Secret Santa:

Anonymity, Signaling, and Conditional Cooperation

David Hugh-Jones and David Reinstein*

August 14, 2009

Abstract

Costly signaling has been proposed as an explanation for participation in religion and ritual. But if the signal's cost is too small, free-riders will send the signal and behave selfishly later. Effective signaling may then be prohibitively expensive. However, if the average level of signaling in a group is observable, but individual effort is not, then free-riders can behave selfishly without being detected, and group members will learn the aggregate level of selfishness in the group at a low cost. We demonstrate this in a formal model, and give examples of institutions for anonymous signaling, including ritual, religion, music and dance, voting, charitable donations, and military institutions. We explore the value of anonymity in the laboratory with a repeated two-stage public goods game with exclusion. When first-stage contributions are *anonymous,* subjects are better at predicting second-stage behavior, and maintain a higher level of overall contributions.

Keywords: signaling, anonymity, public goods, experiments

JEL codes: H41, Z12, D82

*David Hugh-Jones is a post-doctoral researcher at the Max Planck Institute for Economics, Jena. David Reinstein is a lecturer in the Department of Economics at Essex University.

1 Introduction

In many social groups, members take part in joint activities which appear inefficient. In some of these activities – call them "rituals" – a measure of the group's aggregate effort can be seen by all the participants, and often by outsiders as well, while each individual's effort is unobserved. We propose that these rituals evolve not to signal individuals' types or fitness, nor to screen out "bad" types, but to *measure* the overall level of cooperativeness of the group. We offer three motivational examples.

(1) A group of co-workers wonder if they can strike for higher wages. Will they support each other, or will some blackleg their colleagues? To work out the level of mutual goodwill, they think back to the Christmas presents they received from their workmates. But these may have been given to cultivate a reputation for generosity and curry favor with colleagues. On the other hand, if the workplace has the "secret Santa" institution, in which present-givers are anonymous to the receivers, then generous gifts convey real information about co-workers' character.

(2) Members of a church wish to do business with each other, but can they trust their co-religionists? If contributions to the church are made anonymously, the proportion of their fellow congregants who tithe may be a good indicator. On the other hand, if contributions are public, families may be tithing only to assimilate into the community and perhaps defraud them later.

(3) A charity is raising money for a threshold public good, requiring a minimum commitment to succeed. Hence, early donations may help persuade other potential donors that others are "on board" and will continue to contribute. These "seed money" donations may be better signals of real commitment if they are given anonymously – public gifts may be only given only to win credit and publicity for the giver, and may not indicate real long-term commitment. Anonymous donations bring no reputational benefit, so they are better signs of real commitment and elicit higher gifts from others.

In all of these examples, individuals may perform an anonymous altruistic action, from which participants and others can learn about the overall level of altruism. This knowledge may be useful in a future collective action problem (or "basic game") in which the stakes are higher. Assume some individuals are selfish and prefer to shirk in the basic game, while others are conditional cooperators who prefer to contribute only if others' contributions are sufficiently high.[1] Then, the anonymous ritual may reduce asymmetric information and allow conditional cooperators to contribute in the basic game only when they are sure enough others will do so too. If the ritual were not anonymous, selfish individuals could face large incentives to pool with cooperators and avoid ostracism, and this would make the institution either uninformative or prohibitively costly.

The next section discusses the existing literature on costly signaling, and explains the contribution of this paper. Section 3 gives more in-depth examples of institutions for anonymous signaling. We then develop a formal model. Section 5 presents our laboratory experiment testing the model and its predictions. The conclusion discusses the model's implications. All proofs are in the appendix, unless otherwise mentioned.

2 Literature

Cooperation in social dilemmas is sometimes assured by externally imposed sanctions, or by an equilibrium in a repeated game (Fudenberg and Tirole, 1991). But cooperation is often observed even when these solutions are unfeasible (Ledyard, 1993). A recent set of theories involve heterogeneous preferences: some players are self-interested while others are reciprocators who will cooperate if they expect enough others to do so (Kreps et al., 2001; Schram, 2000; Ostrom, 2000). Experimental economists have found considerable evidence for "conditional cooperation" and heterogeneity (Keser and van Winden, 2000; Simpson and Willer, 2008; Fischbacher et al., 2001),[2]

[1]There might also be altruistic, or "unconditionally cooperative" types. If anonymous signaling reveals enough of these, cooperation may be a dominant strategy for conditional cooperators.

[2]In our model, differences in type need not reflect differences in underlying motivations, but could also come from differences in individuals' material benefits from a public good or differences in their options outside the group.

In this environment, conditionally cooperative players will wish to learn the other players' types in advance of the basic game. This might be achieved through a small-scale game, played before the basic game, in which reciprocators have an incentive to cooperate more than selfish players. Thus the small-scale game may serve an important purpose even if the cooperative behavior is wholly symbolic. For instance, ritual and religion have been explained as "costly signaling" practices that allow individuals to show their commitment to a particular group (Weber, 1946; Ruffle and Sosis, 2003; Sosis and Ruffle, 2003; Bird and Smith, 2005; Levy and Razin, 2006). However, even the self-interested may have strong incentives to "play nice" and pool with reciprocators, because being publicly identified as selfish is often harmful. For example, members may exclude selfish types from the basic game to avoid crowding, may simply enjoy punishing bad people (Ostrom et al., 1992; Fehr and Gachter, 2000), or may avoid selfish types in one-to-one interactions (Coricelli et al., 2004). Thus, cooperation in the small-scale setting must be costly enough to deter pooling. But such a signaling institution will impose a large cost on those who take part.[3]

This leads to a time-inconsistency problem. While group members are better off *ex ante* by committing not to punish, allowing types to be separated at lower social cost, *ex post* they may have an incentive to sanction or exclude selfish individuals. We propose that societies can avoid this problem by committing not to punish via an *anonymizing technology* that reveals the overall level of cooperation, but keeps individual behavior hidden, in the smaller-stakes setting.

Previous work on voluntarily provided public goods has focused on the disadvantages of anonymity: generosity and pro-social behavior increase when reputation is at stake (Harbaugh, 1998; Glazer and Konrad, 1996; Milinski et al., 2002; Cooter and Broughman, 2005; Andreoni and Petrie, 2004).On the other hand, the advantages of anonymity have been discussed in the literature on transparency in principal-agent relationships (Holmstrom, 1999), as well as in a legislative context (Prat 2005; Levy 2007b; 2007a). With incomplete contracts

[3]An extreme example of the cost of ritual behavior is provided by Saint Simeon Stylites, who lived for 37 years on top of a pillar. The practice continued for centuries after his death.

and "career concerns" principals may benefit from committing not to learn too much; this will better align incentives and induce the agents to make more productive choices (Acemoglu, 2007). In our model the benefit of anonymity is more indirect: revelation of types can be achieved at a lower cost, leading in some cases to more efficient decisions by other agents. In contrast to the earlier experimental work, our experimental results show for the first time that anonymity can increase public goods contributions. In the anonymous treatment, where exclusion cannot be targeted, subjects' behavior is more informative to fellow participants; this conforms to our theoretical predictions.

3 Examples of Anonymous Signaling

Many charities identify and thank donors, and several papers suggest that public acknowledgement has a positive effect on fundraising (Harbaugh, 1998; Carman, 2003). Nonetheless, in many settings anonymity is the norm, and only overall totals are made public; this has puzzled several economists (e.g., Andreoni and Petrie, 2004). Although the norm of anonymity may reduce giving by reputation-seekers, it may serve other purposes. A high rate of anonymous giving may help to create mutual trust. Furthermore, early anonymous donations may possess greater signaling value in encouraging others to come forward (see List and Lucking-Reiley, 2002; List and Rondeau, 2003; Karlan and List, 2007; Reinstein and Riener, 2009 for experimental evidence). Notably, church collections are often taken in a pocket that conceals the amount of individual donations; in other communities, anonymous tithing is the rule.

The need to preserve anonymity may favor forms of ritual which are of little practical use. Dance has long been recognized as a way of enhancing group solidarity (McNeill, 1997). Hagen and Bryant (2003) discuss song and music as an example of "coalitional signaling". In certain forms of music and dance, each individual's effort is masked while the average level is obvious. Communal singing can be judged by its overall volume, harmony, and enthusiasm, while it is difficult to discover if any individuals are out of tune or keeping quiet. In communal dances, if one person makes a mistake, the entire group may lose the rhythm, and the guilty party can not always be identified. Similarly, applause, cheers and jeers are all reliable signals of collective appreciation, since contribution levels cannot be distinguished.

Armies need to signal, to themselves and their opponents, that their members are committed to winning at any cost. Military music has a long tradition. When infantry march in step, a single person who loses the rhythm will make the whole unit break step. This is an anonymous, weakest-link signaling technology.[4] Modern armies also use uniforms. Sociologists and social psychologists note that uniforms "deindividuate" their wearers, making them more likely to conform to group norms (Rafaeli and Pratt, 1993; Joseph and Alex, 1972). Watson (1973) found that cultures which used warpaint or other anonymizing techniques were more likely to fight to the death and take no prisoners. Anonymous individuals cannot use the battlefield as a stage to display their courage and build their personal reputation; hence an opponent or army that *seems* aggressive probably *is*.

In a modern context, we can judge the culture of an organization or group by small-scale institutions like the aforementioned Secret Santa, or by observing the size of donations to the honesty box for office coffee, or the quality of the food brought to crowded pot luck parties. When it is time to pay a bill at the restaurant and everyone contributes "what they think they owe" does the sum exceed the bill or does it come up short?

Lastly, in the nineteenth century both Britain and the US adopted the secret ballot. For political parties, this had the disadvantage that voters could no longer be bribed into supporting them. The offsetting advantage was that a party's vote share became a reliable signal of public support for its policies (Smirnov and Fowler, 2007; Stigler, 1972; Londregan and Vindigni, 2006).

Although we do not explicitly model this here, anonymity-preserving institutions seem especially suited to signaling to outsiders. Anonymity involves a loss of information: observers learn only the overall level of contributions and not who contributes. When a symbolic col-

[4]Vegetius ([c. 400] 2004) emphasizes the importance of disciplined marching in his handbook for the training of recruits to the Roman legions: "troops who march in an irregular and disorderly manner are always in great danger of being defeated."

lective display is meant to inform people outside the group, this is no longer a disadvantage. For example, to impress an adversary, a war chant only needs to signal the number of fanatical warriors present – their identities are not important.

4 Model

The society consists of N agents, who each own $1/N$ of a unit of wealth. They may contribute it to a collective good,[5] and individual i's utility is

$$w^{\tau}(x_i, X_{-i}, P) \tag{1}$$

where $x_i \in [0,1]$ is the share of his wealth he contributes, $X_{-i} = \sum_{j \neq i} x_j / N$ is the total wealth contributed by others, $P \in [0,1]$ is the proportion of agents included in the collective good, and $\tau \in \{G, E\}$ is i's type, which may be conditionally cooperative (G or "good" for short) or selfish (E, "bad" or "evil" for short).[6] We focus on the good type's welfare rather than on total welfare, as the former is likely to be more important to explaining the persistence of anonymous institutions, and our focus is positive rather than normative.[7] For convenience, we write $w \equiv w^G$, and $w^{\tau}(x,X) \equiv w^{\tau}(x,X,1)$ for welfare when no agent is excluded.

Our core assumptions are as follows, where $\tau \in \{G,E\}$. We set $w^{\tau}(0,0,P) = 0$ for all P. w^{τ} is differentiable, with $w_2^{\tau}(x,X,P) > 0$[8] and $w_3^{\tau}(x,X,P) < 0$ if $x + X > 0$, so that others' contributions increase welfare, while crowding lowers it so long as there are any contribu-

[5] The good has both private and public elements. For some types the benefit will be a function both of own contribution and aggregate contributions (this resembles Andreoni (1990), although the private benefit need not be interpreted as "warm glow"). We will also introduce crowding and exclusion. Still, the full benefits of the good will not necessarily be internalized.

[6] This functional form includes the case when residual wealth is spent on private consumption. In this case, $w^{\tau}(x_i, X_{-i}, P) = u^{\tau}(x_i, X_{-i}, P) + v^{\tau}(1 - x_i)$ where v^{τ} is utility from private consumption. Since we need to model conditionally cooperative preferences in which own and others' contributions are complements, this more specific functional form would not be more tractable.

[7] Our argument here is informal, but motivated by evolutionary theory, in the vein of Frank (1987). Groups full of bad types are not likely to survive as they will produce little surplus. Hence the groups that survive will need to attract good types, and hence increase good type's welfare. Furthermore, *within* an institution's political process changes that favor bad types are not likely to be proposed, as they reveal the advocate as a bad type, making her vulnerable to exclusion.

[8] w_a represents the partial derivative of $w(x,X,P)$ with respect to the a'th argument. For example $w_2^{\tau} \equiv \partial w_2^{\tau}(x,X,P)/\partial X$. Cross-partials are represented by double subscripts.

tions. Our first key assumption is that $w^G_{12} > 0$, so good types are more willing to contribute if others contribute. By contrast, $w^E_1 < 0$: bad types never contribute. A second key assumption is that $w^G_2(x,X,P) > w^E_2(x,X,P)$ over the whole parameter space: bad types simply care less about the public good.[9] Finally we make some technical assumptions to guarantee a unique interior equilibrium: w is strictly concave, $w_1(0,0,P) > 0$ and $w_1(1,1,P) < 0$ for all P.[10] Also, define $z(x,P)$ implicitly on an appropriate subset of $[0,1]$ by $w_1(x,z(x,P),P) = 0$. Since $w_{12} > 0$, z is increasing in x; we assume $\frac{dz}{dx}(x,P) > 1$.[11] Write $z(x) \equiv z(x,1)$.

Agents' types are private information. A type τ agent's prior belief is that there are g other good types with probability $P^\tau(g)$, $g \in \{0,...,N-1\}$. Sometimes we wish to increase N while holding the structure of the problem constant. We assume then for large N, $P^G \approx P^E$, i.e., one's own type is not very informative about others' types, and that the *proportion* of good types is distributed with continuous cdf $F(\gamma)$, supported on $\gamma \in [0,1]$. If agents' types are independent or nearly so, then $P^G \approx P^E$ will also hold for small N.[12]

Before introducing signaling institutions, we examine equilibria in this "basic game", first when g is unknown, second when it is revealed.

4.1 Equilibrium without knowledge

Evil types never contribute. A good type's best response x is defined by the first order condition $\int w_1(x,X)d\Phi(x) = 0$, where $\Phi(\cdot)$ is the (continuous or discrete) cdf of others' contributions. Strict concavity of w ensures that this is unique; hence good types' play is symmetric in equilibrium. When Φ is single-valued at X, write $b(X) \equiv x$ solving $w_1(x,X) = 0$. Since $w_{12} > 0$, $b(X)$ is strictly increasing in X. By definition, $z(x) = b^{-1}(x)$.

In equilibrium when x^* is contributed by good types, $X = gx^*/N$ for a good type, where

[9] This is necessary for any signaling model in which the signal costs the same to different types, as here. If bad types, who will not contribute, care more about the public good than good types, then the signaling game described below will fail to separate the types.

[10] The condition on the cross-partial then gives us $w_1(0,X,P) > 0$ and $w_1(1,X,P) < 0$ for all X and P.

[11] These conditions could be weakened somewhat without harming our results, so long as we assumed that agents coordinated on the largest possible equilibrium in terms of good type contributions.

[12] If types are drawn independently, then γ will become almost certain as N becomes large. But independence may not hold, for example because people's preferences are affected by those around them, or by collective culture or education.

g is the number of *other* good types. Hence, an equilibrium x^* satisfies

$$\sum_{g=0}^{N-1} P^G(g) w_1(x^*, gx^*/N) = 0. \tag{2}$$

Our assumptions guarantee $x^* \in (0,1)$, but not uniqueness of x^*. However, all our proofs hold for any value of x^*. The expected value to a good type of an interior equilibrium is

$$\sum_{g=0}^{N-1} P^G(g) w(x^*, gx^*/N). \tag{3}$$

4.2 Equilibrium with known types

When g other good types contribute x each, their contributions sum to $X = gx/N$. Hence in equilibrium x satisfies

$$w_1(x, gx/N) = 0. \tag{4}$$

Define x_g as the solution to this. By the Implicit Function Theorem x_g is strictly increasing in g. Let $V(g) = w(x_g, gx_g/N)$ represent a good type's equilibrium value when there are g other good types. The good type's expected value is then

$$\sum_{g=0}^{N-1} P^G(g) V(g). \tag{5}$$

Equilibrium is illustrated in Figure 1. The solid line shows the level curve $z(x)$ satisfying $w_1(x, z(x)) = 0$. For a given value of g, the interior equilibrium exists where the straight line $X = gx/N$ crosses $X = z(x)$. Two examples are shown. The thick part of the line shows the set of these equilibria.

[Figure 1 about here]

Since $w_1(0,0) > 0$ and $w_1(1, X) < 0$ for all X, there can be no boundary equilibria. Since $z(x) - gx$ is increasing in x for all g, by $z'(x) > 1$, the interior equilibrium is unique.

We want to know when (5) is greater than (3), i.e., when it is better for the number of good types to be publicly known. This depends on how $w(x,X)$ changes with X. Since equilibrium conditions can be specified in terms of w_1, two w functions can have the same equilibria but different welfare. This gives a simple sufficient condition for knowledge of types to be beneficial: roughly, whenever the conditional "warm glow" from good types' own contributions matters enough to their welfare.

Lemma 1. *Fix w^τ for $\tau \in \{G,E\}$, and set $w^{i\tau}(x,X) = w^\tau(b(X),X) + i[w^\tau(x,X) - w^\tau(b(X),X)]$ for $i > 0$. Then the $w^{\tau i}$'s satisfy our assumptions, and for i high enough, good type welfare is higher when g is known than when g is unknown.*

In the Lemma, as i increases, the importance of one's own contribution to welfare becomes higher for every fixed value X of others' contributions. The intuition is straightforward: knowing g enables good types to better condition on others' behavior. The downside is that other players may do the same, which may lower total contributions; when own contributions matter a lot, the first effect dominates.

4.3 Signaling institutions

While complete information increases total welfare if Lemma 1 holds, individuals would not freely reveal their types via cheap talk: evil types would claim they were good, in order to increase good types' contributions. Suppose instead that either type can pay a cost $\sigma \geq 0$ before the basic game, and this payment is publicly observed.[13] In a separating equilibrium only good types will pay; otherwise bad types could strictly increase their welfare by not paying, thus increasing both x_g and others' prediction for g.

For good types to pay in such a separating equilibrium, we must have the Incentive Com-

[13]We focus on conditions for a separating equilibrium. Under a full pooling equilibrium, signaling institutions are either irrelevant or simply impose a deadweight cost on all agents.

patibility (IC) condition

$$\sum_{g=0}^{N-1} P^G(g)V(g) - \sigma \quad \geq \quad \sum_{g=0}^{N-1} P^G(g)w(\hat{x_g}, gx_{g-1}/N) \tag{6}$$

where $\hat{x_g} = b(gx_{g-1}/N)$, i.e., $\hat{x}_g$ is a best response when the g other good players think there are only g good players in total due to our agent's non-payment.[14] Rearranging:

$$\sigma \leq \sum_{g=0}^{N-1} P^G(g) \left\{ w(x_g, gx_g/N) - w(\hat{x_g}, gx_{g-1}/N) \right\} \tag{7}$$

Now $w(\hat{x_g}, gx_{g-1}/N) < w(\hat{x_g}, gx_g/N) \leq w(x_g, gx_g/N) \equiv V(g)$, the first inequality by $w_2 > 0$ and x_g increasing in g, the second by optimality of x_g. Therefore, there is a $\sigma \geq 0$ satisfying (7). Intuitively, good types discourage others from contributing if they pretend to be bad, so they will be prepared to pay some amount to send an honest signal. The value to a bad type when there are g good types in total is $w^E(0, gx_{g-1})$, implying the IC condition for separation

$$\sigma \geq \sum_{g=0}^{N-1} P^E(g) \left\{ w^E(0, gx_g/N) - w^E(0, gx_{g-1}/N) \right\}. \tag{8}$$

This is clearly a non-negative lower bound, since x_g is increasing and $w_2^E > 0$. A separating equilibrium is possible if (7) and (8) can hold simultaneously, and the minimum signaling cost will satisfy (8) with equality.

Lemma 2. *If $P^E(g) \approx P^G(g)$, or if $w_2^E(0,X) < \varepsilon$ for all X and low enough ε, then a separating equilibrium is possible. Furthermore if $w_2^E(0,X) < \varepsilon$ for all X and low enough ε, or if N is large, the minimum signaling cost approaches 0.*

[14] Define $x_{-1} = x_0$, since if there are no other good types signaling makes no difference to others' contributions.

The intuition is simple. Either type can increase other players' contributions by paying σ and convincing others she is a a good type. But, since good types care more about the collective good, they will be willing to pay more to increase others' contributions.[15] Furthermore, when N is large, any individual's contribution will make little difference to the total and therefore have little effect on others' behavior. Also, the chance of being the "marginal" type who causes a big jump in equilibrium donations will be small. Hence the cost of signaling need not be high to deter bad types.

When the conditions for $\sigma \to 0$ hold, and complete information is welfare-improving (for instance if Lemma 1 holds), then a signaling institution can be beneficial, since it allows learning at a low cost.

4.4 Exclusion and crowding

Suppose that after the signaling mechanism, but before the basic game, it is possible to exclude some or all players. Here crowding becomes relevant. There are many possible exclusion mechanisms that capture our intuition. For simplicity we assume that the exclusion mechanism maximizes good types' welfare, given the information provided by the signaling institution.[16] Thus, if the institution separates the types, bad types will be excluded, since they would increase crowding but make no contributions in the basic game.[17]

[15] If $P^E(g)$ differs substantially from $P^G(g)$, the conclusion may not hold. For example, let $N = 3$. There is 1 good type with probability 1/2, and 0, 2 or 3 good types with probability 1/6 each. Suppose that if there are at least two good types, all good types wish to contribute a large amount a; otherwise they wish to contribute almost nothing. Now, in a separating equilibrium a good type who pays σ will only be decisive when there is 1 other good type, which occurs with probability $P^G(1) = 1/5$ by Bayes' rule. A bad type will influence contributions with a 3/5 probability, $P^E(1) = 3/5$. Then if the relevant payoffs of the two types, $w(a, a/3)$ and $w^E(0, a/3)$, are not too different, the bad type will be prepared to pay more than the good type, and no separating equilibrium will be possible.

[16] For example, suppose a player is chosen at random to exclude or include all players. All players will include themselves, but when N is large we can ignore this; otherwise, all players will include only good types. Alternatively, inclusion might be decided on a case-by-case basis by a majority vote. In this case, if there is a separating equilibrium, there will be unanimous agreement to include every good type, so long as their expected contributions outweigh crowding, and a $N-1$ against 1 vote to exclude every bad type, since even bad types wish to exclude other bad types.

[17] Some good types might also be excluded, if their contributions would not compensate for the increased

We consider two kinds of signaling institution; the difference is only relevant if exclusion is possible, otherwise both lead to the results described in the previous section. In an *anonymous* signaling institution, only the number of agents who paid σ is visible. In a *public* signaling institution each agent's choice to pay σ or not is visible, and refusal to pay the signaling cost will result in exclusion from the collective good and a payoff of $w^{\tau}(0,0,1/N) = 0$.[18] Since sending a public signal means avoiding exclusion, the minimum signaling cost that meets the bad type's incentive constraint must be higher in the revealed case. While anonymous signaling institutions are less costly, as anonymity makes targeted exclusion impossible, public signaling can weed out bad types and reduce crowding; the optimal choice of institution depends on the tradeoff between these concerns.

Separating equilibrium in a public signaling institution

Again, only good types pay the cost σ. The IC constraint for the good type is

$$\sigma \leq \sum_{g=0}^{N-1} P^G(g) w(\tilde{x}_g, g\tilde{x}_g/N, (g+1)/N) \tag{9}$$

where $\tilde{x}_g$ is the equilibrium contribution level given $g+1$ included good types, defined analogously with x_g by

$$w_1(\tilde{x}_g, g\tilde{x}_g/N, (g+1)/N) = 0, \tag{10}$$

noting that not signaling results in exclusion and welfare of 0. The IC constraint for the bad type is

$$\sigma \geq \sum_{g=0}^{N-1} P^E(g) w^E(0, g\tilde{x}_g/N, (g+1)/N). \tag{11}$$

As before, these constraints are compatible when $P^E \approx P^G$. However, as N increases, the minimum signaling cost must still exceed the cost of being excluded: the above inequal-

crowding. This issue introduces cumbersome technicalities and is not central to our argument, so we ignore it by fiat: assume that inclusion decisions cannot be probabilistic. Then either all those who pay the cost are included, or none are. (Excluding all good types never increases welfare, since welfare is always non-negative in equilibrium.)

[18] We assume that exclusion from the collective good also prevents the agent from making contributions herself.

ity approaches $\sigma \geq \int_0^1 w^E(0, \gamma\bar{x}_\gamma, \gamma)dF(\gamma)$ where $\bar{x}_\gamma \equiv \tilde{x}_{\gamma N}$, implying maximum good type welfare of

$$\int_0^1 w(\bar{x}_\gamma, \gamma\bar{x}_\gamma, \gamma) - w^E(0, \gamma\bar{x}_\gamma, \gamma)dF(\gamma). \tag{12}$$

Good types' net benefit is thus just the difference between good and bad types' gross welfare.

Our main result is that in certain cases, anonymous institutions yield higher good type welfare than public ones.

Proposition 1. *When exclusion is possible, and when $w_3(x, X, P) < \varepsilon$ for all x and X, and ε is close enough to 0, the lowest cost anonymous signaling institution that separates types yields higher welfare for good types than any public signaling institution that separates types. The same holds when $P^E \approx P^G$ and $w^E(0, g\tilde{x}_g/N, (g+1)/N) \approx w(\tilde{x}_g, g\tilde{x}_g/N, (g+1)/N)$.*

This formalizes the tradeoff between the cost of a signal that can separate types and the benefit of reducing crowding. When crowding is not very costly, it is better for signaling to be anonymous. Furthermore, when the value of the collective good is similar for both types, public signaling becomes prohibitively expensive, while anonymous signaling remains fairly cheap.

4.5 A minigame

The real world often presents institutions that are more complex than merely "burning money." It seems natural that the signaling institution might resemble the basic game; as good types benefit more from their own contribution, it becomes cheaper for them to signal, and thus easier to separate the types. These institutions may still be costly, since the supply of collective-action projects is limited, and there will ultimately be diminishing returns. Diverting resources to a "signaling" project may be inefficient. We model this form of institution to derive experimental predictions, and leave welfare analysis for future work.

Let the signaling institution be a "minigame" which takes the same form as the main game, but with payoffs multiplied by a constant $\alpha > 0$. We again compare an anonymous institution to a public one, and we derive conditions on α that allow pooling equilibria (where all contribute the same amount in the minigame) and separating equilibria (where good and bad types contribute different amounts, revealing the number of good types). For the next result, we employ some additional, but fairly natural, conditions on w, stated formally below: crowding reduces the benefit of contributions, evil types always get less marginal benefit from their own contributions, and for bad types, own and other contributions are substitutes rather than complements (as in the standard model of a diminishing returns public good with self-interested individuals).[19]

Lemma 3. *(A) Suppose $w_{13} < 0$, $w_{23} < 0$ and $w^E_{12} < 0$, and that $w^E_1 < w_1$ everywhere. Suppose also $P^G \approx P^E$. For intermediate values of α, a separating equilibrium is possible in an anonymous signaling institution but not in a public signaling institution. (B) If for some α a separating equilibrium is played in an anonymous signaling institution, while a fully pooling equilibrium is played under public signaling, then (1) the correlation between a player's minigame and main game contributions is higher under anonymity than with a public minigame; (2) the probability of exclusion decreases in public minigame contributions, but is always 0 under anonymity; (3) agents' main game contributions are perfectly predictable from their anonymous minigame contributions, while contributions in the public minigame are uninformative; (4) under anonymity agents' posterior beliefs about other agents are certain, and are perfectly correlated with actions in the minigame, while with public signaling posteriors are the same as priors and thus are uncorrelated with minigame play (along the equilibrium path); (5) under anonymity, individual good types' main game contributions are positively correlated to total contributions in the minigame.*

[19] These conditions make it easier to derive our result, but could be relaxed somewhat.

Proof. Part (A) is proved in an online Appendix. Part (B) follows from the definitions of equilibrium behavior. In particular, (4) describes beliefs along the equilibrium path. Off-equilibrium we require that the belief after observing $x \in (0, x^*)$ is low enough to justify exclusion. (5) follows since x_g is increasing in g, and g is given by the number of contributions in the anonymous minigame.[20]

□

Our experimental setup is not identical to our theoretical model, and we do not believe there are only two types of subjects, or that our subjects instantly play signaling equilibria. Thus, in the following section, we use the predictions of Lemma 3 to guide our hypotheses, rather than literally to fit the model.

5 Experiment

5.1 Design and implementation

Rather than trying to "manufacture" the equilibrium described above (e.g., by offering heterogeneous monetary payoffs) we prefer to test our model and its predictions in a more ambiguous, but well-studied environment, the public goods game. Previous experimental work has found that subjects bring different "homegrown values" (Harrison, 2002) to such games; some subjects are reciprocators while others are selfish (Fischbacher et al., 2001). Motivated by these findings, we simultaneously test for conditionally cooperative preferences, measure the nature and extent of these preferences, and examine how subjects infer others' likely play and respond strategically. We set up an environment with a signaling institution of a fixed size, and observe how players learn to make use of it. If they can do so in the during a brief

[20] In the revealed minigame, total contributions only take on one value of x^*, so technically the correlation is undefined. But with an independently drawn behavioral error, contributions would be uncorrelated across the two games.

laboratory experiment, this suggests that they could also do so in the field. Similarly, if signaling institutions without anonymity fail in the way we expect, real-world institutions may also face problems of pandering.

Rather than using a pure "burning money" signaling institution, we use a "minigame" as in Subsection 4.5. We expected this to be more intuitive for subjects, and to be better at separating types (as argued above). Our design is as follows. Thirty subjects enter the session; fifteen are randomly assigned to the *Anonymous*, and fifteen to the *Revealed* treatment. Subjects play 15 repetitions (six repetitions in the pilot version), and always remain within the same anonymity treatment, i.e., this is a between-subjects design. For each repetition, subjects are randomly assigned to groups of five – we use a "stranger matching" design.[21] Each repetition consists of two linear public goods games. Players in each group are randomly numbered from 1 to 5. Players 1-3 (the "leaders") play in the first game; players 4-5 (the "followers")[22] then observe their contributions and may choose to exclude one of the leaders. In the Anonymous treatment, followers learn the Stage 1 contributions, but not which player contributed what, so they cannot target leaders who gave too little for exclusion. In the Revealed treatment, player numbers are given along with their respective donations, so leaders can be targeted. This is the only difference between treatments. In the second public goods game, all players make contribution decisions, but an excluded player's decision is ignored.[23] Figure 2 shows the design in more detail.[24]

[Figure 2 about here]

In repetitions 3, 7, 11 and 15, we also elicited incentivized guesses about other players' Stage 2 contributions,using a quadratic scoring rule. At the end of the game, participants received their payoffs from two randomly chosen repetitions (one for the Stage 1 game, one

[21]In even repetitions, first players 1-2 are selected for each group, from a pool made up of the previous repetition's players 4-5; then players 3-5 are selected for each group from the remaining players. This ensures a reasonable balance of leader/follower roles across subjects.

[22]These terms were not used in the experiment itself.

[23]Exclusion is revealed only at the end of each repetition. If a player is excluded the returns from the public good are calculated based on the contributions of the remaining four players and shared among them. The excluded player simply receives her initial second-stage endowment.

[24]Further details are available on request.

for Stage 2) and for their guesses.

5.2 Predictions

We expect a range of types, rather than the model's two types. Nevertheless, in equilibrium, some or all types may pool in the first round. We predict that our chosen payoff sizes cannot sustain full separation when contributions are revealed, but can separate types to a larger extent when they are anonymous.[25] Thus, we use Lemma 3.B to make the following conjectures.

Notation:

c_j^i : Contribution by player $i \in \{1,2,3,4,5\}$ in Stage j.

$C_1 \equiv c_1^1 + c_1^2 + c_1^3$, $C_2 \equiv \sum_{i=1}^5 c_2^i$, $C_j^{-i} \equiv \sum_{k \neq i} c_j^i$.

$\{C_1\} \equiv \{c_1^1, c_1^2, c_1^3\}$, $\{C_2\} = \{c_2^1, c_2^2, ..., c_2^5\}$.

$\{C_j^{-i}\}$: Vector of contributions in Stage j other than i's ; $\{E^i C_j^{-i}\}$: Vector of i's expectations of the elements of $\{C_j^{-i}\}$.

$e^i \equiv 1$ if player $i \in \{1,2,3\}$ is excluded, $e^i \equiv 0$ otherwise.

e^0 : Indicates that *some* player is excluded.

T : Indicates the Revealed treatment.

Conjecture 1. *The correlation between Stage 1 and Stage 2 contributions will be non-negative, and greater in the Anonymous treatment:* $\rho(c_1^i, c_2^i | T = 0) > \rho(c_1^i, c_2^i | T = 1) \geq 0$. *(Lemma 3.1)*

Conjecture 2. *Players will exclude less, given higher Stage 1 contributions:* $E(e^0 | \{C_1\})$ *is non-increasing in the elements of* $\{C_1\}$*, and is strictly decreasing in* $\min\{C_1\}$. *(Cf. Lemma 3.2)*[26]

[25] This prediction does not come from our theoretical model, but is based on our intuition given our choice of payoff sizes.

[26] In the theoretical model, anonymity leads to no exclusion. Here, exclusion is rational if all leader players are likely to contribute little in Stage 2. This is a consequence of the leaders and followers design, which we chose to avoid contamination of exclusion decisions by reciprocity motives from Stage 1.

Conjecture 3. *Lowering one's contribution will increase the risk of one's own exclusion more in the Revealed than in the Anonymous treatment:* $\partial E(e^i|c_1^i, T=1)/\partial c_1^i < \partial E(e^i|c_1^i, T=0)/\partial c_1^i \leq 0$. *(Cf. Lemma 3.2)*

Conjecture 4. *Players' subjective expectations of Stage 2 contributions will respond more to Stage 1 contributions in the Anonymous treatment than in the Revealed treatment:* $0 \leq \partial(E^j(c_2^i|c_1^i, T=1)/\partial c_1^i < \partial(E^j(c_2^i|c_1^i, T=0)/\partial c_1^i$ *for* $i \in \{1,2,3\}$ *and* $j \in \{1,2,3,4,5\}$. *(Lemma 3.4)*

Conjecture 5. *Expectations of Stage 2 contributions will be more accurate in the Anonymous treatment than in the Revealed treatment:* $\rho(c_2^i, E^j(c_2^i|c_1^i)|T=1) < \rho(c_2^i, E^j(c_2^i|c_1^i)|T=0)$ *for* $i \in \{1,2,3\}$ *and* $j \in \{1,2,3,4,5\}$.

Our expectation that, as in the model, at least some players are reciprocators, leads to:

Conjecture 6. *Players' Stage 2 contributions increase with their expectations of others' Stage 2 contributions:* $E(c_2^i|\{E^i C_2^{-i}\})$ *is nondecreasing in – and increasing over some range of – the elements of* $\{E^i C_2^{-i}\}$.

Combining this with Conjecture 4 yields

Conjecture 7. *Under anonymity, players' Stage 2 contributions increase in Stage 1 contributions:* $E(c_2^i|\{C_1^{-i}\}, T=0)$ *is nondecreasing in – and increasing over some range of – the elements of* $\{C_1^{-i}\}$. *(Cf. Lemma 3.5)*

Finally, we predict that the extra information in the Anonymous treatment will result in higher contributions. This is not inevitable in our model, but since public goods experiments have consistently found that contributions decline over repetitions, we suspect that the equilibrium without a successful signaling institution will have low cooperation levels, so that signaling should be an improvement.[27]

[27] This accords with the interpretation of Holt and Laury (2009), who note that, "if cooperative gain-seekers [Brandt and Schram, 1996] systematically overestimate the number of others of this type in their group" this could explain the decline in cooperation over time in VCM contributions. Lemma 3 gives conditions for an anonymous signaling institution to increase contributions.

Conjecture 8. $E(C_2|T=0) > E(C_2|T=1)$.

5.3 Results

Overview

We present summary statistics of subject's choices and predictions in Table 1. For both treatments, Stage 1 investments are within the range typically found in prior work (Ledyard, 1993), although Stage 2 investments in the Revealed treatment are on the low end of this spectrum. Exclusion was common in both treatments, but subjects in the revealed treatment were significantly more likely to vote to exclude someone. Although predictions for others Stage 2 investments were lower in the revealed case, they still were somewhat overoptimistic. The final column pertains to the number of votes to exclude a particular leader; a leader's overall probability of being excluded was roughly 15%. [28]

As figure 3 demonstrates, Stage 1 investments remained fairly constant across repetitions in both treatments. In contrast, Stage 2 investments began higher and declined much less over the 15 repetitions under the Anonymous treatment, remaining at above 30% on average while falling to around 10% in the Revealed treatment. We argue that the anonymous first stage made this persistent cooperation – which is unusual in the absence of punishment (Ledyard, 1993) – possible. Figure 4 shows that these patterns are similar across sessions. Stage 2 average investment in the final stage is strictly lower in the revealed case for *all* sessions. The difference in average Stage 2 investments between treatments was statistically significant in both parametric and nonparametric tests (Wilcoxon rank-sum $p = 0.07$, taking the treatment/session as the unit of observation). This result supports Conjecture 8.

[Table 1 about here]

[28] In the anonymous case, since the selection of whom to exclude is effectively random, we replace the actual number of votes against a player with one third of the total votes (0,1,or 2) to exclude (in the relevant repetition and group).

[Figures 4 and 5about here]

Figure 5 shows frequencies of Stage 2 contributions by Stage 1 contributions for each treatment (bubble width indicates number of observations), for the later repetitions, presumably after some strategic learning has taken place. Conjecture 1 appears to hold: there is a positive correlation between giving in the stages, and the correlation is much stronger in the anonymous treatment. In particular the Revealed treatment shows many high Stage 1 contributions followed by low Stage 2 contributions, which suggests attempts to avoid exclusion. Next we decompose the variance into its explained and unexplained components, reporting marginal and total sums of squares.[29]

[Table 2 about here]

First stage investment explains much of the variance in Stage 2 investment in the Anonymous treatment, while in the Revealed treatment it explains almost nothing. It is not just the *presence* of an anonymous Stage 1 contribution that matters, but also its magnitude; a 1 ecu investment explains little, while larger investments matter a great deal. Almost all the explanatory power of first stage investment is via individual heterogeneity. In the final two columns, after conditioning on subject-specific effects, first stage investment explains little of the remaining variation for either treatment. However, these subject effects are not observable to the subjects, only to the econometrician; hence the informativeness of first stage investment matters.[30]

Econometric Discussion

While observations at the treatment/session level are strictly independent, they do not fully exploit the information in the data. Our design uses “imperfect” stranger matching; subjects play many repetitions in a limited pool. As in all such experiments the per-subject

[29]In a regression framework, the marginal sum of squares can be interpreted as “the reduction in R-sq if you removed that variable only." These add up to the TSS only if the variables are exactly orthogonal.

[30]Regression analysis of Stage 2 contributions (available by request) yielded comparable results: the coefficient on Stage 1 investment is significant and positive in the anonymous case, but significantly smaller, and sometimes insignificant, in the revealed case. Again, a 1 ecu investment explains little, while the coefficients on the larger anonymous investment dummies are significant, suggesting a nonlinear relationship.

observations are not completely independent, and play may be affected by experience in earlier repetitions. To deal with this, we estimate robust standard errors, clustered at either the subject level or the treatment/session level as appropriate. Where noted, we also use a set of four control variables for subject *i*'s experience: (i) C_1^{-i} , (the average Stage 1 investment of other subjects in *i*'s group) for the previous repetition, (ii) the same for C_2^{-i}, and (iii,iv) the means of C_1^{-i} and C_2^{-i}, respectively, over *all* of *i*'s previous repetitions. [31]

Exclusion

Table 3 gives Probit regressions for a follower's decision whether to exclude any leaders. As Conjecture 2 implies, in both treatments the probability of an exclusion decreases in the minimum Stage 1 contribution. For *early* repetitions, the minimum gift has a negative and sometimes significant impact on the probability of an exclusion for both treatments. However, for the Revealed treatment the effect is significantly greater (again following Conjecture 2) and persists even through the later repetitions.[32]

[Tables 3 and 4 about here]

To test Conjecture 3, we examine the impact of a leader's gift on the probability that she is excluded. Table 4 demonstrates that the probability a leader was excluded varied inversely with her Stage 1 investment, and this effect was much stronger in the revealed treatment (see footnote 28). The overall conditional probabilities of exclusion were 78%, 53%, 12%, 6%, and 4% given revealed Stage 1 contributions of 0,1,2,3, and 4 respectively, with an even steeper slope in later stages (probabilities are imputed as one third the total "simulated" votes against a subject).

Beliefs

[Table 5 about here]

---

[31]This specification was chosen for parsimony and based on some preliminary tests. The simple lag term is motivated by the idea that memory has a recency bias. Regressions allowing an intercept for sessions/treatments yielded similar results.

[32]In some regressions anonymous median gifts have a negative and significant coefficient, while the coefficient on revealed median gifts is positive and significant. We speculate that this is because greater contrast between the contributions makes the exclusion decision harder in the former case and easier in the latter..

Table 5 regresses players' predictions about leaders' Stage 2 contributions on the leaders' actual Stage 1 contributions.[33] The coefficient of first-stage investment is somewhat lower (significantly in columns 1 and 5) in the revealed case, although the summed coefficient remains significant. This supports Conjecture 4. In line with Conjecture 5, subjects' guesses were significantly better in the Anonymous than in the Revealed treatment, with correlations to targets' choices of 0.39 and 0.26 respectively (details by request).

Stage 1's effects on others' Stage 2 investments

By Conjecture 6, we expect a player's Stage 2 investments to increase in her expectation of others' investments. Because the prediction itself may be correlated to subject-specific unobservables (e.g., more generous people may be more optimistic about others) we control for a subject-specific effect.[34]

[Table 6 about here]

The first column of table 6 measures the relationship between a player's Stage 2 investment and her guesses of others' contributions, a direct measure of conditional cooperation. We only include followers, to rule out motives such as direct reciprocity for Stage 1 or bitterness from the perceived probability of being excluded.[35] We allow the slope in the minimum guess to vary by treatment, as minimum givers in the revealed treatment are likely to be excluded. As before, these regressions control for a per-treatment time trend; here we also include a "final repetition" dummy to allow for an end-game effect. There is clear evidence of conditional cooperation: the coefficients on each of the guesses (the lowest, middle, and highest guesses for leader subjects' Stage 2 investments, and the guess for the other follower subject) are positive, and these are jointly significant in an F-test at the 5% level.[36]

[33] In the Anonymous treatment predictions were for (e.g.) "the player who contributed 4 ecus."

[34] We do not include lagged controls for a subject's experience in previous repetitions here as we expect the effect of these to be subsumed in the subjects' expectations; the results are not sensitive to this.

[35] Subjects that were followers in one or fewer "guessing" stages are naturally dropped from this fixed-effects estimation; hence this column has fewer observations than column 4.

[36] This interpretation might be critiqued on the ground that the subject may first choose how much to invest and her prediction may be an ex-post rationalization of this choice (see Fehr and Schmidt 2006). In response we first note that our subjects' guesses are financially motivated. We second point to our evidence below that followers also respond to Stage 1 contributions. Finally, instrumental variables regressions (available by

Confirming Conjecture 7, players' Stage 2 investments increase in others' *anonymous* Stage 1 investments. As noted above, the probability of exclusion is nonlinear in Stage 1 investment; hence subjects' inferences may also be nonlinear. In columns 2 and 3 we use the number of leaders who gave 2-4 and 3-4 ecus in the first stage as independent variables.[37] Column 2 shows that these donations had strong effects. As column 3 shows, these persist into the later stages for the anonymous treatment, but disappear in the revealed case.[38] This suggests that followers grow less confident in revealed first-stage investments as predictors of leaders' second-stage behavior, and thus cease to respond positively.

These regressions are "reduced form": in our theory, first stage contributions only affect Stage 2 contributions via their effect on expectations of Stage 2 contributions. Column 4 tests and fails to reject this interpretation by regressing followers' investments on both their guesses and the *actual* leader investment counts.[39] The coefficients on the guesses themselves remain largely positive and significant in net. However, after controlling for guesses, the coefficients on the leaders' investment variables are smaller and are no longer positive and significant (for either treatment).

Discussion

The results presented are consistent with our conjectures. The anonymous ritual appears to have served as an effective coordination device for conditional cooperators. Overall contribution levels in Stage 2 were significantly and substantially higher in the anonymous treatment. The data also support our account of the mechanism behind this: leaders contributed to avoid exclusion under the revealed treatment; thus contributions were more closely cor-

request) using others' Stage 1 investments as instruments for predicted Stage 2 contributions strongly support our results.

[37]Since the leader who gives the least is likely to be excluded in the revealed treatment, we do not include the lowest gift in this count, hence these variables may take values 0,1, or 2..

[38]The combined coefficients for the revealed case are significantly different from the anonymous case, and not significantly different from 0.

[39]We also ran this same regression controlling for a subject fixed effect, for the reasons previously mentioned. This yielded the same qualitative result but relied on a very small number of observations, since the fixed effect could only be identified for subjects who were followers in multiple prediction repetitions and invested a positive amount in each. Hence this data yielded very little conditional variation in the "Num Ldrs" variables, few degrees of freedom, and very wide standard errors for the corresponding coefficients.

related between the stages under anonymity, and as a result beliefs were more accurate. In later stages followers' gifts continued to vary positively with anonymous leaders' gifts, but this relationship disappeared in the revealed treatment. This suggests an explanation for the relative decline in cooperation in the latter case: good types could no longer confidently identify other good types, and this increasing uncertainty reduced Stage 2 investments. The robust positive relationship between subjects' investments and their predictions of others' investments supports our story of conditional cooperation. This laboratory experiment was a demanding setting for our theory, since it required subjects to understand and respond to the strategic incentives to manipulate contributions over a few repetitions in a brief time span. We are correspondingly more confident that over much longer timescales, players could learn from anonymous institutions.

6 Conclusion

In societies without large-scale markets, and in areas that the market does not reach (e.g., inside the firm itself), others' character and intentions towards us may be vital for our success and survival. In these contexts it is crucial to be able to gauge others' character, so we can choose who to interact with and how much to invest in these interactions. This in turn gives some individuals a strong incentive to conceal their true character. Certain rituals can be seen as institutions which provide a forgiving environment, inducing people to act in accordance with their true character by obscuring their identity behind the screen of the collective. In this paper, we developed a model showing how anonymous rituals can foster greater cooperation than revealed ones, and demonstrated in a laboratory public goods game that anonymous contributions can lead to subsequent higher contributions and thus increase participants' welfare.

Further empirical work is needed to establish whether this model explains the survival of *particular* rituals and institutions. Still, we note that cultural forms such as song seem especially well-suited for preserving the anonymity of participants and shirkers. We hope this approach will be of interest to anthropologists and sociologists of religion.

We also believe there are lessons for policy-makers and managers who wish to predict whether cooperation will be sufficient to undertake an ambitious project, or to reduce agents' uncertainty and thus foster voluntary cooperation. In designing these "trust-measuring" and "trust-building" exercises it may be important not to expose individual performance too much, since this can lead to uninformative pooling or pandering behavior. Incentive-compatible mechanisms which allow for collective achievement may work better. Indeed, a large informal literature on team-building emphasizes just this kind of group work (e.g. Newstrom and Scannell 1998). For example, potential borrower groups from micro-credit organizations might be asked to play such anonymous contribution games (or undertake small projects) before receiving the major loan; this might help lenders and aid agencies learn which groups are likely to be successful, and build trust among participants.

References

Acemoglu, D. (2007, December). Incentives in markets, firms and governments. *Journal of Law, Economics and Organization*, 1–34.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal 100*, 464–477.

Andreoni, J. and R. Petrie (2004). Public goods experiments without confidentiality: a glimpse into fund-raising. *Journal of Public Economics 88*(7-8), 1605–1623.

Bird, R. B. and E. A. Smith (2005, April). Signaling theory, strategic interaction, and symbolic capital. *Current Anthropology 46*(2), 221–248.

Carman, K. (2003). Social influences and the private provision of public goods: Evidence from charitable contributions in the workplace. *Manuscript, Stanford University*.

Cooter, R. and B. Broughman (2005). Charity, Publicity, and the Donation Registry. *The Economists' Voice 2*(3), 4.

Coricelli, G., D. Fehr, and G. Fellner (2004, June). Partner selection in public goods experiments. *Journal of Conflict Resolution 48*(3), 356–378.

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

Fehr, E. and K. Schmidt (2006). The economics of fairness, reciprocity and altruism–experimental evidence and new theories. *Handbook of the Economics of Giving, Altruism and Reciprocity 1*, 615–694.

Fischbacher, U., S. Gächter, and E. Fehr (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters 71*(3), 397–404.

Frank, R. (1987). If homo economicus could choose his own utility function, would he want one with a conscience? *The American Economic Review*, 593–604.

Fudenberg, D. and J. Tirole (1991). *Game Theory*. Mit Press.

Glazer, A. and K. Konrad (1996). A signaling explanation for charity. *American Economic Review 86(4)*, 1019–1028.

Hagen, E. H. and G. A. Bryant (2003). Music and Dance as a Coalition Signaling System. *Human Nature 14*(1), 21–51.

Harbaugh, W. T. (1998, May). The prestige motive for making charitable transfers. *The American Economic Review 88*(2), 277–282.

Harrison, G. (2002). Introduction to experimental economics. *At http://dmsweb. badm. sc. edu/glenn/manila/presentations*.

Holmstrom, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies 66*(1), 169–182.

Holt, C. and S. Laury (2009). Forthcoming. Theoretical explanations of treatment effects in voluntary contributions experiments. C. Plott, V. Smith, eds. *Handbook of Experimental Economic Results*.

Joseph, N. and N. Alex (1972). The uniform: A sociological perspective. *American Journal of Sociology 77*(4), 719.

Karlan, D. and J. List (2007). Does Price Matter in Charitable Giving? Evidence from a Large-Scale Natural Field Experiment. *American Economic Review 97*(5), 1774–1793.

Keser, C. and F. van Winden (2000). Conditional Cooperation and Voluntary Contributions to Public Goods. *Scandinavian Journal of Economics 102*(1), 23–39.

Kreps, D. M., P. Milgrom, J. Roberts, and R. Wilson (2001). Rational cooperation in the finitely repeated prisoners' dilemma. *Readings in Games and Information*.

Ledyard, J. (1993). *Public Goods: A Survey of Experimental Research*. Division of the Humanities and Social Sciences, California Institute of Technology.

Levy, G. (2007a). Decision making in committees: Transparency, reputation, and voting rules. *The American Economic Review 97*(1), 150–168.

Levy, G. (2007b). Decision-Making procedures for committees of careerist experts. *American Economic Review 97*(2), 306–310.

Levy, G. and R. Razin (2006). A Theory of Religion: Linking Individual Beliefs, Rituals, and Social Cohesion. *Department of Economics WP, LSE*.

List, J. and D. Lucking-Reiley (2002). The effects of seed money and refunds on charitable giving: Experimental evidence from a university capital campaign. *Journal of Political Economy 110*(1), 215–233.

List, J. and D. Rondeau (2003). The impact of challenge gifts on charitable giving: an experimental investigation. *Economics Letters 79*(2), 153–159.

Londregan, J. and A. Vindigni (2006). Voting as a credible threat.

McNeill, W. H. (1997). *Keeping Together in Time*.

Milinski, M., D. Semmann, and H. Krambeck (2002). Reputation helps solve the'tragedy of the commons'. *Nature 415*(6870), 424–6.

Newstrom, J. W. and E. E. Scannell (1998). *The Big Book of Team Building Games: Trust-Building Activities, Team Spirit Exercises, and Other Fun Things to Do*. McGraw-Hill Professional.

Ostrom, E. (2000). Collective action and the evolution of social norms. *The Journal of Economic Perspectives*, 137–158.

Ostrom, E., J. Walker, and R. Gardner (1992). Covenants with and without a sword: Self-Governance is possible. *The American Political Science Review 86*(2), 404–417.

Prat, A. (2005). The wrong kind of transparency. *The American Economic Review 95*(3), 862–877.

Rafaeli, A. and M. G. Pratt (1993). Tailored meanings: On the meaning and impact of organizational dress. *The Academy of Management Review 18*(1), 32–55.

Reinstein, D. and G. Riener (2009). Reputation and influence in charitable giving: An experiment.

Ruffle, B. J. and R. H. Sosis (2003). *Does it Pay to Pray? Evaluating the Economic Return to Religious Ritual*. SSRN. SSRN eLibrary.

Schram, A. (2000). Sorting out the seeking: The economics of individual motivations. *Public Choice 103*(3), 231–258.

Simpson, B. and R. Willer (2008). Altruism and Indirect Reciprocity: The Interaction of Person and Situation in Prosocial Behavior. *Social Psychology Quarterly 71*(1), 37–52.

Smirnov, O. and J. H. Fowler (2007). Policy-Motivated parties in dynamic political competition. *Journal of Theoretical Politics 19*(1), 9.

Sosis, R. and B. J. Ruffle (2003). Religious ritual and cooperation: Testing for a relationship on israeli religious and secular kibbutzim. *Current Anthropology 44*(5), 713–722.

Stigler, G. J. (1972). Economic competition and political competition. *Public Choice 13*(1), 91–106.

Vegetius, M. D. (2004). *Epitoma rei militaris*. Clarendon Press New York; Tokyo: Oxford University Press, Oxford.

Watson, R. I. (1973, March). Investigation into deindividuation using a cross-cultural survey technique. *Journal of Personality and Social Psychology 25*(3), 342–5. PMID: 4705668.

Weber, M. (1946). *The Protestant Sects and the Spirit of Capitalism*. New York: Oxford.

Appendix

Proof of Lemma 1

Proof. Set $w^i(x,X) = w(b(X),X) + i[w(x,X) - w(b(X),X)]$ for $i > 0$. It is easy to show that $w^i(x,z(x)) = w(x,z(x))$, and that $w^i_1(x,X) = iw_1(x,X)$. Therefore, the sets of solutions to (2) and (4), i.e. of equilibria with and without knowledge, do not vary with i, and expected welfare when g is known is also unchanged. That the other Assumptions for w continue to hold for w^i is also straightforward. (To show $w^i_2 > 0$, differentiate w^i to get $w^i_2(x,X) = iw_2(x,X) + (1-i)w_2(b(X),X) + (1-i)w_1(b(X),X)b'(X)$, and observe that the last term is zero by definition of z.)

We wish to show that

$$\sum_{g=0}^{N-1} P^G(g)[w^i(x_g, gx_g/N) - w^i(x^*, gx^*/N)] > 0 \tag{13}$$

for high enough i. First observe that if $x_g \geq x^*$, then $w^i(x_g, gx_g/N) - w^i(x^*, gx^*/N) \geq 0$. For

$$\begin{aligned}
& w^i(x_g, gx_g/N) - w^i(x^*, gx^*/N) \\
= \; & w^i(x_g, gx_g/N) - w^i(b(gx^*/N), gx^*/N) + w^i(b(gx^*/N), gx^*/N) - w^i(x^*, gx^*/N) \\
= \; & \int_{b(gx^*/N)}^{x_g} \left\{ \frac{d}{dx} w^i(x, z(x)) \right\} dx + \left[w^i(b(gx^*/N), gx^*/N) - w^i(x^*, gx^*/N) \right]
\end{aligned}$$

where the integral is taken along $z(x)$ (since $z(b(gx^*/N)) = gx^*/N$). Now the second term in brackets is non-negative by $b(\cdot)$ a best response (and positive whenever $x^* \neq x_g$. On the other hand, since $w_1^i(x, z(x)) = 0$, $\frac{d}{dx} w^i(x, z(x)) = w_2^i(x, z(x))$ and this is positive, so the integral is non-negative. Thus the whole term is non-negative. (This whole step holds for any w^i including $w^1 \equiv w$.)

Next for $x_g < x^*$ we have

$$\begin{aligned}
w^i(x_g, gx_g/N) - w^i(x^*, gx^*/N) \\
= w(x_g, gx_g/N) - w^i(x^*, gx^*/N) \\
(\text{since } x_g = b(gx_g/N)) \\
= w(x_g, gx_g/N) - w(b(gx^*/N), gx^*/N) + i\,[w(b(gx^*/N), gx^*/N) - w(x^*, gx^*/N)]
\end{aligned}$$

and as $i \to \infty$ the sign of this depends on the sign of the term in square brackets, which is positive since $b(gx^*/N)$ is the unique best response to gx^*/N. Thus for i high enough the terms of (13) are always non-negative, and are positive whenever $x_g \neq x^*$.

□

Proof of Lemma 2

Proof. If $w_2^E(0,X) < \varepsilon$ for all X, then $w^E(0, gx_{g+1}/N) - w^E(0, gx_g/N) < \frac{g(x_{g+1}-x_g)}{N}\varepsilon$ and for ε low enough, the right hand side of (8) approaches 0, meaning the lowest signaling costσ approaches 0, and ensuring that (8) and (7) are compatible.

If $P^E(g) \approx P^G(g)$ then the right hand side of (8) is approximately

$$\sum_{g=0}^{N-1} P^G(g) \left\{ w^E(0, gx_g/N) - w^E(0, gx_{g-1}/N) \right\}. \tag{14}$$

Comparing this with the right hand side of (7),

$$\begin{aligned}
\sum_{g=0}^{N-1} P^G(g) \left\{ w(x_g, gx_g/N) - w(\hat{x_g}, gx_{g-1}/N) \right\} &\geq \sum_{g=0}^{N-1} P^G(g) \left\{ w(\hat{x_g}, gx_g/N) - w(\hat{x_g}, gx_{g-1}/N) \right\} \\
&\geq \sum_{g=0}^{N-1} P^G(g) \left\{ w(0, gx_g/N) - w(0, gx_{g-1}/N) \right\},
\end{aligned}$$

the first inequality by optimality of x_g, and the second by $w_{12} > 0$. But the last expression is at least as great as (14) by our assumption that $w_2^E < w_2$. Thus (8) and (7) can hold simultaneously.

Lastly, as N grows large, by assumption P^G approaches P^E. Then the RHS of (8), which defines the lowest possible signaling cost σ, approaches

$$\int_0^1 w^E(0, \gamma\dot{x}_\gamma) - w^E(0, \gamma\dot{x}_{\gamma-1/N})\, dF(\gamma)$$

where we have written $\dot{x}_\gamma = x_{\gamma N}$ for equilibrium good type contributions when a proportion γ of agents are good. x_g is continuous by standard arguments, and so therefore is $\dot{x}_\gamma$. Thus $w^E(0, \gamma\dot{x}_\gamma) \to w^E(0, \gamma\dot{x}_{\gamma-1/N})$ as $N \to \infty$, so that the expression above goes to 0.

□

Proof of Proposition 1

Proof. We take each case in turn.

1. Exclusion may still occur even with an anonymous signaling institution, since, after the number of good types is revealed, it might maximize social welfare to exclude a fraction of the players. However, if w_3 is close enough to 0, this will not be the case. For then we have for all $g \in \{0,...,N-1\}$, and $k \in \{0,...,N\}$:

$$w(x_g, gx_g/N, 1) \geq \frac{k}{N} \sum_{j=0}^{g} B_k(j) w(y_{k,g}, jy_{k,g}/N, k/N)$$

where k is the number of included players, k/N is the ex ante probability of being included, $B_k(j)$ is the probability of having j other good types from a total of $k-1$ other players who each have probability $g/(N-1)$ of being a good type, and $y_{k,g}$ is the largest equilibrium contribution under these circumstances. Proof: by assumption, $w(x,X,P) \approx w(x,X,1)$ for all x, X. Then $y_{k,g}$ is a maximizer of w given the distribution of other included good types. Clearly this is bounded above by g; then by $w_{12} > 0$ we have $y_{k,g} \leq x_g$. But then for all $j \in \{0,...,g\}$, $w(y_{k,g}, jy_{k,g}/N, k/N) \approx w(y_{k,g}, jy_{k,g}/N, 1) \leq w(y_{k,g}, gx_g/N, 1) \leq w(x_g, gx_g/N, 1)$, the first inequality since $w_2 > 0$, the second by optimality of x_g. Thus it will maximize welfare to exclude no one. In this case, then, anonymous signaling gives welfare of

$$\sum_{g=0}^{N-1} P^G(g) w(x_g, gx_g/N, 1) - \sigma, \tag{15}$$

with x_g defined as before. Once again the signaling cost σ satisfies (8).

On the other hand, public signaling gives welfare of

$$\sum_{g=0}^{N-1} P^G(g) w(\tilde{x}_g, g\tilde{x}_g/N, (g+1)/N) - \sigma_X. \tag{16}$$

Here σ_X is the signaling cost when exclusion is possible, and $\tilde{x}_g$ is defined as in (10). Now when w_3 is small, $w(x,X,(g+1)/N) \approx w(x,X,1)$ for all x, X and g; then by continuity of w and the Maximum Theorem, $\tilde{x}_g \approx x_g$.[40] The first terms of (15) and (16) thus are approximately the same value: the benefit from exclusion is negligible. The proof then follows from comparing the lower bounds on the signaling costs:

$$\begin{aligned}\sigma_X &\geq \sum_{g=0}^{N-1} P^E(g) w^E(0, g\tilde{x}_g/N, (g+1)/N) \\ \sigma &\geq \sum_{g=0}^{N-1} P^E(g) \left\{ w^E(0, gx_{g+1}/N, 1) - w^E(0, gx_g/N, 1) \right\}\end{aligned}$$

from (11) and (8); and since $w^E(0, g\tilde{x}_g/N, (g+1)/N) \geq w^E(0, g\tilde{x}_g/N, 1)$ and $x_g \approx \tilde{x}_g$ it is easy to see that the lowest σ is less than the lowest possible σ_X.

2. Suppose $P^E \approx P^G$ and

$$w^E(0, g\tilde{x}_g/N, (g+1)/N) \approx w(\tilde{x}_g, g\tilde{x}_g/N, (g+1)/N) \tag{17}$$

for all g. [41] Then, substituting into (11), the lowest signaling cost in a public institution is close to

$$\sum_{g=0}^{N-1} P^G(g) w(\tilde{x}_g, g\tilde{x}_g/N, (g+1)/N)$$

and this is exactly the expression for good type welfare in the main game. Thus welfare in the lowest cost public signaling institution is approximately 0. On the other hand, the lowest signaling cost in an anonymous institution is again

$$\sum_{g=0}^{N-1} P^G(g) w(x_g, gx_g/N, 1) - \sum_{g=0}^{N-1} P^G(g) \left\{ w^E(0, gx_g/N, 1) - w^E(0, gx_{g-1}/N, 1) \right\}$$

[40] More technically: fix w and take a series of functions w^i with $w^i(x,X,1) = w(x,X,1)$, $\forall x,X$, and with $w^i_3(\cdot,\cdot,\cdot)$ approaching 0 uniformly. Then the proof holds when i is high enough. Fixing w is required since, if w varied very little with x and X, even small differences between $w(x,X,(g+1)/N)$ and $w(x,X,1)$ might cause large differences between x_g and $\tilde{x}_g$.

[41] The latter condition requires that $w(x,X,P)$ does not vary very much in x, since $w^E(0, g\tilde{x}_g/N, (g+1)/N) < w(0, g\tilde{x}_g/N, (g+1)/N)$ by $w^E_2 < w^G_2$. Technically, we are taking two series of functions $\{w^{Ei}\}_{i=1}^{\infty}$ and $\{w^i\}_{i=1}^{\infty}$, so that $w^{Ei}(0, g\tilde{x}_g/N, (g+1)/N) \nearrow w^i(0, g\tilde{x}_g/N, (g+1)/N)$ for all g, and $w^i(0, g\tilde{x}_g/N, (g+1)/N) \nearrow w^i(\tilde{x}_g, g\tilde{x}_g/N, (g+1)/N)$ for all g.

where the second term comes from substituting P^G for P^E in(8). Rewriting this as

$$\sum_{g=0}^{N-1} P^G(g)\left\{w(x_g, gx_g/N, 1) - w^E(0, gx_g/N, 1)\right\} + \sum_{g=0}^{N-1} P^G(g) w^E(0, gx_{g-1}/N, 1),$$

the first term is positive since $w(x_g, gx_g/N, 1) > w(0, gx_g/N, 1) \geq w^E(0, gx_g/N, 1)$ by optimality of x_g and $w_2 > w_2^E$; hence the whole is positive. Thus the anonymous signaling institution gives strictly greater welfare than 0 which is the maximum for any public signaling institution, as was required.

$\square$

Tables and Figures

Table 1: Summary statistics

	Stage 1 investment	Stage 2 invt.	Voted to exclude [a]	Min guess (Target: leader) [a]	Med. guess (Tgt: ldr) [a]	Max. guess (Tgt: ldr) [a]	Avg. guess (Tgt: follower) [a]	Votes against [b],[c]
Anonymous Treatment								
Mean	2.14	4.34	.377	2.64	4.36	5.79	4.32	.251
SD	1.21	3.24	.485	2.1	2.26	2.59	2.37	.222
P25	1	1	0	1	2.5	4	2	0
Median	2	5	0	2	4	6	5	.333
P75	3	7	1	4	6	8	6	.333
Min	0	0	0	0	0	0	0	0
Max	4	10	1	8	9	10	10	.667
Obs.	585	975	390	108	108	108	270	585
Revealed Treatment								
Mean	2.42	2.89	.495	2.21	3.49	4.67	3.46	.33
SD	1.02	2.7	.501	1.95	2.09	2.27	2.32	.61
P25	2	0	0	0	2	3	2	0
Median	2	2	0	2	3	5	3.5	0
P75	3	5	1	4	5	6	5	1
Min	0	0	0	0	0	0	0	0
Max	4	10	1	8	8	10	9	2
Obs.	585	975	390	108	108	108	270	585
Overall								
Mean	2.28	3.61	.436	2.43	3.93	5.23	3.89	.291
SD	1.13	3.07	.496	2.04	2.22	2.49	2.38	.46
P25	2	1	0	1	2	3.5	2	0
Median	2	3	0	2	4	5	4	0
P75	3	6	1	4	5	7	5.5	.333
Min	0	0	0	0	0	0	0	0
Max	4	10	1	8	9	10	10	2
Obs.	1170	1950	780	216	216	216	540	1170

[a] Observations for followers only, [b] ... for leaders only

[c] Anonymous case: set to $\frac{1}{3}$ the total votes (0,1,or 2) to exclude (for group/rep).

Table 2: Analysis of variance of Stage 2 contributions by Stage 1 contributions

Partial (marginal) sums of squares:						
	Anon.	Revealed	Anon. later	Rvld. later	Anon.	Rvld.
1 ecu invt.	14	14	35	7	16	21
2 ecu invt.	211	20	199	9.1	37	24
3 ecu invt.	605	42	428	15	36	28
4 ecu invt.	1008	88	497	19	75	44
Subject effects					2038	1560
Model Degrees of freedom	4	4	4	4	78	78
Observations	585	585	288	288	585	585
Model SS	1937	133	849	27	3975	1694
Total SS	6172	3997	2846	1519	6172	3997
R-sq.	.31	.033	.3	.018	.64	.42

’Later’ refers to stages 8-15

Table 3: Probit regressions: decision to exclude someone

Dependent variable = Dummy: subject chooses to exclude someone.								
	(1)		(2)		(3)		(4)	
	All Repetitions		All Repetitions		Repetitions 8-15		Repetitions 8-15	
Minimum St. 1. Invt.	-0.075	(0.049)	-0.13*	(0.056)	-0.012	(0.065)	-0.10	(0.067)
Rvld $\times$ Min St. 1 Invt	-0.17*	(0.081)	-0.18*	(0.088)	-0.20+	(0.11)	-0.15	(0.11)
Median St. 1. Invt.	-0.011	(0.039)	-0.079+	(0.041)	-0.058	(0.053)	-0.11*	(0.054)
Rvld $\times$ Med. St. 1 Invt.	0.16*	(0.067)	0.19*	(0.080)	0.22*	(0.11)	0.21+	(0.12)
Range St 1. Invts.	-0.011	(0.041)	-0.040	(0.044)	0.028	(0.053)	-0.0027	(0.054)
Rvld $\times$ Range St. 1 Invts.	0.062	(0.062)	0.055	(0.071)	0.13	(0.092)	0.13	(0.093)
Revealed Treatment	-0.054	(0.17)	-0.018	(0.19)	-0.24	(0.22)	-0.072	(0.24)
History & Lag Var's	No		Yes		No		Yes	
Observations	780		660		384		384	

+ p<0.10, * p<0.05, ** p<0.01

Marginal effects at means reported. Std. errors (clustered by subject) in parentheses.

Table 4: Poisson regressions: exclusion votes against a player

Dependent variable = Number of votes [*] to exclude subject (in single repetition).						
	(1)		(2)		(3)	
	All Repetitions		Repetitions 8-15		All Repetitions	
St. 1 Investment	-0.13**	(0.049)	-0.17*	(0.071)		
Rvld × Invt.	-0.80**	(0.13)	-0.72**	(0.20)		
Repetition	0.015	(0.012)	-0.0047	(0.025)	0.014	(0.012)
Rvld × Reptn.	-0.012	(0.020)	0.035	(0.036)	-0.010	(0.022)
Invested 1 ecu					-0.24*	(0.12)
Rvld × Invt. 1 ecu					-0.29	(0.18)
Invested 2 ecu's					-0.32+	(0.17)
Rvld × Invt. 2 ecu's					-1.45**	(0.44)
Invested 3 ecu's					-0.56**	(0.20)
Rvld × Invt. 3 ecu's					-2.28**	(0.57)
Invested 4 ecu's					-0.52**	(0.17)
Rvld × Invt. 4 ecu's					-2.73**	(0.24)
Constant	-1.09**	(0.12)	-0.64+	(0.39)	-1.01**	(0.15)
Session/Trtmt. Dummies	Yes		Yes		Yes	
Observations	1170		576		1170	

+ p<0.10, * p<0.05, ** p<0.01

[*] Conditional expectation simulated for anonymous case; see footnote.

Poisson coef's: marginal effects. Robust (clustered by session/trtmt) SE's in parens.

Table 5: Poisson regressions: predictions for leaders

Dependent variable = Prediction of target's stage 2 investment						
	(1)	(2)	(3)	(4)	(5)	(6)
	All reps, Poisson	...	Reps 11,15	...	Rep 15	...
Target St. 1 Inv.	0.34**	0.25**	0.36**	0.24**	0.37**	0.27**
	(13.61)	(9.34)	(10.61)	(5.74)	(8.63)	(5.65)
Tgt. St. 1 Inv × Rvld.	-0.069+	-0.038	-0.091	-0.017	-0.22*	-0.15
	(-1.69)	(-0.97)	(-1.21)	(-0.21)	(-2.35)	(-1.55)
Repetition	-0.0077	0.0071				
	(-1.26)	(1.07)				
Rvld. × Repetition	-0.050**	-0.041**				
	(-4.65)	(-3.76)				
Dummy: revealed ritual	0.34*	0.33*	-0.25	-0.14	-0.044	0.0080
	(2.42)	(2.39)	(-1.12)	(-0.61)	(-0.17)	(0.03)
Constant	0.70**	0.19	0.55**	0.14	0.50**	0.079
	(7.87)	(1.58)	(5.63)	(1.23)	(4.21)	(0.48)
History & Lag 1 Var's	No	Yes	No	Yes	No	Yes
Observations	1152	1152	576	576	288	288
Sum: invt. & rvld.	.27**	.21**	.27**	.20**	.15+	.13+

+ p<0.10, * p<0.05, ** p<0.01

Session 1 excluded due to error in prediction instructions.

Poisson coef's: marginal effects. Robust s.e. (clustered by subject) in parens.

In anonymous treatments predictions were for (e.g.,) "the guy who contributed 4 ecus."

Table 6: Determinants of followers' Stage 2 investment: Measuring Conditional Cooperation

	(1)	(2)	(3)	(4)
	All reps	All reps	Reps 8-15	All reps
	(Poisson regressions of follower subjects' investments)			
Min guess (target: leader)	0.11*			0.054
	(0.056)			(0.052)
... × Rvld. Trtmt.	0.065			0.080
	(0.087)			(0.062)
Med. guess (target: leader)	0.021			-0.029
	(0.068)			(0.060)
Max. guess (target: leader)	0.084			0.155**
	(0.061)			(0.044)
Guess (target: follower)	0.053			0.075**
	(0.038)			(0.022)
Num. Ldrs. Inv. 2+ [a]		0.205*	0.257+	0.104
		(0.092)	(0.139)	(0.180)
... × Rvld. Trtmt.		-0.081	-0.736*	-0.187
		(0.161)	(0.333)	(0.673)
Num. Ldrs. Inv. 3+ [b]		0.125+	0.149	-0.187+
		(0.066)	(0.103)	(0.112)
... × Rvld. Trtmt.		0.087	0.020	0.153
		(0.105)	(0.167)	(0.162)
Repetition	0.026	-0.035**	-0.036	-0.006
	(0.024)	(0.010)	(0.027)	(0.024)
Rvld × Reptn.	-0.030	-0.046**	-0.115**	0.013
	(0.029)	(0.015)	(0.038)	(0.025)
Final Rep.	-0.255	-0.115	-0.023	-0.157
	(0.210)	(0.170)	(0.171)	(0.246)
Dummy: Rvld. Trtmt.		-0.037	2.141**	-0.347
		(0.336)	(0.765)	(1.303)
Constant		1.184**	1.079*	0.119
		(0.196)	(0.492)	(0.392)
Subject-Fixed Effects	Yes	No	No	No
Observations	125	780	384	192
P-val: F test, guesses	0.020			0.000
Sum coef: Num. 3+ Ldrs., Anon		0.331**	0.407*	-0.083
Sum coef: Num. 3+ Ldrs. Rvld		0.338*	-0.309	-0.116

+ p<0.10, * p<0.05, ** p<0.01

Standard errors in parentheses

"Target" indicates the target of subject's prediction of round 2 gift for that repetition (3,7,11, or 13).

Pilot session dropped from regressions with subjects' prediction variables.

[a] Count of leaders (in group/rep.) who invested 2 or more in st. 1; excludes min. invt.

[b] ... 3 or more

Figure 1: Equilibria when g is known

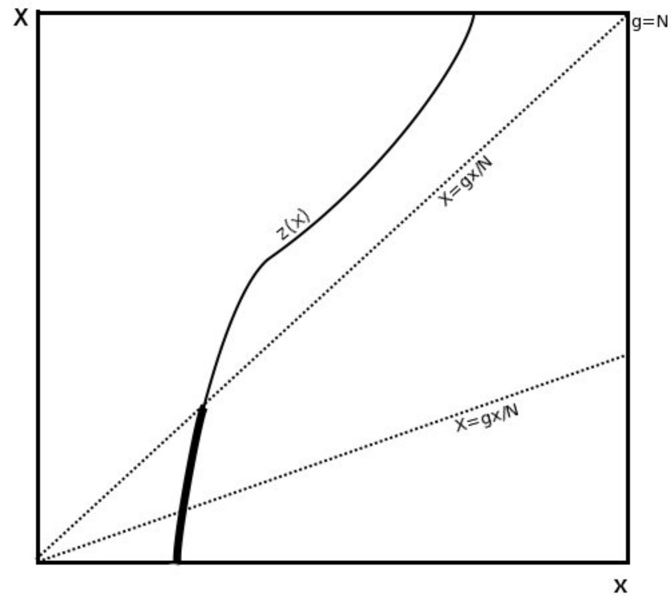

Figure 2: Experimental design

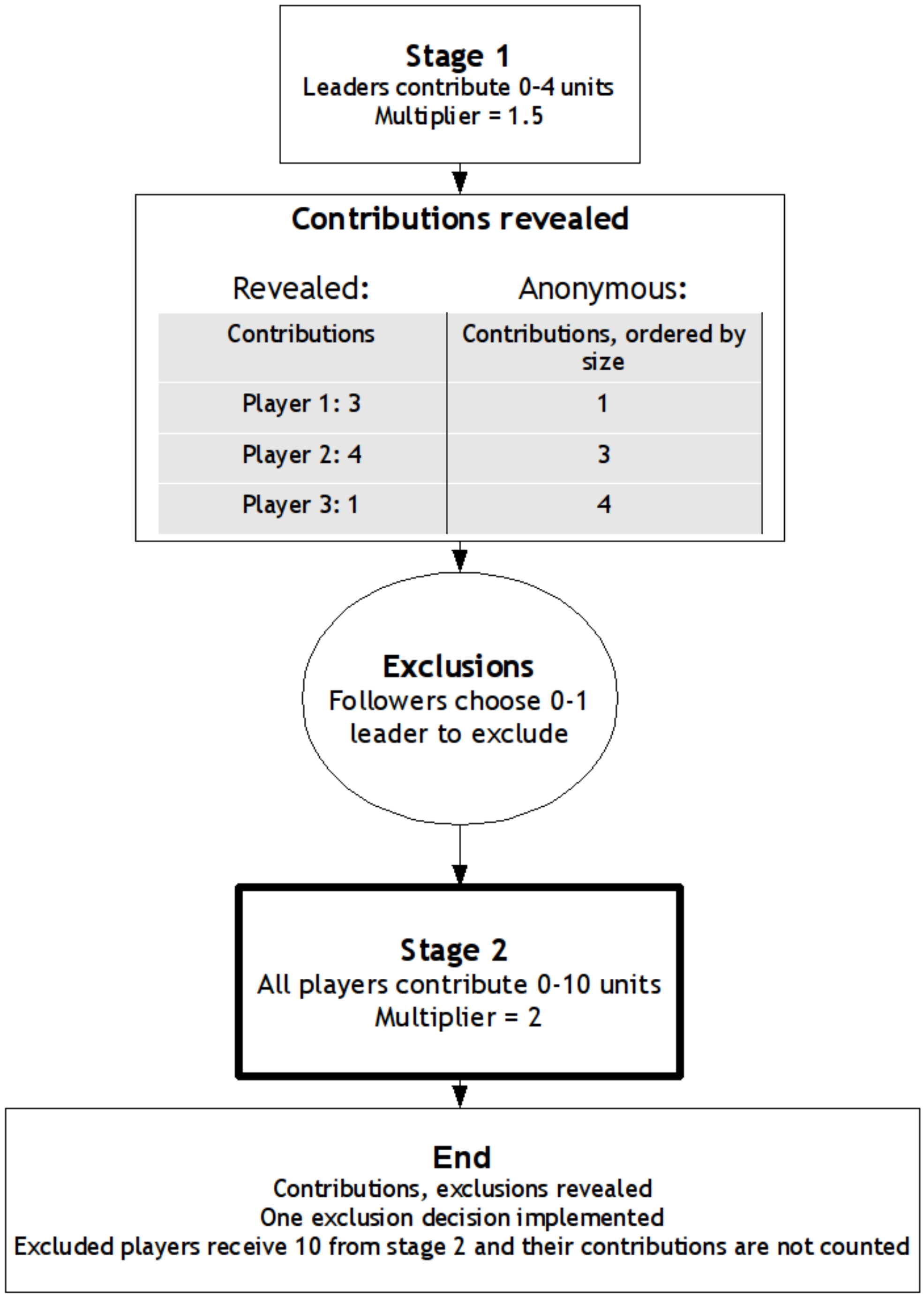

Figure 3: Mean Contributions by Repetition by Treatment

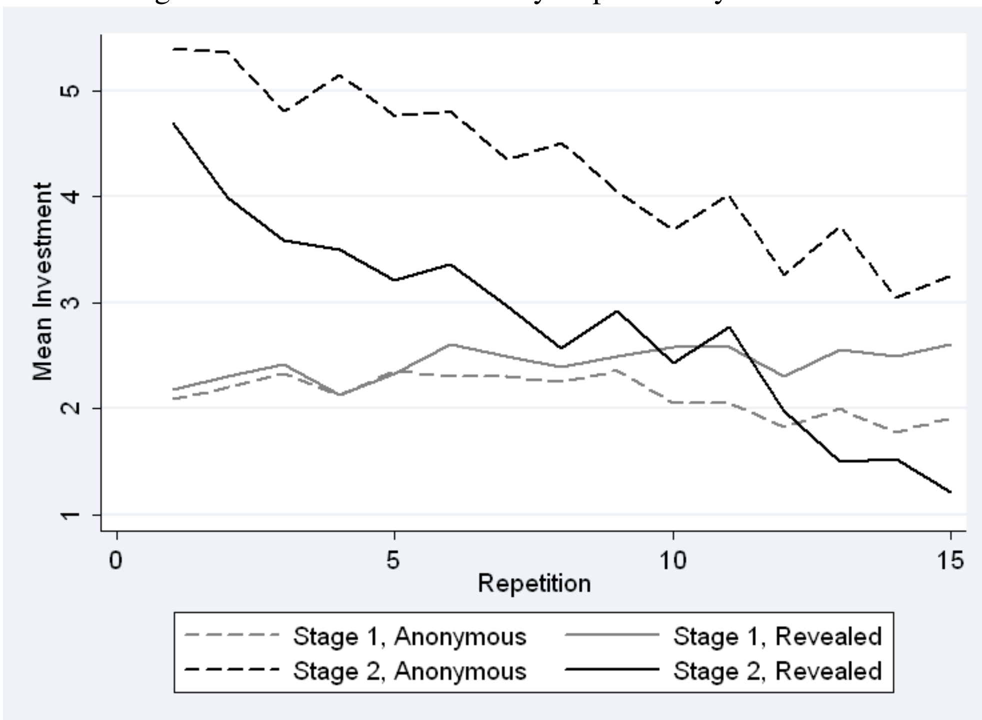

Figure 4: Mean contributions by session and treatment

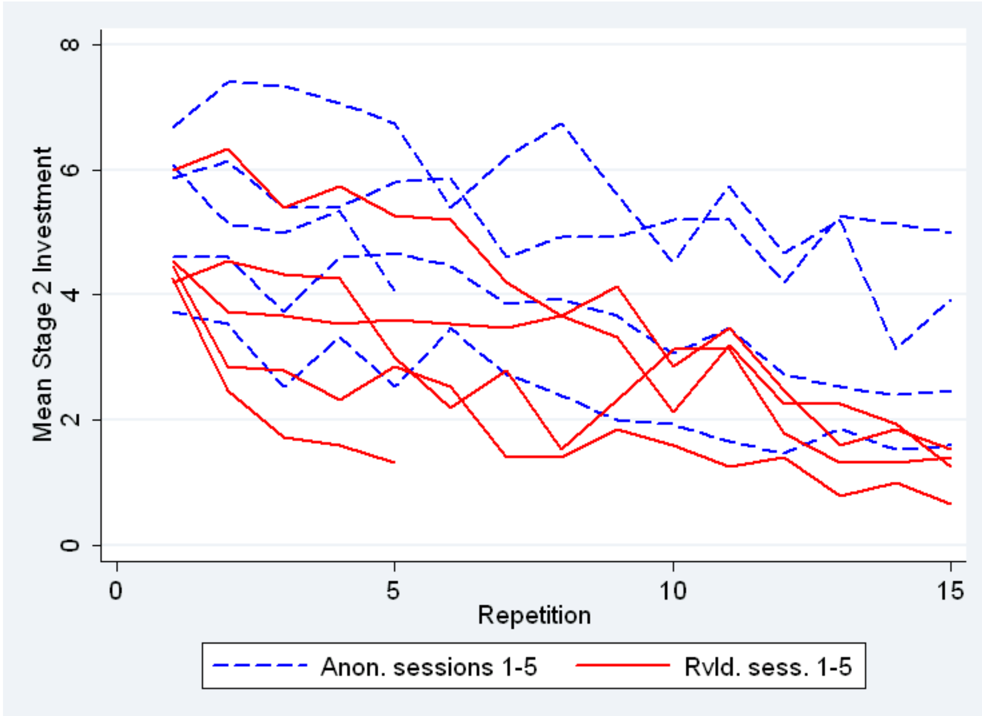

Figure 5: Stage 1 investment by Stage 2 investment (repetitions 8-15). Left=Anonymous, right=Revealed