators to cooperators. If the random drift is large enough, the proportion of cooperators will be large enough so that within-party defectors will no longer be contained by reciprocators. The party will collapse and a new party of reciprocators will quickly take over the population.

When the cost of intergroup competition is *high*, there is strong selection for "nonpartisan" defection and no selection on any of the reciprocating types, partisan or not. Increasing the cost of intergroup competition is substantively analogous to removing the intergroup competition from the model altogether. This illustration implies that once intergroup competition is taken away strong reciprocity cannot be sustained. Similarly, research in the life sciences shows that altruistic punishment can evolve in the presence of group selection (Boyd et al., 2003; Sober & Wilson, 1998). These simulations indicate that strong reciprocity depends on group competition (and therefore partisanship) and not the other way around.

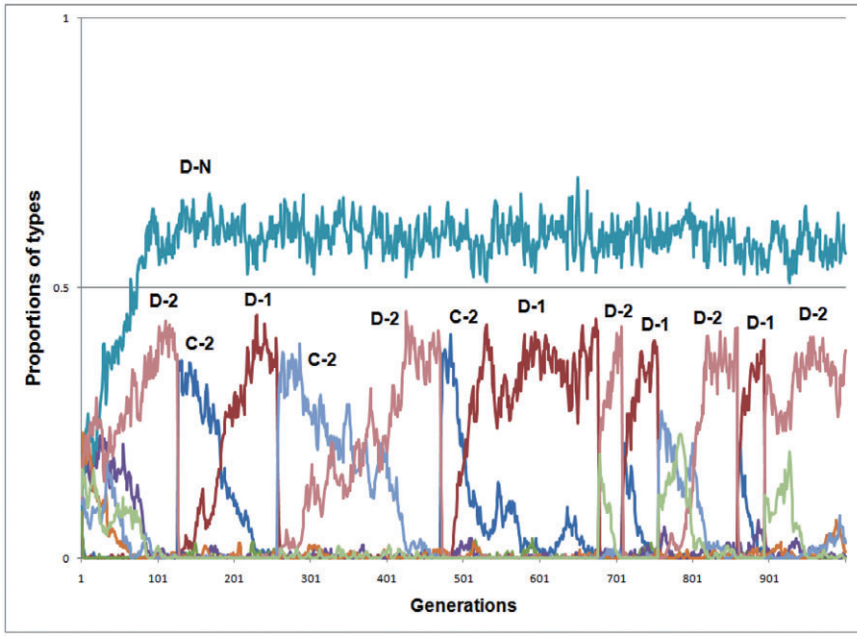

Figure 2. High cost of intergroup competition.

Experimental Design

Given a theoretical expectation that group competition will spur group members to engage in strong reciprocity, we wanted to determine whether partisans are more likely to cooperate and punish free riders than their nonpartisan counterparts. Here we focus on the behavior of a population of college students primarily because they are a commonly used convenience sample. However, young people also provide an advantage in terms of statistical efficiency—since partisan strength tends to increase as people age, the number of partisans and nonpartisans tends to be more unbalanced in older populations, requiring larger samples to achieve equivalent discriminatory power. And in spite of this change over time, the best predictor of partisanship in later years is partisanship in early years (Niemi & Jennings, 1991). Although we have no reason to expect the results would differ substantially in samples drawn from a wider population, we do note here that the convenience sampling method necessarily limits the generalizability of our results.

We recruited 120 undergraduate students from the University of California at Davis to play two games: a random income game that measures aversion toward income inequality and a public goods game with punishment. The design and procedures of the random income game and the public goods game closely

replicate a widely cited public good experiment (Fehr & Gächter, 2002). One hundred ($n = 120$) students from the University of California at Davis volunteered to participate in the random income experiment and public goods game. Subject recruitment was conducted in several different departments to maximize the chance that subjects did not know one another; any student who was at least 18 years old was eligible to take part in the study. Twenty subjects attended each of the six sessions in either of the experiments, and each session involved five periods. The show-up fee was $5 and each subject earned a total of about $10 on average.

Combining a random income game that measures aversion toward income inequality and a public goods game with punishment allows us to distinguish norm enforcement (punishment driven by strong reciprocity) from egalitarian reductions ("punishment" driven by egalitarian motives and/or spite) and, as well, to examine the contribution behavior of participants. Some subjects played the random income game first while others first participated in the public goods game with punishment. As we will show below, the *order of treatments* did not have a significant effect on the individual behavior.

At the end of the experiment we conducted a survey, in which subjects were asked to identify themselves along a 7-point partisanship scale: Strong Democrat—Democrat—Independent/Closer to Democrat—Independent—Independent/Closer to Republican—Republican—Strong Republican. We define *partisans* as those who identified themselves as Strong Democrat, Democrat, Republican, and Strong Republican. For the sake of clarity and simplicity, we treat partisanship as a dichotomous variable, but in the online appendix we also replicate all regression results using a four category variable that differentiates between independents, leaners, partisans, and strong partisans (see the online appendix). Given our dichotomous definition, our subject pool consisted of 51 partisans (42.5%) and 68 nonpartisans (56.7%); one subject did not answer the survey question.

Before describing the experimental results, we briefly outline the random income game and the public goods game with punishment. In the random income game, subjects are divided into groups of four anonymous members each. Each player receives a sum of money randomly generated by a computer. To maintain comparability with other public goods games, random payoffs were drawn from the empirical distribution of payoffs in the first stage of a widely cited public goods game with punishment (Fehr & Gächter, 2002). Subjects are shown the payoffs of other group members for that round and are then provided an opportunity to alter others' incomes by giving them "negative tokens" or "positive tokens." Each negative token reduces the sender's payoff by 1 monetary unit (MU) and decreases the payoff of the targeted individual by 3 MUs. Each positive token reduces the sender's payoff by 1 MU but increases the payoff of the targeted individual by 3 MUs. Groups are randomized after each round to prevent reputation from influencing decisions; interactions between players are strictly anonymous and subjects know this. The total number of rounds is five. At the end of each period, subjects learned the amount of negative tokens they received and their new payoff. The

experiment lasted 30 minutes. Since income reduction cannot be normatively justified in this experiment (that is, punishment cannot promote cooperation in this case), we use variations of the term "costly reduction" to describe "punishment" in the random income game.

In the public goods game with punishment, subjects are also divided into groups of four anonymous members. In our specific implementation of the game, those subjects are the same as those playing the random income game. Each player is endowed with an amount of money (20 MUs) and is given an opportunity to contribute some or all of this money to a common pool. Once contributions to the common pool have been made, the value of the common pool is multiplied by 1.6, which is the same as in Fehr and Gächter (2002). Multiplication of the common pool makes it so that the group income is maximized when all contribute, but personal income is maximized by withholding contributions regardless of the behavior of other group members. After multiplication, the common pool is distributed equally among all members of the group regardless of their contributions. Thus, the experiment models the classical public goods game dilemma. Subjects are then given the opportunity to reduce others' incomes by distributing "negative tokens." As in the random income game, each negative token reduces the purchaser's payoff by 1 monetary unit (MU) and decreases the payoff of a targeted individual by 3 MUs. The total number of rounds is five. This experimental design matches the procedures used in the well-known experiment by Fehr and Gächter (2002).

All activity in the experiments was completely anonymous. Group composition changed every period so that no one played with the same person more than once. The subjects were ignorant of other players' experimental history: neither past payoffs nor past decisions were known. Different group composition each period and the absence of any history of play ensured that subjects could neither develop reputations nor target other subjects for revenge.

At the beginning of each session subjects were asked to read experiment instructions on their individual computer screens, and they also had a paper copy available for reference. The instructions explained all features of the experiment, including how payoffs are determined, how group composition is altered every period, and how anonymity of individual decisions and payoffs in the experiment is preserved. In order for the experiment to start, subjects had to answer *correctly* several test questions designed to ensure full understanding of how choices in the game generate payoffs. At the end of the experimental session, subjects were asked to complete a survey about their demographic characteristics. The experiment was programmed using GameWeb software written by Richard McElreath.

Experimental Results

Subjects spent money to reduce others' incomes, and they did so at high levels (cf Dawes et al., 2007). Among participants, 68% reduced another player's

income at least once, 28% did so five times or more, and 6% did so 10 times or more. As well, 74% of participants increased another player's income at least once, 33% did so five times or more, and 10% did so 10 times or more. Most (71%) negative tokens were given to above-average earners in each group, whereas most (62%) positive tokens were targeted at below-average earners in each group. Unexplained punishment of below-average earners (29%) may be attributed to spite whereas unexplained reward of above-average earners (38%) may be attributed to altruism.

The size of income alterations varied with the relative income of the recipient. Individuals who earned considerably more than other members of their group were heavily penalized. Subjects who earned 10 MUs more than the group average received a mean of 8.9 negative tokens compared to 1.6 for those who earned at least 10 MUs less than the group. In contrast, individuals who earned considerably less than other group members received sizable gifts. Subjects who earned 10 MUs more than the group average received a mean of 4 positive tokens compared to 11.1 for those who earned at least 10 MUs less than the group. Individual spending decisions also suggest that subjects were influenced by concerns for inequality. On average, the bottom earner in each group spent 96% more on negative tokens than the top earner and the top earner spent 77% more on positive tokens than the bottom earner (both differences significant, t test, one-tailed, $p < 0.0008$).

Income reduction was also frequent in the public goods game, similar to previously reported experiments (Fehr & Gächter, 2002; Johnson et al., 2007), with punishments targeted to low contributors. 62% reduced another player's income at least once, 29% did so five times or more, and 12% did so 10 times or more. Most (84%) punishment was targeted at below-average contributors in each group. The size of punishments varied with the target's relative contribution to the public good. Individuals who contributed considerably less than other members of their group were heavily penalized. Subjects who contributed at least 10 MUs less than the group average received a mean punishment of 12.1 MUs compared to 0.5 MUs for those who contributed at least 10 MUs more than the group.

Aggregate results from these experiments suggest that most subjects are generally willing to pay to reduce inequality and to contribute to the provision of public goods, both directly and indirectly via altruistic punishment. The goal of our research design here is to identify the degree to which partisans are more likely than nonpartisans to engage in these behaviors. Table 1 shows that, in general,

Table 1. Average Behavior in the Random Income and Public Goods Games, by Partisanship

	Average MUs Spent on		
	Contributions	Income Alterations	Punishment
Partisans (n = 51)	39.0	6.55	2.91
Non-Partisans (n = 68)	22.6	4.14	1.57

partisans appear to be more willing to engage in behaviors that characterize strong reciprocity. For example, in the public goods game, partisans spent 73% more on contributions to the public good than did nonpartisans. A Wilcoxon rank sum test (a nonparametric test related to the parametric *t* test) of the total amount spent by each subject suggests that this difference is significant ($p = 0.002$). Thus, in our convenience sample drawn from a college student population, partisans are significantly more willing to engage in prosocial behavior that benefits fellow group members. They are also more willing than nonpartisans to reduce the incomes of high-income and low-contributing group members. Partisans spent significantly more (58%, $p = 0.03$) on income alteration in the random income game when cooperation was not at stake, suggesting that they have a greater willingness to influence group distributional outcomes. Partisans also spent 85% more ($p = 0.03$) on the punishment of other group members in the public goods game, indicating that they may also care more about enforcing cooperation.

Income and Contribution Sensitivity

Analysis of contributions to a group project is fairly straightforward, since it does not rely on other group members' behavior. Subjects make their decisions in each newly formed group before they see anyone else's decision. However, the analysis of income alteration behavior must incorporate both the target's and the sender's behavior. Ideally, we would like to have a measure for the willingness to assign positive and negative tokens as a function of the target's relative income. To create such a measure, we use the 15 observations for each subject (since subjects play a total of five rounds and in each round they make decisions concerning three other group members) and plot the subjects' purchases of positive and negative tokens against the target's income deviance. We then fit a regression line to the data produced by each subject (Figure 3a). The height of each line (or intercept) indicates the subject's general willingness to alter others' incomes; the slope represents the degree to which the subject is motivated to reduce the incomes of those who earn the most and/or increase the income of those who earn the least. We denote this slope *income sensitivity*; it serves as a simple measure of egalitarian preferences. Subjects with steeper, positive slopes tend to be more responsive to the income of the target when they decide whom to help and whom to harm. In the same way, we can create an index of altruistic punishment. In Figure 3b (public goods game) we plot a line that describes, for each subject, the relationship between a target's contribution behavior and the subject's willingness to punish the target. We denote the slope *contribution sensitivity*.

Figure 3 shows that most subjects behave in a way that is consistent with retrieving equality in the random income game and punishing noncontributors in the public goods game: 65% have positive income sensitivity and 58% have positive contribution sensitivity. Only a few subjects have negative income

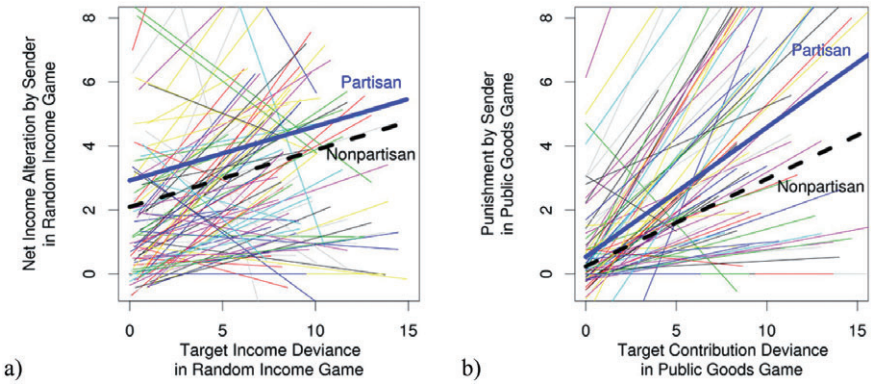

Figure 3. (a) Best fitting lines for each subject. The lines describe the effect of target relative income on net income reduction in the random income game. (b) Best fitting lines for each subject that describe the effect of target (negative) contribution deviance on net income reduction in the public goods game with random payoff. Most lines have a positive slope, indicating a general desire to reduce the income of the wealthy and to punish free-riders. The solid lines in both panels show the average sensitivity for all partisan (solid blue) and nonpartisan (dashed black) subjects. Both income and contribution sensitivity are significantly different from zero ($p < 0.0001$).

sensitivity (21%) or negative contribution sensitivity (3%). This figure also shows differences in the mean behavior of partisan (solid blue line) and nonpartisan (dashed black line) subjects. As already noted, partisans generally appear to spend more on both income alteration and punishment (the blue line is higher in both panels). However, sensitivity to the income of the target is almost identical for partisans and nonpartisans (equal slopes) in the random income game. A Wilcoxon rank sum test of the difference in the individual slopes confirms that they are statistically indistinguishable ($p = 0.15$). Similarly, although partisans appear to be somewhat more sensitive to contributions than nonpartisans (steeper slope) in the public goods game, the difference is not significant ($p = 0.69$).

These initial results suggest that partisans and nonpartisans share similar norms regarding the extent to which group members should be penalized or rewarded based on their income and contribution behavior, but partisans are willing to spend more of their own money to contribute to collective goods and to influence group outcomes. Thus, people who identify with the Republican and Democratic Party are more likely to be strong reciprocators than those who do not. However, there are several potential confounds. For example, it is important to control for gender and income since these have previously been shown to correlate with partisanship (Chaney, Alvarez, & Nagler, 1998). Also, since distributional concerns may underlie observed behavior, we also control for the direction of partisanship—i.e., to which parties partisans belong—to be sure that differences in policy preferences are not driving the results.

Analyzing Contribution Behavior

As is standard in the literature (Fehr & Gächter, 2002; Dawes et al., 2007; Johnson et al., 2007), we analyze contributions, income alteration, and punishment behavior using Tobit regression. Tobit regression explicitly accounts for the fact that subjects may desire choices that are not available to them. For example, in the experiment we limit contribution to the public good to fall between 0 and 20. Some individuals may have wanted to give even more of their money, while others might have wanted the opportunity to actually take money away from the group project. Similarly, we limit the number of positive and negative tokens purchased in the random income game to range between 0 and 20 and the number of negative tokens purchased in the public goods game to range between 0 and 10. In addition, since we have multiple observations from each subject, we estimate *clustered standard errors* for each individual. This permits us to control for the fact that *within-subject observations are not independent*. We also conducted Lagrange multiplier tests (Beck & Katz, 1995; Engle, 1984) that suggest serial correlation is not significant after controlling for clustering.

The results of these regressions appear in Tables 2–5. In Table 2, we show the results of Tobit regressions that model contributions in the public goods game. Notice that partisans contribute significantly more to public goods provision than nonpartisans, even when we control for demographic characteristics in Model 2.2. Notice that we do not control for education since our subjects (undergraduate students) had the same level of education. Overall, these results suggest that partisans have a stronger desire to cooperate with fellow group members than nonpartisans.

To what extent is the cooperative nature of partisans linked to punishment? Past work has shown that contributions in the public goods game are sensitive to the

Table 2. Effect of Partisanship on Contributions in the Public Goods Game

	Dependent Variable: Amount Contributed to Public Good					
	Model 2.1			Model 2.2		
	Coef	SE	p	Coef	SE	p
Partisan	3.15	1.48	0.03	2.53	1.33	0.05
Male				–3.01	1.40	0.03
Age				0.00	0.30	0.99
Republican				–0.79	0.48	0.11
Income				–0.27	0.27	0.33
Constant	3.37	0.86	0.00	12.43	7.66	0.11
Log scale	2.22	0.08	0.00	2.18	0.09	0.00
Log likelihood		–1585.3			–1550.3	
Null likelihood		–1593.1			–1571.8	
N (5 rounds * 120 subjects—missing data)		595			590	

Table 3. Effect of Punishment and Partisanship on Change in Public Goods Contributions

	Dependent Variable: Change in Amount Contributed to Public Good					
	Model 3.1			Model 3.2		
	Coef	SE	p	Coef	SE	p
Partisan	–0.02	0.31	0.95	0.38	0.41	0.36
Punishment Received in Previous Round	0.40	0.10	0.00	0.57	0.12	0.00
*Partisan*Punishment Received in Previous Round*				–0.30	0.17	0.07
Constant	–1.13	0.25	0.00	–1.36	0.28	0.00
Log scale	1.65	0.09	0.00	1.65	0.09	0.00
Log likelihood		–1444.7			–1443.6	
Null likelihood		–1452.1			–1452.1	
N (4 changes * 120 subjects—missing data)		476			476	

Table 4. Effect of Partisanship on Income Alteration in the Random Income Game

	Dependent Variable: Income Alteration in the Random Income Game								
	Model 4.1			Model 4.2			Model 4.3		
	Coef	SE	p	Coef	SE	p	Coef	SE	p
Partisan	0.54	0.49	0.27	0.60	0.58	0.30	0.66	0.42	0.12
Target Income Deviance	0.14	0.03	0.00	0.14	0.04	0.00	0.13	0.03	0.00
*Partisan*Target Income Deviance*				–0.01	0.06	0.82			
Male							–0.83	0.47	0.08
Age							0.07	0.12	0.59
Republican							0.02	0.15	0.87
Income							–0.09	0.10	0.35
Constant	–1.34	0.38	0.00	–1.36	0.42	0.21	–0.98	2.77	0.73
Log scale	1.19	0.08	0.00	1.19	0.08	0.00	1.17	0.08	0.00
Log likelihood		–2644.1			–2644.1			–2576.8	
Null likelihood		–2658.0			–2658.0			–2606.9	
N (*120 subjects * 3 decisions per round * 5 rounds* — missing data)		1785			1785			1770	

presence of punishment (Fehr & Gächter, 2002) and individual subjects who were previously punished tend to increase their contributions in the following round (Johnson et al., 2007). To model the link between cooperation and punishment, we treat change in the contribution from the previous to the current round as the

Table 5. Effect of Partisanship on Punishment in the Public Goods Game

	Dependent Variable: Punishment in the Public Goods Game								
	Model 5.1			Model 5.2			Model 5.3		
	Coef	SE	p	Coef	SE	p	Coef	SE	P
Partisan	1.08	0.50	0.03	1.04	0.56	0.06	1.19	0.45	0.00
Target Negative Contribution Deviance	0.39	0.05	0.00	0.38	0.07	0.00	0.33	0.04	0.00
*Partisan*Negative Contribution Deviance*				0.01	0.07	0.90			
Income Alteration From Random Income Game							0.04	0.02	0.03
Contributions From Public Goods Game							0.00	0.00	0.81
Male							–0.01	0.48	0.99
Age							0.01	0.13	0.97
Republican							0.15	0.14	0.28
Income							0.00	0.09	0.99
Constant	–4.11	0.62	0.00	–4.09	0.65	0.00	–5.14	2.90	0.08
Log scale	1.10	0.10	0.00	1.10	0.10	0.00	1.02	0.10	0.00
Log likelihood		–1459.4			–1459.4			–1377.5	
Null likelihood		–1584.9			–1584.9			–1534.7	
N		1785			1785			1770	

dependent variable and examine to what extent partisanship and punishment in the previous round influence these changes. In Model 3.1 we see that punishment has a big effect on changes in contributions—each punishment point received increases subsequent contributions by about 0.4 MUs on average. This finding underscores the importance of decentralized punishment in the maintenance of cooperation.

However, partisans exhibit no general tendency to increase their contributions over time relative to nonpartisans. In Model 3.1 we interact the partisan variable with past punishment to see if partisans give more because they are more sensitive to punishment. In fact, what the model suggests is that, if anything, partisans are somewhat less likely to respond to punishment ($p = 0.07$). Thus punishment appears to play little role in the general difference between partisan and nonpartisan contribution levels.

Analyzing Costly Income Alteration and Altruistic Punishment

Model 4.1 in Table 4 shows that the simple relationship between partisanship and the amount spent on income alteration in the random income game is positive but not significant. This result contrasts with the Wilcoxon tests because it incorporates the uncertainty associated with observations that occur at the boundary of the available range of income alteration. Several subjects spent no money on

positive or negative tokens for their fellow players, and some of these might have been much stronger in their desire to avoid paying for income alteration than others. Model 4.1 also shows that subjects tend to spend more on income alteration when the target's income is far (higher deviance) from the group income. In Model 4.2 we add an interaction term to test the hypothesis that partisans and nonpartisans have different sensitivity to the target's income. However, the coefficient is not significant and very close to zero, suggesting partisans and nonpartisans care equally about target income when they are choosing who to help and who to harm. In Model 4.3 we add a number of control variables that indicate demographic characteristics of the subject, including gender, age, direction of partisanship, and household income. Although this model improves the likelihood that the partisanship variable is significant, the *p*-value still falls below conventional thresholds of significance. Thus, the evidence suggests only a weak relationship between strength of partisanship and willingness to pay to redistribute incomes within one's group. Moreover, partisanship does not appear to mediate the content of the egalitarian norm that is present in this game. Partisans and nonpartisans alike tend to exhibit egalitarian behavior in their income alteration decisions.

What, however, is the impact of partisanship on punishment in the public goods game? Model 5.1 in Table 5 shows that there is a strong relationship between partisanship and punishment. Partisans spend about 1.1 MU more on punishment than their nonpartisan peers. Model 5.1 also includes target contribution deviance, the amount the target contributed minus the average amount contributed by other group members (the value is set to 0 if the target contributed more than the average group member).

Fehr and Gächter (2002) show that this is an important factor in punishment decisions and that the tendency to punish low contributors is a key support for cooperation. We test the hypothesis that partisans are more sensitive to this contribution deviance than nonpartisans by adding an interaction term in Model 5.2. The coefficient on the interaction is near zero and not significant, suggesting that partisans and nonpartisans are equally sensitive to changes in the contribution level of the target.

To further test the partisan effect on willingness to punish, we add a number of demographic factors to Model 5.3. If the difference between partisan and nonpartisan punishment behavior is due to differences in type of partisanship (Republican vs. Democrat), gender, age, or income, then adding these variables to the model would cause the coefficient on partisanship to become insignificant. We also add two terms to this model that test the extent to which other motives may intermediate the partisan desire to punish more. First, we add the total number of positive and negative tokens bought by the subject in the random income game. We do this because we want to see if partisan behavior in the public goods game can be explained by a general willingness to punish fellow group members under any circumstances. If so, then adding this variable to the model should cause the coefficient on partisanship to become insignificant. Second, we add the total

amount of contributions by the sender in the public goods game. If partisan behavior can be explained by a general willingness to help other group members, we would expect that adding this variable to the model would cause the coefficient on partisanship to become insignificant. The results in Model 5.3 show that partisanship is still strongly significant and the size of the estimate remains about the same, despite the addition of these factors. This finding indicates that—holding demographic status and distributional preferences constant—partisans punish others at greater rates than nonpartisans. Thus, partisans appear to be uniquely interested in contributing more to the decentralized maintenance of cooperation by incurring costs to punish free riders.

Conclusion

Partisanship remains one of the most studied subjects in political science. From the early days of *The American Voter* (Campbell et al., 1960) until the present, scholars have studied the origin (Brader & Tucker, 2001; Converse, 1969; Cassel, 1993) and prevalence (Aldrich, 1995; Bartels, 2000) of partisanship, not to mention its dynamic (Fiorina, 1981; Franklin, 1984; Franklin & Jackson, 1983), cognitive (Gant & Luttbeg, 1987; Lodge & Hamill, 1986), and psychological (Greene, 1999) aspects. In this study, we contribute to this literature by looking into another—previously unexamined—"facet" of partisanship: partisan behavior in social dilemmas other than voting and elections. Albeit unstudied, that aspect of partisan behavior rests at the center of partisan activity. As first noted by Olson (1965)—and more recently professed by Aldrich (1995)—most partisan activity involves collective action. Yet, whereas Olson (1965) and Aldrich (1995)—among others—propose that altered incentive structures compel partisan activity, we examine whether the behavior of partisans defies the model of rational self-interest and, thus, enables partisan activity without incentive-altering institutions.

Our findings indicate that the behavior of partisans is, in fact, different from nonpartisans. Partisans act as strong reciprocators contributing to public goods and punishing noncontributors at higher rates than nonpartisans. Theoretical evidence suggests that constant competition between political parties places pressure on within-group cooperation and, in turn, promotes strong reciprocity. That is, since political success is largely a function of within-party cooperation, parties with strong reciprocators will likely prosper at greater rates than parties without strong reciprocators; thus, over time, the population of partisans will consist of strong reciprocators since the party attachments of other individuals will be nullified by the failure of their respective parties. On the other hand, existence of strong reciprocity is not a necessary condition for the existence of intergroup competition. Assessing the empirical plausibility of this assumption might serve as a prosperous line of research for future scholars.

We also note that our finding provides evidence that might inform past research on partisanship. For instance, Abramson and Aldrich (1982) explain

1960–80 turnout decline by pointing to aggregate decreases in the strength of partisanship during that period. Leaving the empirical validity of the finding aside, we note that the logic of the result is consistent with the evidence we present. If the population of strong reciprocators decreased during that time period then there would be less contribution to the public good (i.e., voting), and there would be less punishment to stem this decline. Another relevant example is the explanation of partisanship through the prism of social identity theory (Fowler & Kam, 2007; Greene, 1999), which holds that people maintain an "us" versus "them" portrait of the social world. Theory suggests that the prosocial elements of strong reciprocity will likely be geared toward ingroup members since the selective pressures yielding that behavioral type involve inter-group competition (Boyd et al., 2003; Fowler & Kam, 2007).

In sum, identifying the partisan tendency toward strong reciprocity yields insight into two important subjects in political science—collective action as manifest in partisan activity and partisanship itself. Our evidence shows that a behavioral model separate from that of rational self-interest captures the actions of partisans and, in so doing, resolves fundamental questions about why certain political organizations—parties—exist despite the problem of collective action. Our evidence also raises questions about political science's blind and convenient reliance on certain behavioral assumptions. Most political scientists assume that individuals act strategically and with self-interest, and—to bolster that assumption—they often regard evidence of alternative behavioral dispositions (see, e.g., Henrich et al., 2006) as irrelevant because political science focuses on political agents whose sophistication and Machiavellianism differs from the regular population. Here we show that such an assumption is dangerous, if not false. Individuals who show greater involvement in political activity actually defy the model of strategic self-interest ("rationality") and appear to accord more closely with a model indicating that individuals act to promote and maintain a common good.

ACKNOWLEDGMENTS

The authors thank the Center for Adaptive Behavior and Cognition at the Max Planck Institute for Human Development and the UC Davis Institute of Government Affairs for generous research support. Correspondence concerning this article should be sent to Oleg Smirnov, Assistant Professor, Department of Political Science, Stony Brook University, Stony Brook, NY 11794, USA. Phone: 786-566-1185. E-mail: oleg.smirnov@stonybrook.edu

REFERENCES

Abramson, P. R., & Aldrich, J. H. (1982). The decline of electoral participation in America. *American Political Science Review, 76*(3), 502–521.

Aldrich, J. H. (1995). *Why parties?* Chicago: The University of Chicago Press.

Bartels, L. M. (2000). Partisanship and voting behavior, 1952–1996. *American Journal of Political Science, 44*(1), 35–50.

Beck, N., & Katz, J. N. (1995). What to do (and not to do) with time-series cross-section data. *American Political Science Review, 89*(3), 634–647.

Bowles, S., & Gintis, H. (2002). Social capital and community governance. *Economic Journal, 112*, 419–436.

Boyd, R., Gintis, H., Bowles, S., & Richerson, P. J. (2003). The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences USA, 100*(6), 3531–3535.

Brader, T., & Tucker, J. A. (2001). The emergence of mass partisanship in Russia, 1993–1996. *American Journal of Political Science, 45*(1), 69–83.

Camerer, C., & Fehr, E. (2006). When does "economic man" dominate social behavior? *Science, 31*(5757), 47–52.

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter*. New York: Wiley.

Cassel, C. A. (1993). A test of Converse's theory of party support. *Journal of Politics, 55*(3), 664–681.

Chaney, C. K., Alvarez, R. M., & Nagler, J. (1998). Explaining the gender gap in US presidential elections, 1980–1992. *Political Research Quarterly, 51*(2), 311.

Converse, P. E. (1969). Of time and partisan stability. *Comparative Political Studies, 2*, 139–171.

Converse, P. E. (1976). *The dynamics of party support: Cohort-analyzing party identification*. Beverly Hills, CA: Sage.

Converse, P. E., & Pierce, R. (1987). Measuring partisanship. *Political Methodology, 11*, 143–166.

Dawes, C. T., Fowler, J. H., McElreath, R., Johnson, T., & Smirnov, O. (2007). Egalitarian motives in humans. *Nature, 446*, 794–796.

Engle, R. (1984). Wald, likelihood ratio and Lagrange multiplier tests in econometrics. In Z. Griliches & M. Intriligator (Eds.), *Handbook of econometrics* (Vol. 2, pp. 775–826). Amsterdam: North-Holland.

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

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

Fiorina, M. P. (1981). *Retrospective voting in American national elections*. New Haven, CT: Yale University Press.

Fowler, J. H. (2006). Altruism and turnout. *Journal of Politics, 68*(3), 674–683.

Fowler, J. H., Johnson, T., & Smirnov, O. (2005). Egalitarian motive and altruistic punishment. *Nature, 433*, doi:10.1038/nature03256.

Fowler, J. H., & Kam, C. (2007). Beyond the self: Altruism, social identity, and political participation. *Journal of Politics, 69*(3), 811–825.

Franklin, C. H. (1984). Issue preferences, socialization, and the evolution of party identification. *American Journal of Political Science, 28*(3), 459–478.

Franklin, C. H., & Jackson, J. E. (1983). The dynamics of party identification. *American Political Science Review, 77*(4), 957–973.

Gant, M. M., & Luttbeg, N. R. (1987). The cognitive utility of partisanship. *Western Political Quarterly, 40*(3), 499–517.

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology, 206*, 169–179.

Greene, S. (1999). Understanding party identification: A social identity approach. *Political Psychology, 20*(2), 393–403.

Henrich, J., McElreath, R., Barr, A., Ensimger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., & Ziker, J. (2006) Costly punishment across human societies, *Science, 312*, 1767–1770.

Johnson, T., Dawes, C. T., Fowler, J. H., McElreath, R., & Smirnov, O. (2007). The role of egalitarian motives in altruistic punishment. Unpublished manuscript. Available at http://jhfowler.ucsd.edu.

Key, V. O., Jr. (1966). *The responsible electorate: Rationality in presidential voting, 1936–1960*. Cambridge, MA: Harvard University Press.

Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. H. Kagel & A. E. Roth (Eds.), *The handbook of experimental economics* (pp. 111–194). Princeton, NJ: Princeton University Press.

Lodge, M., & Hamill, R. (1986). A partisan schema for political information processing. *American Political Science Review, 80*(June), 399–419.

Miller, W. E. (1991). Party identification, realignment, and party voting: Back to the basics. *American Political Science Review, 85*(2), 557–568.

Niemi, R. G., & Jennings, M. K. (1991). Issues and inheritance in the formation of party identification. *American Journal of Political Science, 35*, 970–88.

Olson, M. (1965). *The logic of collective action*. Cambridge, MA: Harvard University Press.

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge: Cambridge University Press.

de Quervain, D. J.-F., Fischbacher, U., Treyer, V., Schellhammer, M., Schnyder, U., Buck, A., & Fehr, E. (2004). The neural basis of altruistic punishment. *Science, 305*, 1254–1258.

Sethi, R., & Somanathan, E. (1996). The evolution of social norms in common property resource use. *American Economic Review, 86*, 766–788.

Smirnov, O. (2007). Altruistic punishment in politics and life sciences. *Perspectives on Politics, 5*(3), 489–501.

Sober, E., & Wilson, D. S. (1998). *Unto others: The evolution and psychology of unselfish behaviour*. Cambridge, MA: Harvard University Press.

Verba, S., Schlozman, K. L., & Brady, H. E. (1995). *Voice and equality, civic voluntarism in American politics*. Cambridge: Harvard University Press.