THE ROLE OF TRUST, ENDOGENOUS INSTITUTIONS, AND THE POSSIBILITY OF GRACE IN PUBLIC GOODS PROVISION

DOUGLAS A. NORTON and R. MARK ISAAC

Both authors are:

Department of Economics
Florida State University
Tallahassee, FL 32306-2180

Isaac is the corresponding author:
misaac@fsu.edu
850-644-7081

ABSTRACT

We report a series of intentionally sequenced, four stage, laboratory public goods experiments to test a phenomenon we call "reverse crowd out". Motivated by the Gruber and Hungerman field study on faith-based charity crowd out and the New Deal we ask whether private provision of a public good can ever cause a reduction in public provision. Our first stage uses the standard voluntary contributions mechanism followed by a second stage that imposes an exogenous tax. Subsequent stages replace the mandatory tax with an endogenous tax and some experiments induce more charitable preferences for a subset of the group (we call these agents of grace). The results from these treatments are diverse and different than would be expected in a standard voluntary contributions game. Like past experiments and field studies we detected intermediate levels of crowd out, but, did not observe the phenomenon of reverse crowd out. One strong pattern that emerged from the aggregate data, even amid its idiosyncratic nature, was the influence of initial levels of trust on later taxation decisions. The agents of grace had a suggestive, but not definitive, influence on the other participants.

Key Words: Public Goods, Charity, Crowding-out, Endogenous, Tax

JEL Classifications: H41, A13, C92

1 . INTRODUCTION

When government provision of public goods reduces the level of private provision of those same goods, a process of "crowding out" has taken place. The theoretical channel by which crowding out may occur broke ground many years ago (Bergstrom, Blume, and Varian [1986]). And, even though there seems little doubt as an empirical matter that crowding out exists, the empirical evidence differs widely in scale. Field studies include Payne (1998), Hungerman (2005), and Hungerman (2007). Laboratory economics experiments (beginning with Andreoni [1993] and including Chan, et al. [1998] and Eckel, et al. [2005]) also have detected intermediate levels of crowding out. The experimental design presented here, however, tests a different phenomenon we call "reverse crowding out" which is the reduction of government-provided public goods as a result of increased private provision.

The recent paper by Gruber and Hungerman (2007), regarding New Deal social service expenditures, provided the motivation for this research. Gruber and Hungerman note that religious contributions for social services remained steady through the worst early years of the Great Depression, but began to tumble in the period of 1933-1934, coinciding with FDR's implementation of relief and works projects. Eventually religious contributions settled in at about 30 percent less than their earlier level. Hungerman and Gruber estimate that, even though the rate of crowding out was not large, the massive increase in federal social service expenditures can account for nearly all of the reduction in private religious transfers. Is such a historical level of crowding out reversible? By abstracting from the complex environment of the field into the laboratory we investigate whether there are conditions under which an exogenous and inefficient tax is abandoned in favor of a more efficient voluntary process.

In the dire circumstances of the Great Depression and with the continued influence of movements such as the "Social Gospel" the private sector and churches transferred their roles as the chief providers of charitable and compassionate public goods to the government. We wish to set the stage to ask whether or not the historical crowding out of religious provision of public goods is reversible. And, if reverse crowd out could happen, what conditions would such a reversal require?

The experimental design here represents a major departure from the seminal experimental literature on public goods crowding-out conducted by Andreoni. We report a series of linear public goods experiments, very intentionally ordered in four stages, to test two changes in the economic environment which we identified as possible enhancers of the private provision of public goods. Our participants first experience a fully voluntary cycle of the provision of public goods, followed by a cycle with provision through mandatory taxation. We then introduce, in the third stage, institutional endogeneity by offering participants a choice of using either voluntary contributions or the tax scheme (or both). Therefore, our first research question is the following: will participants, become so acclimated to an exogenous taxation regime that they never choose to return to lower levels of taxation? Or, can reverse crowding out be observed in our simple laboratory environment?

The second research focus is whether the likelihood of returning to voluntary provision of the public goods at a lower tax rate while maintaining higher levels of provision can be increased by consistent cooperation from a subset of the group. We investigate this question by a treatment condition across groups. We implement a fourth

stage of the experiment in which we increase the marginal per-capita return (MPCR, see Isaac and Walker, 1984) of some group members in half of our sessions. We refer to these group members as "agents of grace." The agents of grace MPCRs are increased so that contribution to the private provision of the public good in the voluntary process becomes a single-period dominant strategy.

The following is a preview of our experimental results. We conducted ten sessions with a standard subject pool. The introduction of both the exogenous and endogenous tax treatments produce a time path of aggregate provision of public goods that has significantly less decay than would be expected from a typical voluntary contributions mechanism (see Davis and Holt, 1993). The data were quite idiosyncratic by group. Moreover, the phenomenon of substantial reverse crowd out is never observed. However, one unexpected result was the power "trust" had in explaining the variation in tax rates.

Turning to the individual results we chose to focus on residual giving and voting behavior. First, we define "residual giving" as the number of tokens contributed to the group exchange as a share of the tokens available after tax. The results point to an influence of the agents of grace on non-agents. But, the statistical support for this effect is of boundary statistical significance, which causes us to make the case with caution and note the need for further research. We present a statistical analysis of residual giving that examines other possible determinants, such as the level of the tax and the subject's divergence of revealed preferences from the results of the tax vote. We also analyze the individual votes to test the hypothesis of a "ratchet effect" but find no statistical significance for such voter behavior.

The remainder of the paper is organized as follows. The second section discusses the details of our experimental design and some hypotheses as to what we might expect to observe. The third section presents the experimental results. Finally, we offer some conclusions and directions for further research. In particular, we discuss two additional sessions in which we made multiple design changes in an attempt to create an environment in which we could observe a case of substantial reverse crowding-out.

2 . EXPERIMENTAL DESIGN AND HYPOTHESES

As noted, Andreoni's seminal experimental design was constructed to focus on warm glow and the possibility of incomplete crowding out. He used a relatively small group (three subjects) with a concurrent "payoff chart" technology and incorporated an interior optimum and equilibrium. Indeed, the term "crowding out" in Andreoni's sense is an exercise in comparing equilibrium levels of contribution. In fact, it requires that there be a non-cooperative equilibrium in which each individual has a positive contribution to the public good. Furthermore, in the standard environment, there is no difference in the productivity of private and government provision. Government provision is observationally equivalent to a required "minimum contribution" in the voluntary contributions mechanism (indeed this is how it was explained to the subjects). Many other researchers, for purposes of comparability, have used this same structure.

Because we are asking different questions, we begin down a new path. First, we want a larger group dynamic, so our experiments consist of seven persons. Secondly, we are interested in a world of strong incentives to free ride and thus we drop the interior

payoff structure and return to the equally traditional linear public goods model, where underprovision of the public good is a dominant strategy in the stage game.

Thirdly, we wanted to model a world in which participants had a preference for private provision of the public good, so we made government provision less efficient than private provision. Thus, in our experiments, provision through government taxation is made to be moderately inefficient (twenty percent of tax receipts are lost). This could represent actual physical and organizational inefficiency in the government provision of social services, or population preferences that value government provision less. Either story fits our design.

Fourthly, while the role of government and the levels of taxation are exogenous in most of the existing literature, we are interested in endogenous mechanism choice.[1]

Finally, we are interested in how endogenous institutions respond to changes in individual preferences. Presumably, religious institutions will be among those likely to achieve reverse crowding out if individuals in the community/society become more receptive to personal action. Therefore we introduce a final experimental stage in which some individuals become more receptive to providing the public good.

In our design, the voluntary contributions equilibrium is defined by complete free riding. The empirical reality, as repeatedly demonstrated in laboratory experiments, is that of positive, but non-optimal, levels of public goods provision.[2] In such an environment, a standard research question has been to examine those mechanism features that can maximize the efficiency of the system, (an example being face to face non-binding communication [Isaac and Walker, 1988]).

We look at a specific mechanism sequence: purely voluntary contributions, exogenously determined forced government provision (taxation) combined with voluntary contributions, and then endogenously determined taxation combined with voluntary contributions.[3] To implement this, our baseline is an [A, B, C, D] sequence . For the first seven periods (Stage 1), seven subjects participate in a standard VCM design. Each subject has the task of investing 100 tokens in either an individual or a group exchange. Investment in the individual exchange returns a one token payment with certainty, while the group exchange operates as a VCM process with a marginal per-capita return (MPCR) = .5.

In the second round of seven periods (Stage 2), we come the closest to matching the Andreoni design. We impose a tax of 30 tokens upon every participant to fund a minimum level of the public good. The tax efficiency of 80 percent means that 24 tokens are actually transferred to the group exchange for each person; six tokens disappear.[4] Voluntary contributions are allowed beyond the minimum through the VCM process.

In the third treatment (Stage 3), participants are allowed to choose the tax level through a voting mechanism, with VCM operating on their residual tokens. We used a

---

[1] One part of our design is addressed by Sutter and Weck-Hannemann (2004) in which subjects are allowed to vote yes/no by majority vote on a exogenously suggested tax scheme.

[2] We use the term "equilibrium" only in the sense of individual-income maximizing preferences.

[3] Another stream of research into crowding-out has been to look at individual decisions in dictator game experiments (see Eckel, et al. [2005]).

[4] The 20 percent inefficiency figure is not intended to be an advertisement for an actual field scale. It is one that is neither so small as to be trivial nor so large as to seem to be overwhelming to the subjects.

median voter process. Each participant submits a number from 0 – 100 (the per-period endowment of tokens) and the experimenter chooses the median tax.[5]

In our fourth treatment (Stage 4) we introduce our agents of grace. Two of the seven participants are randomly chosen in advance to have their MPCRs increased to 1.43 (accomplished by lowering their return on their individual exchange).[6] The fact that a change in incentives is possible, but not the exact form nor the distribution, is publicly announced to the individuals.

These experiments contain only this sequencing because it is of interest in the historical evolution of government involvement in charitable activity. We will return later to the issue of controls to this sequencing.

The underlying contributions game admits the standard theoretical properties. In the stage game, full investment in the individual exchange is the dominant strategy, while the group income-maximizing outcome is for each participant to invest all tokens in the group exchange. Nothing changes when an exogenous tax is enacted. However, with the median voter tax mechanism in Stage 3, there are a multiplicity of equilibria.

Consider that the choice of a subject can be represented as the pair $<t_i , p_i|T>$, meaning that each subject chooses a vote for a tax, t_i , and, conditional on the tax rate actually chosen, T, a proportion, p_i , of his untaxed funds to contribute to the group good. In the stage game, whatever tax is chosen, it will remain a dominant strategy in the subgame to set p_i, what we will call the "residual contribution," equal to zero. The only remaining question is "What vector of tax choices support an equilibrium?" The following provide a partial characterization of the outcomes in Stage 3:[7]

Let V* be any vector ($<t_1 , 0> , \ldots, <t_7 , 0>$) with the properties that a) Each person contributes no residual tokens to the group exchange regardless of the value of T, and b) that no one person could have altered the median vote. Then V* is an equilibrium.

The following is one category of outcomes that are not equilibria: For any vector V = ($<t_1, p_1(T)> , \ldots, <t_7 , p_7(T)>$), if V has the property that the t_is are strongly distinct, then V is not an equilibrium.

The first statement has the following corollary. The set of equilibrium vectors so characterized includes both what we call the "full tax outcome" through the vector (<100, 0>, …, <100, 0>) and the "full free riding outcome" through the vector (<0,0>,…,<0,0>). The former Pareto dominates the latter. The idea of finding the maximum tax seems very focal (even though the 20 percent inefficiency means that it is not the group optimum). The first statement also suggests the manner in which this mechanism could become "stuck" at intermediate levels of taxation, perhaps at the initial level of 30 tokens per individual.

---

[5] See Holcombe and Kenney (2007) for information on actual field implementations of this voting mechanism in Florida school finance elections. We choose this mechanism simply so as to avoid the need to structure complicated agenda procedures such as would be needed in majority rule (see Capra, et al. [2008] who used a random-proposer process with majority rule in a coordination game).

[6] In the parlance of some public goods research, this would be known as raising only the "inside" MPCR.

[7] The Stage 4 outcomes need only be adjusted by the change in the dominant strategies in the stage game.

The second statement is important because it suggests a way in which the full tax outcome could be achieved. If intermediate levels of taxation occur with distinct votes, then one person can change the level of the tax. By assumption $T < 100$, and a single unit increase in the tax will make that pivotal person better off. This might provide for a ratcheting up of the initial, exogenously imposed tax.

What, then, is the possible theoretical basis of what we call "reverse crowd out?" In Stage 3, the participants have access to an endogenously determined level of taxes. The discussion above demonstrates that both the full-tax outcome and the zero tax outcomes have stability properties. Within the zero tax outcome, the traditional free-riding outcome remains the only equilibrium. However, there is at least the possibility that the endogenous process of choosing a tax level might promote contributions, not unlike what has been observed with other "endogenous" features in VCMs.[8] Furthermore, in Stage 4 the two agents of grace will have a dominant strategy to contribute all their tokens to the group exchange. Logically, they, like everyone else, would prefer that whatever contributions are made be made through the VCM process because it is more efficient; but, they may also have a strong preference for requiring the contributions of all other group members via the tax. The most complete manifestation of reverse crowd out would occur if tax rates fell and total contributions to the group exchange increased.

We can now use Table 1 to summarize the twelve session conducted in this experiment.

In a traditional experimental economics design, we might desire to unravel sequencing effects among the four treatments. However, our stage sequencing is integral to the underlying historical motivation. Stage 1 is analogous to the pre-New Deal period with a reliance on the voluntary provision of charitable goods. Stage 2 captures a change to a non-trivial level of tax-based support for the public good. Stage 3 may be viewed as a modification in which the political system can endogenously move to higher or lower taxes, given that the participants have had experience in both a no-tax and a fixed tax regime. Finally, Stage 4 is not historical but forward-looking. This treatment, in which endogenenity is combined with transformed preferences (for a subset of the population) towards increased private provision of the public good, models new theories of the development of limited government (Norton [2008]).

Despite the rationale for this specific sequencing, there are still some control possibilities open to us. First, because voluntary contributions public goods environments have a robust pattern, we can use the data in Stage 1 to make sure that nothing unique to our sessions is causing any untold influences on the results. Given this control, the Stage 1 results provide a good benchmark for the institutional changes which follow.

In addition, the effects of the introduction of the agents of grace can be compared to five sessions in our "Control" design: [STAGE 1, STAGE 2, STAGE 3, STAGE 4~] where Stage 4~ represents periods 21-28 conducted without any agents of grace.

[8] Such features include cheap talk (Isaac and Walker), and endogenous group formation (Ahn, et al. [2008]).

	BASELINE (5 Sessions)	CONTROL (5 Sessions)
Traditional VCM	Stage 1	Stage 1
Exogenous Tax	Stage 2	Stage 2
Endogenous Tax	Stage3	Stage 3 and Stage 4~
Endogenous Tax Plus Agents of Grace	Stage 4	N.A.

Table 1

We say that we conducted a "Stage 4~" segment in the Control experiments rather than one 14-period Stage 3 segment because in the Control experiments we included the same "break" in the subjects' decisions as was needed in the Baseline experiment to implement the change in MPCRs for the agents of grace. We did this in order to control for any possible "restart effects" inherent in the mechanics of introducing a fourth stage. Specifically, we replicated in the Control sessions a similar intervention in the introduction of the final seven periods.[9]

The standard local recruitment protocol using an undergraduate subject pool was followed. Subjects were guaranteed $10 (US) as a show-up fee, and total earnings were typically in the $20 (US) to $40 (US) range for a time commitment of between one and two hours. However, we also conducted two sessions with a special undergraduate subject pool from a group formed endogenously outside of this experiment. A full motivation and discussion of this treatment is delayed until section 3.4 .

Finally, the following information conditions were maintained through all sessions. When voting opened the period, individuals made their tax votes simultaneously. The median voter result (if applicable) was announced before participants made their contribution decisions. At the end of the period, individuals were informed of their earnings, including a specific report of total tokens invested in the public good. The pattern of individual contributions was not reported. The subjects did not receive the text of the changes in the instructions until they occurred.

[9] See Appendix II for a complete copy of the baseline instructions, and Appendix III for a full discussion of the intervention in the control series. As will be seen in a later figure, there was a very small restart effect in period 22.

3. EXPERIMENTAL RESULTS

3.1 Aggregate Results

We begin by examining the aggregate results, first on the provision of the public good, then on the aggregate decisions from the endogenous tax periods. We will organize the presentation by stages.

Stage 1: Standard VCM

The sessions begin with the seven periods of Stage 1, which are a standard linear VCM environment. Based upon the substantial literature on VCMs (see Davis and Holt), we would expect that groups in Stage 1 would begin with levels of contribution between the social optimum and the free-riding equilibrium, with a process of decay setting in. Indeed, this is the case in our ten sessions. Table 2 presents the results of a simple regression on the first seven periods' results for each session. These regressions predict total contributions to the group exchange with an intercept and a coefficient on period (1-7). It can be seen easily that the intercepts are all approximately mid-way between zero (the free-riding prediction) and 700 (the social optimum). The coefficients on period provide a measure of the decay tendencies in each session. All of the decay coefficients are indeed negative, and even with only seven observations all but one of the coefficients are statistically significant.

	Intercept	Decay Coefficient
Baseline 1	532.57	-23.00 (-4.37)†††
Baseline 2	591.71	-28.75 (-2.39)††
Baseline 3	498.71	-28.57 (-1.62)†
Baseline 4	488.57	-18.89 (-2.64)††
Baseline 5	569.29	-13.75 (0.90)
Control 1	356.14	-32.82 (-4.60)†††
Control 2	456.14	-31.07 (-4.74)††
Control 3	599.14	-16.86 (-1.72)†
Control 4	482.86	-20.5 (-1.75)†
Control5	412.43	-24.43 (-2.35)††

t-values are in parentheses. The significance levels (one-tailed) are: † (10 percent), †† (5 percent), ††† (1 percent)

Table 2

We conclude that nothing in our computer interface, instructions, or local subject pool leads to any unexpected results in a standard VCM environment.

Stages 2 and 3: Addition of the Tax Institutions

Here, we examine whether and how the two alternative tax institutions affect provision of the public good. Stage 2 implements an imposed 30-token per-person, per-period tax that operates at 80 percent efficiency. Stage 3 replaces the 30 token exogenous tax institution with an endogenously determined tax institution (still 80% efficient). Figure 1 presents aggregate provision of the public good averaged across all ten sessions

for Stages 1, 2, and 3.[10] The lower line is an extrapolation of the time paths of contributions as estimated in the regressions in Table 2 above. This line serves as one type of intra-group calibration of the effects of the two tax institutions.[11]

Figure 1 Here

A visual examination of Figure 1 indicates that in *both* Stage 2 (exogenous tax) and Stage 3 (endogenous tax) the time path of provision of the public good has stabilized, with little or no decay. This can be checked against two simple statistical analyses presented in Table 3. The first regression (left column) explains provision of the public good as a constant with dummy variables for both Stage 2 and Stage 3. This regression should capture the simple level effects of the two tax treatments. If the tax institutions are having no effect, then the dummy variables should be negative and significant, reflecting the expected decay of a standard VCM process. Instead, the estimated coefficients for Stage 2 and Stage 3 are highly insignificant, indicating that the level of provision is not decaying across these two stages.

The second regression (right column) focuses upon the decay process by period, both across all 21 periods in Stages 1, 2, and 3, and also within the seven periods each of the Stages. This regression includes the following variables.

Period (This is a time trend across all periods 1 – 21.)

*Stage 2^; Stage 3^* (These are constant-shift variables for the two stages. The ^ signs are added because the stage dummy variables are interpreted differently than in the previous regression.)

Stage2^xPeriod (This is an interacted variable of *Period* times *Stage 2^*. It will estimate the per-period decay within Stage 2 as compared with Stage 1).

Stage3^xPeriod (This is an interacted variable of *Period* times *Stage 3^*. It will estimate the per-period decay within Stage 3 as compared with Stage 1).

[10] In presenting the total level of the provision of the public good, the figure combines the voluntary contributions plus the 80 percent of the tax receipts that are applied to the group exchange. Also note that because we are considering only Stages 2 and 3, there are no treatment differences between the Baseline and Control sessions, and so they can be combined.

[11] When the extrapolations from any of the ten sessions predicted a negative level of contributions to the group exchange, we censored those predictions at zero in producing the line in Figure 1.

	Simple Stage Effects	Stage Effects on Decay
Intercept	403.30 (.0000)***	498.8 (.0000)***
Stage 2 Dummy	5.14 (.8026)	
Stage 3 Dummy	-12.67 (.5397)	
Period (1-21) (Overall Decay)		-23.86 (.0000)***
*Stage 2^* Dummy		27.07 (.4854)
Stage2^xPeriod (Change in Decay in Stage 2)		13.19 (.0096)***
*Stage 3^* Dummy		65.39 (.2822)
Stage3^Period (Change in Decay in Stage 3)		14.22 (.0060)***
Adjusted R-sq.	-0.063	.7956
(Estimated Change in Provision of Public Good Between Periods 8 and 7)		+107.47
(Estimated Total Decay Stage 2)		-10.67
(Estimated Change in Provision of Public Good Between Periods 14 and 15)		+43.1
(Estimated Total Decay Stage 3)		-9.64

P-values in Parentheses *** Significant at 1%; **Significant at 5%; *Significant at 10%

Table 3

The results in Table 3 can be interpreted as follows. The underlying estimated time trend beginning in Stage 1 (*Period)* is, not surprisingly, negative and highly significant. On the other hand, the estimated coefficients on both *Stage2^xPeriod* and *Stage3^xPeriod* are both positive and highly significant, again indicating that both tax regimes flatten the decay initiated in Stage 1. The net total decay coefficients for both Stages 2 and 3 are shown in the bottom of the table. [12]

It may not be surprising that the exogenous tax institution serves to increase the provision of the public good when compared to the default (standard VCM). Consider the field evidence on crowding out. These studies typically demonstrate that a) crowding-out is incomplete but b) nevertheless positive. Increases in government funding of charitable activity raise the total level of the charitable good, but with an associated reduction in private provision. (This is, for example, reflects the Gruber and Hungerman conclusions about private charity and the Great Depression). The results from Stage 2 indicate that we have captured at least the first component (crowding-out is incomplete) in our laboratory environment: the exogenous tax has raised total provision of the public good from its expected level. The individual results section will address whether the overall increase in provision from the tax is accompanied by reductions in the residual gifts of the individuals.

Endogenous Tax Choices in Stage 3

Here we consider the aggregate tax levels chosen by the participants during Stage 3. Figure 2 displays the average tax in Stage 2 (when the tax is exogenously set by the

[12] Recall that in the second regression, the Stage constants represent the shifts at the origin. The "jump" at the beginning of each stage are calculated using the stage and decay dummies.

experimenters) and Stage 3 (when it is chosen by the participants through the median voter process).

Figure 2 Here

Figure 2 indicates that there is a slight increase in the average tax level as Stage 3 proceeds. However, these average results mask a great deal of idiosyncrasy in the behavior of the various sessions. The next figure, Figure 3, displays three such sessions. The top panel shows a session in which the subjects essentially abandoned the tax, with provision of the public good decaying from a moderate level. The middle panel shows a session in which taxes varied around the level that we imposed in Stage 2, with public goods provision dropping but then remaining steady at around 300 units.. The bottom panel shows a session in which the group converged to full taxation.[13] Three of our ten sessions had chosen full taxation at least once by the end of Stage 3. None of the sessions produced a robust demonstration of reverse crowding out in which the participants abandoned the tax but maintained provision of the public good at high levels.

Figure 3 Here

These idiosyncratic results indicate that conditions specific to the composition of each group are capable of yielding very different outcomes. In the process of examining these data, we discovered an unexpected pattern that partially predicted what level of taxation different groups would choose. Namely, we found that the endogenously chosen levels of taxation in Stage 3 could be partially explained by the levels of cooperation in Stage 1. Groups that were relatively *less successful in providing the public good in Stage 1* were likely to choose *relatively higher levels of taxation in Stage 3.* This can be seen in Figure 4, which displays the results of a simple regression of average Stage 3 taxes on average Stage 1 contributions to the group exchange.[14]

Figure 4 Here

To summarize, both the exogenous and endogenous tax regimes produced levels of provision of the public good that were unexpectedly high, both in terms of expectations from the existing literature and in terms of the patterns set by the groups themselves in Stage 1. However, the tax choices in the endogenous tax condition were highly idiosyncratic across groups, with none of the groups achieving what we would call robust reverse crowding out. We turn now to a consideration of the individual data.

3.2 Individual Results

We focus on two types of individual data. First, we examine the *residual cooperation* of each individual. We measure this as the percentage of the tokens remaining after the tax that the individual contributes to the group exchange. (This

[13] Recall that the figures display effect levels of taxation, which is total tax times .8 .

[14] We conducted several alternate versions of this regression such as including two additional sessions discussed in the conclusions of this paper and using such variants as “end point” and “mid-point” to see if the results from Figure 4 are robust. The qualitative results are identical in every case.

measure obviously requires excluding those observations from subjects in a group that has chosen a 100 percent tax). This investigation will lead us to consider results from all the stages, including Stage 4 in which the existence of the agents of grace is the treatment in five of our ten sessions. Secondly, we examine individual votes on the endogenous tax.

Residual Contributions: Descriptive Statistics

Figure 5 graphs the residual contributions in the fourteen periods of Stages 3 - 4 of three separate groups of subjects. One group consists of those subjects who were in the Baseline treatment and became agents of grace in period 22. The second group consists of those subjects who were in the Baseline treatment and were not agents of grace in Stage 4. The final group consists of all of those subjects who were in the Control sessions in which there no agents of grace.

Figure 5 Here

A couple of patterns are obvious from Figure 5. First, the residual cooperation of the agents of grace strongly reacts to the induced change in incentives in Stage 4. This is born out by simple descriptive statistics. Six of the ten agents of grace in the five baseline sessions followed the dominant strategy of full contributions in every period. A seventh individual fully contributed in four periods and at least 80% elsewhere. Two other individuals alternated between four periods of full contribution with low contributions elsewhere. Only one of the ten agents of grace showed no tendency towards any pattern of full contributions. These data are broadly consistent with other findings of how frequently individuals exactly choose or come close to choosing a dominant strategy (Isaac and James [2000]).

Secondly, there is a suggestive, but not overwhelming, indication that individuals who were in the Baseline sessions but were not themselves agents of grace were more residually cooperative in Stage 4 than the subjects in the Control sessions (which had no agents of grace). We will examine these effects in more detail in the following statistical analysis of the individual data.

Residual Contributions: Statistical Analysis

We have a total of 1806 observations. The variables in our basic estimation are described in Table 4. We report, for the sake of examining the robustness of the results, both a pooled OLS clustered at the group level and a random effects Tobit analysis with residual cooperation as the dependent variable. Both regressions contain the same constant and period control variables we utilized above. The most important variables for our analysis are the following:

Tax: This is the level of tax chosen by the group in the particular period. A positive estimated coefficient on "*Tax*" would indicate that individuals become more residually cooperative as tax levels increase, perhaps even to the point of attempting to keep their private contribution levels constant (no crowding out).

Grace and GracexPeriod: These are dummy variables relating to individuals who are agents of grace in Stage 4 of the Baseline sessions. *Grace* is a simple dummy variable equal to 1 if the individual is an agent of grace in that period. *GracexPeriod* is the interaction of the *Grace* variable and *Period.* Because the incentives of the agents of grace are radically transformed, we would expect, especially in light of Figure 5, for at least one of these variables to be statistically significant.

GroupAOG and GroupAOGxPeriod: These two variables relate to individuals who are in Stage 4 of the baseline experiments but who are not themselves agents of grace. *GroupAOG* is the basic dummy variable for this classification. However, because our interest lies in determining whether altered behavior by the agents of grace affects other participants in the group, we restrict the assignment of the *GroupAOG* dummy variable. First, it only applies beginning in period 23, after the agents of grace have had one period to alter their behavior. Secondly, the variable is set equal to 1 only for subjects in those groups in which the agents of grace contributed at least 90 percent of their tokens in period 22 (that is, we restrict the classification to those groups in which the agents of grace are, in fact, changing their behavior). *GroupAOGxPeriod* is the interaction of the simple dummy variable with the period variable.

	OLS	R.E. Tobit
Intercept	.639 (.000)***	.576 (.000)***
Period (1-28)	-.015 (.000)***	-.022 (.000)***
Stage 2 Dummy	.093 (.108)	.198 (.071)*
Stage 3 Dummy	.080 (.397)	.155 (.359)
Stage 4 Dummy	.311 (.012)**	.456 (.024)**
Stage2xPeriod	-.007 (.136)	-.014 (.176)
Stage3xPeriod	-.002 (.778)	-.004 (.690)
Stage4xPeriod	-.011 (.024)	-.015 (.082)*
Tax	-.000 (.884)	-.000 (.272)
Grace	-.313 (.582)	-.717 (.400)
GracexPeriod	.034 (.096)*	.071 (.039)**
GroupAOG	-.645 (.023)**	-1.15 (.107)
GroupAOGxPeriod	.027 (.024)**	0.047 (.097)*
Wald chi2	265.62	322.06

*p-values in parentheses. ***= significant at 1%, ** = significant at 5%, * = significant at !0% (two-tailed tests).*

Table 4

We note the following implications from the results in Table 4. First, the coefficient on "*Tax"* is negative, small in magnitude, and highly insignificant. This addresses the second part of our earlier discussion on crowding out. If individuals wished to maintain a given absolute level of private provision as the tax-funded provision increased, then the coefficient on "*Tax*" would have to be positive. Clearly, this is not the case. Individuals are not becoming more residually cooperative as taxes increase, and at least some crowding out is taking place, as individuals are not behaving to maintain the absolute level of their private contributions as taxes increase. Thus, combined with the

results from the aggregate behavior described above, we note that our environment mirrors the same general behavior as previous field estimates: there is positive, but incomplete, crowding-out.

Furthermore, given that we observed previously that tax levels in Stage 3 were inversely related to initial levels of cooperation, the essentially zero coefficient on "*Tax*" leads us to conclude that subjects use the taxes primarily as a commitment device.

Secondly, and not surprisingly, the positive and statistically significant coefficient on "*GracexPeriod"* indicates that agents of grace changed their own behavior. Specifically, this coefficient captures the fact that agents of grace significantly stopped their decay in contributions in Stage 4. This result is consistent with the descriptive result that most of the agents of grace began following their new dominant strategy.

Finally, the results on the two coefficients relating to the subjects in the Baseline experiments offer tentative support to the data in Figure 5 that these "non agents of grace" changed their behavior in Stage 4. Working through the combined effects of the coefficients does not lead to any major jump by these subjects in period 23, but there is a statistically significant reduction in the decay in periods 23- 28. It should be noted, however, that not even the addition of the agents of grace in our Baseline sessions yielded any evidence of robust reverse crowd out.

We obtained substantially similar results with other definitions of "*GroupAOG."* However, we were surprised when we tried a different specification of the "*Tax"* variable. We included in the regression what we call our "Goldilocks" variables: dummy variables for individuals who found that the tax chosen by their group was not high enough ("*NHE*") or not low enough ("*NLE*") when compared to their own votes. The results from this random effects Tobit estimation are presented in Table 5.

	R.E. Tobit
Intercept	.576 (.000)***
Period (1-28)	-.021 (.000)***
Stage 2 Dummy	.119 (.085)*
Stage 3 Dummy	.179 (.289)
Stage 4 Dummy	.465 (.020)**
Stage2xPeriod	.010 (.143)
Stage3xPeriod	-.004 (.673)
Stage4xPeriod	-.015 (.081)*
Tax	-.000 (.660)
NHE	-.055 (.187)
NLE	-.081 (.053)**
Grace	-.622 (.464)
GracexPeriod	.066 (.052)*
GroupAOG	-1.11 (.119)
GroupAOGxPeriod	.044 (.111)
Wald chi2	327.87

*p-values in parentheses.*** = significant at 1%, ** = significant at 5%, * = significant at !0% (two-tailed tests).*

Table 5

The coefficients on both goldilocks variables are negative, and "*NLE*" is statistically significant. This indicates that dissatisfaction with the group's tax vote never helps, and may hurt, the level of cooperation among group members.[15]

Individual Voting Behavior

Until this point, we have examined the effects of our experimental treatments of endogeneity only in terms of aggregate provision of the public good and of individual's residual contributions to the group exchange. This section shifts the focus to a subject's votes on the endogenous tax in periods 15-28.

Given the widely varying time paths of the winning tax level, modeling individual votes in that process may not add much insight. Instead, we look at two special questions about the voting decision.

First, our theoretical model suggests that it is the unique median voters who have an incentive to "ratchet" up voting levels. We were curious as to whether that pattern was reflected in the individual voting data. Do the median voters increase their votes with a greater frequency than the "Goldilocks" voters (those who think the tax level is "not high enough" or "not low enough")? We created a contingency table (Table 6) to investigate that question. The data are all from Stage 3, periods 16-21. The table crosses whether the individual's previous vote was higher than, equal to, or lower than the median with whether the individual raised, didn't change, or lowered his vote.[16]

	Individual Lowered Vote	Individual Kept Vote the Same	Individual Raised Vote
Previous Vote Was Lower Than Median ("*NLE")*	27	25	37
Previous Vote Was At the Median	24	13	30
Previous Vote Was Greater Than Median ("*NHE")*	40	23	43

Table 6

Three chi-squared tests among the rows each indicate that the frequencies are not statistically different from one another. This result suggests the time path of winning tax

[15] It should also be noted that in this new specification, the coefficient on "*GroupAOGxPeriod"* flips from being just barely significant to just barely insignificant. This is another reason we are guarded in our conjecture about the influence of the agents of grace on the other participants.

[16] We removed all boundary data where a subject could not raise or lower his vote. Also, we included all median voters if there were multiple median voters because median voters did not know if they were unique.

levels results from all individuals, including those off the median, changing their votes from one period to the next with approximately the same frequency.

A second question about individual voting was suggested by the incentives of the agents of grace. Because they have a dominant strategy to contribute to the group good, we wondered if they would be more likely to go to the extremes on voting --- going either full a full-tax outcome or complete crowding out. And, were they different than the other voters? We constructed the distribution of votes from the baseline experiments in Stage 4, for agents of grace and non-agents-of- grace. The distribution of those votes is given in Figure 6.

Figure 6 Here

Even though the simple mean of the votes of the agents of grace is higher than the other individuals (46 percent versus 38 percent) a Kolmogorov-Smirnov test indicates no statistical difference between the two distributions.[17]

4. Discussion

We designed a set of intentionally ordered experiments to test for the phenomenon of "reverse crowd out" in which private provision of a public good crowds out existing tax-supported provision of that good. Our design was able to capture the intermediate levels of crowding out typical of field studies. But, we never observed a robust example of reverse crowding out in our ten Baseline and Control Sessions, even after the introduction of the agents of grace in Stage 4 of the five Baseline Sessions.

If a phenomenon exists in the naturally occurring economy, in principle we ought to be able to capture some manifestation of it in the laboratory. There is a long tradition in experimental economics of the search for such "existence results."[18] At some point, when the goal of observing a previously unobserved phenomenon has failed, it is appropriate to ask what *combination* of changes in an experimental design might produce the elusive result. We decided to conduct an artefactual field experiment (Harrison and List, 2004) in which we conducted two additional sessions using the experimental design reported here but using subjects from a local student ministry community. In addition to changing the subject pool, we matched each individual's total earnings to an orphanage familiar to all the subjects. As explained in Appendix I, this matching had the potential for altering the subjects' incentives towards providing the public good. One session evolved to full taxation while the other group provided the first evidence of reverse crowd out by providing a large amount of the public good with low taxation (jokingly we dubbed the two groups Rauschenbusch and Rothbard). To see the data from these two groups see the appendices. This lays the groundwork for an agenda of further research. Which of the two effects was most critical in obtaining reverse crowd out? How robust could this phenomenon be with more replication?

Even though we did not observe reverse crowd out with the standard subject pool, the sessions yielded several interesting observations. The averaged data on Periods 1 – 21

[17] http://home.ubalt.edu/ntsbarsh/Business-stat/otherapplets/ks.htm

[18] See for example, Kim and Walker (1984) and Isaac, McCue and Plott (1985) on free riding; Isaac and Smith (1985) on predatory pricing; Smith, Suchanek, and Williams (1988) on asset bubbles; and Lynch, et al. (1991) on the "lemons" phenomenon.

demonstrated that pattern in these sessions was significantly different from the robust decay we would expect from the VCM alone. We also noted that these averaged data mask many idiosyncracies in the paths of the voluntary contributions. But, this led to a striking discovery: a common thread in these experiments was that the level of the tax could be predicted well by the level of "trust". Specifically, later period decisions about taxation can be explained by how cooperative the groups were in the non-tax world of Stage 1. For groups with low levels of trust, group members bypassed the voluntary contributions mechanism and sought higher tax rates for themselves and their group members. This coincides somewhat with an observation from Adam Smith in The Theory of Moral Sentiments, "What institution of government could tend so much to promote the happiness of mankind as the general prevalence of wisdom and virtue? All government is but an imperfect remedy for the deficiency of these." (Smith [1759]) The inefficient tax was an imperfect remedy for groups that could not trust each other, but, it worked in securing more of the public good than the projected trend.

The individual results are not based on the total provision of the public good; instead, they are based on the residual contributions of the subjects. A primary concern for us was the potential for the agents of grace to influence the behavior of other subjects in their group. Both the descriptive and statistical analysis suggests that they did exert some influence, the magnitude and statistical significance of the results means that we consider the conclusion suggestive rather than definitive. A potentially productive area for further research would be a different experimental design isolating the issue of the agents of grace. Specifically, perhaps seven periods was not long enough for the agents of grace to have full influence on the other subjects.

We also found interesting individual results with regards to issues of endogenous taxation. While the level of the tax did not explain residual gifts, a subject's vote during the endogenous periods was much more explanatory. We constructed so-called "goldilocks" variables to explain subject's votes relative to the selected tax: not high enough ("*NHE*"*)*, not low enough ("*NLE*"), or just right. When taxes were not low enough this had a significant difference in subject's residual gift. Both "*NHE*" and "*NLE*" had negative signs suggesting a negative behavioral reaction to not having their preferred tax as the one selected by the median vote.

Taxation seems like a very useful commitment device when there is a lack of trust among group members. But, what changes in our experimental design could cause a greater relative provision of public goods by voluntary giving and less by an endogenous tax? First, subjects might be less attracted to the tax if there were greater inefficiency associated with the tax. Our design has standardized on an 80% efficient tax across both treatments, but, that level of efficiency could be varied to locate a breaking-point that leads to more voluntary provision. Secondly, more experiments with tighter social networks and intentionally formed groups may reveal that naturally formed groups, presumably exhibiting higher levels of trust, do not need to rely on the tax to provide public goods. This would certainly be consistent with the theory that people provide more of a public good voluntarily when they see the beneficiaries as more like themselves (Hungerman [2007]). Thirdly, the observations of a negative influence on residual cooperation when subjects had preferences for either a higher or lower tax level suggest extensions of the current experimental design to a Tiebout-like environment, in which individuals can vote on tax levels *and* change groups. Research with endogenous group

formation research in different contexts is a fresh new area already underway (Cinyabuguma, Page and Putterman [2005]; Ahn, Isaac, and Salmon [2008]; Kosfeld, Okada, and Riedl [2009]).

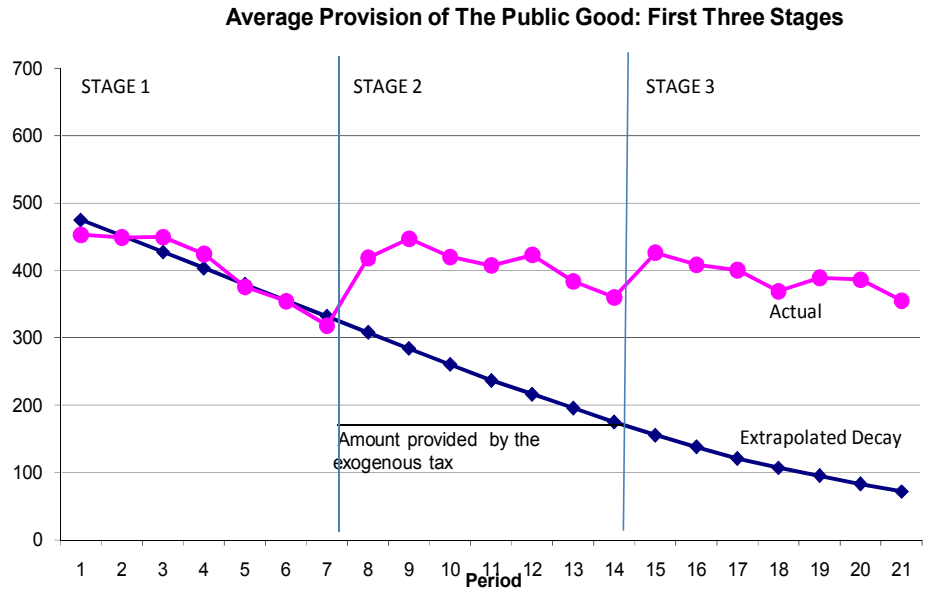

Average Provision of The Public Good: First Three Stages
700
600
500
400
300
200
100
0
STAGE 1
STAGE 2
STAGE 3
Actual
Amount provided by the exogenous tax
Extrapolated Decay
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Period

FIGURE 1

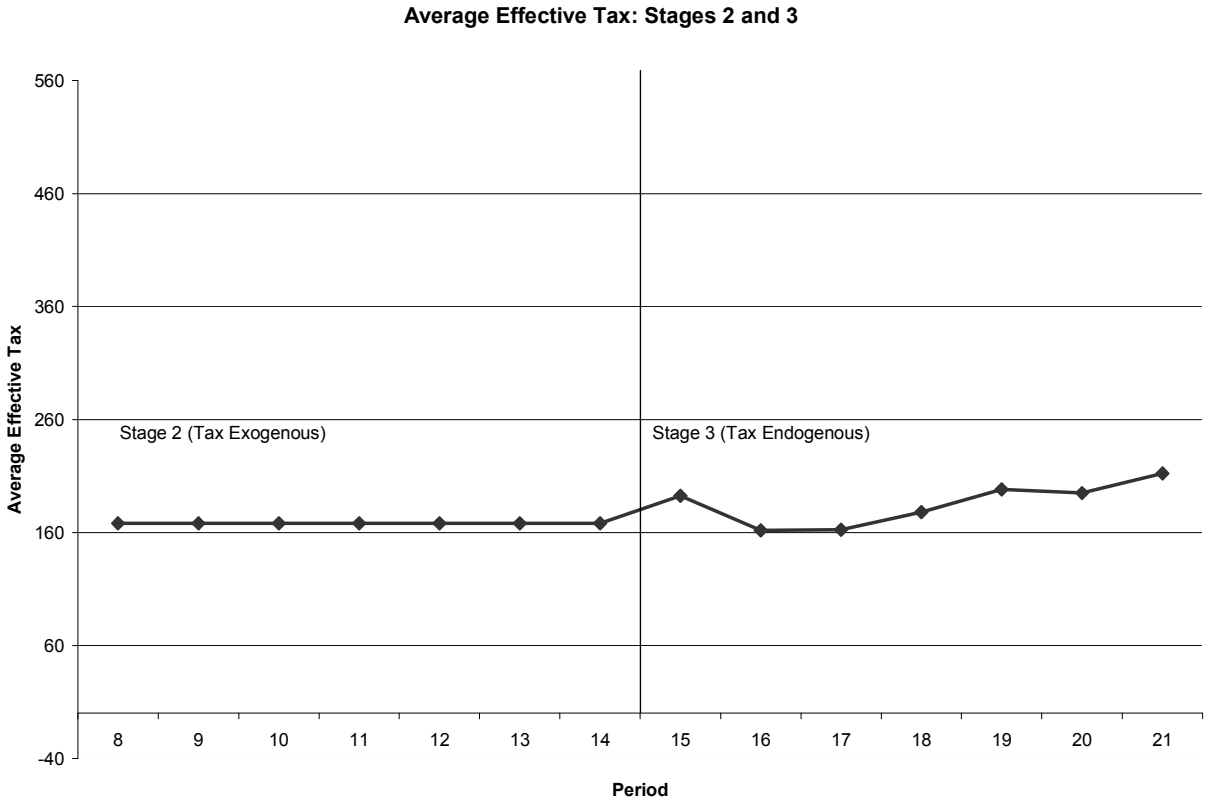
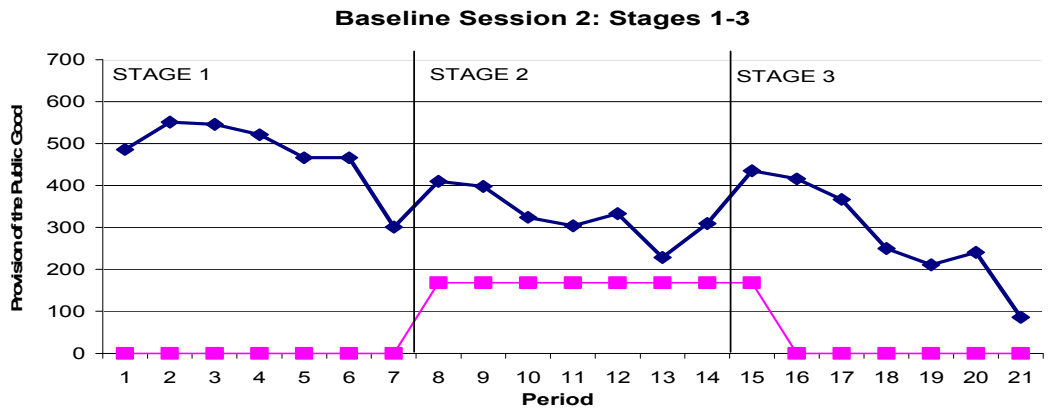
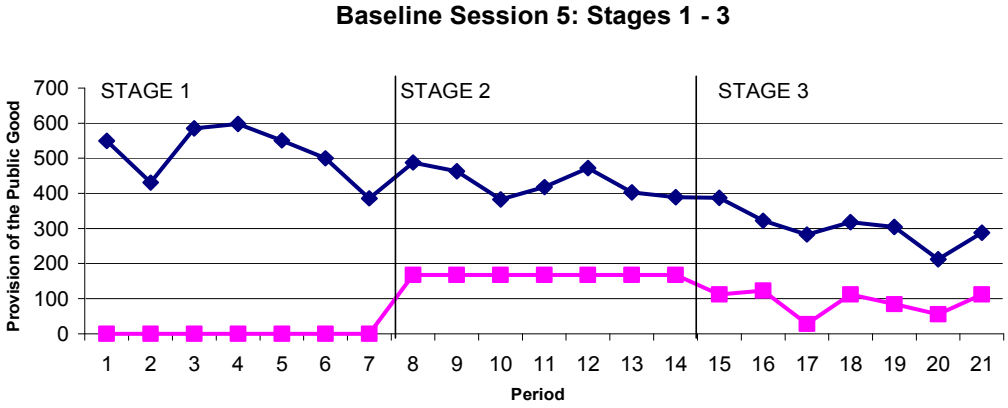

Average Effective Tax: Stages 2 and 3
560
460
360
260
160
60
-40
Average Effective Tax
Stage 2 (Tax Exogenous)
Stage 3 (Tax Endogenous)
8
9
10
11
12
13
14
15
16
17
18
19
20
21
Period

FIGURE 2


Baseline Session 2: Stages 1-3
STAGE 1
STAGE 2
STAGE 3
Provision of the Public Good
Period


Baseline Session 5: Stages 1 - 3
STAGE 1
STAGE 2
STAGE 3
Provision of the Public Good
Period

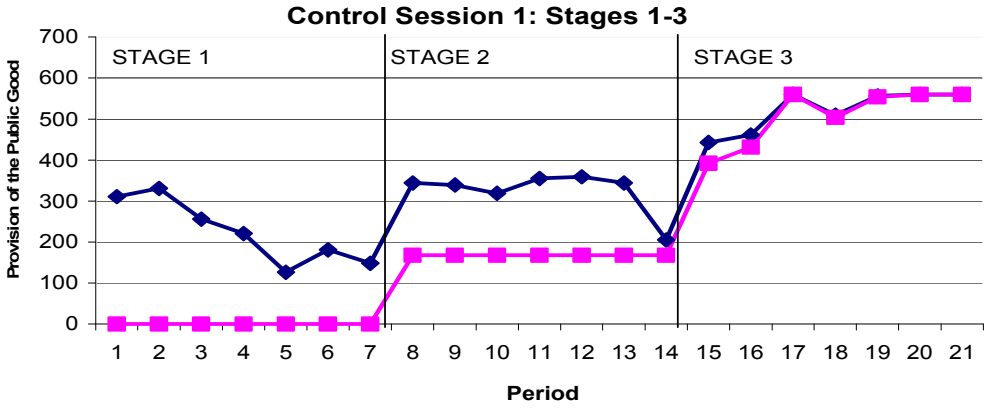

Control Session 1: Stages 1-3
STAGE 1
STAGE 2
STAGE 3
Provision of the Public Good
Period

FIGURE 3

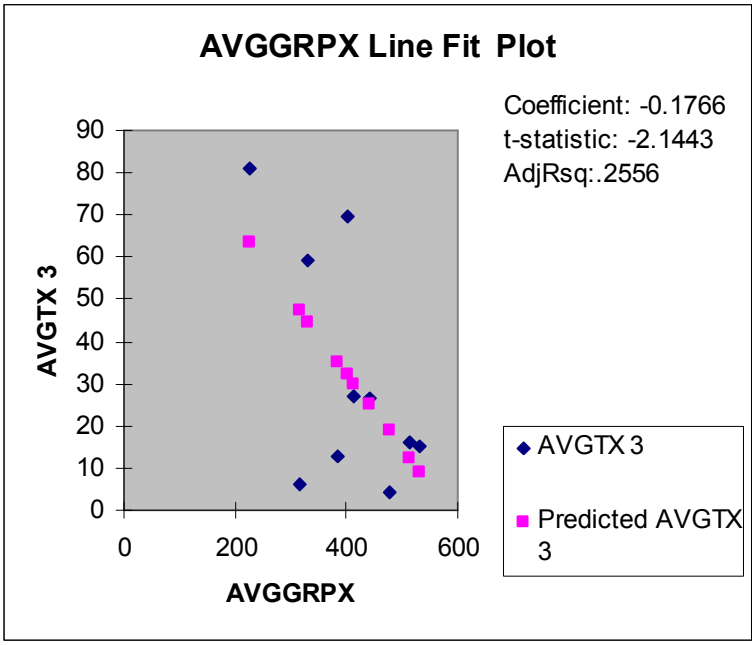

AVGGRPX Line Fit Plot
Coefficient: -0.1766
t-statistic: -2.1443
AdjRsq:.2556
AVGTX 3
AVGGRPX
0
10
20
30
40
50
60
70
80
90
0
200
400
600
AVGTX 3
Predicted AVGTX 3

FIGURE 4

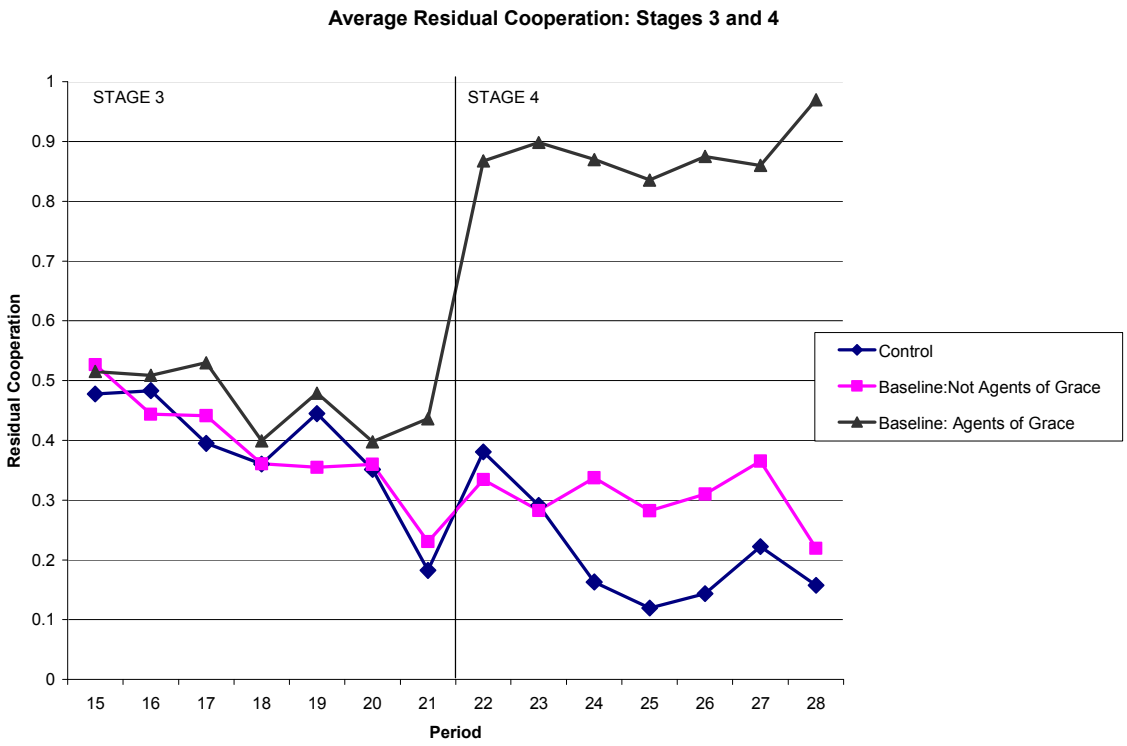

Average Residual Cooperation: Stages 3 and 4
STAGE 3
STAGE 4
Residual Cooperation
Period
Control
Baseline:Not Agents of Grace
Baseline: Agents of Grace
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
15
16
17
18
19
20
21
22
23
24
25
26
27
28

FIGURE 5

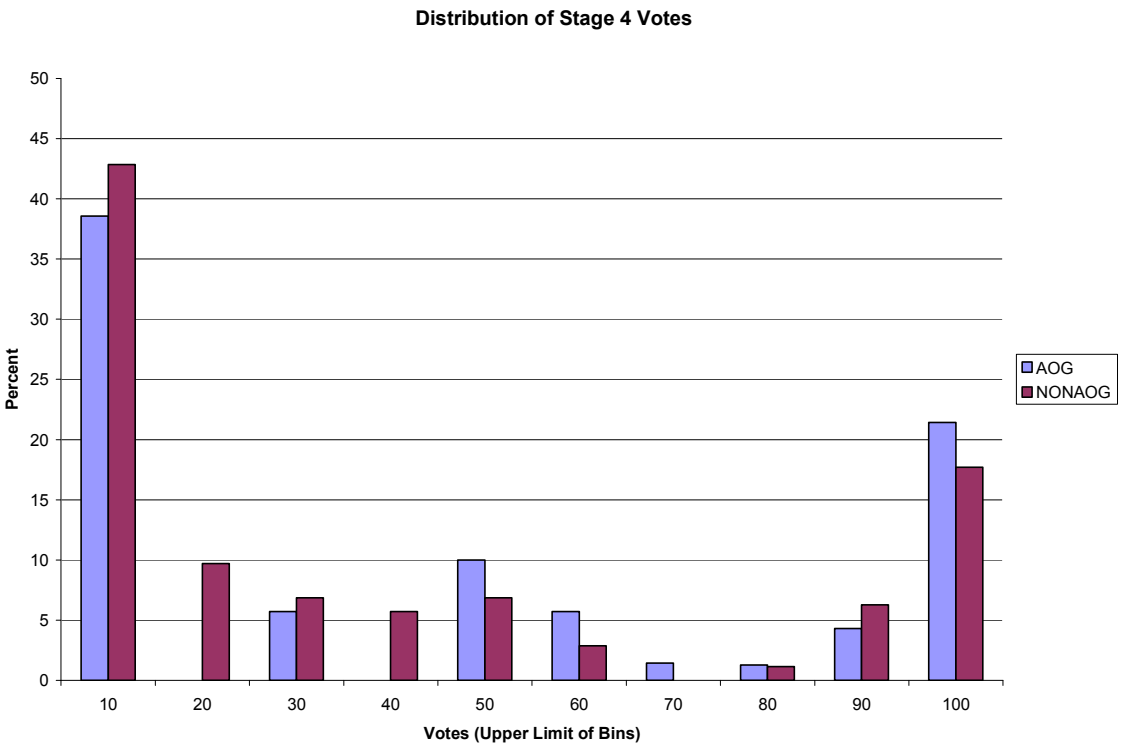

Distribution of Stage 4 Votes
Percent
50
45
40
35
30
25
20
15
10
5
0
10
20
30
40
50
60
70
80
90
100
Votes (Upper Limit of Bins)
AOG
NONAOG

FIGURE 6

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the John and Hallie Quinn Eminent Scholar Chair, the Charles Koch Foundation, and the National Science Foundation (SES-0451970). We would like to thank participants at the University of Arizona conference in honor of Vernon Smith, the 2008 Conference on Social Dilemmas (held at Florida State University) and the 2008 APEE and SEA meetings for many helpful comments. We are indebted to Anna Gunnthorsdottir, Laura Razzolini, and Andrew Smyth for many detailed comments and to Glenn Dutcher for the programming, using Z-Tree (Fischbacher [2007]), and to Sean Collins for assistance in conducting the experiments. We are alone responsible for any remaining errors.

REFERENCES

Ahn, T.K., R. Mark Isaac, and Tim Salmon. "Endogenous Group Formation," Journal of Public Economic Theory 10 (2008): 171- 194.

Andreoni, James. "An Experimental Test of the Public Goods Crowding-Out Hypothesis." American Economic Review 83 (1993): 1317-1327.

Bergstrom, Theodore, Lawrence Blume, and Hal Varian. "On the Private Provision of Public Goods." Journal of Public Economics 29 (1986): 25-49.

Capra, C. Monica, Tomomi Tanaka, Colin F. Camerer, Lauren Munyan, Veronico Souero, Lisa Wang, and Charles Noussair, "The Impact of Simple Institutions in Experimental Economies with Poverty Traps." Emory University, (2008).

Chan, Kenneth, Rob Godby, Stuart Mestelman, and R. Andrew Muller. "Crowding-out Voluntary Contributions to Public Goods." Journal of Economic Behavior and Organization 48 (2002): 305-317.

Cinyabaguma, Matthias, Talbot Page and Louis Putterman. "Cooperation Under the Threat of Expulsion in a Public Goods Experiment." Journal of Public Economics 89 (2005): 1421-1435.

Davis, Douglas D. and Charles A. Holt, *Experimental Economics* (Princeton, NJ: Princeton University Press, 1993).

Eckel, Catherine, Philip Grossman, and Rachel Johnson. "An Experimental Test of the Crowding out Hypothesis." Journal of Public Economics 89 (2005): 1543-1560.

Espinola-Arrendondo, Ana and Fleix Munoz-Garcia. "The Importance of Foregone Options: Generalizing Social Comparisons in Sequential-Move Games." Working Paper. Department of Economics, University of Pittsburgh. (2007).

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

Gruber, Jonathan and Daniel Hungerman, "Faith Based Charity and Crowd Out During

the Great Depression." Journal of Public Economics. 91 (2007):1043-1069.

Harrison, Glenn W. and John A. List. "Field Experiments" Journal of Economic Literature 42 (2004): 1009-1055

Holcombe, Randall, and Lawrence Kenny. "Evidence on Voter Preferences From Unrestricted Choice Referendums." Public Choice 131.1-2 (2007): 197-215.

Hungerman, Daniel. "Are Church and State Substitutes? Evidence From the 1996 Welfare Reform." Journal of Public Economics 89 (2005): 2245-2267.

Hungerman, Daniel. "Diversity and Crowd-Out: A Theory of Cold-Glow Giving." NBER Working Paper No. 13348. (2007).

Isaac, R. Mark and Duncan James. "Just Who Are You Calling Risk Averse?" Journal of Risk and Uncertainty 20 (2000).

Isaac, R. Mark, Svetlana Pevnitskaya, and Tim Salmon, "Do Preferences for Charitable Giving Help Auctioneers?" Forthcoming (currently available online), Experimental Economics (2009).

Isaac, R. Mark, Kenneth F. McCue and Charles R. Plott. "Public Goods Provision in an Experimental Environment." Journal of Public Economics 26: 51-74. (1985)

Isaac, R. Mark and Vernon L. Smith. "In Search of Predatory Pricing." Journal of Political Economy 93 (1985): 320-345.

Isaac, R. Mark, James Walker, and Susan Thomas. "Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations." Public Choice 43 (1984): 113-149.

Isaac, R. Mark, and James Walker. "Communication and Free-Riding Behavior: The Voluntary Contribution Mechanism." Economic Inquiry 26 (1988): 585-608.

Kim, Oliver and Mark Walker. "The Free Rider Problem: Experimental Evidence." Public Choice 43 (1984): 3-24.

Kosfeld, Michael, Akira Okada, and Arno Riedl. "Institution Formation in Public Goods Games." American Economic Review 99 (2009):1335-1355.

Lynch, Michael, Ross M. Miller, Charles R. Plott, and Rusell Porter. "Product Quality, Informational Efficiency, and Regulation in Experimental Markets." R. M. Isaac, ed., Research in Experimental Economics, Vol. 4. (Greenwich, Conn.: JAI Press, Inc.) (1991).

Norton, Douglas A. "Incentives Matter Less: Standing at the Crossroads of the Coercive

State and Transforming Preferences." Working Paper, Florida State University (2008).

Payne, A. Abigail. "Does the Government Crowd Out Private Donations? New Evidence from a Sample of Non-Profit Firms." Journal of Public Economics. 69 (1998): 323-345.

Saijo, Tasuyoshi and Hideki Nakamura. "The 'Spite' Dilemma in Voluntary Contributions Mechanism Experiments," *Journal of Conflict Resolution* 39 (1995): 535-560.

Smith, Adam. *The Theory of Moral Sentiments*. (London: A. Millar, 1790 {1st ed. 1759}). Reprinted in The Library of Economics and Liberty, Liberty Fund, www.econlib.org/Library/Smith/smMS1.html .

Smith, Vernon, Gerry L. Suchanek, and Arlington W. Williams. "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets." Econometrica 56 (1988): 1119-1152.

Sutter, Matthias, and Hannelore Weck-Hannemann. "An Experimental Test of the Public-goods Crowding-out Hypothesis When Taxation is Endogenous." FinanzArchiv 60 (2004): 94-110.

APPENDIX I: SPECIAL SUBJECT POOL SESSIONS

The results from the five baseline and five control sessions were different than we had expected. The pattern was idiosyncratic, suggesting that group composition and/or path dependence were determining factors. In particular, we did not observe any session with robust tendencies towards reverse crowding out.

We became convinced that issues of the level of group expectations might be an important additional factor. Because these expectations could be different when home-grown among naturally-occurring groups rather than groups created randomly in the laboratory, we decided to investigate our design with subjects drawn from a naturally occurring intentional community with slightly different incentives to generate strong common values and expectations.

To do this, we recruited enough subjects for two sessions from a campus ministry community of students. The experiments were otherwise conducted as already noted, with the following exception. We announced that at the end of the experiment we would match each person's earnings with a contribution to an orphanage in Guatemala. The chosen campus ministry has been active in supporting the orphanage, and it is likely that every subject in the experiment either had participated in a mission project at the orphanage, knew someone who had participated in a mission project at the orphanage, or had plans to go on a mission project to the orphanage.

We emphasize that we did not tie payments to the orphanage to the provision of the experimental group exchange, but rather to the *total* earnings of the subjects. This is because in these sessions we were not interested in home-grown preferences *per se*; rather, we were interested in home-grown expectations.

If we had tied payments to the orphanage to the provision of only the group good, then we would have substituted the subjects' home-grown preferences for the orphanage for the induced preference structure of our experimental design. Suppose we had found, for example, that contributions to the public good had increased. This could quite logically have been interpreted as deriving from a de facto increase in the "own marginal per capita return" for the public good. But the fact that increases in the MPCR drive behavioral increases in provision of the public good has been known since the 1980s.

Instead, by tying payments to the orphanage to total earnings, at one level the incentive structure of our design is preserved. A subject wanting to provide more of his funds to the orphanage would have to consider the tension between the stage game dominant strategy and the social optimum of full group contributions. But our design had the potential to change the incentive structure of the game in other ways.

Consider the choice of an individual in the stage game. If he increases his contribution to the group exchange by one token, the following happens:

His personal token earnings decrease by .5 tokens
His personal contributions to the orphanage decrease by .5 tokens

At this point, incentives from the contributions to the orphanage exactly mirror those of the incentives in cash from the existing game: there is a tension between individual

maximization and group maximization. However, this person should also consider how his additional token affects contributions to the orphanage via other subjects' earnings. Suppose we adopt the natural conjectural variation that this subject believes that the *token allocations of other subjects do not change.* Then, in addition, this subject should calculate that increasing his contributions to the group exchange by one token increases contributions to the orphanage via others' earnings by a total of 3 tokens.

So, there is a well defined trade-off of a decrease in personal earnings of .5 tokens with an increase in token earnings for the orphanage of 2.5 tokens, assuming that the individual values tokens contributed to the orphanage the same regardless of whether they come from him or from the other individuals.[19] As long as this person values the contributions to the orphanage "enough" the stage game admits a dominant strategy to contribute to the public good.[20] But what could go wrong with this scenario?

One, this dominant strategy in the stage game could fail if an individual does not value contributions to the orphanage "enough."

Two, this dominant strategy in the stage game could fail if individuals value contributions that they make much more than contributions that others make. (This is essentially the same distinction as the "Basic Charity" vs. "Bonus From Winning" preferences in Isaac, Pevnitskaya, and Salmon, 2009).

Three, there could be unanticipated attributes of the incentives beyond the one-period stage game. Unlike in the standard VCM, subjects in this experiment carry over from period to period a portfolio of their own earnings, their own contributions to the orphanage, and the total contributions to the orphanage. Preferences over these items may not be separable. This may allow for the existence of more complicated equilibria. In a limit, the Nash/Cournot assumption on other subjects' tokens allocation might fail.

Thus, while there is good reason to believe that this design alters incentives towards greater contribution, it is by no means certain and is therefore testable. What we valued in these subjects was not their home-grown preferences for contributions to the orphanage but rather their common expectations that everyone else in the room also had strong reasons to want to maximize the total earnings of the group, and that they had strong reasons to believe that everyone else knew this to be true, and so forth. We believed that all of the conditions of a dominant strategy in the stage game above are likely to be met, and we believed that this was known with something approaching common expectations.[21]

The incentives that we induced in the lab are strong, but they are ones that are may be overlooked (compared with standard models) when economists consider how naturally occurring groups actually make decisions about public goods. When there have been unexpectedly large amounts of cooperation in a standard VCM experiment, research conjectures have centered either on models that essentially retain the concept of the public good as an instrument to individual maximization (reciprocity) or on models that

---

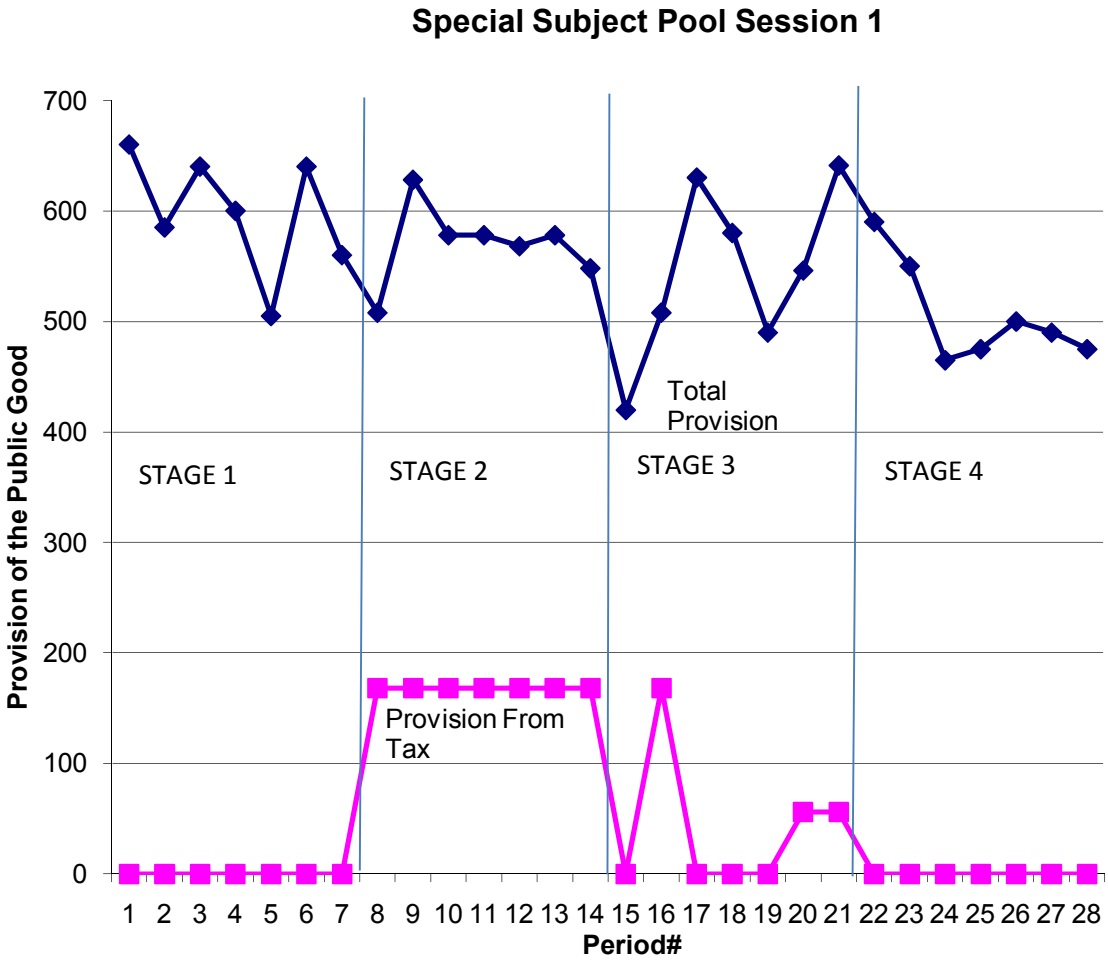
[19] The trade off between a decrease in personal earnings of .5 and an increase in the earnings for the orphanage of 2.5 tokens is rooted in the Stage 1 and VCM portion of the experiment. For the stages with taxation the tradeoff is different for the public good provided through the tax. The trade off is .5 to 1.8 tokens because of the tax efficiency parameter.

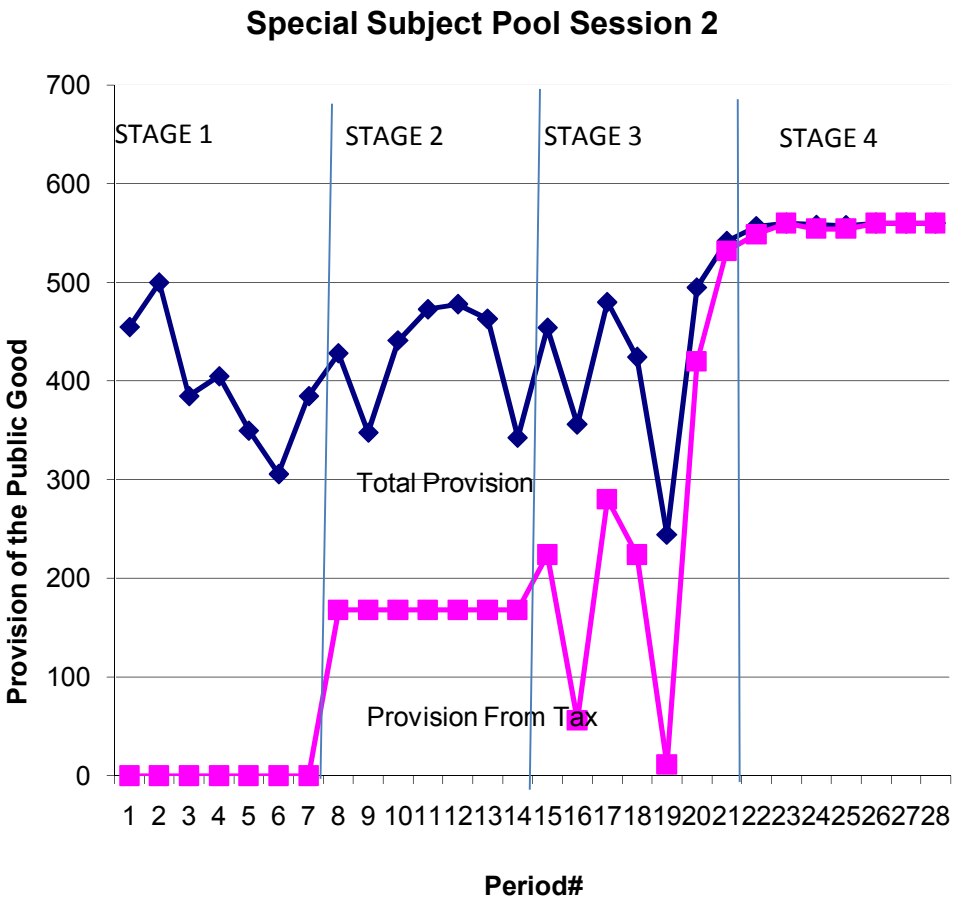
[20] Recall that because these are the home grown preferences of the subjects, they need not be additively separable, so "enough" may be a more complex relationship than a ratio of one value to another.

[21] In fact, the Isaac, Pevnitskaya, and Salmon paper uses the same subject pool and concludes that their choices in charity auctions arenot well explained by the "See and Be Seen" model.

presume that these strangers care about each others' earnings (altruism). What we propose here is a model in which individuals' preferences and expectations are aligned in a way so that everyone recognizes that it is in everyone's interests to provide the public good. While this does not rule out altruism or reciprocity, we believe that it is a distinct social phenomenon.

The results of these two sessions are striking. In the two sessions with the special subject pool and the altered incentives, we observed one each of the two the boundary outcomes: a high-tax outcome and a plausible "reverse crowding out" outcome that we had failed to see in the other 10 sessions. In Session 1, the tax system was "crowded out" and replaced by a purely voluntary regime of substantial, stable, although not quite optimal, levels of contribution. In Session 2, on the other hand, the subjects embraced the tax system and ratcheted the tax level up to the full tax regime. (Recall that the assignment of the two groups was made randomly by the computer from the 14 participants). Charts of tokens to the group and of the tax levels are presented below.




Special Subject Pool Session 2
700
600
500
400
300
200
100
0
Provision of the Public Good
STAGE 1
STAGE 2
STAGE 3
STAGE 4
Total Provision
Provision From Tax
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28
Period#

APPENDIX II: BASELINE INSTRUCTIONS

INTRODUCTION

This is an experiment in the economics of decision making. If you follow the instructions carefully you could earn a considerable amount of money which will be paid to you in a check at the end of the experiment. In today's experiment, all of your earnings on the computer will be denominated in experimental dollars, which will be translated into U.S. dollars at the end of the experiment. The translation rate today will be that 400 experimental dollars equals one U.S. dollar. From this point on, everything on your screen is your own private information and there should be no communication with any other participant except as you are directed in these instructions.

You are one of seven persons in a group. The decision that you must make during each decision period is how to divide a stock of tokens between an individual and a group exchange. There are four stages in today's experiment, each consisting of seven periods, for a total of 28 periods. Because the rules of the exchanges are somewhat different, we will go over first the rules for the first stage (the first seven periods). We will continue with the instructions for the second stage and the end of the first seven periods and so forth.

INSTRUCTIONS FOR STAGE 1

In stage one, each of you receives, at the beginning of each period, 100 tokens. You must decide how to invest the tokens: either in your own Individual Exchange, or in a Group Exchange. The tokens earn money differently in the two exchanges, so we will go over them in turn.

The individual exchange: Each token that you invest in your own individual exchange earns exactly 1 experimental dollar for that period.

The group exchange: The rule for earnings in the group exchange is different. After each of you has made your investment decision, the computer sums all of the tokens that each of you has invested in the group investment exchange. It then pays each one of you one-half times the number of summed tokens invested in the group exchange *regardless of how many tokens you have invested.* For example, if the seven of you invest a total of 200 tokens in the Group Exchange, each one of you would earn 100 experimental dollars from the Group Exchange. If the seven of you invested a total of 500 tokens in the Group Exchange, each of you would earn 250 experimental dollars from the Group Exchange, and so forth. In either of these two examples, your total earnings would equal your earnings from the Group Exchange plus your earnings from tokens invested in your Individual Exchange. We have provided a payoff chart for other values of possible investments in the Group Exchange.

You must use all of your tokens in each period, that is, the number of tokens invested in your Individual Exchange added to the number of tokens you invest in the Group Exchange must equal 100. You may not carry tokens over from one period to another.

If you will now look at your computer screen, you will see that we are displaying a practice round in which you can familiarize yourself with the process of investing your tokens. This practice round has no impact at all on your final earnings. Please go to your screen and try making different types of investments. When you are done, please hit the CONTINUE button.

If you will look at your screen now, you will see the results for this practice round. The screen shows you

Your earnings from your individual exchange
The total number of tokens invested in the group exchange
Your earnings from the group exchange
Your total earnings

To summarize, tokens invested in your Individual Exchange earn a certain return of one experimental dollar. Earnings from the Group Exchange depend upon the total of the investments in the Group Exchange from your group. Are there any questions?

Finally, please notice that we are handing each of you a sealed envelope. This contains some additional information for later in the experiment. Please do not open the envelope before we instruct you to do so.

INSTRUCTIONS FOR STAGE 2

The earnings rules of the Individual Exchange and the Group Exchange are the same in Stage 2 as they are in Stage 1. However, in this stage we begin each round with a computer enacted tax of 30 tokens on each one of you. Of this tax, 80 percent will be invested in the group exchange for you, while 20 percent will disappear. In other words, of the 30 tokens that are taken from you as a tax, 24 will be invested in the group exchange, 6 disappear. That means that across all seven of you, 210 tokens will be taken as a tax, and 168 tokens will be automatically invested in the group exchange.

At the beginning of each period, each of you then has the decision of how to invest your remaining 70 tokens. As before, you may divide them between your Individual Exchange and the Group Exchange. Your Individual Exchange continues to earn 1 experimental dollar per token, and the group exchange continues to earn one-half times the sum of all tokens invested in the group exchange. Remember that this sum includes the 168 tokens automatically invested from the tax plus any additional amount that you decide to invest.

On the next page of the handout, you will see a screen print of how your token investment screen will look.

Are there any questions?

INSTRUCTIONS FOR STAGE 3

In the next stage, the rules for investment in the individual and group exchanges will be the same as in Stage 2 except that the tax will no longer automatically be 30 tokens. Instead you and the other six members of your group will determine the tax for your group. This will be done by a voting process. As before, whatever tax you as a group choose, 80 percent will be invested in the group exchange and 20 percent will disappear.

Here is how the voting will proceed for you to choose your tax rate. You will submit to the computer what level you would like the tax to be. The computer will rank all the nominations from highest to lowest and choose the median, or middle, amount. In your case, this will be the 4th highest number. That tax rate will be the one chosen for all seven members of your group.

As an example, suppose that your group nominated the tax levels of : 98 tokens, 77 tokens, 60 tokens, 40 tokens, 39 tokens, 18 tokens, and 0 tokens. Then the winning tax would be 40 tokens because it is the median, or 4th highest. (Notice that 40 is NOT the average). Using the same rule as before, if the tax is 40 tokens then 80 percent (that is, 32 tokens) would be invested in the group exchange from each of you, for a total investment of 224 tokens.

As before, after the tax has been determined and you have been informed of its value for this period, you will have the opportunity to make an additional investment decision with your remaining tokens. You will have the opportunity to vote on a new tax level in each period in this stage.

Investment in your individual exchange continues to earn one experimental dollar per token. Investment in the group exchange continues to earn for each person .5 * (the sum of all tokens invested in the group exchange) including the 80 percent of the total from the tax.

On the following page is a print of the screen in which you are asked to submit your preferred tax level.

Are there any questions?

INSTRUCTIONS FOR STAGE 4

In Stage 4, all of the rules are the same as in stage 3. The only difference is that some of you will have a different return from the individual exchange than you did before, while others of you will continue to earn one experimental dollar per token invested in your individual exchange. To the extent that some participants' values have changed, the identities of these participants were chosen anonymously and at random by the computer before any of your decisions were made. You may now open the envelope we gave you at the beginning of the experiment which contains a piece of paper telling you what your return from the individual exchange will be for the next seven periods. Please do not share this information with anyone.

The rules for choosing the group's tax level do not change.

The rule that the tax will be 80 percent invested in the group exchange does not change.

The payoff to each person from the group exchange stays at one-half times (the sum of all tokens invested in the group exchange).

Are there any questions?

APPENDIX III : INSTRUCTIONS FOR STAGE 4 IN CONTROL SESSIONS

In today's experiment, the rules for Stage 4 are exactly the same as they were for Stage 3. The rules for voting on the tax are the same. The rule that 80 percent of the total tokens from any tax are invested in the group exchange remains the same. The rule for calculating your earnings on your individual exchange (one experimental dollar per token) has not changed. The rules for calculating earnings from the group exchange have not changed (.5 * the sum of all tokens invested in the group exchange, including the 80 percent of the total tokens from the tax). Are there any questions?

APPENDIX IV: ANNOUNCEMENT FOR SPECIAL SUBJECT POOL

In today's experiment, in addition to paying you your final earnings from today's experiment, the experimenters will also make a contribution to the New Life Children's Home in Guatemala equal to each of your final earnings in today's experiment (excluding the $10 show up fee). That is, to use some arbitrary examples, if you make $1.00 in earnings in the experiment, you will receive that $1.00 in earnings, *and* we will make a contribution of $1.00 to New Life Children's Home. If you make $5.00 in earnings in the experiment, you will receive that $5.00 in earnings *and* we will make a $5.00 contribution to New Life, and so forth for any amount of earnings that you make.

Please note that this applies to the total earnings you make in the experiment. This specifically includes your earnings from the group account and your earnings from your individual account. We match your *total* earnings at the end of the experiment (excluding the show-up fee). Another way of saying the same thing is that at the end of experiment we will add-up each of your earnings (excluding show-up fees) and write a single check to New Life Children's Home for that amount. You will be shown this check.

New Life Children's home operates a home in Guatemala. From their web-page the following information is obtained: "NLCH is truly a 'home' for these needy children, where their physical, emotional, and spiritual needs are addressed." Each child is provided a "sound academic Christian education." Members of the [the campus ministry] of FSU have participated in mission trips to the New Life Children's home in Guatemala.

Are there any questions?