The Good of the Few: Reciprocity in the Provision of a Public Bad

Jason Delaney
Department of Economics and Finance
University of Arkansas at Little Rock
322 Reynolds Business Center
2801 S. University Ave.
Little Rock, AR 72204
jjdelaney@ualr.edu
Phone: 501-569-8874
Fax: 501-569-8871

Sarah Jacobson[1]
Department of Economics
Williams College
24 Hopkins Hall Dr.
Williamstown, MA 01267
Sarah.A.Jacobson@williams.edu
Phone: 413-597-2476
Fax: 413-597-4045

June 2012

Abstract:

People trade favors when it is efficiency-enhancing to do so; will they also trade favors when it reduces efficiency? This dilemma may arise, for example, in a series of wasteful public projects where each project benefits an individual. We introduce the "Stakeholder Public Bad" game to study this question. In each round, contributions to a common fund increase the earnings of one person (the "Stakeholder") but reduce the earnings of the rest of the group so much that overall efficiency is reduced. The Stakeholder position rotates through group members and the promise of high Stakeholder rewards provides a lever for reciprocal actions. We hypothesize that some people will help a current Stakeholder by contributing in hopes of being rewarded later with a reciprocal gift. In a lab experiment, we find evidence of such favor trading. Favor trading does not increase public bad provision, but it could if the parameters or population were different. We show that information provision can enable efficiency-decreasing reciprocal acts. We also find that Stakeholders seem quite willing to sacrifice the good of the group to reap their own personal rewards, even when contribution decisions are public.

JEL codes: C91, D01, D62, D64, D72, H41
Keywords: social preferences, reciprocity, externalities, public bad, public good, logrolling

[1] Corresponding author. The authors are grateful for funding from NSF Award SES-0752754 and support from Williams College. For helpful input and comments we also thank participants in various seminars and presentations, including at Williams College, University of Arkansas, and the Economic Science Association.

1. Introduction

Groups must frequently decide on the provision of projects that have both winners and losers. A defense contract may benefit one constituency while incurring large tax-funded expenses; a factory siting decision brings jobs to one area but may have broad environmental consequences; an appropriations bill may fund wasteful "bridges to nowhere" to the benefit of individual committee members. Further, these groups must often repeat the decision process with different stakeholders for different projects. This structure may give rise to reciprocal behavior: you support my project and I'll support yours. Existing evidence shows that this kind of trading may work in favor of pro-social projects. Does the same dynamic occur when projects are overall anti-social—when the harms they cause are greater than the benefits they generate? More generally, can reciprocity, a force championed for pro-sociality in settings from interpersonal interactions to the macroeconomy to evolution, be destructive?

We study this question using a model that allows reciprocal behavior to enter into a group's provision of a common project with heterogeneous costs and benefits. We create a game called the "Stakeholder Public Bad" game. In each round, members of a group decide how much to contribute to a common fund. These contributions determine the provision of an overall efficiency-reducing project in which one member has a stake (i.e., a strong financial interest) while other members' payoffs are reduced by project provision. This Stakeholder role rotates so that each group member will periodically be the beneficiary from the project. In some settings, public information makes reciprocal acts possible, but in others information is hidden so that targeted reciprocal acts are impossible. Behavior in these settings depends crucially on the existence and type of social preferences and on agents' expectations of reciprocation by others.

In a lab experiment implementing this model, we find evidence of anti-social reciprocal behavior, despite the result that such behavior remains unprofitable. We also find that subjects contribute nearly fully in the role of Stakeholder, even though this is an anti-social act.

2. Favor-Trading in Public Good and Public Bad Provision

Agents' behavior in a public bad game with rotating high return has, to our knowledge, not yet been studied. We take lessons from the extensive literature on public goods (useful surveys of which include Chaudhuri, 2011; Ledyard, 1995), noting particularly that subjects in public goods experiments contribute much more than selfish rational models predict. Suggested motives for this cooperation include fairness (e.g., Marwell and Ames, 1981), altruism (e.g., Dawes, 1980), and conditional cooperation (Gächter, 2007), although some suggest that confusion plays a role (e.g., Andreoni, 1995a; Ferraro et al., 2003; Houser and Kurzban, 2002). Many institutions for project provision have been examined, the most relevant of which is the linear voluntary contributions mechanism game in which individuals' contributions to the common fund have constant returns to each member of the group. This game is the basis for our Stakeholder Public Bad game.

Projects that reduce overall efficiency have received less attention. This is in part because many models treat public bads as dual to public goods. For example, one can argue that preferences against (the public bad of) pollution are the same as preferences for (the public good of) pollution abatement. The theoretical and experimental literatures agree that treating bads as isomorphic to goods causes problems, however. First, because the Nash equilibrium can lead to unbounded amounts of a public bad, it is unclear how the dual of a public good should be treated theoretically (Shitovitz and Spiegel, 2003). Second, there are systematic differences in the extent of observed cooperation when isomorphic games are framed as public good provision as

compared to public bad reduction. Schwartz-Shea (1983), Andreoni (1995b), and Sonnemans et al. (1998) demonstrate that people are more pro-social under a public goods framing than a public bad framing. Relatedly, DeScioli et al. (2011) find that antisocial acts of omission occur more frequently and receive less punishment than acts of commission. Thus, public bad settings merit separate study. The existing public bad literature indicates that the magnitude of social costs, the existence of social norms, and the order of play all affect decisions (Barr and Serra, 2009, 2010; Moxnes and van der Heijden, 2003; van der Heijden and Moxnes, 2003) but has not yet studied reciprocal behavior.

Reciprocity has been shown to be important in pro-social settings in general, and reciprocal acts are central to our model. Sobel (2005) provides a useful discussion of the related literature. He classifies "intrinsic reciprocity" as favor-trading rooted in other-regarding preferences, as opposed to "instrumental reciprocity" which is favor-trading that is simply strategic in the sense of seeking a future reward. Robust theories of reciprocal preferences have been developed (e.g., Cox et al., 2008; Dufwenberg and Kirchsteiger, 2004; Rabin, 1993; Wilson, 2008) and experimentalists have provided extensive evidence (e.g., Berg et al., 1995; Charness and Rabin, 2002; Cox, 2004; Fehr et al., 1993) of the existence of reciprocal behaviors and of their link to social preferences. Instrumental reciprocity has also been shown to be important (e.g., List, 2006).

Reciprocal behavior in the form of conditional cooperation is important in public good games (Gächter, 2007). However, information conditions and payoff asymmetries can provide additional leverage for reciprocity of all types within a group provision setting. Information provision alone may increase giving (e.g., Andreoni and Petrie, 2004; Sell and Wilson, 1991). But information can also enable reciprocity: it is impossible to reciprocate without knowing who

has been kind to you (Wilson, 2008). Asymmetric returns have mostly been studied to look at either responsiveness to returns to self and others (e.g., Goeree et al., 2002), or to look at leadership (e.g., Brandts et al., 2007; Glöckner et al., 2011), sometimes in the spirit of Olson's (1965) "privileged groups" (Reuben and Riedl, 2009). Isaac et al. (2011) study an asymmetry in that common projects benefit some people and hurt others, but do not provide opportunities for favor trading. In a study of the role of asymmetric returns and information in favor trading in support of a public good, Jacobson and Petrie (2012) demonstrate that other-regarding preference-based (intrinsic) reciprocity does boost support of a pro-social common project. We ask whether this will also occur in support of an anti-social common project.[2]

We extend the existing literature by examining reciprocal acts in a project provision setting where the common project is anti-social. We do not seek the kind of "negative reciprocity" (i.e. spite) examined in work like Abbink et al. (2000), where reciprocal preferences cause people to reduce one another's payoffs through punishment. In such settings, negative reciprocity can be socially positive because it may enforce cooperative norms. In our setting, reciprocal acts are socially harmful.

3. Model

We model a "Stakeholder Public Bad" game. In this game, members of a group make repeated simultaneous individual decisions to contribute to a common fund. These contributions generate a public project with asymmetric returns: some group members benefit from provision while others are hurt, and the socially optimal level of provision is zero. One can interpret contributions in this setting as either public bad provision or common pool resource extraction.

[2] Separate work examines behavior when the project affects people with no power to decide on provision of the project, e.g., "bystanders" in Engel and Rockenbach (2011) and "outsiders" in Delaney and Jacobson (2012).

Agents belong to groups of size *N*. Group membership is fixed and the interaction continues for a finite number of periods. Each member has an endowment of *z* tokens each period to allocate between a private fund and a common fund. The private fund represents the opportunity cost of support for the common fund and provides a return of $a > 0$ per token. Agent *i* earns return b_{it} to for every token contributed by any group member to the common fund. This return varies across roles and may be negative for some group members. Each agent also earns a role-specific baseline return G_{it} from the "status quo" (no contributions) level of public project provision.[3]

Agent *i*'s payoff in period *t* is given by equation 1.

$$\pi_{it} = G_{it} + b_{it}\left(\sum_{j=1\ldots N} g_{jt}\right) + a\left(z - g_{it}\right) \qquad (1)$$

In each period *t*, agents are exogenously assigned roles. The values of G_{it} and b_{it} vary according to agent *i*'s role in period *t*. Each agent then chooses his contribution g_{it}. The simple net return to *i* for any token he contributes to the public fund is $b_{it} - a$.

In each period, one member of the group has the role of Stakeholder (role *S* if $Stakeholder_t = i$). The Stakeholder strongly prefers the common project to the status quo: the Stakeholder return from the common fund is $b_S > a > 0$, thus making this a privileged group (Olson, 1965). The Stakeholder role rotates through all group members from period to period.[4]

[3] This fixed status quo is similar to the "alternative public project" in Isaac et al. (2011). There, however, agents can choose to contribute to this alternative and such contributions actually reduce provision of the main project.

[4] It is worth noting that in some situations, a Stakeholder in a potential project that would be anti-social may be able to (in addition to withholding his own contributions) "bury" his project so that no-one has opportunity to contribute. In other situations, as in our model, a Stakeholder may have no such power.

The remaining $N-1$ group members in each period are Non-Stakeholders (role *NS if* $Stakeholder_t \neq i$). Non-Stakeholders prefer the status quo to positive project provision; their per-token return from the public fund is $b_{NS} < 0 < a$. Agent *i*'s payoffs for period *t* are given by equation 2.

$$\pi_{it} = \begin{cases} G_S + b_S\left(\sum_{j=1\ldots N} g_{jt}\right) + a\left(z - g_{it}\right) & \text{if } Stakeholder_t = i \\ G_{NS} + b_{NS}\left(\sum_{j=1\ldots N} g_{jt}\right) + a\left(z - g_{it}\right) & \text{if } Stakeholder_t \neq i \end{cases} \tag{2}$$

This project is a public bad if the total return from a token contributed is negative. This happens if the total losses of Non-Stakeholders combined with the opportunity cost of the token are larger than the gains of Stakeholders. Thus, the project is a public bad if $b_S + (N-1)b_{NS} < a$.[5]

We can make some theoretical predictions based on the bounds we have placed on parameters. Because payoffs are linear in own-contribution, each role has a dominant strategy if all agents are rational and purely self-regarding: each agent contributes fully to the common fund when he is Stakeholder and contributes nothing when he is Non-Stakeholder. Because there is a known end-period, rational agents should not attempt strategic cooperation because they expect unraveling.

Other-regarding agents face a dilemma. Contribution helps one member of the group at others' expense and reduces efficiency. Altruism or efficiency-seeking may cause Stakeholders to reduce their contributions to the common fund. Altruistic Non-Stakeholders should generally not contribute to the common fund unless they have preferences that privilege the current

[5] The "public good" version of this game examines favor-trading in support of a public good, where full contribution is socially optimal but privately suboptimal for Non-Stakeholders. This was examined in Jacobson and Petrie (2011).

Stakeholder above other group members, and efficiency-seeking Non-Stakeholders should never contribute.

If agents are inequity-averse, their behavior depends on their expectations of others' actions. Because the Stakeholder role rotates through all group members, if other agents contribute as Stakeholder and refrain as Non-Stakeholder, then an inequity-averse agent should do the same to ensure equal payoffs. If group members deviate from that pattern, an inequity-averse agent should match that deviation.

Untargeted conditional cooperation should have little effect on contributions. Conditional contributions by Non-Stakeholders intended to spur or respond to contributions by the group at large are unlikely because increased contributions would reduce all agents' earnings.

However, group members may use the rotating Stakeholder position to alternately "help" one another in a targeted way. Imagine that in three sequential periods, first Adam is Stakeholder, then Beatrice, then Cynthia. In the first period, Beatrice contributes a large amount while Cynthia contributes nothing. Beatrice's contribution was personally costly in that she sacrificed her own payoff to increase Adam's. Cynthia made no such sacrifice. If Adam exhibits reciprocal behavior, these acts may affect his contributions in the following periods. He may contribute a large amount when Beatrice is Stakeholder and less when Cynthia is Stakeholder.

With reciprocal other-regarding preferences (intrinsic reciprocity), this discrimination happens because Adam's preferences for Beatrice's and Cynthia's payoffs are changed by their previously kind and unkind, respectively, acts. On the other hand, reciprocation may be instrumental. Adam might seek future rewards by strategically contributing when a likely

reciprocator is Stakeholder. He might guess from her past generosity that Beatrice is reciprocator. He could then mimic a reciprocating type in pursuit of a higher payoff.[6]

Stakeholder contributions may also be affected by intrinsic or instrumental reciprocity. A Stakeholder who wishes to earn higher contributions in future Stakeholder stints or to reward the kindness of past benefactors may reduce his common fund contributions now since those contributions hurt all of his group members.

Both kinds of reciprocity require that agents know one another's history of actions and schedule of Stakeholder timing. Without this information, reciprocity in the sense of targeted rewards for individuals' past actions cannot influence contributions. We model two information conditions. In the Public condition, group members know one another's contribution history and roles in each period. In the Private condition, group members know only their own role and history; they learn the individual amounts contributed in past periods but cannot associate them with any particular group member and know nothing about the timing of others' Stakeholder stints.

To summarize, given either self-interest or other-regarding preferences, we expect Non-Stakeholder contributions in the Private information condition to be low—some baseline level near zero. In the Public condition, Non-Stakeholder contributions might increase because of targeted reciprocity. This outcome depends crucially on initial contributions: if Non-Stakeholders reduce contributions (relative to the baseline) to previously unkind Stakeholders, this reduction may offset gains from higher contributions to previously kind Stakeholders so that the effect on Non-Stakeholder contributions of Public information is ambiguous. Regardless, direct favor-trading can occur in the Public condition only. Additionally, Stakeholder

[6] If there is common knowledge that no-one is a reciprocator and everyone is fully forward-looking, then this kind of "cooperation" would wholly unravel. If agents are fully self-interested but myopic the unraveling might be incomplete and might generate strategic contributions in early periods of the game.

contributions may decrease in the Public as compared to the Private condition because of instrumental or intrinsic reciprocity.

4. Experiment

We implement the Stakeholder Public Bad game in a laboratory experiment based on a linear public bad game with rotating asymmetric payoffs. We use two treatments corresponding to the Private and Public information conditions, which we describe in detail below.

In each treatment, subjects are randomly assigned to fixed four-person groups for eight rounds. In each round, each subject is endowed with $z = 10$ tokens. He then must choose how many tokens ($g_{it} \in \{1,\ldots,10\}$) to invest in the Group Fund, while the remaining tokens are kept in the Personal Fund. Each group has one Stakeholder and three Non-Stakeholders in each round. The Stakeholder role rotates through all members of each group so that each subject is Stakeholder twice and Non-Stakeholder six times in each eight-round treatment.

The per-token payoff from the Personal Fund is $a = \$0.02$ for all group members. For Stakeholders, the Group Fund yields no base payment ($G_S = 0$) but the per-token payoff from the Group Fund is $b_S = \$0.10$.

For Non-Stakeholders, the Group Fund pays a base payment of $G_{NS} = \$2.00$. This base payment ensures that no Non-Stakeholder can ever earn a negative amount in any round; the Stakeholder base payment G_S is zero because this concern does not exist for them. This base payment is reduced by $b_{NS} = -\$0.05$ per token in the Group Fund.

Given these parameters, the net marginal social return to each token in the Group Fund is $\$0.08 - 3*\$0.05 = -\$0.07$ (or $-\$0.09$, considering the \$0.02 opportunity cost). If all subjects are purely self-regarding and myopic, there is a dominant strategy equilibrium in each round:

Stakeholders contribute fully and Non-Stakeholders do not contribute. In this case, the total group investment is 10, the Stakeholder earns \$1.00, and the Non-Stakeholders each earn \$1.70.[7] The total group payoff is then \$6.10. The socially optimal outcome is for no tokens to be invested in the Group Fund, in which case $\pi_S = \$0.20$ and $\pi_{NS} = \$2.20$, and total group earnings are \$6.80. If all agents contribute all tokens, the total group investment is 40, $\pi_S = \$4.00$, $\pi_{NS} = \$0.00$, and total group earnings are \$4.00.

In the Public information condition, each subject is assigned a letter code. Subjects see a table in which the timing of Stakeholder position for all group members is reported and in which each group member's contribution history is displayed. In the Private condition, subjects' contributions to the Group Fund are reported in a disaggregated list (it has been noted, e.g., Sell and Wilson, 1991, that disaggregated reporting of group member contributions may affect giving). Because contributions are listed in a random order that is reshuffled each round, norms may be established and subjects may follow one another. However, reputations cannot be established and Stakeholder timing is private information so targeted reciprocity is impossible.

The experimental procedure is outlined in Figure 1. The experiment interface is computerized using software written in z-Tree (Fischbacher, 2007). Subjects enter the lab and are given general instructions.[8] They are told that they will make decisions in two sets of eight rounds with two different groups and that they will then make one unrelated decision, but are not told the exact nature of the decisions they will make in each treatment until directly before the treatment begins. The design is within subject—each subject participates in both the Public and

[7] It is not impossible, given these values, that inequity aversion (rather than self-interest) could explain full Stakeholder contributions. Given the rotating nature of the role, the level of myopia required for inequity-aversion to generate the contributions observed seems implausible.

[8] Instructions are available on the corresponding author's website.

Private information treatments—and the final unrelated task is a risk preference elicitation in the style of Holt and Laury (2002).

FIGURE 1 GOES HERE.

The first treatment begins with instructions that explain the roles and the information condition for that treatment. The software randomly assigns subjects into four-person groups. The subjects then play through all of the rounds of the treatment. After the first treatment, subjects are randomly assigned into new four-person groups. The second treatment features the complementary information condition and proceeds in much the same way, with treatment-specific instructions read first. After both treatments are complete, subjects receive instructions for and perform the additional task. Finally, subjects complete a questionnaire and receive payment anonymously. Each subject's total earnings is the sum of his earnings in each treatment, which in turn are the sum of his earnings in each round plus his earnings from the final task.

5. Results

The experiment was run at the Experimental Economics Center (E*x*CEN) at Georgia State University in March 2010 in four separate 20-subject sessions, for a total of 80 subjects. All subjects played two treatments of eight rounds, one in a Public and one in a Private information condition. Half of the sessions ran the Public treatment first, and half ran the Private treatment first. Some small order effects are detectable, so all analysis includes only data from subjects' first treatment.[9] The protocol was double anonymous: subjects could not identify which subjects were in their group, and the experimenters could not identify which subject made any set of

[9] Order effects are: Non-Stakeholder contributions are higher in a subject's first treatment; and Stakeholder contributions are higher in both conditions if the Private treatment is first. Results change little when the full data set is used: Stakeholder contributions are significantly greater in the Public than in the Private treatment; and the difference-in-difference test comparing favor-trading between the Public and Private treatments is significant.

decisions. Of the 80 subjects, 40 (50%) were female, and the average age was 20.1. Each session lasted about 90 minutes, and average earnings were $23.38 (standard deviation $1.85).[10]

Contribution Level Results

Figure 2 shows the path of contribution decisions across the rounds of each treatment. Stakeholder decisions in both treatments are close to the endowment, which is consistent with the selfish dominant strategy of full contribution by Stakeholders. This should be viewed in light of the fact that such contributions are now social costly on net because of the harm they cause to Non-Stakeholders. Non-Stakeholder contributions are low but positive in all rounds. Contributions show the downward trend usually seen in public goods games, even though this public investment is actually a public bad.[11]

FIGURE 2 GOES HERE

Table 1 shows summary statistics of the distribution of individual contributions across rounds. The majority of contributions follow the selfish dominant strategy, but many subjects deviate from perfect adherence. In particular, Non-Stakeholders have a greater tendency to give nonzero amounts than do Stakeholders tend to give less than full endowment.

TABLE 1 GOES HERE.

We compare mean contributions by role and treatment in Table 2. Both Stakeholders and Non-Stakeholders contribute slightly less in the Public treatment as compared to the Private treatment, but this is not significant. If subjects act reciprocally, it may be the case that subjects'

[10] We do not report results from the risk elicitation task or the questionnaire, although we make some discussion of the minor points of interest with regard to these in a Reviewer's Appendix. We note that twelve of the 80 subjects incorrectly answered a comprehension question on the questionnaire. The results reported in the body of the text include their choice data, but all results presented are robust to their exclusion.

[11] Obviously, interpretation of our results depends on the assumption that subjects understand the game they are playing and in particular understand the public bad nature of the common fund. While it is impossible to prove this with any certainty, we find at least some questionnaire responses to imply an understanding of this feature.

increased contributions to kind Stakeholders are offset by decreased contributions to unkind Stakeholders. In the next sub-section we present evidence that reciprocity does indeed occur in the Public treatment even though average contribution levels do not change.

TABLE 2 GOES HERE.

Total (summed across all group members) group contributions also do not differ between the Public and Private conditions. Figure 3 shows the trend across rounds for the two treatments. Across-round average total contribution is 41.77% of the maximum possible group contribution as compared to 40% of the maximum in the Public treatment, and these numbers are not statistically different (Wilcoxon rank-sum test *p*-value 0.626).

FIGURE 3 GOES HERE.

Because of the novel structure of induced preferences here, it is difficult to compare contributions in this experiment to contributions in other experiments. In Jacobson and Petrie (2012), with a similar asymmetric payoff setup but a public good instead of a public bad, Stakeholders gave slightly more (95-97% of endowment) than they do here, Non-Stakeholders gave much more (33-38%) than they do here, and both roles show trends similar to the trends shown here. This seems like a surprisingly small reduction in contributions given the public bad nature of the public project in the Stakeholder Public Bad game as compared to the public good nature of the public project in Jacobson and Petrie (2012). While some Non-Stakeholder giving may be driven by social preferences, Stakeholders' willingness to follow self-interest to the detriment of their group appears to be uninhibited by social preferences; in this setting, self-interest and reciprocal behavior appear to overwhelm both altruism and a desire for efficiency.

Presence of Reciprocal Behavior

Reciprocal contributions may be caused by intrinsic or instrumental reciprocity. We test for reciprocal acts, without distinguishing the kind of reciprocity, by comparing how much a Non-Stakeholder contributes in two different conditions. The first condition is that the Non-Stakeholder is facing a Stakeholder who contributed generously (more than half his endowment) in a past period in which this person was Stakeholder. The alternative condition is that the Non-Stakeholder is facing a Stakeholder who contributed ungenerously (less than half of his endowment) when this person was Stakeholder.[12]

As shown in Table 3, we test whether subjects discriminate between these generous and ungenerous Stakeholders in both the Private and Public treatments. We do not detect evidence of reciprocity in the Private treatment, where anonymity of contributions makes targeted reciprocal acts impossible. In the Public treatment, however, Non-Stakeholders do respond to the current Stakeholder's past kindness, i.e., they reciprocate: they contribute more when the current Stakeholder was previously generous than when he was previously ungenerous. The within-subject difference is significant even though the sample size is greatly reduced.[13] The difference in this sort of discrimination across the Private and Public treatments is not significant (Wilcoxon rank-sum *p*-value 0.241), although again, the sample size is very low. Still, some subjects' tendency to discriminate seems quite strong.

TABLE 3 GOES HERE.

To explore this further, we use regression methods to examine the relationship between previous generosity and Non-Stakeholder contributions. For the regression analysis, we assume

[12] The results that follow hold for alternative specifications (30% - 90%) of the threshold for generosity.

[13] Most people do not face both a previously-kind and previously-unkind Stakeholder. Each subject experiences only 3-6 rounds as Non-Stakeholder after his first Stakeholder stint. Since the majority of Non-Stakeholder contributions are zero in both treatments, is it not surprising that fairly few subjects face both types of Stakeholders after their first Stakeholder stint in this small number of rounds.

that subject *i*'s Non-Stakeholder contributions in period *t*, denoted g_{it}^{NS}, can be modeled as a constant baseline level with a time trend. Further, g_{it}^{NS} may be affected by a within-group norm, which we measure as the cumulative average of other group members' Non-Stakeholder contributions and denote $\bar{g}_{it}^{NS}$, and by "favor" contributions by the current Stakeholder *j* in previous periods when subject *i* was Stakeholder, denoted $\bar{f}_{ijt}$. To examine differences across treatments, we define indicator variable d_p, which is 1 for the Public treatment and 0 otherwise. We first examine a full-interaction model (noting that d_p must be omitted if we include individual fixed effects since each subject is observed in only one treatment). The regression model is given in equation 3.

$$g_{it}^{NS} = \beta_0 + \beta_1 \bar{f}_{ijt} + \beta_2 \bar{g}_{it}^{NS} + \beta_3 t + \beta_4 \left(d_p \times \bar{f}_{ijt} \right) + \beta_5 \left(d_p \times \bar{g}_{it}^{NS} \right) + \beta_6 \left(d_p \times t \right) + c_i + \varepsilon \qquad (3)$$

Parameter estimates from this model are given in specification I in Table 4. The interaction terms between the Public treatment dummy and the time trend and group norm terms are not significant, and we had no a priori reason to expect these terms to be meaningful. Therefore, we consider a reduced model, given in equation 4, retaining only the interaction on previous "favor" contributions from the current Stakeholder.

$$g_{it}^{NS} = \beta_0 + \beta_1 \bar{f}_{ijt} + \beta_2 \bar{g}_{it}^{NS} + \beta_3 t + \beta_4 \left(d_p \times \bar{f}_{ijt} \right) + c_i + \varepsilon \qquad (4)$$

Parameter estimates from this model are presented in specification II in Table 4. We note that the coefficient on $\bar{f}_{ijt}$ (the "favor" contribution by the current Stakeholder) is positive and significant, although small. This implies the existence of some reciprocal behavior in the Private treatment. On further examination, this appears to be a spurious correlation driven by low last-period contributions coinciding with Stakeholders who (like other subjects) had steadily reduced contributions.

More importantly, the size and significance of the interaction term coefficient indicates that reciprocal giving is occurring in the Public information condition much more than in the Private condition: previous generosity by the current Stakeholder has a significant effect on Non-Stakeholder contributions. The Public condition results are robust to the exclusion of the last period. Thus we have strong evidence of favor-trading in support of a public bad, enabled by the provision of information.

Does trading favors pay, i.e. is the sacrifice worthwhile for a Non-Stakeholder? For each token that the current Stakeholder contributed in previous rounds, Non-Stakeholders in the Public treatment contribute an additional 2.220 + 0.376 = 2.596 percent of their endowment, or 0.260 tokens (p = 0.017). The net private cost of contributing a token when Non-Stakeholder is -7 cents, while Stakeholders earn 10 cents from each token donated other subjects. The reciprocity-fueled net return to a token contributed by a Non-Stakeholder is thus $(10 \times 0.260) - 7 = 2.6 - 7 = -4.4$ cents. Even at the upper bound of the confidence interval for any reasonable confidence level, Non-Stakeholder contributions remain unprofitable in expectation, inclusive of the reciprocity effect.

TABLE 4 GOES HERE.

Anti-social favor trading, which is clearly a socially harmful activity, is enabled by information on subjects' histories and the timing of their interest in the common fund. In this case, reciprocal behavior and the information that enables it serve as a force for evil. We contrast this with the result in Jacobson and Petrie (2012), who show that providing information that allows reciprocity in a similar game (a "Stakeholder Public Good" rather than a "Stakeholder Public Bad" game) increases efficiency by leveraging pro-social favor trading.

6. Conclusion

In this paper, we examine reciprocal behavior in the provision of a public bad. Individuals may have competing motivations when members of a group must decide on the provision of a common project. We model a sequence of inefficient projects, each of which is nonetheless privately desirable to a booster within the group. Natural analogies include political logrolling or influence peddling. Reciprocity has been shown to be a force for social good in many settings; our results show that it can also cause social harm.

We develop a novel "Stakeholder Public Bad" model in which asymmetric returns create a public bad that is privately beneficial to a single constituent. In our model, we show that intrinsic or instrumental reciprocity could lead to favor-trading in the provision of these projects which, in turn, could increase overall provision. In an experiment implementing this model, we find that some subjects discover the opportunity to trade favors and actively engage in this favor trading when information renders favor trading possible. Reciprocal behavior does not increase the level of public bad provision in this case. Given the existence of reciprocity, however, overall public bad provision might increase or decrease by favor trading, depending on early contributions and the distribution of agent types (e.g., reciprocators, etc.) in the population.

In this way, reciprocal acts—and the information that makes reciprocity possible—may reduce efficiency. This potential for anti-social use of information has been noted in the context of campaign finance reform in the United States. Ackerman and Ayres (2002) argue that all campaign contributions should be anonymous to render political favor trading impossible.

More broadly, our results demonstrate observe how people behave when they can take privately beneficial actions at others' expense. The effects of both self-interest and reciprocal behavior dominate social preferences like altruism and efficiency-seeking. This is a subtly but

fundamentally different issue than people's failure to provide a public good. When subjects benefit directly from a public bad, as in our Stakeholder role, they contribute nearly fully. Even some subjects who bear a private cost from the public bad (Non-Stakeholders) contribute a positive amount to the public bad. Our results show that some of this behavior is caused by subjects' hopes of garnering future rewards when their "pet project" is the one being provided (i.e., when they are Stakeholder).

The direct rewards reaped by a person with a stake in a common project appear to be so tempting that they can counterbalance a person's inherent social preferences. Despite the negative social effects of provision of the public bad, subjects from both roles contribute to the provision of the public bad to their mutual detriment. Altruism and pro-social reciprocity are real and have been proven repeatedly. In this setting, however, individuals attempt to harness reciprocity and information as a force for the good of the few but against the good of the many.

Appendix A: Correlates of Behavior

Additional insights into contribution behavior come from subjects' choices in a risk preference elicitation task (in the style of Holt and Laury, 2002) and from subjects' questionnaire responses. It should be noted that neither of these could have primed subjects, as both were completed after the previous experimental tasks were concluded. We considered some individual characteristics, including self-identified race, gender, religion, previous experience with economics, and charitable giving behavior, as well as several measures of political affiliation and attachment. The characteristics that vary with contribution decisions are described in Table A-1.

TABLE A-1 GOES HERE

Subjects who identify with the Democratic Party tend to give less as Stakeholders in the Public treatment as compared to subjects who do not identify with that party. In the Private treatment the same pattern exists but is not statistically significant (Democrats give 89.71 as compared to 96.74 percent of endowment, Wilcoxon rank-sum *p*-value 0.122). This accords with the results from other social sciences that Democrats are more oriented toward duties toward society and a feeling of obligation to help others (Coffé and Bolzendahl, 2011).[14]

We also look at risk aversion as measured by a subject's lottery switch point. An early switch point indicates that the person is less risk-averse. We consider those who are risk-seeking to mildly risk-averse as one group and compare their contributions to those of the more risk-averse subjects.[15] Less risk-averse subjects give more as Non-Stakeholder than more risk-averse subjects do in the Private treatment. This is sensible because giving as Non-Stakeholder in the Private treatment is very risky, in the sense that it is extremely unlikely to be reciprocated. In the

[14] Subjects who identify as black are more likely to identify as Democrats, as are subjects who have not taken economics classes. Stakeholder contributions do not vary significantly by whether a subject has taken economics classes. Black subjects do give less as Stakeholder in the Public treatment, but not the Private treatment.

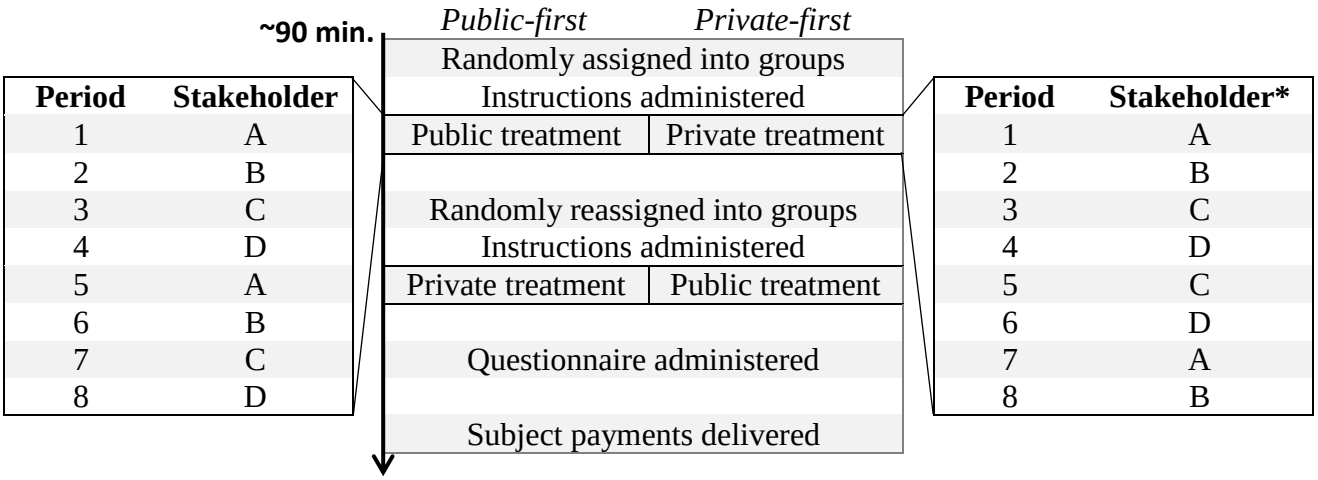
[15] This result is robust to other thresholds of the "more risk-averse" classification.

Private treatment, any "payback" would rely wholly on the weaker force of group-level conditional cooperation, while in the Public treatment targeted reciprocal acts can generate returns.

Finally, we included a simple comprehension test in the questionnaire. This gave a simple scenario with two funds with different returns, and asked the subject how many tokens he should put into the fund with the higher return to maximize his profit. Out of the 80 subjects in the experiment, 68 (85%) answered this question correctly. A tendency to make non-dominant contributions—to contribute high amounts as Non-Stakeholder and to contribute low amounts as Stakeholder—is correlated with tendency to answer this comprehension question incorrectly.[16] As noted in the text, all of our main results are robust to the exclusion of the subjects who answered this question incorrectly.

[16] Subjects who answered the comprehension question incorrectly were much more likely to have reported taking no economics classes, so this characteristic is also associated with non-dominant contributions.

Figures



*The Stakeholder position in the Private treatment follows this pattern, but letter codes are not revealed to subjects in this treatment. Half of the sessions began with the Public treatment first, while half began with the Private treatment first.

Figure 1. Experiment design.

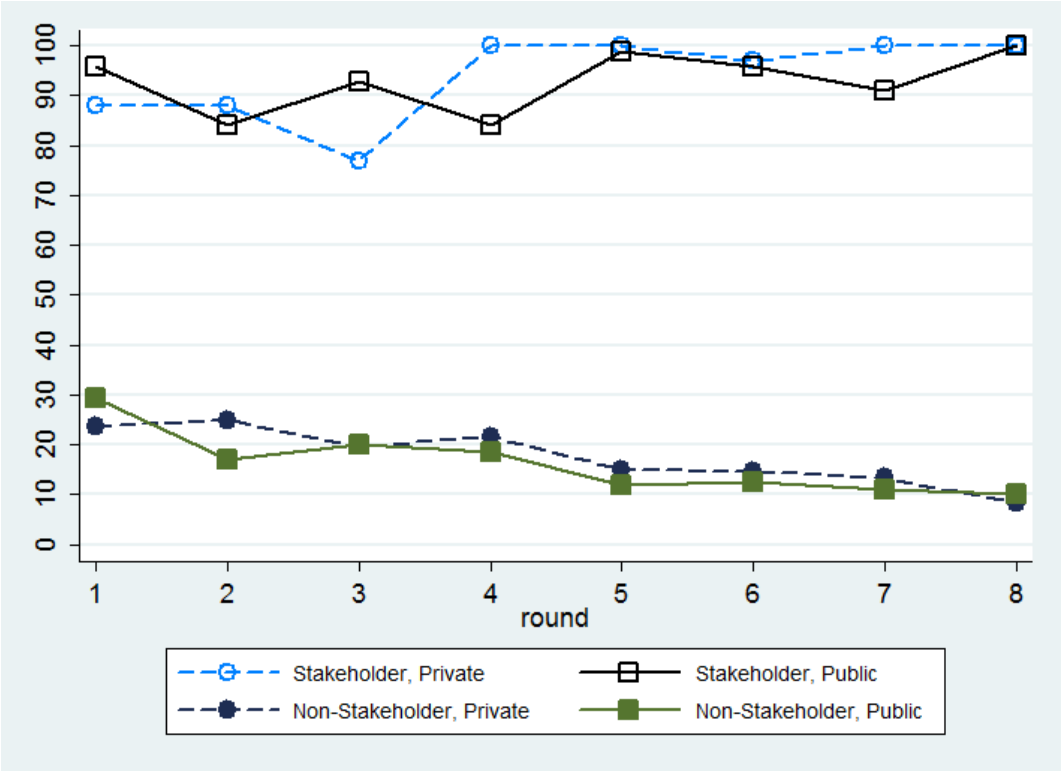


Figure 2: Contributions across rounds by role and treatment (in percent of endowment)

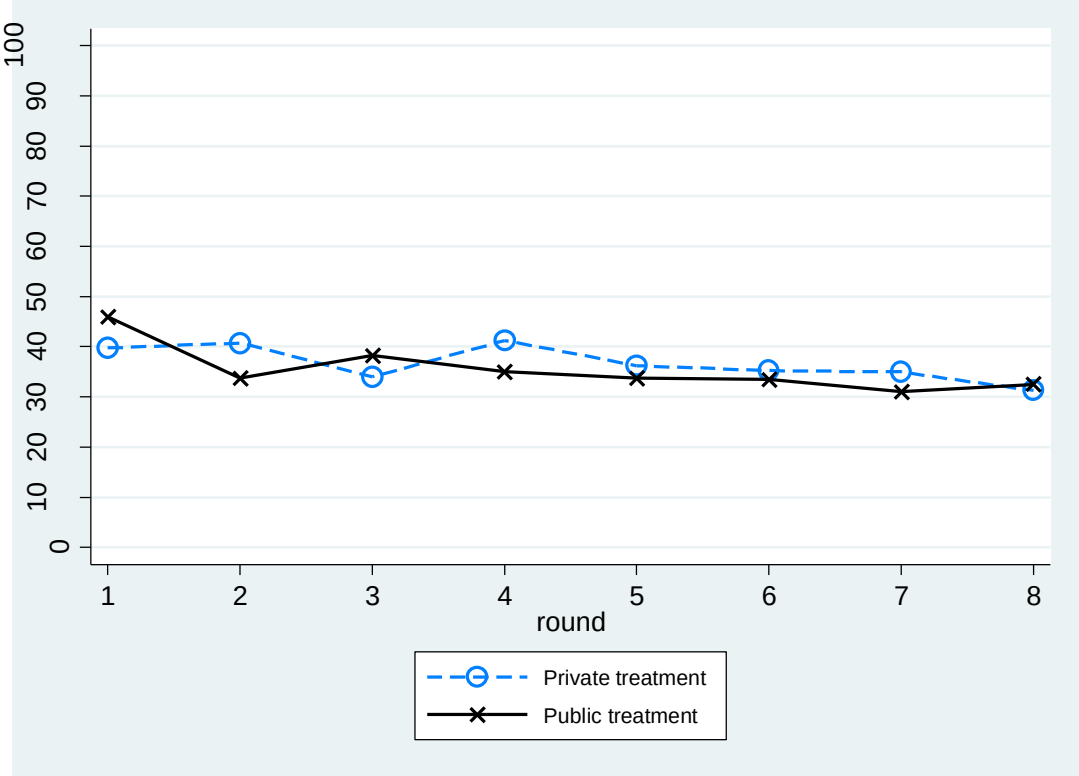


Figure 3: Group average public bad provision by treatment (in percent of total group endowment)

Tables

Table 1: Distribution of contribution amounts by treatment and role

	Non-Stakeholder		Stakeholder	
	Private	**Public**	**Private**	**Public**
Gave 0% of endowment[a]	161 (67.08%)	140 (58.33%)	3 (3.75%)	2 (2.5%)
Gave intermediate amount[a]	60 (25%)	86 (35.83%)	8 (10%)	12 (15%)
Gave 100% of endowment[a]	19 (7.92%)	14 (5.83%)	69 (86.25%)	66 (82.5%)
Median contribution	0	0	Endowment	Endowment
Subjects who always follow selfish dominant strategy[b]	13 (32.5%)	14 (35%)	29 (72.5%)	27 (67.5%)
Number of contributions	240	240	80	80
Number of subjects	40	40	40	40

[a] Cells contain number of contributions with percent of contributions in parentheses.
[b] Selfish dominant strategy is to contribute 0 as Non-Stakeholder and 100% of endowment as Stakeholder.

Table 2: Mean contributions by role and treatment (in percent of endowment)

	Non-Stakeholder	Stakeholder
Private	17.67	93.75
	(24.58)	(13.95)
Public	16.33	92.88
	(21.79)	(15.34)
Wilcoxon rank-sum test *p* value	0.992	0.680

N = 40 in each cell. Standard deviations in parentheses.

Table 3: Evidence of reciprocal contributions

	Private	Public
Stakeholder extracted > 50% when I was Stakeholder in past	25.00	29.32
	(9.64)	(10.02)
Stakeholder extracted ≤ 50% when I was Stakeholder in past	19.23	9.06
	(8.42)	(3.22)
N	13	11
Wilcoxon signed-rank test *p*-value	0.420	0.058

N < 40 in each cell because subjects who did not face both "nice" and "mean" Stakeholder were dropped. Standard deviations in parentheses.

Table 4. Fixed-Effects Panel Regression of Non-Stakeholder Contributions on History

	I	II
Stakeholder average past contributions to me	0.461***	0.376**
	(0.166)	(0.163)
Average past Non-Stakeholder contributions (excluding self)	-0.032	0.020
	(0.129)	(0.105)
Period	-2.289**	-1.824***
	(1.030)	(0.662)
Public indicator × Stakeholder average past contributions to me	2.283**	2.220**
	(1.088)	(1.076)
Public indicator × Average past Non-Stakeholder contributions (excluding self)	0.231	
	(0.189)	
Public indicator × Period	0.876	
	(1.334)	
Observations (rounds)	360	360
Number of subjects	80	80
F	3.34	4.26
R^2 (overall)	0.0304	0.1092

Robust standard errors in parentheses; errors are clustered on groups and all models are individual fixed-effects models.
Significance levels: *: 10%, **: 5% ***: 1%

Table A-1: Correlates of contribution decisions

Characteristic	Subjects with Characteristic	Treatment	Role	Difference
Democrat	21 of 40	Public	Stakeholder	Democrats give less (88.81 vs. 97.37, p=0.009)
Risk averse	17 of 40	Private	Non-Stakeholder	Less risk averse give more (27.94 vs. 10.07, p=0.026)
Risk averse	17 of 40	Private	Stakeholder	Less risk averse give less (87.65 vs. 98.26, p=0.054
Comprehension question wrong	6 of 40	Private	Non-Stakeholder	Wrong answer give more (42.50 vs. 13.28, p=0.022)
Comprehension question wrong	6 of 40	Public	Non-Stakeholder	Wrong answer give more (35.00 vs. 13.04, p=0.008)

Amounts given in percent of endowment. *P*-values are for Wilcoxon rank-sum test

References

Abbink, Klaus, Irlenbusch, Bernd and Renner, Elke, 2000. "The moonlighting game: An experimental study on reciprocity and retribution." *Journal of Economic Behavior & Organization*, 42(2), 265-277.

Ackerman, Bruce A. and Ayres, Ian, 2002, *Voting with dollars: a new paradigm for campaign finance*. Yale Univ Pr.

Andreoni, James, 1995a. "Cooperation in Public-Goods Experiments: Kindness or Confusion?" *The American Economic Review*, 85(4), 891-904.

Andreoni, James, 1995b. "Warm-Glow versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments." *Quarterly Journal of Economics*, 110(1), 1-21.

Andreoni, James and Petrie, Ragan, 2004. "Public goods experiments without confidentiality: A glimpse into fund-raising." *Journal of Public Economics*, 88(7-8), 1605-1623.

Barr, Abigail and Serra, Danila, 2009. "The effects of externalities and framing on bribery in a petty corruption experiment." *Experimental Economics*, 12(4), 488-503.

Barr, Abigail and Serra, Danila, 2010. "Corruption and culture: An experimental analysis." *Journal of Public Economics*, 94(11-12), 862-869.

Berg, Joyce, Dickhaut, John W. and McCabe, Kevin A., 1995. "Trust, Reciprocity, and Social History." *Games and Economic Behavior*, 10(1), 122-142.

Brandts, Jordi, Cooper, David J. and Fatas, Enrique, 2007. "Leadership and Overcoming Coordination Failure with Asymmetric Costs." *Experimental Economics*, 10(3), 269-284.

Charness, Gary and Rabin, Matthew, 2002. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics*, 117(3), 817-869.

Chaudhuri, Ananish, 2011. "Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature." *Experimental Economics*, 14(1), 47-83.

Coffé, Hilde and Bolzendahl, Catherine, 2011. "Partisan Cleavages in the Importance of Citizenship Rights and Responsibilities*." *Social Science Quarterly*, 92(3), 656-674.

Cox, James C., 2004. "How to Identify Trust and Reciprocity." *Games and Economic Behavior*, 46(2), 260-281.

Cox, James C., Friedman, Daniel and Sadiraj, Vjollca, 2008. "Revealed Altruism." *Econometrica*, 76(1), 31-69.

Dawes, R M, 1980. "Social Dilemmas." *Annual Review of Psychology*, 31(1), 169-193.

Delaney, Jason and Jacobson, Sarah, 2012, "Those Outsiders: How Downstream Externalities Affect Public Good Provision." Williams College Economics Department Working Papers. Williams College, Williamstown, MA.

DeScioli, Peter, Christner, John and Kurzban, Robert, 2011. "The Omission Strategy." *Psychological Science*, 22(4), 442-446.

Dufwenberg, Martin and Kirchsteiger, Georg, 2004. "A theory of sequential reciprocity." *Games and Economic Behavior*, 47(2), 268-298.

Engel, Christoph and Rockenbach, Bettina, 2011. "We Are Not Alone: The Impact of Externalities on Public Good Provision." *SSRN eLibrary*.

Fehr, Ernst, Kirchsteiger, Georg and Riedl, Arno, 1993. "Does fairness prevent market clearing? An experimental investigation." *Quarterly Journal of Economics*, v108(n2), p437(423).

Ferraro, Paul J., Rondeau, Daniel and Poe, Gregory L., 2003. "Detecting other-regarding behavior with virtual players." *Journal of Economic Behavior & Organization*, 51(1), 99-109.

Fischbacher, Urs, 2007. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments." *Experimental Economics*, 10(2), 171-178.

Gächter, Simon, 2007, Conditional cooperation: Behavioral regularities from the lab and the field and their policy implications. In: Bruno S. Frey and Alois Stutzer (Eds.), *Economics and Psychology. A Promising New Cross-Disciplinary Field*. CESifo Seminar Series. The MIT Press, Boston, MA.

Glöckner, Andreas, Irlenbusch, Bernd, Kube, Sebastian, Nicklisch, Andreas and Normann, Hans-Theo, 2011. "Leading with(out) Sacrifice? A Public-Goods Experiment with a Privileged Player." *Economic Inquiry*, 49(2), 591-597.

Goeree, Jacob, Holt, Charles and Laury, Susan, 2002. "Private Costs and Public Benefits: Unraveling the Effects of Altruism and Noisy Behavior." *Journal of Public Economics*, 83(2), 255-276.

Holt, Charles A. and Laury, Susan K., 2002. "Risk Aversion and Incentive Effects." *American Economic Review*, 92(5), 1644-1655.

Houser, Daniel and Kurzban, Robert, 2002. "Revisiting Kindness and Confusion in Public Goods Experiments." *The American Economic Review*, 92(4), 1062-1069.

Isaac, R. Mark, Norton, Douglas A. and Pevnitskaya, Svetlana, 2011. "Polarized Demands for Public Goods and the Generalized Voluntary Contributions Mechanism." Florida State University

Jacobson, Sarah and Petrie, Ragan, 2012, "Favor Trading in Public Good Provision." Williams College Economics Department Working Papers. Williams College, Williamstown, MA.

Ledyard, John O., 1995, Public goods: A survey of experimental research. In: John H. Kagel and Alvin E. Roth (Eds.), *The Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ, USA.

List, John A., 2006. "The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions." *Journal of Political Economy*, 114(1), 1-37.

Marwell, Gerald and Ames, Ruth E., 1981. "Economists free ride, does anyone else? : Experiments on the provision of public goods, IV." *Journal of Public Economics*, 15(3), 295-310.

Moxnes, Erling and van der Heijden, Eline, 2003. "The Effect of Leadership in a Public Bad Experiment." *The Journal of Conflict Resolution*, 47(6), 773-795.

Olson, Mancur, 1965, *The logic of collective action: public goods and the theory of collective action*. Harvard University Press, Cambridge, Mass.

Rabin, Matthew, 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review*, 83(5), 1281-1302.

Reuben, Ernesto and Riedl, Arno, 2009. "Public Goods Provision and Sanctioning in Privileged Groups." *Journal of Conflict Resolution*, 53(1), 72-93.

Schwartz-Shea, Peregrine, 1983. Normative rhetoric and the definition of cooperation. Dissertation Thesis, University of Oregon, Eugene, OR.

Sell, Jane and Wilson, Rick K., 1991. "Levels of information and contributions to public goods." *Social Forces*, 70(1), 107-124.

Shitovitz, Benyamin and Spiegel, Menahem, 2003. "Cournot-Nash and Lindahl Equilibria in Pure Public 'Bad' Economies." *Economic Theory*, 22(1), 17-31.

Sobel, Joel, 2005. "Interdependent Preferences and Reciprocity." *Journal of Economic Literature*, 43(2), 392-436.

Sonnemans, Joep, Schram, Arthur and Offerman, Theo, 1998. "Public Good Provision and Public Bad Prevention: The Effect of Framing." *Journal of Economic Behavior and Organization*, 34(1), 143-161.

van der Heijden, Eline and Moxnes, Erling, 2003. "Leading by example? Investment decisions in a mixed sequential-simultaneous public bad experiment." Tilburg University, Center for Economic Research

Wilson, Bart J., 2008. "Language games of reciprocity." *Journal of Economic Behavior & Organization*, 68(2), 365-377.