ELSEVIER

Journal of Economic Behavior & Organization
Vol. 37 (1998) 315–331

JOURNAL OF
Economic Behavior
& Organization

An experimental test of the crowding out hypothesis: The nature of beneficent behavior

Gary E. Bolton*, Elena Katok

Department of Management Science & Information System, Smeal College of Business Administration, 310 Beam, Penn State University, University Park, PA 16802, USA

Received 11 May 1996; received in revised form 13 January 1998; accepted 11 February 1998

Abstract

An extensively studied model of public goods provision implies that government donations to charity crowd out private donations dollar-for-dollar. Field studies fail to verify this result. Several analysts argue that the problem lies with the specification of donor preferences. We report on a new experiment that provides a direct test of donor preferences free of the strategic factors that can confound tests in the field, and in other experimental settings. Our method involves the dictator game. We find extensive but incomplete crowding out – direct evidence that donor preferences are incorrectly specified by the standard model.

JEL classification: C72; C91; H30; H41

Keywords: Public goods; Crowding out; Experiment

1. Introduction

The standard model of public goods provision assumes that charitable giving is motivated by 'a desire to improve the general well-being of recipients' (Becker, 1974). One of the well-known, and to some, disturbing implications of this model is that government donations to charity, funded by lump-sum tax transfers, should crowd out private donations dollar-for-dollar. Field studies fail to verify this result. In fact, they generally find that crowding out is quite small.

One hypothesis, advanced by several analysts, argues that the standard model misspecifies the donor's preferences (or motives) for giving. Andreoni (1989) makes a

* Corresponding author. Tel.: +1 814 8650611; fax: +1 814 8632381; e-mail: geb3@psu.edu

0167-2681/98/$ – see front matter
PII: S0167-2681(98)00091-2

useful distinction. *Pure altruism* refers to the motive posited by the standard model: Improving the well-being of recipients. *Impure altruism* refers to being motivated by the act of voluntary gift giving – giving to do the 'right thing.' Basically (we elaborate in Section 2), the donor preference critique argues that the presence of impure altruism prevents complete crowding out.

It may be practically impossible to test for impure altruism on the basis of field data alone, because in the field, preferences can be observed only indirectly, and are easily confounded with uncontrollable phenomena. The laboratory offers a controlled testing environment. But even in the laboratory, the public goods games most closely associated with the crowding out issue raise obstacles to a clean test. In particular, donor preferences are subject to confounding with player expectations along with other strategic considerations. As we explain in Section 2, previous experimental studies have looked at the crowding out issue, but are inadequate to distinguish between pure and impure altruism.

In this paper, we report on a laboratory experiment that tests for impure altruism using the *dictator game*. In the dictator game, one individual (the donor) decides how to distribute a sum of money between him-or-herself and one other (the recipient). The dictator game is an ideal vehicle for studying the question of what motivates people to redistribute income to others. To begin with, previous dictator studies have demonstrated that a substantial number of people do in fact give to the recipients.[1] Second, the 'game' label is a misnomer; the dictator game is really a one-person decision task. Consequently, there are no strategic factors to confound our reading of the donor's preferences. A paper closely related to the charitable giving literature, Hochman and Rodgers (1969), used what is essentially a dictator game to theoretically examine preferences for income distribution.

The new experiment provides a test of pure versus impure altruism unencumbered by either strategic confounding or a priori restrictions on admissible preference profiles. We find direct evidence that donor preferences have an impure altruism component. Before introducing the new experiment, we provide a more detailed discussion of the issues and literature.

2. Background issues and literature

The standard public goods model assumes that people have preferences over two entities: own consumption and a public good. These are taken to be imperfect substitutes for one another, so income is generally allocated for both. It has been shown by several analysts using various versions of the model that government donations to the public goods, funded by lump-sum taxes, should crowd out private donations dollar-for-dollar.[2] The argument is compelling in its simplicity: Since an individual cares only about the final allocation between private and public consumption, he or she should be indifferent to whether the allocation results from contributions in the form of a voluntary private gift

[1] See, for example, Forsythe et al. (1994); Hoffman et al. (1994); and Bolton et al. (1997b).
[2] These include Warr (1982); Roberts (1984); Bergstrom et al. (1986).

or from an involuntary tax transfer. We might then expect a person to neutralize a new tax transfer by subtracting the tax amount from his private gift. In fact, if each person expects that all others will act to neutralize, he is effectively presented with the same choice set after tax transfer as before, making neutralizing the optimal action; dollar-for-dollar crowding out turns out to be the (unique) equilibrium.

The argument we sketch assumes an individual tax transfer not greater than the individual pre-tax gift, but Sugden (1982) shows that the model implies nearly complete crowding out even without this assumption. Bernheim (1986) shows that the argument can be extended to certain distortionary taxes. Bernheim and Bagwell (1988) show that government transfers can have neutral effects on private goods as well.

Field studies of charitable giving, however, find that the amount of crowding out is actually quite small, with estimates ranging from 5 to 28 percent.[3] There are several possible explanations for the inconsistency. A variety is discussed by Sugden (1982, 1984). The explanation that has received the most attention is impure altruism, the argument that people receive some value from the act of voluntary gift-giving. Public transfer is an imperfect substitute for voluntary giving, and hence, if giving is motivated by impure altruism, crowding out would be incomplete. We refer to this explanation for incomplete crowding out as the *donor preference critique*.[4]

2.1. Testing for complete crowding out in the laboratory

Several previous experiments have looked at issues surrounding altruism.[5] An experiment by Andreoni (1993) focused specifically on the crowding out issue, and it will help us frame some of the experimental design issues in distinguishing pure altruism from impure altruism. There were two treatments. In the *no tax* treatment each player in a three-person public goods game was allocated seven tokens, and had to decide how many to contribute to a public good. The payoff matrix is reproduced here in Fig. 1. The matrix was generated from the utility function,

$$U_i = A[(w - g_i)^{1-\alpha} G^{\alpha}]^{\gamma} \tag{1}$$

where g_i is the number of tokens contributed by i, $G=g_1+g_2+g_3$, and A, w, α, and γ are constants fixed by the experimenter. The standard Nash equilibrium – based on the assumption that subjects care solely about own consumption – has each subject contributing three tokens. Essentially the same game was played in the *tax* treatment, except that each player was now constrained to contribute at least two tokens, meant to simulate a two-token lump-sum tax. The standard Nash equilibrium now has each player *voluntarily* contributing one token.

[3] Abrams and Schmitz (1978, 1984); Clotfelter (1985); Kingma (1989); Ribar and Wilhelm (1995).

[4] Margolis (1982); Cornes and Sandler (1984); Stark (1985); Steinberg (1987); Palfrey and Rosenthal (1988); Andreoni (1989). Also see Ledyard (1995).

[5] Cadsby and Murray (1991) report on an experiment that tests Ricardian equivalence, a hypothesis which is in theory, closely related to the charitable crowding out issue. Dawes and Thaler (1988) survey various experiments having to do with characterizing altruism; also see Palfrey and Prisbrey (1995); Cooper et al. (1996); Offerman et al. (1996); Bolton et al. (1997a).

No Tax Game

tokens played

total tokens played by the other two players

	0	1	2	3	4	5	6	7
0	0	1	3	6	9	10	11	10
1	1	4	8	11	14	15	15	14
2	5	9	14	18	20	21	20	17
3	12	17	22	26	28	28	25	22
4	21	28	33	36	37	35	32	27
5	34	40	45	48	47	44	39	32
6	49	56	60	61	59	54	47	38
7	68	74	77	76	72	64	55	44
8	90	95	96	93	86	76	64	51
9	115	118	117	111	102	89	74	58
10	143	144	140	131	119	103	85	66
11	175	173	166	153	137	118	97	75
12	210	205	193	177	157	134	109	84
13	248	239	223	203	178	151	122	93
14	290	276	256	230	201	169	136	103

Tax Game

tokens voluntarily played

total tokens voluntarily played by the other two players

	0	1	2	3	4	5
0	33	36	37	35	32	27
1	45	48	47	44	39	32
2	60	61	59	54	47	38
3	77	76	72	64	55	44
4	96	93	86	76	64	51
5	117	111	102	89	74	58
6	140	131	119	103	85	66
7	166	153	137	118	97	75
8	193	177	157	134	109	84
9	223	203	178	151	122	93
10	256	230	201	169	136	103

Fig. 1. Andreoni (1993): game payoff matrices (payoff in cents).

The experiment tests the crowding out hypothesis under the assumption that the game induces preferences for a public good. The argument is that, if subjects enter the lab caring only about their own monetary payoff from the experiment, then their preferences over game outcomes will be equivalent to those implied by (1). Comparing the equilibria for the two games, one defines total crowding out as voluntarily contributing an average of two tokens less in the tax game than in the no-tax game. Using this benchmark, Andreoni found crowding out to be incomplete.

Andreoni makes clear that impure altruism is not the only explanation for his results. One reason other explanations cannot be ruled out has to do with the assumption, implicit in the experiment's design, that laboratory subjects care only about their own consumption. As we noted, dictator game studies indicate that lab subjects do at least sometimes give gifts.[6] This behavior per se is not inconsistent with pure altruism, and hence not inconsistent with the standard public goods model. It turns out that when we permit the possibility that subjects enter the laboratory with purely altruistic preferences, it is unclear whether we should actually expect complete crowding out. To illustrate with a very simple example, consider the utility function

$$U_i = \min[x_i, P_i/2] \tag{2}$$

where x_i is i's own monetary payoff, and the public good P_i is the sum total monetary payoff for the *other* players. These pure altruism preferences have the key characteristic that a tax transfer is a perfect substitute for a voluntary gift, conforming to the standard public goods model.[7] Inspection of Fig. 1 verifies that if all players have preferences given by (2), then each player identically (and voluntarily) contributing either 3, 4, or 5 tokens is a symmetric Nash equilibrium in both no-tax and tax games. This is consistent with any crowding out level from -2 to 2 tokens, depending on the combination of equilibria in play. So even though the preferences conform to the standard public goods model, the crowding out hypothesis may fail to apply.[8]

The reason the hypothesis may not apply can be gleaned from Fig. 1: if subjects come to the lab with some (pure) altruism motive, an additional token played is not always a monetary transfer *from* private consumption *to* the public good; sometimes an additional token increases both own consumption as well as that of others. Absent the preference equivalence implied by (1), the two-token shift on which the experiment pivots is not equivalent to a tax transfer. In sum, while the experiment produces data inconsistent with the straightforward application of the total crowding out argument, strategic considerations prevent us from saying whether the motive for contributions is pure or impure altruism.

---

[6] In fact a very recent model by Bolton and Ockenfels (1997) treats this type of giving behavior as the basis for a theory of behavior observed in a wide variety of laboratory games.

[7] Use of the Leontief utility allows us to make our point with a transparent example, although a standard public goods model usually assumes that utility is strictly increasing in each argument. Inspection will probably convince the reader that our examples are robust to this consideration. More technically, consider the CES utility function,

$$V_i = (x_i^{\rho} + (P_i/2)^{\rho})^{1/\rho} \tag{3}$$

Function (3) converges point-wise to (2) as ρ goes to $-\infty$ (see Varian (1984), pp. 30–31). Taking ρ small enough, and substituting (3) for (2), all our examples remain valid. In addition, there is some ambiguity as to the correct way to define the public goods, but see Footnote 8.

[8] For example, if the 3-token equilibrium is played in both the tax and no-tax games, there is a crowding out of 0 tokens; if the 3-token equilibrium is played in the no-tax game and the 5-token equilibrium in the tax game, there is a crowding out of -2 tokens. Another example: define P_i as the sum of *all* player payoffs (i.e., as the joint payoff) and substitute $P_i/3$ for $P_i/2$ in Eq. (2) of the text. If all subjects have these preferences, then voluntarily contributing 3 tokens is a symmetric Nash equilibrium in both tax and no-tax conditions, consistent with a crowding out level of zero.

3. The new experiment

The design of the new experiment is based on two considerations. First, the donor preference critique is a *non*-strategic argument concerning impurely altruistic preferences. The discussion of the Andreoni experiment illustrates that, when studied in a strategic context, preferences can be confounded with other issues. The dictator game provides a test of pure versus impure altruism in a context where the strategic aspect of the public goods model is stripped away. The issue reduces to whether the donor neutralizes a forced re-distribution when the donor is the sole contributor.

Second, the driving distinction between the two hypotheses we wish to test concerns whether initial distribution matters. Pure altruism implies that people care solely about the final distribution of consumption, whereas impure altruism implies that the initial distribution also matters (the initial distribution determines the feasible set of *voluntary* gifts). The new experiment focuses on determining whether changes in the initial distribution of money influences giving in the dictator game.

The new experiment had two treatments. In both, subjects were pair-wise matched. In the 15-5.*a* treatment, $15 was distributed to the subject in role A (the donor role) and $5 was distributed to the subject in role B (the recipient role). A was then given the opportunity to redistribute some of A's money to B. The 18-2 treatment was identical, save for the initial distribution, now $18 to A and $2 to B.

The pure altruism hypothesis predicts that the additional $3 allocated to the recipient in the 15-5.*a* treatment should crowd out up to $3 of giving per donor relative to the amount given in the 18-2 treatment. Specifically, a donor who would give $3 or more in the 18-2 treatment, should give $3 less in the 15-5.*a* treatment, while a donor who would give less than $3 in the 18-2 treatment should give nothing in the 15-5.*a* treatment. We test this against the impure altruism implication that crowding out is incomplete.

3.1. Laboratory protocol

The laboratory protocol includes a description of all lab procedures as well as written directions given to the subjects. The monitor read all written instructions to subjects directly from the protocol; the only monitor–subject communication not included are answers to individual subject questions. The complete protocol appears in the Appendix A. Here is a synopsis:

For each session, subjects were assembled in a single room, and separated into two groups, one on each side of the room. Subjects were seated so that they could not observe others' choices. Written directions described the dictator game and explained it would be performed twice. Each subject would be anonymously paired with someone from the other group, a different pairing for each game, and would have the role of A in exactly one game. After the monitor read the directions aloud and answered any questions, each subject wrote on a form how he-or-she wished to distribute the money for the game in which they had the role of A. A monitor checked the completed forms to see that the rules had been followed and that amounts added up correctly.[9] A coin was then flipped to choose one game for payoff. In order to preserve anonymity, those having the role of B for the chosen game were paid separately in another room.

All subjects were Penn State University students, and were recruited through billboards posted around the University Park campus. Participation required appearing at a special time and place. Cash was the only incentive offered. Each subject participated in a single session, and was paid his-or-her earnings from the game chosen for payoff.[10]

The experiment has a 'no feedback' design: The game is run twice, but no information about the play of one game is made available to subjects before the play of the other. We wanted a treatment with a 15-5 allocation to allow for a better comparison with previous work. Pre-experiment power simulations, based on a one-game procedure and financially feasible sample sizes (about 25, with 50 subjects) suggested we might not be able to reject the null hypotheses, even if false. Based on these considerations and further simulations, we chose the two-game, no feedback format, and an 18-2 allocation for the second treatment with a target sample size of 46 per treatment. In order to gauge the effect of the format, we ran the 15-5.*a* treatment first and compared with previous data. We then ran the 15-5.*b* treatment (discussed below), and this shows that the no feedback format has no discernible effect on dictator choices.

Our experiment involves no double blind procedure. The double blind hypothesis, that subjects behave differently when the experimenter can associate actions with individual identities, has now been subject to several tests, and most are negative. Roth (1995) summarizes much of the research, and suggests an alternative, non-double blind interpretation for what positive evidence there is. In addition, Laury et al. (1995) found no evidence for double blind in the context of a public goods game.

4. Results from the new experiment

Fig. 2 displays the data from the 15-5.*a* and the 18-2 treatments. Subjects were free to choose any dollar and cent amount. With one exception, however, all gave in whole dollars.[11] Observe that both distributions are bi-modal, with lower modes of 0 and upper modes at the amounts of giving that lead to donor and recipient each receiving \$10.

4.1. Explaining the shape of the distributions

We were struck by the fact that giving in Fig. 2 is noticeably higher than in some other dictator experiments reported in the literature (A stark example is presented in Section 5). There are two features that distinguish the present experiment. One is how we frame the game to subjects. It is well known that the frame can influence an experiment's results. The language used to describe the game is similar (but not identical) to that used by

[9] If not, and only if not, the task was privately re-explained to the subject, who then re-did the form. There were three such cases; in all of these the amounts did not add up correctly.

[10] In all sessions, the same monitor read the protocol and answered all subject questions. All sessions were run in the fall of 1994. Each lasted no more than 30 min. Average total earning was \$10.

[11] Forsythe et al. (1994) report the same phenomenon.

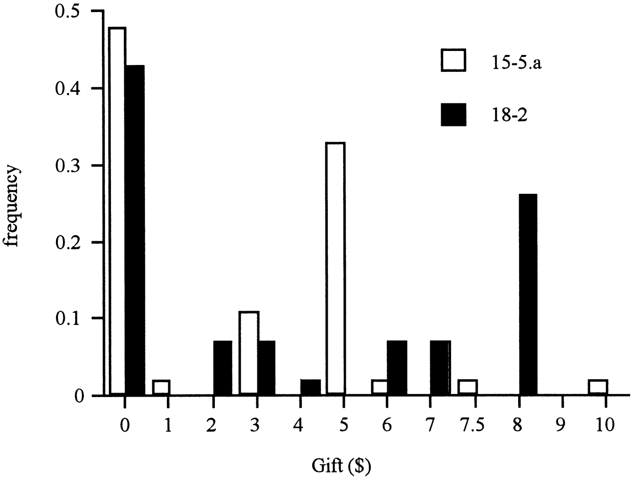


Fig. 2. Comparison of giving: distribution for 15-5.*a* (n=46) versus 18-2 (n=42).

Kahneman et al. (1986); Forsythe et al. (1994).[12] Comparison of their results with other dictator game studies suggests that this language produces somewhat higher generosity, although not quite as much as in our game. The second feature is that subjects did not know whether their decision would count in the payoff until after they had completed the game. In most other studies, the donor role was established before a decision was made. If this second feature were the source of the observed generosity, it might indicate that the coin flip lessens the salience of the game.[13] We were particularly concerned by this second explanation, since it would impinge on our interpretation of giving as a beneficent act.

We tested for the influence of the coin flip with a third treatment, 15-5.*b*, in which gift decisions were made *after* the coin flip determining subject roles; otherwise, the treatment design, and in particular the frame of the game, were the same as in 15-5.*a* (the Appendix A gives the exact changes). Fig. 3 compares the results. Note that the bi-modal nature of the data is completely unchanged; a contingency table test provides no evidence that the timing of the coin flip matters (p value of 0.6537).[14] We conclude that it is the frame of the game that distinguishes our data from that of some other studies. (Only the 15-5.*b* decisions chosen for payment are considered here, although including the unpaid decisions would not alter the conclusions.) The implications are discussed in the final section of the paper.

---

[12] Our instructions are similar in that we speak of the allocation being ‘provisionally distributed.’ The language is different in that all of our treatments explicitly allocate some sum to both players, rather than giving the entire sum to the dictator.

[13] Sefton (1992) reports some evidence that the probability of being paid is inversely related to donor generosity.

[14] The p value is the average of five 20 000 trial samplings of the actual distribution associated with the test statistic. This technique yields a more accurate p value than any table approximation of the statistic's distribution.

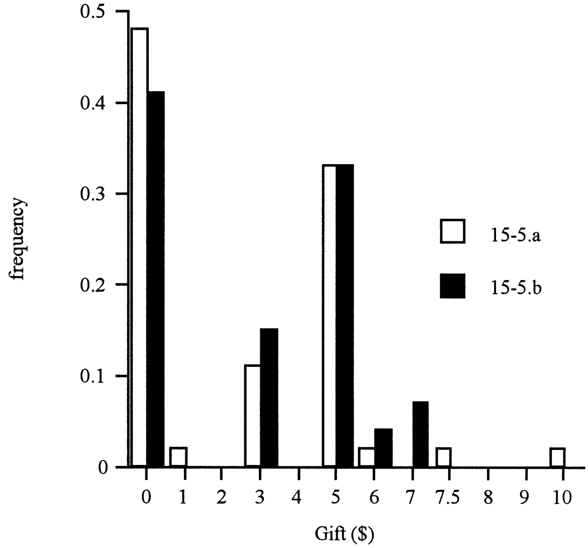


Fig. 3. Comparison of giving: coin flip after (15-5.*a*, n=46) versus before (15-5.*b*, n=27).

4.2. Evidence on the crowding out hypothesis

We divide testing into two steps. We first look for evidence that the $3 transfer crowds out *some* giving. Upon finding in the affirmative, we then examine evidence on whether the crowding out is *complete*. Due to the nature of the hypotheses, we focus on comparing the distribution means. The testing is complicated by the bi-modality, particularly that of the 18-2 data, because the accompanying high variances weaken test sensitivity. This consideration leads to the following approach. Both distributions can be expressed as

$$P(G=g) = p \times d(g,0) + (1-p) \times P(G=g \mid G>0)$$

where g is the size of the gift, and p is the frequency of a contribution of 0; $d(g,0)$ is equal to 1 when g=0 and equal to 0 otherwise. The mean of the distribution can then be expressed in terms of two components,

$$\mu[G] = (1-p) \times \mu[G \mid G>0]$$

where $\mu[G \mid G>0]$ is the mean among those in the donor role who actually give. Note that each component involves a single mode of the overall distribution. In this way, the overall variance of distribution is partitioned. Testing the separate components increases the sensitivity of the statistical analysis. It also provides specific information about the nature of any shifts in giving.

The effect of the high variances is apparent from a comparison of the means of the two distributions in Fig. 2. A test of the hypothesis that average giving is lower in 15-5.*a* than it is in 18-2, is only weakly significant (p value of 0.0712), in spite of the fact that a rather large shift is apparent from an examination of the figure. A contingency table test on the proportion of donors who give nothing, p, indicates no significant shift in the proportion of non-givers (p value of 0.6401). However, when we compare the means conditional on actually giving, we find strong evidence for crowding out (one tail p value of 0.0047). So while the proportions of non-givers and givers are not significantly different, those who do give tend to give less in 15-5.*a* than do those in 18-2. The $3 transfer then does lead to some crowding out.[15]

To test whether the crowding out is complete, we shift each gift, g, in the 18-2 data to the value max [0, g-3]. If crowding out is complete, we would expect the *shifted distribution* to have the same location as the 15-5.*a* distribution. We test this against the hypothesis that crowding out is incomplete. A test on the overall averages of the distributions is not quite able to reject equality (p value of 0.1078); nor is either component test significant (p values of 0.1302 for the conditional mean and 0.382 for the proportion p).

Recall, however, that we found no significant differences between giving in 15-5.*b* and 15-5.*a*, so it is appropriate to pool the 15-5 distributions to increase the power of the test.[16] The comparison of the pooled, shifted 15-5 distribution with that of 18-2 is displayed in Fig. 4. Testing the overall average using pooled data indicates a significant

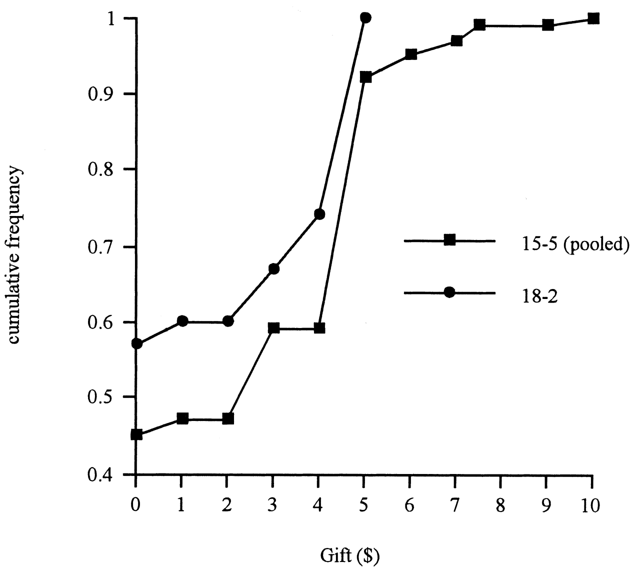


Fig. 4. Comparison of giving: shifted 15-5 distribution (pooled, n=73) versus 18-2 (n=42).

[15] In this case, considering the pooled 15-5 treatments in place of 15-5.*a* does not significantly affect p values.

[16] Another way to check whether the timing of the coin flip has an influence: Recall that we found no detectable statistical difference between 15-5.*a* and 18-2, the two treatments in which the coin was flipped at the end of the session. We might then pool these two samples, and test whether the pooled sample comes from a different population than 15-5.*b*. Of course, if the test rejects, we would not be able to say whether this was due to the influence of the coin flip, or to the difference in giving patterns between 18-2 and the other two treatments. But if we cannot reject, it gives us extra confidence that the coin flip makes no difference. In fact, a χ^2 test cannot reject (p=0.4232).

shift (one-tail p value of 0.045). We, therefore, have evidence that crowding out is incomplete. It is interesting to note that, although some difference in both components is evident from Fig. 4, neither difference is quite significant (p values of 0.1057 for the conditional mean and 0.2180 for the proportion p). The component evidence combined, however, is strong enough to be so.

The reader has probably noticed that the modal donations in Figs. 2 and 3 correspond to the simple fairness norm of splitting the money 50–50 ($10 each for donor and recipient). This behavior is *consistent* with standard public goods theory; donors that split equally give $8 in 18-2 but $3 less in 15-5, consistent with the model and the total crowding out prediction. In fact, the fairness norm can be written as a Leontief utility function. So it is not the modal donors that are responsible for the failure of the theory.

5. Concluding remarks

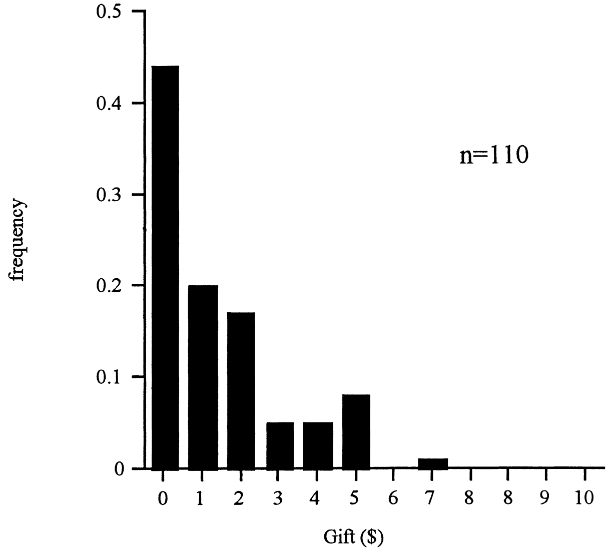
Comparing the averages of the two distributions displayed in Fig. 4, 73.7 percent of private giving is crowded out by the $3 transfer from donors to recipients. Our major finding, then, is that the $3 transfer is extensively but incompletely crowded out. The experiment provides direct evidence that the problem with the crowding out argument lies with the underlying specification of donor preferences.[17]

The donor preference critique states that people are motivated to give, at least in part by impure altruism. That is, they derive some satisfaction from the actual act of private giving. So by this account, subjects in the 15-5.*a* treatment fail to fully neutralize the transfer because an involuntary transfer does not provide the same satisfaction as voluntary giving. How do we explain that field studies find crowding out to be between 5 and 28 percent (references in Footnote 3), much smaller than in the experiment? One obvious possibility is that giving to impoverished children (or even public radio) evokes a relatively higher degree of giving-satisfaction than does giving to a fellow student in a laboratory experiment.

Consider, however, Fig. 5, which compares the gift-giving distribution for the pooled 15-5 treatments to the gift-giving distribution reported in Bolton and Katok (1995). Both experiments involved the dictator game. Both were performed by roughly the same group of experimenters using essentially the same subject pool. The amount and initial allocation of money was also the same.[18] The only difference is how the dictator's task was framed to subjects. In Bolton and Katok (1995), the task was described as a 'game'

[17] A meaningful comparison with the amount of crowding out Andreoni (1993) observes is very difficult. The similarity between the value we observe and that reported by Andreoni can be no more than coincidence: To calculate the amount of surplus giving in the tax treatment (i.e. the token giving not crowded out by the token transfer), Andreoni adds two tokens to the average number of tokens played in the tax treatment, and then subtracts the average of the no-tax treatment. This may overstate surplus giving because it assumes that the those who played zero tokens in the tax treatment would play two tokens in the no-tax treatment, even though they would be free to play less. Our technique, described in Section 4, avoids this difficulty, and provides a strictly tighter upper bound.

[18] In the Bolton et al. experiment, the dictator split $10. In addition, dictator and recipient were each paid a $5 show-up fee. So the total amount of money, as well as the recipient's initial endowment were the same as here. While the dictator could not give his-or-her show-up fee, Fig. 4 shows this was a non-binding constraint.



15-5 Treatments (pooled)

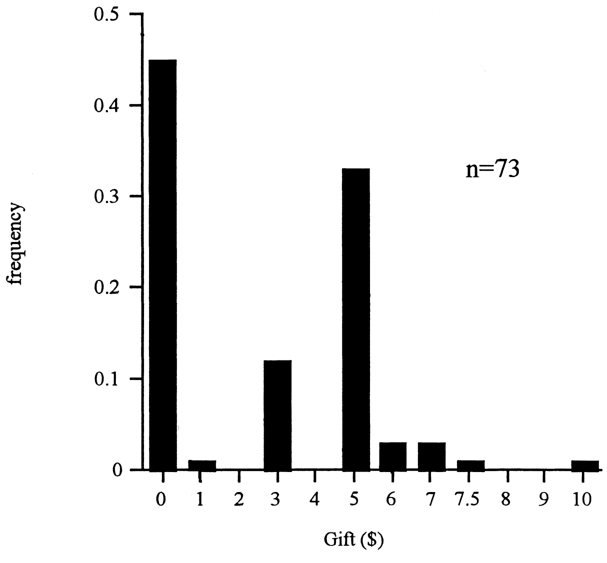


Fig. 5. Comparing giving: pooled distribution Bolton and Katok (1995) versus pooled 15-5.

concerning ‘Player A’ and ‘Player B.’ The present experiment described a ‘distribution task’ in which the money is ‘provisionally distributed’ subject to redistribution by A (see the protocol in the Appendix A). We chose the later language because it is closer in spirit to the public goods model than is the competitive game language. As is apparent from Fig. 5, the difference this change in framing makes is quite large. Framing effects have been reported by other experimenters for various experiments (for example, Croson, 1993; Schotter et al., 1994; Andreoni, 1995).

The strong influence of the frame is not predicted by the public goods model which assumes fixed pure altruistic preferences. (Our experiment shows something stronger, that the model's crowding out result fails to hold even when the frame is held fixed.) Thinking about why the frame matters suggests an alternative explanation for the crowding out pattern we observe: Framing is about perceptions. In the eyes of private donors, a larger government contribution might strengthen the charity's claim to being a worthy cause. The accompanying increase in willingness to give would tend to offset the neutralizing tendency identified by the standard public goods model.

Why would the size of the government's contribution matter in this way? Part of the reason could be that government has more resources to devote to gathering and assessing information than do most private individuals and organizations; a larger government donation might signal information about the charity's needs. It seems less likely that subjects in an experiment would be influenced by the amount initially allocated by the experimenter to the recipient. Most subjects probably consider themselves at least as good a judge as the experimenter of the charitable worthiness of their peers. Unfortunately, this 'information' interpretation of why government giving might matter is insufficient to explain the shift in Fig. 5; both were lab studies, and both gave the recipient an identical initial endowment.

Alternatively, we might attribute part-or-all of the influence of a larger government donation to 'moral suasion'; that is, a larger government contribution might heighten the donor's sense of obligation by signaling a new and greater collective social concern for the particular charitable cause. The precise amount the experimenter allocated to the recipient has less moral weight than a collective social decision, and so would have only a small influence on the donor's sense of obligation. But it also seems plausible that changing the frame of the game from a distribution task to a competitive game might have a big influence on how a person views their obligations, since behavior that is socially acceptable in a competitive situation is not acceptable in other contexts. So the moral suasion interpretation explains both why there is more crowding out in our experiment than in the field as well as the large shift displayed in Fig. 5. This explanation does not contradict the impure altruism interpretation of giving. On the contrary, it attributes impure altruism to the easing one feels from having fulfilled an obligation.

Acknowledgements

We thank Jordi Brandts and an anonymous referee for comments on earlier versions of the manuscript. We thank Keith Ord for statistical advice, and Heather Smith for assistance in running the experiments. This project was funded by a grant from the Center for Research in Conflict and Negotiation, Penn State University.

Appendix A

Laboratory protocol

This section contains a description of all procedures as well as verbal and written directions given to subjects in the 15-5 treatments. The monitor read all verbal instructions directly from the protocol. The only monitor–subject communication not included in the protocol are answers to individual subject questions. The 18-2 treatment differed solely in the numbers for initial allocations given in the written instructions.

Seating. *Upon entering the room, participants are seated. When all are seated (an even number), each is given a randomly chosen folder. Half the folders are red and half are blue*. May I have your attention please. We are ready to begin. Thank you for coming. With the exception of the folder, please remove all materials from your desk. Open your folder and take out the sheet marked ‘Instructions’ together with the one marked ‘Consent Form’. At this time please first read the instructions and then the consent form. *Participants read silently. Written instructions given to participants begin here*:

General. The purpose of this session is to study how people make decisions when given the task of distributing money. Feel free to ask a monitor questions as they arise. From now until the end of the session, unauthorized communication of any nature with other participants is prohibited.

During the session you will make money. Upon completion of the session the amount you make will be paid to you in cash. Payments are confidential: no other participant will be told the amount of money you make.

Half the participants were given blue folders and half red folders. During the session, you will be paired with a person having a different color folder than yours. No one, however, will know the identity of the person they are paired with. Nor will these identities be revealed after the session is complete.

Distribution task. In each pair, one person will have the role of A, and the other will have the role of B. The sum of $15 has been provisionally distributed to A and the sum of $5 has been provisionally distributed to B. A will have the opportunity to redistribute some of A’s money to B. A does this by completing a paper form. At the end of the session, A gets $15 minus the amount redistributed to B. B gets the amount A redistributed to B plus $5.

Conduct of the session. You will participate in two distribution tasks, called Task 1 and Task 2. Both tasks are identical to the description in the previous paragraph. For each task, you will be paired with a different person. You will have the role of A in one task, and the role of B in the other. In Task 1, those with red folders will have the role of A, and those with blue folders will have the role of B. In Task 2, blue folders are A, and red folders are B. You will receive a decision form only for the task in which you have the role of A. Both tasks will be done simultaneously, so the results of one task will not be known before the other has been completed.

You will actually be paid for just one of the two tasks. The one for payment will be chosen by a coin flip after the tasks have been completed [*In 15-5.b, underlined text replaced with:* prior to doing the tasks]. In order to preserve anonymity, the participants who are *B* for the chosen task will be taken to a second room and paid separately. Once you are paid, you may leave.

Consent forms. If you wish to participate in this session, please read and sign the accompanying consent form. Please note: In order to collect your payment you must stay until the end of the session which will last about 30 min. *Written instructions end here*.

Are there any questions?...We will now come around to collect the consent forms. *Collect*. [*In 15-5.b only: Coin Flip*. We will flip a coin to determine which decision task will be the one for payment before actually doing the tasks. If the coin comes up heads, decision Task 1 will be the one. If tails, decision Task 2 will be the one. *Flip*.]

Performing the distribution task. *Pass out distribution forms*. The person in role A indicates how they wish to distribute the money by filling out the form that is now being handed out. Those with red folders will receive the form for distribution Task 1, and those with blue folders will receive the form for distribution Task 2. [*In 15-5.b only*: Due to the results from the coin flip, only task (1 or 2) will be paid for. We would ask that those doing task (*other one*) also fill out the form, even though this task will not be paid for.] Please read the form along with me. *Read*. *Decision form for person in the role of A begins here*:

Distribution Task 1

You have the role of A. The person you have been paired with has the role of B. You must decide how to distribute $15 between the pair. Once you have decided, complete lines 1 and 2.

1.Of the $15, A receives – – – – – – – – – – .
(Must be a money value from a minimum of $0 to a maximum of $15.)
2.Of the $15, B receives – – – – – – – – – – .
(Must be a money value from a minimum of $0 to a maximum of $15. Lines 1 and 2 must sum to $15.)

Use the information you gave in lines 1 and 2 to complete lines 3 and 4.

3. For the session, A gets (repeat line 1) – – – – – – – – – – .
4. For the session, B gets (line 2 plus $5) – – – – – – – – – – .

If this distribution task is chosen for payment, line 3 is the total amount you, A, will get for the session. Line 4 is the total amount the person you are paired with, B, will get for the session. When finished, please turn this sheet over and wait quietly. *Written decision form ends here*.

Are there any questions?...Please fill out the form. Again, when you are finished, turn the form over and wait quietly. *Wait until all are finished*. A monitor will now come around to check that the numbers on the forms add up correctly. *Inspect*. *Subjects are then asked to complete the forms*. *A monitor inspects them to see that responses comply with instructions*. *In 15-5.a only*: *A coin is flipped to determine roles*. *Those in role* B *are taken to a different room and paid separately after those in role* A.

References

Abrams, B.A., Schmitz, M.A., 1978. The crowding out effect of government transfers on private charitable contributions. Public Choice 33, 29–39.

Abrams, B.A., Schmitz, M.A., 1984. The crowding out effect of government charitable contributions: Cross sectional evidence. National Tax Journal 37, 563–568.

Andreoni, J., 1989. Giving with impure altruism: Applications to charity and Ricardian equivalence. Journal of Political Economy 97, 1447–1458.

Andreoni, J., 1993. An experimental test of the public-goods crowding-out hypothesis. American Economic Review 83, 1317–1327.

Andreoni, J., 1995. Warm-glow versus cold-pickle: The effects of positive and negative framing on cooperation in experiments. Quarterly Journal of Economics 110, 1–22.

Becker, G.S., 1974. A theory of social interactions. Journal of Political Economy 82, 1063–1093.

Bergstrom, T.C., Blume, L.E., Varian, H.R., 1986. On the private provision of public goods. American Economic Review 29, 25–49.

Bernheim, B.D., 1986. On the voluntary and involuntary provision of public goods. American Economic Review 76, 789–793.

Bernheim, B.D., Bagwell, K., 1988. Is everything neutral? Journal of Political Economy 76, 789–793.

Bolton, G.E., Katok, E., 1995. An experimental test for gender differences in beneficent behavior. Economics Letters 48, 287–292.

Bolton, G.E., Brandts, J., Katok, E., 1997a. A simple test of explanations for contributions in dilemma games. Penn State University Working Paper.

Bolton, G.E., Katok, E., Zwick, R., 1997b. Dictator game giving: Rules of fairness versus acts of kindness. International Journal of Game Theory, forthcoming.

Bolton, G.E., Ockenfels, A., 1997. ERC: A theory of equity, reciprocity and competition. Penn State Working Paper.

Cadsby, C.B., Murray, F., 1991. Experimental tests of Ricardian equivalence. Economic Inquiry 29, 645–664.

Clotfelter, C.T., 1985. Federal Tax Policy and Charitable Giving. University of Chicago Press, Chicago.

Cooper, R., DeJong, D.V., Forsythe, R., Ross, T.W., 1996. Cooperation without reputation: Experimental evidence from prisoner's dilemma games. Games and Economic Behavior 12, 187–218.

Cornes, R., Sandler, T., 1984. Easy riders, joint production, and public goods. Economic Journal 94, 580–598.

Croson, R., 1993. Information in Ultimatum Games: An Experimental Study. Wharton School, University of Pennsylvania Working Paper.

Dawes, R.M., Thaler, R.H., 1988. Cooperation. Journal of Economic Perspectives 2, 187–197.

Forsythe, R., Horowitz, J.L., Savin, N.E., Sefton, M., 1994. Fairness in simple bargaining experiments. Games and Economic Behavior 6, 347–369.

Hochman, H.M., Rodgers, J.D., 1969. Pareto optimal redistribution. American Economic Review 59, 542–557.

Hoffman, E., McCabe, K., Schachat, K., Smith, V., 1994. Preferences, property rights and anonymity in bargaining games. Games and Economic Behavior 7, 346–380.

Kahneman, D., Knetsch, J.L., Thaler, R.H., 1986. Fairness and the assumptions of economics. Journal of Business 59, 285–299.

Kingma, B., 1989. An accurate measure of the crowd-out effect, income effect, and price effect for charitable contributions. Journal of Political Economy 97, 1197–1207.

Laury, S.K., Walker, J.M., Williams, A.W., 1995. Anonymity and the voluntary provision of public goods. Journal of Economic Behavior and Organization 27, 365–380.

Ledyard, J.O., 1995. Public goods: A survey of experimental research. In: Kagel, J., Roth, A.E. (Eds.), Handbook of Experimental Economics. Princeton University Press, Princeton.

Margolis, H., 1982. Selfishness, Altruism and Rationality. Cambridge University Press, New York.

Offerman, T., Sonnemans, J., Schran, A., 1996. 'Value orientation, expections and voluntary contributions to public goods, The Economic Journal 106, 817–845.

Palfrey, T.R., Prisbrey, J.E., 1995. Anomalous behavior in public goods experiments: How much and why. 1997 American Economic Review 87, 829–846.

Palfrey, T.R., Rosenthal, H., 1988. Private incentives in social dilemmas: The effect of incomplete information and altruism. Journal of Public Economics 35, 309–332.

Ribar, D., Wilhelm, M., 1995. An empirical test of altruistic and joy-of-giving motivations in charitable behavior. Presented at the AEA meeting, Washington, DC.

Roberts, R.D., 1984. A positive model of private charity and public transfers. Journal of Political Economy 92, 136–148.

Roth, A.E., 1995. Bargaining experiments. In: Kagel, J., Roth, A.E. (Eds.), Handbook of Experimental Economics. Princeton University Press, Princeton.

Schotter, A., Weigelt, K., Wilson, C., 1994. A laboratory investigation of multi-person rationality and presentation effects. Games and Economic Behavior, 445–468.

Sefton, M., 1992. Incentives in simple bargaining games. Journal of Economic Psychology 13, 263–276.

Stark, O., 1985. On private charity and altruism. Public Choice 46, 325–332.

Steinberg, R., 1987. Voluntary donations and public expenditures in a federalist system. American Economic Review 77, 24–36.

Sugden, R., 1982. On the economics of philanthropy. Economic Journal 92, 341–350.

Sugden, R., 1984. Reciprocity: The supply of public goods through voluntary contributions. Economic Journal 94, 772–787.

Varian, H.R., 1984. Microeconomic Analysis, 2nd ed. W. W. Norton, New York.

Warr, P.G., 1982. Pareto optimal redistribution and private charity. Journal of Public Economics 19, 131–138.