YOUNG, SELFISH AND MALE: FIELD EVIDENCE OF SOCIAL PREFERENCES*

John A. List

This study examines social preferences in three distinct field environments. In the first field setting, I allow consumers of all age and education levels to participate in one-shot and multiple-shot public goods games in a well-functioning marketplace. The second field study, an actual university capital campaign, gathers data from mail solicitations sent to 2,000 Central Florida residents. In the third field experiment, I examine data from an uncontrolled environment, a television gameshow, which closely resembles the classic prisoner's dilemma game. Several insights emerge; perhaps the most provocative is that age and social preferences appear linked.

At least since the seminal work of Samuelson (1954), many scholars have conjectured that private provision of public goods is inefficient because of the tendency for individuals to free-ride. The bulk of empirical evidence indicates that free-riding is a real concern, as underprovision of public goods in comparison to the social optimum occurs in most experimental settings; see Davis and Holt (1992) for a review. The general nature of the experimental results has been broadly consistent across different cultures, mechanism design,[1] number of players, and levels of pre-play communication, with most authors noting that the level of contributions is greater than the dominant strategy Nash equilibrium predicts but less than the symmetric Pareto efficient allocation.[2] Positive contributions in such environments have been interpreted as evidence that people have preferences for others' well-being.

Recently, a literature has developed that examines the connection between demographic factors and social preferences experimentally. For instance, Eckel and Grossman (1997) and Andreoni and Vesterlund (2001) devise dictator game experiments that provide insights into the link between gender and altruism. While the gender/altruism relationship has garnered increasing attention, an area that has gone largely unexplored is the correlation between age and other regarding behaviour. Although there is some anecdotal and reduced-form empirical evidence that suggests a link exists between age and altruism (Feldstein and Taylor, 1976; Reece, 1979), to date there has not been a systematic analysis examining the effect of age on social preferences.

[1] A recent exception that has performed remarkably well in the lab can be found in Falkinger *et al.* (2000).

[2] But see the work of Fehr and Gaechter (2000), who show that in the presence of a punishment opportunity, almost complete cooperation can be achieved.

In this paper, I report data from two quite distinct field experiments and one natural experiment to test whether a link exists between age and other regarding behaviour. My data-gathering approaches are unique in that I

  (i) recruit individuals on the floor of a sportscard show, allowing subjects of all age, income, and education levels to participate in public goods games,

 (ii) solicit donors in a way that matches, as closely as possible, the current state of the art in fundraising, and

(iii) investigate data drawn from the game show *Friend or Foe?*, which is remarkably similar to the classic prisoner's dilemma tale.

Accordingly, in all three cases I take advantage of the natural exogeneity inherent in the field, which permits a test of the main hypothesis of interest. Using the field as my laboratory also allows me to examine if the results of the voluntary contribution mechanism extend beyond the walls of the academic lab: regardless of the consistency of laboratory results, sceptics within the profession and the policy community may contend that nuances such as subject pools, lack of social distance between experimenter and subject, and subtleties associated with the laboratory setting compromise the generality of the empirical findings.

Beside Bohm (1984) and Marwell and Ames (1981), who had very different motivations and experimental designs from the present study, little has been done to extend the lab results systematically from these sorts of games to the general public.[3] Two additional treatment variables vary the manner in which subjects receive their initial endowments: in one treatment subjects earn their initial endowments and in the other they use goods purchased with out-of-pocket resources they personally bring to the sportscard show. These design parameters were used to capture the distinction between those subjects investing 'found money' and those investing personal funds.

The main results of the study fall into three categories. First, the 'found money' characteristic of lab experiments does not alter fundamental results in public goods games significantly: data across the various treatments are statistically indistinguishable. Second, in the multiple-shot treatments, the level and temporal nature of the contributions are consistent with data reported in lab experiments. This is good news in that the important laboratory results documented in the large literature of private provisioning of public goods appear to be smoothly transferable to the general public. In addition, this result carries over to one-shot games, where the level of provisioning closely accords to received results.

The third major result is that subject age and behaviour are correlated. In the public goods games, younger and middle-aged subjects tend to contribute at rates consistent with extant laboratory data, whereas older subjects contribute larger amounts of their endowment to the public good – in both multiple-shot and one-shot games. These results spill over into the realm of charitable contributions: even after I control for individual-specific factors, the data suggest that both the

---

[3] Casual evidence suggests that many public goods are successfully provided by private agencies. From groups of green to groups of aid, the private sector supplies a wide range of public goods. Funding for these non-profit agencies typically comes from a vast array of individuals: Andreoni (1988*a*) reports that 85% of US households contribute to charitable organisations.

probability of giving and the gift size are related to age. The results are also consonant with data drawn from prisoner's dilemma games, where cooperation rates among older subjects are greater than cooperation rates among younger cohorts, even after controlling for individual-specific factors. While it is important to stress that the data are unable to distinguish between age and generation effects unequivocally, supplemental tests suggest that the observed effects are not purely generational but are indeed related to age.

Overall, these results should have practical significance for economic theorists, empirical researchers, fundraisers, policymakers, and the growing body of scientific research that uses experimental methods. For example, the finding that 'found money' and the lab environment do not appear to influence the major results in voluntary contributions games unduly suggests that results from other classes of games, such as bargaining and market games, may not be considerably influenced by liberal changes in the experimental environment. Methodologically, this is comforting, but of course more research needs to be done to verify this conjecture. For empirical researchers and economic theorists, experimental results on age differences should influence views on normative and positive modelling approaches and research methodologies (e.g., data analyses). Fundraisers and policymakers should also find the results of interest for several reasons – manners in which charitable deduction rates, strategies to optimise charitable contributions and intergenerational transfer rules are formed may all be influenced by such findings.

The remainder of this study proceeds as follows. Section 1 discusses the theory and experimental design for Experiment 1. Section 2 presents the empirical results for Experiment 1. Section 3 discusses the experimental design and results for Experiment 2. Section 4 broadens the scope of the study by examining whether the age phenomenon extends beyond the experimental markets explored. Section 5 concludes.

1. Theory and Experimental Design I

In the spirit of the traditional public goods game, assume that in each period each subject in the group receives an endowment of tokens. A subject chooses whether to invest the tokens in the group project or in a private account. Decisions of group members are made simultaneously and the homogeneous payoff function is given by

$$P_i = e - g_i + \beta \sum_n g_j \qquad (1)$$

for each period, where e represents initial endowment, g_i is the level of tokens that subject i places in the group account, β is the marginal payoff of the public good, and $\sum_n g_j$ is the sum of the n individual contributions to the group project. The total experimental payoff is the summation of the period payoffs P_i over ten periods. By making $0 < \beta < 1 < n\beta$, strong free-riding ($g_i = 0$) is the dominant strategy since $\partial P_i / \partial g_i = \beta - 1 < 0$. Competing with this dominant strategy Nash equilibrium is the symmetric Pareto efficient allocation, which is characterised by

each group member placing all tokens in the group account ($g_i = e$). This follows from $\partial \sum_n P_i / \partial g_i = n\beta - 1 > 0$.

In what could reasonably be considered one of the most robust results from the lab, the usual finding from this general framework is that contributions to the public good begin at around 50% of initial endowments and decrease substantially in later periods, approaching the dominant strategy Nash equilibrium.[4] To provide a strict test of the lab results from this simple linear public goods framework, I follow List (2001) and recruit subjects from a well-functioning marketplace – a sportscard show – and examine whether these lab results spill over to one-shot and multiple-shot games in the field.

In the current set of treatments, however, I make use of a varying level of control. In the one-shot treatments, I examine individual behaviour in an actual marketplace. In this sense, I am observing behaviour within a well-functioning marketplace: an examination that provides a useful middle ground between the tight controls of the laboratory and the vagaries of completely uncontrolled field data. Since it is necessary to have dynamic interaction in the multiple-shot treatments, instead of running these treatments on the floor of the sportscard trading show, I use an adjacent room in the same building. Hence, unlike the one-shot treatments, the multiple-shot treatments should not be considered field experiments in the strict sense. Rather, they should be considered a laboratory experiment with a non-standard subject pool.[5]

1.1. One-Shot Treatments

The one-shot treatments were run on the floor of a sportscard show in a large southern US city.[6] Each participant's experience followed four steps:

(1) inspection of the good,
(2) learning the rules,

---

[4] The results are sensitive to group size, parameter values, number of periods etc. but the general results hold across various designs.
[5] Other experiments using non-standard subject pools exist (Fehr *et al.*, 1998; Cadsby and Maynes, 1998; Bohm, 1994) but do not address the same fundamental issues examined herein.
[6] Many readers may be unfamiliar with sportscard shows. With the rise in popularity of collector sportscards and memorabilia across the US in the past two decades, markets have naturally arisen that organise buyers and sellers. Temporal assignment of the physical marketplace is typically done by a professional association or local sportscard dealer who rents a large space, such as a gymnasium or hotel conference centre, and allocates six-foot tables to dealers for a nominal fee. When the market opens, consumers mill around the marketplace, higgling and bargaining with dealers, who have their merchandise prominently displayed on their six-foot table. The duration of a typical sportscard show is a weekend and a lucrative show may provide any given dealer hundreds of exchange opportunities (buying, selling, and trading of goods). I should note that collector markets, such as the sportscard market, are much more than fringe activity in the US: the Beckett Fact Sheet notes that their collector sportscard magazines have a paid circulation of over 889,000 copies. This circulation rate is comparable to popular magazines such as *Gourmet*, *Jane* and the *New Yorker*; see http://www.beckett.com/publications/ and http://asme.magazine.org (American society of magazine). Considering collectors broadly, according to the US Mint 2000 Annual Report, 'An estimated 125,000,000 American adults collect Q50 Quarters from pocket change with one-third collecting more than 25 of each state from pocket change, and the popularity of the program only increases with time.' See http://www.mint/gov). Relatedly, according to a March 13, 2000, article, the universe of stamp collectors includes 6,830,000 people; see http://www.linns.com/print/archives/20000313/editor.asp).

(3) allocating the cards, and
(4) conclusion of the transaction.

In Step 1, a potential subject approached the experimenter's table and enquired about the sale of the 1989 Michael Jordan Hoops basketball cards displayed on the table – note that a 1989 Michael Jordan Hoops basketball card has a book value of $3–$5 and is highly liquid at $1. The experimenter then invited the potential subject to take about five minutes to participate in an experiment. If the individual agreed to participate, he could pick up and visually examine each card (in sealed cardholders, with the graded card condition clearly marked – all cards were graded near mint/mint). The experimenter worked one-on-one with the participant, and imposed no time limit on his inspection of the cards.

In Step 2, the experimenter gave the participant an instruction sheet that consisted of two parts:

(5) a rules sheet, which adjusted Andreoni's (1988b) multiple-shot instructions to facilitate a one-shot game, and
(6) an allocation sheet.

The instructions explained that

(1) each subject was to divide her 10 Michael Jordan *Hoops* basketball cards between a private account and a group account, where only integers between 0–10 (inclusive) could be used;
(2) the marginal return to the group account was 0.40; and
(3) any card fractions would be paid in cash – e.g., 10.5 experimental units would yield 10 Jordan cards and $0.50.

In the instructions, subjects were informed that the experiment would be one-shot, payoffs would be private information, and that four-person groups would be determined randomly. Several examples of the public goods game were meticulously carried out to ensure that the wide range of subjects, in not only age (7 to 76) but also education (grade school to Ph.D.), would understand the rules of the game. No decisions were made until the subject fully understood the rules and her task.

After having her questions answered, the participant placed her allocation on the sheet provided (Step 3). Finally, in Step 4 the experimenter explained that the subject should return at a specified time to find out the results of the game and receive her goods. Transactions took place at 6pm on Sunday. If a subject did not return for the specified transaction time, she would receive her payoff within three days via standard postal service (postage paid by the experimenter).

I completed two treatments using this one-shot methodology. The first treatment, denoted *Baseline*, is similar in spirit to the standard public goods game in the received literature: subjects received their initial endowment of 10 Jordan cards for their participation and proceeded to allocate these cards across the two accounts. The second treatment, denoted *Own-Resources*, is similar to the *Baseline* treatment in every respect, except the participants first had to purchase the 10 Michael Jordan cards with their own out-of-pocket resources. To recruit subjects for this

treatment, I placed a sign at my dealer's table advertising 'FIRE SALE – 10 1989 MICHAEL JORDAN *Hoops* BASKETBALL CARDS FOR $5'. As aforementioned, since the Jordan card is very liquid at $1, recruitment was not difficult. When subjects approached my table enquiring about the cards, I informed them that they could purchase the 10 Jordan cards if they participated in the experiment – to allay their fears I oftentimes had to note that in the experiment they had a chance to earn even more Jordan cards.

At this point, I should mention a few noteworthy items. First, no subjects participated in more than one treatment. Second, I randomised subjects into treatments by changing the treatment every four hours; hence subjects' treatment type was based on the time they visited the table at the card show. Third, to gather an appropriate cardstock, I hand-graded more than 1500 Michael Jordan *Hoops* basketball sportscards to provide enough cards that varied only marginally across important attributes.[7] Fourth, I chose sportscards, rather than money, to be allocated across the two accounts to increase the level of realism. Subjects at sportscard shows commonly trade sportscards and memorabilia, hence it seems natural to have them allocate 'currency' that they commonly use in this particular market. Since 1989 *Hoops* Jordan cards are quite liquid on the floor of sportscard shows, this choice probably does not unduly affect decisions; yet if there are differences in preferences for the sportscards, this usage of physical commodities represents a potential serious loss of control. To ensure robustness of these results, I ran an identical baseline with 10 one-dollar bills replacing the 10 Jordan cards to provide an explicit link between these treatments and the extant literature, which uses currency. While the sample sizes are small – only 44 subjects in total – results for the pooled and cohort data are never qualitatively different from the *Baseline* treatment using sportscards as the medium of exchange. Finally, the treatments took approximately sixteen hours to complete (10:00am to 6:00pm on Saturday and Sunday). The average subject earned nearly 13.39 experimental units in the *Baseline* treatment and 13.32 units in the *Own-Resources* treatment (i.e., 13 Jordan cards and $0.39 and 13 Jordan cards and $0.32, respectively).

1.2. *Multiple-Shot Treatments*

In typical laboratory experiments, more than one trial is often required before subjects understand the nature of certain mechanisms. In the spirit of providing a robustness test to the one-shot treatments, I used a series of multiple-shot treatments at a similar sportscard show in a large southern city. Each participant's experience in the multiple-shot treatments typically followed two steps:

(1) consideration of the invitation to participate in an experiment that would take about 1 to 1½ hours, and
(2) participation in the experiment.

In Step 1, the monitor approached potential subjects entering the trading card show and inquired about their interest in participating in an experiment that

---

[7] I obtained the 1,500 Michael Jordan cards by purchasing bulk lots on the open market.

would take about 1 to 1½ hours. If the individual agreed to participate, the monitor briefly explained that in return the subject would receive a complimentary 1982 Nolan Ryan *Topps* baseball card (displayed on my dealer table) and the chance to earn a considerable amount of money. The monitor explained that at a pre-specified time (either 1pm or 3pm, Saturday or Sunday) the subject should enter an adjacent room to take part in the experiment. Directions to the room were provided and the subject was informed that she would receive the 1982 Nolan Ryan *Topps* baseball card and instructions for the experiment when she arrived.

Step 2 began when subjects entered the room and signed a consent form in which they acknowledged their voluntary participation in the experiment and agreed to abide by the rules of the experiment. Subjects in the 3pm Saturday and 1pm Sunday sessions were placed in the *Non-Earnings* treatment (e.g., given their initial endowments each round). Subjects in the other two sessions (1pm Saturday and 3pm Sunday), denoted *Earnings*, first took a quiz in which they earned their initial endowments (experimental money) and the right to participate in the second part of the experiment. With the goal of presenting participants with a difficult task to ensure sufficient effort on their part, I composed a quiz of 10 sports-related questions that would require significant cognitive effort – some subjects took nearly 30 minutes to answer the 10 questions.[8] Subjects were informed that a simple rule would determine their probability of participating in the experiment: the top 16 (out of 20) subjects would earn the 1982 Nolan Ryan *Topps* card and $10 to use in the experiment. If the subject did not rank in the top 16, he would receive only the 1982 Nolan Ryan *Topps* card and would be unable to participate in the experiment.

After the monitor answered all pertinent questions, subjects were notified that they had 30 minutes to complete the exam, after which the exam would be graded and the second stage of the experiment would begin. Upon completion of the exam, which took slightly less than 30 minutes in each session, administrators graded the exams and ranked the subjects. Those not in the top 16 were given a 1982 Nolan Ryan *Topps* card and informed that they should quietly exit the room. The other 16 subjects proceeded to play the public goods game for ten trials.

In total, I conducted four sessions (2 with earnings, 2 without earnings); each had 16 subjects split into 4 equal-sized groups. Each of the four sessions was run in a medium-sized room and subjects were permitted to talk only with administrators. In the spirit of previous public goods' laboratory studies, I followed Andreoni's (1988*b*) instructions verbatim in each session, with two slight changes:

(1) each subject was allocated 100 tokens per round, where one token = $0.01 (consistent with the statement that subjects would receive a $10 reward for performing well on the quiz), and
(2) the marginal return to the public good was $0.40.

In the instructions, subjects were informed that the experiment would last 10 rounds, payoffs would be private information, four-person groups would be

---

[8] Copies of the survey and quiz are available on http://www.res.org.uk/economic/ta/tahomt.asp

determined randomly each round and no group would be repeated.[9] As in the one-shot treatments, several examples of the public goods game were carried out to ensure that the wide range of subjects, in not only age (7 to 65) but also education (grade school to Ph.D.), would understand the rules of the game. None of the experiments was started until all subjects fully understood the rules and their task.

The non-earnings treatments took approximately 1 hour to complete and the earnings treatments took approximately 1½ hours to complete. No subject participated in more than one treatment. The average subject earned nearly $13 and all subjects that showed up for the experiment received a 1982 Nolan Ryan baseball card. I invited 30 subjects to each session because I anticipated in this setting that a large number of absences would occur. This premise was realised: on average 22 people showed up.

Table 1 provides a statistical description of the subject characteristics. In total, I observed 80 one-shot allocation decisions and 640 multiple-shot allocation decisions: 64 subjects across 10 rounds. First and second moments of the variables

Table 1
Selected Characteristics of Participants

	Mean (Std. Dev.)	Minimum	Maximum
(a) One-Shot Treatments			
Income	3.93	1	8
	(2.16)		
Age	39.05	7	76
	(19.47)		
Gender (% male)	0.88	0	1
	(0.33)		
Education	3.75	1	6
	(1.67)		
n	80	–	–
(b) Multiple-Shot Treatments			
Income	3.98	1	8
	(2.38)		
Age	32.32	7	65
	(15.92)		
Gender (% male)	0.91	0	1
	(0.29)		
Education	3.19	1	6
	(1.71)		
n	64	–	–

Notes:
1. *Income* denotes categorical variable (1–8): (1) Less than $10,000, (2) $10,000 to $19,999, (3) $20,000 to $29,999, (4) $30,000 to $39,999, (5) $40,000 to $49,999, (6) $50,000 to $74,999, (7) $75,000 to $99,999 (8) $100,000 or more.
2. *Age* denotes actual age in years.
3. *Gender* denotes categorical variable: 0 if female, 1 if male.
4. *Education* denotes categorical variable: (1) Eighth grade or less, (2) High School, (3) 2-Year College, (4) Other Post-High School, (5) 4 Year College, (6) Graduate School Education.

[9] This provides the sense of ten one-shot games; see Andreoni (1988*b*); Fehr and Gaechter (2000).

reveal that the current experimental pool has a much richer assortment of subjects than previous experimental studies were afforded. Unfortunately, the variation is only across education, income, and age. Variation across gender is minimal since few women participate in the sportscard market. Nevertheless, the current subject pool includes quite a diverse group of subjects: ranging from annual incomes of less than $10,000 to over $100,000 and education levels spanning from grade school to Ph.D. (although the questionnaire did not ask about Ph.D., some subjects circled 'Graduate School Education' and wrote in Ph.D.).

2. Experimental Results I

A quick summary of the empirical results is that consistent with the bulk of past experimental evidence, there are signs of significant contributions to the public good in the aggregate data. This result lends support to previous experimental efforts and offers a reassurance that lab results readily spill over to the general public. Across all consumer types in both one-shot and multiple-shot treatments, however, there seems to be a tendency for more mature subjects (age > 49) to give more to the group fund. These and other empirical findings are described more fully below.

2.1. *Result I: A Found Money Effect is not Evident*

Giving subjects initial endowments is a common practice in lab experiments. While this practice has important advantages, such as saving time to run the experiment, inducing participation and allowing subjects to make decisions without leaving the experiment in debt, there are potential consequences that could compromise the interpretation of results. For example, evidence in different circles suggests that individuals behave quite differently with 'found' or 'windfall' money than with ordinary income; (see Keeler *et al.*, 1985, and Battalio *et al.*, 1990). Findings from these studies imply that subjects are more risk-seeking when given endowments. In the standard public goods game, subjects may be more willing to 'buy' altruism or induce reciprocal fairness through strategic play when playing with found money.[10]

Table 2(*a*), (*b*) presents sample statistics from the one-shot treatments and across the ten rounds of both multiple-shot treatments.[11] One important aspect of these data is that within the one-shot treatments and within the ten trials of the multiple-shot treatments, the data across treatments essentially mirror one another. This observation is substantiated when one considers results from a Mann-Whitney rank-sum test of treatment differences. The rank-sum test is a standard nonparametric test that has a null hypothesis of no treatment effect, or that the

---

[10] Besides a strategy of maximising income, it is commonly asserted that subjects bring motives, such as altruism, to the experiment. There is lively research being carried out to more fully understand the environments that induce these other regarding preferences (Fehr and Schmidt, 1999). The contribution of this section could also be thought of as a piece of evidence within this class of issues.

[11] Data presented in Table 2(*a*)(*b*) and 3(*a*)(*b*) can be read as follows: (*a*) is the mean individual contribution (out of 10); (*b*) figures are out of 100 (since 100 was the initial endowment, one can consider these as percentages).

Table 2

Average Contributions across Treatments

	Mean (Std. Dev.)
(a) One-Shot Treatments	
Baseline	5.65
($n = 40$)	(3.77)
Own-Resources	5.28
($n = 40$)	(3.90)
Difference	0.37
Mann-Whitney Rank-Sum Test ($n = 80$)	z = 0.37

	Trial mean (Std. Dev.)									
	All	1	2	3	4	5	6	7	8	9	10
(b) Multiple-Shot Treatments											
Earnings	34.9	48.6	49.2	43.4	32.2	27.7	31.4	29.7	32.8	26.5	27.2
($n = 320$)	(30.9)	(35.6)	(34.8)	(28.1)	(21.4)	(25.3)	(25.7)	(25.7)	(33.5)	(28.1)	(34.9)
Non-Earnings	33.1	50.9	43.3	39.6	35.3	36.0	26.2	27.3	24.8	23.7	23.9
($n = 320$)	(26.7)	(29.1)	(27.5)	(27.7)	(25.6)	(30.4)	(21.9)	(21.3)	(17.7)	(26.2)	(25.0)
Difference	1.8	−2.3	5.9	3.8	−3.1	−8.3	5.2	2.4	8.0	2.8	3.3
Mann-Whitney Rank-Sum Test ($n = 64$)											z = 0.20

Notes:
Data in (*a*) are the individual contribution (out of 10); (*b*) figures are out of 100 (since 100 was the initial endowment, one can consider these as percentages).
Mann-Whitney rank-sum test of treatment differences is a standard nonparametric test that has a null hypothesis of no treatment effect, or that the two samples come from identical populations. The test statistic is normally distributed.

two samples are derived from identical populations. In the one-shot treatments the test is straightforward and yields a value of z = 0.37. To construct the test statistic in the multiple-shot treatments, I first calculated the individual mean contribution levels across the ten rounds and then ranked subjects via these means in the joint earnings/non-earnings sample. The test statistic is normally distributed, and takes on a value of z = 0.20. Neither of these values is significant at conventional levels. Although in the multiple-shot data there is the typical pulsing in contributions that is inherent in laboratory data, there is no evidence of a significant treatment effect. Overall, these findings provide evidence that the 'found money' characteristic of lab experiments does not significantly alter fundamental results in public goods games.[12] Given that the earnings treatments do not significantly influence allocation decisions, I pool the data for the analyses below.

2.2. *Result II: Sportscard Market and Lab Data Match Quite Closely*

The top rows in in Table 3(*a*), (*b*) report summary statistics of the aggregate data across each of the experiment types. Statistics in the pooled samples suggest a

---

[12] One caveat to this finding in the multiple-shot treatments concerns the way in which participants were chosen. If those who are apt at sports trivia give systematically more or less to public goods, then this result is potentially confounded.

Table 3

Average Contribution Levels across Age Cohorts

	Mean (Std. Dev.)
(a) One-Shot Treatments	
Pooled Data ($n = 80$)	5.59 (3.79)
Young Cohort ($n = 25$)	4.88 (3.70)
Middle Cohort ($n = 27$)	4.74 (3.83)
Mature Cohort ($n = 28$)	7.00 (3.48)
Mann-Whitney Rank-Sum Test Young vs. Middle ($n = 52$)	0.01
Mann-Whitney Rank-Sum Test Young vs. Mature ($n = 53$)	−1.91
Mann-Whitney Rank-Sum Test Middle vs. Mature ($n = 55$)	−2.27

	Trial Mean (Std. Dev.)										
	All	1	2	3	4	5	6	7	8	9	10
(b) Multiple-Shot Treatments											
Pooled Data	33.9	49.8	46.3	41.5	33.8	31.8	28.8	28.5	28.8	25.1	25.5
($n = 640$)	(28.6)	(32.3)	(31.3)	(27.7)	(23.4)	(28.0)	(23.8)	(23.4)	(26.9)	(27.0)	(30.1)
Young Cohort	35.4	42.5	44.7	40.6	32.5	35.6	25.8	25.9	29.7	38.4	37.8
($n = 160$)	(27.3)	(25.4)	(34.1)	(30.3)	(18.0)	(25.6)	(21.0)	(16.9)	(31.3)	(35.8)	(28.4)
Middle Cohort	29.8	46.9	44.4	37.4	32.2	28.9	25.0	27.6	25.4	13.8	16.6
($n = 340$)	(29.5)	(37.2)	(32.3)	(27.5)	(27.4)	(28.5)	(24.9)	(25.6)	(25.8)	(19.6)	(28.3)
Mature Cohort	42.5	65.0	52.5	52.5	38.9	34.6	41.4	33.6	36.1	37.4	33.2
($n = 140$)	(25.8)	(21.4)	(26.4)	(24.1)	(18.6)	(30.7)	(21.1)	(25.1)	(24.5)	(18.8)	(31.2)
Mann-Whitney Rank-Sum Test Young vs. Middle ($n = 50$)											$z = 1.39$
Mann-Whitney Rank-Sum Test Young vs. Mature ($n = 30$)											$z = -1.41$
Mann-Whitney Rank-Sum Test Middle vs. Mature ($n = 30$)											$z = -2.57$

Notes:
Data in (*a*) are the individual contributions (out of 10); (*b*) figures are out of 100 (since 100 was the initial endowment, one can consider these as percentages).
Young cohort: age < 19; middle cohort: age 19–49; mature cohort: age > 49.
Mann-Whitney rank-sum test of treatment differences is a standard nonparametric test that has a null hypothesis that the two samples come from identical populations. The test statistic is normally distributed.

good deal of individual investment in the group accounts. Sample means in the one-shot treatments are in the range of previous one-shot games (Marwell and Ames, 1981), which find that contributions to the public good are in the 40–60% range. Likewise, the multiple-shot data are similar to those in previous public good lab experiments: investment in the public good account begins around 50% and pulsates until the later rounds, where it approaches the dominant strategy Nash prediction of 0. This is perhaps illustrated most efficiently in Figure 1, which plots the current set of pooled data against lab data from other related studies (Fehr and Gaechter, 2000; Isaac *et al.*, 1994; Andreoni, 1988*b*).[13] While group size and marginal returns affect public good outcomes

[13] These studies were selected to provide a sense of temporal variability while maintaining consistency across important parameters. Note that for these data, Fehr and Gaechter (2000) and Andreoni (1988*b*) used random matching (stranger) treatments, while Isaac *et al.* (1994) used fixed matching (partner) treatments. The qualitative results are unchanged if other published studies are used.
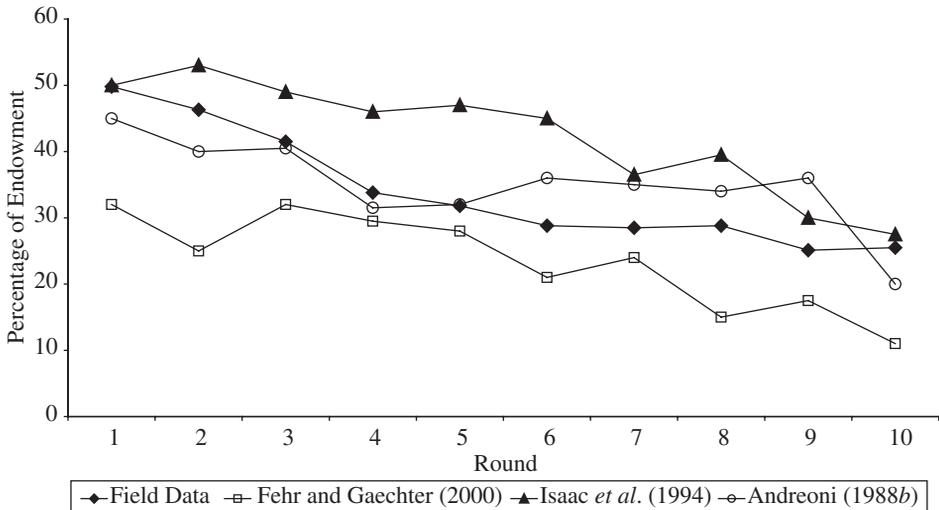
Fig. 1. *Public Goods Provision across Multiple-Shot Experiments*

Fehr and Gaechter (2000): $n = 4$, MR $= 0.4$;
Isaac *et al.* (1994): $n = 10$, MR $= 0.3$;
Andreoni (1988*b*): $n = 5$, MR $= 0.5$,

the data herein are remarkably similar to data from these laboratory studies. Although there is the typical pulsing in contributions, trajectories across studies are similar, as trial means from the field commingle with trial means from the lab. However, one difference is that the data gathered from the general public do not show signs of a significant decrease between trials 9 and 10.

2.3. *Result III: A Link Exists between Age and Cooperation Rates*

While measuring and comparing contributions in aggregate provides important insights, the richness of the current data set permits a much more detailed analysis. In this regard, to examine the main conjecture of interest, I split the sample according to subject age and examine whether age and contributions are systematically related. Rows 2 to 4 in Table 3(*a*), (*b*) present allocation decisions in split subsamples across age categories of participants. The results are compelling. While subjects in the younger cohort (age < 19) tend to contribute about the same as subjects in the middle-age cohort (age from 19 to 49), more mature subjects (age > 49) invest a higher proportion of their endowment in the public good account across both one-shot and multiple-shot treatments. These general results are robust across liberal changes in the age categorisations and are supported statistically via Mann-Whitney rank-sum tests. In the one-shot treatments, the mature cohort contributes significantly more to the group account than both the younger cohort ($z = -1.91$) and the medium age cohort ($z = -2.27$) at conventional levels. In addition, examining data from the multiple-shot treatments,

results from Mann-Whitney rank-sum tests yield similar insights: young cohort vs. middle cohort: z = 1.39; young cohort vs. mature cohort: z = −1.41; middle cohort vs. mature cohort: z = −2.57. All of these results are similar when I use parametric tests.

Another interesting aspect of the data relates to contribution trends across age cohorts. I provide Figure 2, which plots the temporal path of contributions for each of the three age cohorts, to show the disparate paths. Figure 2 illustrates that more mature subjects tend to contribute more than their younger counterparts across the majority of decision periods. Although younger and middle-aged subjects invest similar amounts in the group fund in early periods, rates of giving do not deteriorate over time for younger subjects but actually increase. As a result, near the end period younger and more mature subjects invest much more in the public good than middle-aged subjects contribute. This general result was masked in the aggregate data displayed in Figure 1, which implied that subjects at the sportscard show behaved similarly to those in the lab, with the exception of contributions between periods 9 and 10. Interestingly, the observation that younger subjects contribute about the same as middle-aged subjects but greater in later periods in multiple-shot experiments, is very similar to findings in Harbaugh and Krause (2000), who present results that suggest six to twelve-year-old subjects have contribution distributions similar to young adults until the end periods of the experiment.

While examining aggregate cohort data can lend important insights into the underlying data-generation process, much is lost via the aggregation. For example, data summaries in Tables 2 and 3 might be masking the fact that individuals within a particular cohort are choosing dichotomously, whereby a non-trivial subset are strong free-riders (contribute zero to the group account), while another group is socially optimising (contribute 100% of their endowment to the group fund). To examine individual decisions more carefully, I provide Figures 3 and 4. Figure 3
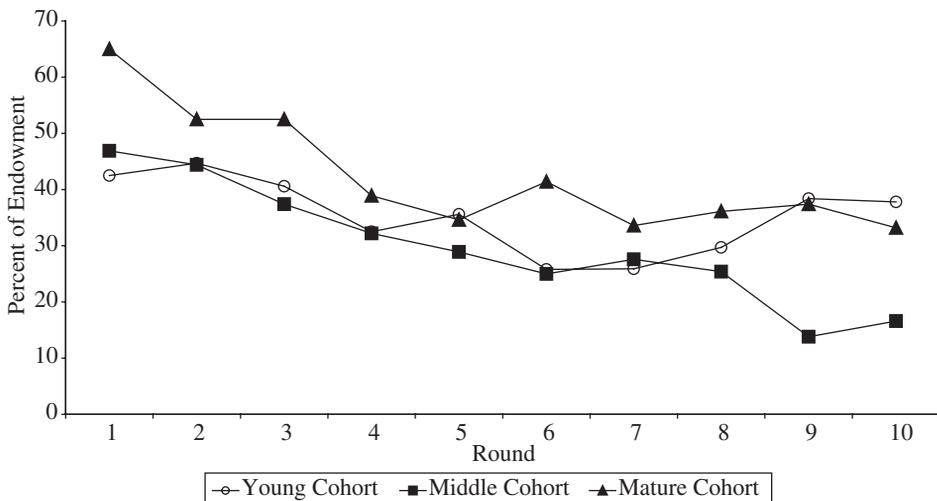
Fig. 2. *Contributions Across Age Cohorts in Multiple-Shot Treatments*

presents the distribution of contributions across age cohorts in the one-shot treatments. Here, we can readily see that a large percentage of individuals in the mature cohort play the socially optimising choice (over 35%). This compares favourably to the 17% (12%) of individuals in the middle (young) cohort that choose this option. Indeed, making use of a Pearson chi-square test, which examines the null hypothesis of H_o: $p_{Mature} = p_{Middle} = p_{Young}$ where p_i are the parameters of three independent binomially distributed random variables, I find that the null hypothesis of no age effect on socially optimising play should be rejected at the $p < 0.07$ level (χ^2 (2 df) = 5.18). Moreover, testing the null hypothesis of no age effect on the pairwise cohorts using a Fisher's exact test, which has a hypergeometric distribution under the null, I find that the null hypothesis
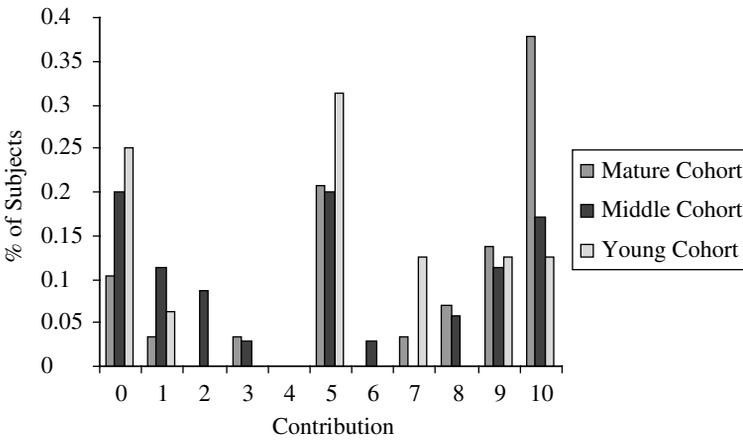
Fig. 3. *Distribution of Contributions Across Age Cohorts in One-Shot Treatments*
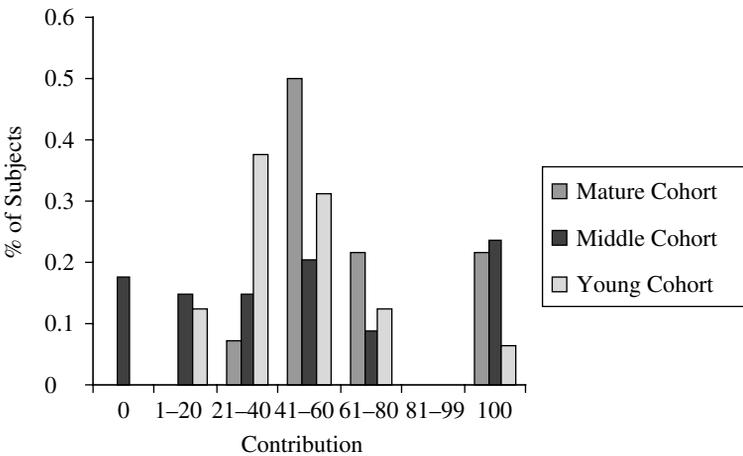
Fig. 4. *Distribution of Contributions Across Age Cohorts in First Round of Multiple-Shot Treatments*

should be rejected at the p < 0.06 level for both (i) mature cohort versus medium age cohort and (ii) mature cohort versus young cohort. Although this same pattern of behaviour is exhibited when we consider strong free-riding, the null cannot be rejected at conventional levels (χ^2 (2 df) = 1.80).

Figure 4 presents the relevant distributions across age cohorts in the first round of the multiple-shot treatments. In this case, what readily emerges is the lack of a lower tail in the mature cohort. Whereas nearly 20% of the middle cohort plays the strong free-rider strategy, and nearly one-third of the middle cohort gives less than 20% of their endowment to the public good, the mature cohort has no representation in these cells. Accordingly, using a Pearson chi-square test, the null of homogeneous strong free-riding across age cohorts can be rejected at the p < 0.06 level (χ^2 (2 df) = 5.84), implying that in the first round of the multiple-shot treatments, age had an effect on the strong free-riding decision. In addition, pairwise Fisher's exact tests yield similar insights, as statistically both the younger and medium age cohorts play the strong free-riding strategy significantly more than individuals in the mature cohort. Interestingly, unlike the one-shot data, the mature and middle cohorts had roughly equal representation in the socially optimising cell, leading to nonrejection of the null hypothesis in this case.

Analysis of the raw data provides evidence that supports the notion that age is related to allocation decisions, but there has been no attempt to control for other factors that may affect the contribution decision. If income and contributions to the public good are positively associated, insights from an unconditional analysis may send erroneous signals since age and income are positively correlated in the data (r ≈ 0.50). Alternatively, this same effect could be confounding significant differences between contributions across the middle cohort and the younger cohort. Income levels, and other subject-specific factors, including education and gender, can be adequately accounted for in a well-specified econometric model. To condition on these factors, I estimate variants of the following Tobit model:

$$Contribution = \varphi + \boldsymbol{\beta} \boldsymbol{X} + \alpha_i + u_t + \varepsilon_{it}, \tag{2}$$

where $E[\alpha_i] = 0$, $E[u_t] = 0$, $E[\alpha_i^2] = \sigma_\alpha^2$, $E[u_t^2] = \sigma_u^2$, $E[\alpha_i \alpha_j] = 0$ for $i \neq j$, $E[u_t \alpha_z] = 0$ for $t \neq z$, and u_t, α_i, ε_{it}, are orthogonal for all i and t.

In (2), *Contribution* is the level of investment in the public good; $\boldsymbol{X}$ includes subject-specific variables that may influence the level of investment. Variables in $\boldsymbol{X}$ are listed in Table 1 and include annual income, age, gender, and education. Age enters (2) via two dichotomous variables: *Young Cohort* = 1 if age < 19, 0 otherwise; *Mature Cohort* = 1 if age > 49, 0 otherwise. Accordingly, the middle-aged cohort (age from 19 to 49) is the baseline (omitted) category. To control for subject-specific effects in the multiple-shot data, I include α_i as random subject effects, which control for unobservable subject characteristics such as propensity to contribute. Also, I include u_t to capture potentially nonlinear trial effects; u_t indicate whether subjects' behaviour changes over trials, and therefore capture learning, amongst other dynamic factors. ε_{it} is the contemporaneous well-behaved error term. Maximum likelihood estimates of the random effects Tobit model in (2) treat unmeasured characteristics as error components, economise on degrees of freedom and yield coefficients that are not conditioned on unmeasured subject and time effects.

Summary estimates of (2) are presented in Table 4. I include ordinary least squares regression estimates, Tobit estimates, and estimates from a Tobit random effects model. Given that a chi-squared test rejects equality of the random effects (χ^2 (1) = 81.03), the Tobit random effects (columns 5 and 6) estimates are preferred. Regardless of estimation technique, however, each of the models is statistically significant and coefficient estimates suggest similar insights. In particular, after controlling for income, gender, education, trial, and individual unobserved heterogeneity, I find that age remains an important determinant in the level of contributions to the public good in both one-shot and multiple-shot treatments. In each specification, the *Mature Cohort* regressor is positive and significantly different from zero at conventional levels. For example, estimates from the OLS specification in the one-shot treatments suggest that individuals in the *Mature Cohort* contribute approximately 2.14 more Jordan cards to the public fund than individuals in the *Middle Cohort* baseline, which is significantly different from zero at the $p < 0.05$ level. This estimate, which is consonant with the response coefficient from the Tobit model in column 2, represents a 45% increase from the average middle-aged contribution level of 4.74 Jordan cards.

Likewise, Tobit random effects estimates of the coefficient of *Mature Cohort* in column 6 suggest that, *ceteris paribus*, subjects in the mature cohort contribute approximately 15 more tokens to the group account than subjects in the

Table 4

Estimation Results for Contribution Function

Variable	One-Shot		Multiple-Shot			
	OLS	Tobit	OLS	Tobit	Tobit RE	Tobit RE
Constant	4.66**	4.05*	28.44**	18.58	19.15	5.99
	(2.07)	(2.39)	(8.37)	(7.69)	(18.90)	(20.5)
Young Cohort	0.41	0.64	5.99	10.52**	10.25	9.85
	(1.65)	(1.91)	(4.85)	(5.07)	(13.35)	(14.89)
Mature Cohort	2.14**	2.34**	12.24**	16.77**	17.86**	17.77*
	(0.92)	(1.06)	(4.60)	(3.97)	(8.72)	(9.65)
Income	−0.37	−0.40	0.42	0.83	0.68	0.63
	(0.26)	(0.30)	(1.18)	(1.03)	(2.14)	(2.32)
Gender	−0.85	−1.37	1.04	3.54	3.55	3.43
	(1.24)	(1.42)	(5.43)	(5.04)	(10.91)	(11.94)
Education	0.60*	0.79*	−0.39	−0.60	−0.75	−0.79
	(0.35)	(0.41)	(1.56)	(1.23)	(2.29)	(2.43)
F	2.29**	–	4.18**	–	–	–
Subject Effects	No	No	No	No	Yes	Yes
Period Effects	No	No	No	No	No	Yes
N	80	80	640	640	640	640

Notes:
Tobit RE denotes Tobit with random subject-specific effects. Dependent variable is the level of contributions to the public good. *Gender* = 1 if male, 0 otherwise; *Young Cohort* = 1 if age < 19, 0 otherwise; *Mature Cohort* = 1 if age > 49, 0 otherwise.
Standard errors are in parentheses beneath coefficient estimates.
**Denotes significant at the $p < 0.05$ level. *Denotes significant at the $p < 0.10$ level.

© Royal Economic Society 2004

middle-aged cohort.[14] This marginal effect is significantly different from zero at the $p < 0.05$ level using a two-sided alternative and represents a 50% increase from the average middle-aged contribution level of 30 tokens. Alternatively, the empirical evidence suggests that the *Younger* and *Middle Cohorts* contribute statistically similar amounts to the public good account. Overall, these empirical results support the unconditional results in Table 3 and provide evidence consistent with the notion that more mature individuals tend to provide greater voluntary contributions to public goods than younger cohorts. Interestingly, these estimates are very close to reduced-form charitable giving estimates published in the literature. For example, using 1970 Treasury Tax File data, Feldstein and Taylor (1976) find that citizens over 65 years old give 56% more to charities than younger taxpayers, independent of income.

I should stress that the general empirical results are robust to alternative definitions of the age variable. For example, to focus exclusively on the effects of a continuous age variable, I also ran regression models making individual contributions a function of age, and subject and period-specific effects (when applicable) in the estimation of (2). The observed relationship across the models again suggests that more mature subjects tend to be more altruistic than younger cohorts. In the nonlinear models, the age effect has a U-shape and begins to increase around the age of 30. The spirit of these results is in accord with the intuition of Hannan *et al.* (1999), who observe that people bring their life experiences into the lab. This notion is entirely consistent with remarks I received from more mature subjects after the experiment: this cohort of subjects made it clear that they were intuitively more aware of the necessity to cooperate given their life and job experiences.[15]

3. Experimental Design and Results II

The results above paint a consistent portrait concerning the relationship of age and other regarding behaviour; in this Section, I explore the robustness of these results by examining behaviour in a much different setting – a university fundraiser. Besides a sensitivity check, I can also examine if the age effect found in the public goods games spills over to the realm of charitable giving, which could invoke much different preferences than those called upon in simple linear public goods games.[16]

---

[14] This marginal effect is computed by using a scale factor of 0.85 and a conditional mean of 31.52. Note the only difference between results in columns 5 and 6 in Table 4 is that the column 6 model includes period and subject-specific effects whereas the empirical model in column 5 includes only subject-specific effects.

[15] Some commentators have suggested that repeated play in laboratory public goods games is meant to be a super-fast, *mach7*-like, analogy to life-long learning (akin to time lapse photography of a flower blossoming). Viewed in this light, the received lab results may be interpreted as subjects learning the dominant strategy of contributing nothing over time (note that other theories may also explain the finding that contributions shrink over time, such as reciprocity not mattering in the latter stages of an experiment). As a whole, the results herein suggest that caution should be taken when interpreting results from laboratory public goods games in such a manner, as other types of effects, such as life-cycle, life experiences, etc., may overwhelm these 'learning effects'.

[16] Note that in certain situations charitable giving and provisioning of public goods are quite different – e.g., one may be a pure transfer and the other an efficiency-enhancing exercise. Accordingly, it is important to recognise that the nature of the potential other-regarding behaviour may be quite different across these two institutions.

Also, this particular analysis permits an exploration of whether the apparent age effects observed above are purely generational or if they are indeed age effects.

To provide initial insights into these issues, I took advantage of a unique opportunity I was provided at the University of Central Florida (UCF). While a faculty member at UCF, I was approached to spearhead a capital campaign to fund a new Center for Environmental Policy Analysis (CEPA). Having received permission to design the fundraising campaign as an experiment, I solicited contributions from 2000 Central Florida residents to help fund the start-up of CEPA. In carrying out the field experiment, I wished to solicit donors in a way that closely matched current fundraising standards.[17] With advice from fundraising companies *Donnelley Marketing* (and associates) in Englewood, Colorado, and *Caldwell* in Atlanta, Georgia, I followed generally accepted rules believed to maximise overall contributions. First, I purchased names and addresses of households in the Central Florida area that met two criteria:

(i) annual household income greater than $70,000, and
(ii) the household had given to a charity in the previous year.

From *Donnelley Marketing* I purchased the names and home addresses of 2,000 Central Floridians who met both criteria.

Second, I designed an attractive brochure describing the new centre and its purpose.

Excerpts from the brochure read as follows:

The primary objective of The *Center for Environmental Policy Analysis* (CEPA) will be to improve the quality of Florida's public and private decisions that have environmental, economic, and resource-use implications. In addition, the CEPA will propose economically efficient solutions to national and international problems ranging from endangered species protection to global issues such as climate change and sustainable development.

The CEPA will accomplish these tasks through an integrated program of communications, publications, and education, designed to lead from awareness through knowledge to action. Through these programs, the CEPA will improve communication between the public, including various governmental branches, and the business community.

The CEPA will also offer courses, seminars, and an opportunity for students to conduct research under the guidance of some of the nation's leading scholars in environmental and resource economics. The CEPA's current faculty have served on government advisory bodies, editorial boards, and have been visiting scholars at prominent universities around the globe.

---

[17] This discussion closely follows List and Lucking-Reiley (2002). I should note that none of these data are reported in List and Lucking-Reiley (2002), however.

Third, I constructed a personalised letter of solicitation that noted CEPA's role within the Central Florida community and the number of solicitations sent out. I also noted the tax deductibility of the contribution, and closed the letter with contact information in case the donors had questions. The solicitation letter is available at www.res.org/economic/ta/tahome.asp.

3.1. Experimental Results II

Table 5(a) reports summary statistics. Of the 2,000 solicitation packets sent, 54 were 'returned to sender', leaving a total sample of 1,946. Of the 1,946, 55.2% were mailed to male head of households, while 44.8% were delivered to female head of households. Average age was 47.2, average income was $92,619, and average education was 4.76, where education is again measured categorically (see the notes to Table 5). In total, about 4.5% of individuals donated money to CEPA. From these 90 donors, CEPA received a total of $1,895 in donation

Table 5

Charitable Giving Summary Estimates

			Mean (Std. Dev.)
(a) Descriptive Statistics			
Gift			$0.97
			(5.41)
Gift|Giving			$21.06
			(14.5)
Income			$92,619
			(154,780)
Age			47.2
			(14.05)
Gender (% male)			0.552
			(0.50)
Education			4.76
			(1.47)
n			1,946

	No. of Givers	Average Gift	Gift|Giving
(b) Age/Gender Cells			
Pooled Data (*n* = 1,946)	90 (4.6%)	$0.97 (5.4)	$21.06 (14.5)
Women (*n* = 871)	45 (5.2%)	$0.99 (5.0)	$19.22 (14.5)
Young Women (*n* = 552)	24 (4.3%)	$0.70 (3.9)	$16.04 (10.5)
Mature Women (*n* = 319)	21 (6.6%)	$1.50 (6.4)	$22.86 (12.2)
Men (*n* = 1,075)	45 (4.2%)	$0.96 (5.7)	$22.89 (16.8)
Young Men (*n* = 586)	12 (2.1%)	$0.26 (2.3)	$12.92 (10.1)
Mature Men (*n* = 489)	33 (6.8%)	$1.79 (8.0)	$26.52 (17.4)

Notes:
Education denotes categorical variable (1) Eighth grade or less, (2) High School, (3) Vocational, (4) 2-Year College, (5) Other Post-High School, (6) 4-Year College, (7) Graduate School Education.
In (*b*), following the previous definitions, 'young' women and men are those who are less than 50 years old. 'Mature' is defined as subjects 50 years old and above.
Gift|*Giving* denotes average gift conditional on giving.
Figures to the right of 'Average Gift' and 'Gift|Giving' are standard deviations.

© Royal Economic Society 2004

cheques. This amounted to roughly $1 for every solicitation packet sent. The average gift[18] conditional on giving, denoted by Gift|Giving, was $21.06.

Table 5(b) presents information about various subsamples of the data, with special attention paid to age and gender. The first row provides summary information for the pooled sample that corresponds with data in Table 5(a). Upon parsing the sample along gender and age dimensions, and using the above criteria for denoting 'young' and 'mature' (young: age < 50; mature: age > 49), I find results consonant with the research hypothesis. For example, empirical results in the first column suggest that the proportion of both mature women and men who give is higher than comparable proportions among the younger cohort – women: 6.6% versus 4.3%; men: 6.8% versus 2.1%. Using a test of proportions, I find that the difference in rates of giving among men is significant at the $p < 0.01$ level, whereas the difference is not significant at conventional levels for women ($z = 1.44$).

Columns 2 and 3 in Table 5(b) provide summary measures for the actual contributions. Figures in column 2 suggest that for every solicitation packet mailed, CEPA received between $0.26 and $1.79 from the subsamples, with the lowest return rates coming from young men and the highest return rates from mature men. According to a large sample t-test, average gift rates are significantly higher in the *Mature Cohort* in both gender subsamples (women: $t = 2.02$; men: $t = 4.09$).[19] Yet, it should be noted that again behavioural differences among men appear to be greater than behavioural differences among women – the average gift difference among young and mature men is $1.53 (588%), whereas the average gift difference among women is only $0.80 (114%). Data in column 3 suggest similar insights: conditional on giving, results from a small-sample t-test imply that both mature women and mature men give more than their younger counterparts (women: $t = 2.02$; men: $t = 2.54$). The observed differences are again larger in the male cohort.

These unconditional results strengthen findings from the public goods games and open up the possibility that in the realm of charitable contributions, age and giving are correlated. Given that charitable giving profiles appear to have different temporal aspects across gender, with men's rates of giving and gift size showing much larger increases over time than women's, these results may also indicate that the age effects observed in the public goods game are indeed age effects rather than purely generational effects. While I cannot completely rule out generational effects with the data herein, evidence in favour of the age effect is reinforced if one considers further differences across gender cohorts in of Table 5(b): even though

---

[18] Akin to List and Lucking-Reiley (2002), this level of giving is much higher than fundraising handbooks would have predicted – Mixer (1993) reports that direct-mail solicitations typically yield a response rate of only about 1 percent from 'cold' lists of potential donors. Perhaps the high level of giving herein is due to (i) the nature of the cause and/or (ii) the targeted subsample (high income Central Floridians who have recently given to a charity). Interestingly, soliciting another 3,000 people from the same mailing list as in the present paper, I attempted to raise money for another Central Florida charity, Soldiers to Scholars, which helps decommissioned armed-services veterans go to college and mentor inner-city children. Despite putting up $2,400 in lottery prizes for donors, I managed to receive a total of only $188 in contributions from five people.

[19] Similar to the data from the public goods games above, these estimated age effects are in the range of Feldstein and Taylor's (1976) findings on charitable gifts across age cohorts.

the proportion of females that donate in the *Young Cohort* (4.3%) is significantly larger than the proportion of male donators in the *Young Cohort* (2.1%) at the $p < 0.01$ level ($z = 2.21$), there is no statistical difference across gender in the *Mature Cohort* ($z = 0.09$). This finding holds for the average gift (young women versus young men: $z = 2.30$; mature women versus mature men: $z = 0.56$), and directionally, but not significantly, for the conditional gift amount (young women versus young men: $z = 0.85$; mature women versus mature men: $z = 0.84$).

While these unconditional results provide evidence that is in line with findings in the public goods game, other subject-specific factors that are left uncontrolled, such as income and education, could be the factors driving the results. To condition on these factors, I estimate variants of the following specifications:

$$Donate \,=\, \varphi + \boldsymbol{\beta X} + \varepsilon_i, \tag{3}$$

$$Gift \,=\, \alpha + \boldsymbol{\Omega X} + e_i, \tag{4}$$

where *Donate* $= 1$ if individual i donated, 0 otherwise; *Gift* represents individual i's contribution gift to CEPA. Variables in vector $\boldsymbol{X}$ follow from above and include age, income, gender, and education. I allow age to again enter as a dichotomous variable: *Mature Cohort* $= 1$ if age > 49, 0 otherwise, and in separate specifications as a continuous regressor. Also, since the sample is now comprised of a fair number of men and women, I include an interaction term (gender × age or gender × mature cohort) to explore if heterogeneous temporal contribution paths exist across gender. Before discussing the estimation results, I should briefly mention a few noteworthy aspects of (3) and (4). First, I estimated (3) assuming both normal and logistic error structures and found very few differences across these specifications, hence I report binomial probit estimates. Second, since 95.4% of the sample reported a zero *Gift* value, estimation of (4) must account for this censoring – I use a Tobit model. Third, sample sizes are slightly smaller for these estimation models since some variables, such as income and education, were missing from the data file. Finally, all specifications explain a significant portion of the dependent variable, and reported point estimates are marginal effects computed at the conditional sample means.

The general pattern of results in Table 6 is in line with findings in Table 5. For example, both in the participation empirical estimates in columns 1 and 2 and the gift magnitude results in columns 3 and 4, I find evidence that the mature cohort exhibits a greater concern for the charity than the younger cohort. This effect is consistent across age measures: in the specifications measuring age dichotomously, the marginal effect is significant at the $p < 0.10$ level using a one-sided alternative in both specifications (column 1 and 3), and in the specifications using a continuous age variable the marginal effect is significant in (4) but not (3). There is also evidence that the interaction terms, male × age and male × mature cohort, are important – in each case the marginal effect is different from zero at the $p < 0.10$ level using a one-sided alternative. The response coefficients suggest that the probability of participating and the gift level both increase over time for men much more than for women. In terms of other parameter estimates, the control variables are typically insignificant, but one robust result is the finding that women

Table 6

Estimation Results for Campaign Fundraiser

	Equation (3)		Equation (4)	
Variable	(1)	(2)	(1)	(2)
Constant	−0.15**	−0.16**	0.59	−0.18
	(0.19)	(0.03)	(0.48)	(0.83)
Mature Cohort	0.02	–	0.89**	–
	(0.01)		(0.40)	
Age	–	0.4E-3	–	0.03*
		(0.4E-3)		(0.01)
Male × Mature Cohort	0.03	–	0.63	–
	(0.02)		(0.50)	
Male × Age	–	0.1E-2**	–	0.03*
		(0.7E-3)		(0.02)
Male	−0.03**	−0.08**	−0.43	−1.59*
	(0.02)	(0.03)	(0.33)	(0.88)
Income	−0.2E-7	−0.2E-7	−0.3E-6	−0.4E-6
	(0.5E-7)	(0.5E-7)	(0.8E-6)	(0.8E-6)
Education	−0.002	−0.002	0.03	0.02
	(0.003)	(0.003)	(0.08)	(0.09)
N	1,886	1,886	1,886	1,886

Notes:
Dependent variable is *Donate* in equation (3). *Gift* in (4). *Donate* = 1 if the individual donated, 0 otherwise, thus (3) is a probit model; *Gift* is actual gift level. Gender = 1 if male, 0 otherwise; *Mature Cohort* = 1 if age > 49, 0 otherwise. *Income* is annual income, *Education* is defined in Table 5 footnotes. Point estimates are marginal effects calculated at the sample means. Standard errors are in parentheses beneath point estimates.
**Denotes significant at the $p < 0.05$ level. *Denotes significant at the $p < 0.10$ level.

opt to participate more than men (columns 1 and 2), and the marginal return from sending a solicitation letter to women is higher than the marginal return from solicitations sent to men (columns 3 and 4). This effect, which is also observed in the raw data displayed in Table 5, provides insights into an important gender difference that should be more fully explored in the future.[20]

4. Other Empirical Evidence

Even though the data in both field environments reveal similar insights, it is important to determine if there is evidence of such a relationship in other environments. In this section, I summarise results from a natural experiment of the prisoner's dilemma game by examining 240 individual decisions from the gameshow *Friend or Foe?* The gameshow provides the simplest and sharpest form of the dilemma and strikingly parallels the classic prisoner's dilemma story: teams are endogenously determined, players work together to earn money, after which the agents play a one-shot prisoner's dilemma game over large stakes – varying from

[20] Note that in these specifications I use personal income. If I use household income I find marginally significant and positive point estimates on the income variable. Unfortunately, I do not have access to data on household wealth. This aspect of the study should not be downplayed, as future studies that control for wealth can lend invaluable insights into the relationship between age and social preferences.

$200 to (potentially) more than $22,000. If one were to conduct such an experiment in the laboratory, the cost to gather the 240 data points would be more than $350,000.

Television game shows provide a natural avenue to observe real decisions in an environment with high stakes. For example, Berk *et al.* (1996) and Tenorio and Cason (2002) study contestants' behaviour on *The Price Is Right* to investigate rational decision theory and whether subjects play the unique subgame perfect Nash equilibrium. Gertner (1993) and Beetsma and Schotman (2001) make use of data from *Card Sharks* and *Lingo*, respectively, to examine individual risk preferences. Finally, Metrick (1995) uses data from *Jeopardy!* to analyse behaviour under uncertainty and players' ability to choose strategic best-responses. Each of these studies has provided a fresh look at important economic phenomena over large stakes that would be difficult to investigate outside of a gameshow environment.

The data used in this Section were taken from original *Friend or Foe?* programmes broadcast between May 2002 and September 2002 (May 2002 was the first airing of the show). There are 40 total shows yielding decisions made by 240 subjects. The demographic data provided in the game show includes age, race, gender, and current area of residence. The availability of these demographic characteristics allows a richer analysis into prisoner's dilemma games than heretofore available.

For the sake of parsimony, I provide only a brief description of the gameshow, but direct the reader to List (2002), which more fully explains the show and a broader set of findings with these data. Each show begins with six perfect strangers, who at the beginning of the show are randomly split into two groups of three: group 1 and group 2. Each group 1 agent privately selects one player from group 2 to be his or her partner. After the three groups of two are determined, each team is separated into 'isolation chambers' where all trivia rounds are played. The newly formed teams work together and agree on answers to trivia questions in order to build a 'bank account'. In theory, since each team is initially endowed with $200, a team's bank account can range from $200 to $22,200; however, in practice, the largest trivia earnings in the data is $16,400. At the end of each round (there are 3 rounds in total) the lowest scoring team is eliminated. Before the team is dismissed, the players must decide how their winnings are divided.

In making his or her decision on division of the winnings, each player has a button, which no one else can see. The division depends on whether the player depresses his or her button. There are three possible outcomes:

(1) 'Friend-Friend' (i.e., cooperate-cooperate) – if both players choose 'friend' (do not depress the button), then the total winnings are divided equally between the two;

(2) 'Friend-Foe' (i.e., cooperate-not cooperate) – if only one player depresses the button, he or she will walk away with the entire amount, leaving the other player with nothing;

(3) 'Foe-Foe' (i.e., not cooperate-not cooperate) – if both players press the button, then all of the money is lost, and both players walk away with nothing. Thus, this final stage of the game naturally sets up a prisoner's

dilemma with a weakly dominant strategy: each player has an incentive to play 'Foe' because she is never worse-off monetarily when playing 'Foe'.

While there are several interesting behavioural tendencies one can explore in this setting, herein I summarise the observed relationship between age and cooperation rates. The top panel of Table 7 provides summary statistics of the players; of the 240 players, 49% were men and 73% were white. The contestants had an average age of 31, and 38% of the contestants were from California. This mixture of subjects provides the necessary variation to examine whether important demographic and geographic characteristics are correlated with prisoner's dilemma outcomes. The bottom panel of Table 7 presents an overview of monies earned and unconditional cooperation rates across a few broad classes of players. These figures show that on average the teams earned $3,690 during the trivia portion of the game, suggesting that the average prisoner's dilemma game was played over $3,690. Teams that included men earned the most in the trivia portion

Table 7

Selected Characteristics of Game Show Participants

	Mean (Std. Dev.)	Minimum	Maximum
Gender (% male)	0.49 (0.50)	0	1
Age	31.0 (8.70)	18	61
Race (% white)	0.73 (0.44)	0	1
% Californians	0.38 (0.49)	0	1
n	240	–	–

Outcomes	Trivia Earnings	Cooperation Rate	Take-Home Earnings
Overall	$3,690	0.51	$1,460
($n = 240$)	(2,959)	(0.50)	(2,308)
Men	$4,267	0.44	$1,856
($n = 117$)	(3,275)	(0.50)	(2,740)
Women	$3,143	0.57	$1,082
($n = 123$)	(2,517)	(0.50)	(1,734)
White	$3,936	0.54	$1,382
($n = 177$)	(3,131)	(0.50)	(2,350)
Non-White	$3,001	0.41	$1,676
($n = 63$)	(2,297)	(0.49)	(2,191)
Young (age < 31)	$3,617	0.40	$1,639
($n = 134$)	(2,787)	(0.49)	(2,489)
Mature (age ≥ 31)	$3,783	0.64	$1,231
($n = 106$)	(3,174)	(0.48)	(2,045)

Notes:
1. *Age* denotes actual age in years.
2. *Gender* denotes categorical variable: 0 if female, 1 if male.
3. 'Trivia earnings' denotes amount earned in the trivia part of the show; 'Cooperation Rate' denotes the percentage of subjects that chose cooperate; 'Take-Home Earnings' denotes the amount of money taken home by the subject.

of the game, on average ($4,267), while teams including non-whites earned the least, on average ($3,001).

In terms of cooperation rates, one glaring result in the bottom panel of Table 7 is the cooperation rates across age cohorts. Simply splitting the sample along the first moment of the age variable shows that more mature agents (31 years and older) chose to cooperate much more than younger agents: 64% versus 40%. Constructing a simple 2×2 matrix across the age variable yields Table 8. Table 8 provides the level of trivia earnings (denoted stakes), cooperation rates, average take-home earnings, efficiency rates (take-home earnings divided by trivia earnings), and sample size. Insights contained in Table 8 complement Table 7 and suggest that older agents were more likely to cooperate than younger agents, regardless of partner's age.

Rather than belabour the raw data further, to provide insights into conditional cooperation rates at the individual level, I estimate the following logistic model:

$$cooperate_i = g(\alpha + \boldsymbol{\beta}'\mathbf{Z} + \beta_S \ stakes + \boldsymbol{\beta}'_c \ \textbf{selection}), \tag{5}$$

where $cooperate_i$ equals 1 if agent i chooses 'Friend' (cooperate), 0 otherwise; $g(\cdot)$ is the standard logit function; $\mathbf{Z}$ is a vector of observables, which includes agent i's race, gender, age, and geographic place of residence; $stakes$ is a continuous measure for the amount of money at risk; **selection** is a vector of dichotomous indicator variables of whether the agents formed a partnership by choice and whether the partnership was formed in the first selection period (as described above).

Empirical results from estimation of (5) are contained in Table 9. These results correspond well to the empirical results from the public goods games and the charitable contribution drive above. For example, first note that age is significant

Table 8

Summary Gameshow Outcomes

	Young	Mature
Young (less than 31 years old)	Stakes: $3,714 Cooperation rate: Young: 0.39 Avg. take-home: $1,307 Efficiency: 0.70 $N = 68$	Stakes: $3,516 Cooperation rates: Young: 0.40 Mature: 0.63 Avg. take-home: Young: $1,982 Mature: $881 Efficiency: 0.81 $N = 132$
Mature (31 years old or older)	Stakes: $3,516 Cooperation rates: Young: 0.40 Mature: 0.63 Avg. take-home: Young: $1,982 Mature: $881 Efficiency: 0.81 $N = 132$	Stakes: $4,225 Cooperation rate: Mature: 0.65 Avg. take-home: $1,810 Efficiency: 0.86 $N = 40$

Table 9

Gameshow Regression Estimates

	Individual Logit (Prob. coop.)	Individual Logit (Prob. coop.)
Male	−0.15*	−1.94[†]
	(0.07)	(1.15)
White	0.12[†]	0.54[†]
	(0.08)	(0.32)
Age	0.02*	0.05*
	(0.004)	(0.02)
Male × Age	–	0.05[†]
		(0.03)
Californian	−0.15*	−0.56[†]
	(0.07)	(0.29)
Stakes	0.1E-4	0.1E-4
	(0.1E-4)	(0.1E-4)
1st Selected	0.05	0.05
	(0.08)	(0.08)
1st Selected (Tie)	0.10	0.10
	(0.08)	(0.08)
Constant	−0.56*	−1.69*
	(0.16)	(0.79)
n	240	240

Notes:
In the individual logit model, the dependent variable equals 1 if player chose cooperate, 0 otherwise. Figures in parentheses beneath estimated marginal effects are standard errors. Some marginal effects are denoted in scientific notation (e.g., E-4).
*Significant at the $p < 0.05$ level; [†]Significant at the $p < 0.10$ level.

at the $p < 0.05$ level and positive in both models. This result, which suggests older agents cooperate more than younger agents *ceteris paribus,* holds for specifications measuring age dichotomously as well. Second, similar to the data in the charitable fundraiser, there is evidence that the interaction term, male × age, is important – the marginal effect is different from zero at the $p < 0.10$ level. The response coefficient also has the same sign as in the charitable contribution equation and suggests that the probability of cooperating increases over time for men much more than for women.

As in most of the literature examining the relationship between gender and altruism, this study finds that age and social preferences are inextricably linked. In the spirit of providing a roadmap to model the relationship observed, one could extend the model of impure altruism in Andreoni (1990) to a dynamic setting whereby the altruistic component (G) and/or the warm glow component (g_i) in the following utility function increase over time: $U_i = U_i(x_i, g_i, G)$, where x_i and g_i are person i's consumption of a composite private good and their gift to the public good, and G is the total amount of the public good provided. Using this theory, the results herein would suggest that, independent of income levels, more mature citizens tend to value public goods (G) to a greater extent than younger citizens and/or over time people tend to place higher value on their own gifts to the public good (g_i). Likewise one could envision a dynamic learning theory in the spirit of Sugden (1984) having considerable predictive power. Sugden (1984) proposes a

model of the voluntary sector based on the assumption that people believe free-riding is morally wrong; if this belief strengthens over time, free-riding may be observed less often for older individuals.[21]

5. Concluding Comments

As society grows older the demand to understand preferences and values of more mature individuals becomes increasingly important. This study attempts to take this step by gathering data from a variety of field experiments, which allows tests of several research hypotheses. The data reveal two major insights. A first result that should be of interest to the general economics community is that laboratory results in the popular voluntary contributions game are readily transferable to more universal populations: neither the 'found money' characteristic nor other nuances associated with the laboratory setting (e.g., narrow subject pools, lack of social distance between experimenters and subjects, and other subtle character-istics associated with the tight controls of the lab) unduly influence observed behaviour. This result is comforting since a re-evaluation of a good deal of results from laboratory experiments would be necessary if important findings gained via the lab did not readily spill over to more hostile environments. Accordingly, the empirical findings herein may permit sceptics to view results from the lab less suspiciously.

A second result that has not been observed in the lab or theoretically explained is that age and social preferences appear linked. A positive relationship is observed across several diverse settings – from a simple linear public goods game at a sportscard show to a university fundraising drive to a prisoner's dilemma game on television. While further research is necessary before any solid policy prescriptions can be advanced, this result should be of broad interest to scholars, policymakers, and practitioners in the design of fundraising campaigns. For scholars, this result represents a call to take account of such factors when building theoretical and empirical models. Policymakers may be interested in such findings because an important source of finance for private nonprofit organisations remains individ-uals, who undoubtedly respond heterogeneously to tax reforms, for example. This result has obvious implications for charity fundraisers, as it potentially can have an important influence on fundraising practices.

University of Maryland

---

[21] The natural question that follows from this line of inquiry is why are agents' preferences tem-porally variant? One natural avenue of speculation involves evolutionary factors. While several plausible stories readily emerge in an evolutionary framework, I will leave this modelling exercise for future research, but point the reader to related research in the field of social preferences; see Camerer (2002) and Sobel (2001) for recent reviews.

References

Andreoni, J. (1988*a*). 'Privately provided public goods in a large economy: the limits of altruism', *Journal of Public Economics*, vol. 35(1), pp. 57–74.

Andreoni, J. (1988*b*). 'Why free ride? Strategies and learning in public goods experiments', *Journal of Public Economics*, vol. 37(3), pp. 291–304.

Andreoni, J. (1990). 'Impure altruism and donations to public goods: a theory of warm-glow giving', ECONOMIC JOURNAL, vol. 100, pp. 464–77.

Andreoni, J. and Vesterlund, L. (2001). 'Which is the fair sex? Gender differences in altruism', *Quarterly Journal of Economics*, vol. 116, pp. 293–312.

Battalio, R., Kagel, J. and Jiranyakul, K. (1990). 'Testing between alternative models of choice under uncertainty: some initial results', *Journal of Risk and Uncertainty*, vol. 3, pp. 25–50.

Beetsma, R. M. W. J. and Schotman, P. C. (2001). 'Measuring risk attitudes in a natural experiment: data from the television game show lingo', ECONOMIC JOURNAL, vol. 111(474), pp. 821–48.

Berk, J. B., Hughson, E. and Vandezande, K. (1996). 'The price is right, but are the bids? An investigation of rational decision theory', *American Economic Review*, vol. 86(4), pp. 954–70.

Bohm, P. (1984). 'Revealing demand for an actual public good', *Journal of Public Economics*, vol. 24, pp. 135–54.

Bohm, P. (1994). 'Time preference and preference reversal among experienced subjects – the effects of realpayments', ECONOMIC JOURNAL, vol. 104, pp. 1370–8.

Cadsby, B. B. and Maynes, E. (1998). 'Choosing between a socially efficient and free-riding equilibrium: nurses versus economics and business students', *Journal of Economic Behavior and Organization*, vol. 37, pp. 183–92.

Camerer, C. (2002). *Behavioral Game Theory: Experiments on Strategic Interaction*, Princeton, NJ: Princeton University Press.

Davis D. D. and Holt, C. A. (1992). *Experimental Economics*, Princeton, NJ: Princeton University Press.

Eckel C. C. and Grossman, P. (1997). 'Are women less selfish than men? Evidence from dictator experiments', ECONOMIC JOURNAL, vol. 108, pp. 726–35.

Falkinger, J., Fehr, E., Gaechter, S. and Winter-Ebmer, R. (2000). 'A simple mechanism for the efficient provision of public goods: experimental evidence', *American Economic Review*, vol. 90(1), pp. 247–64.

Fehr, E. and Gaechter, S. (2000). 'Cooperation and punishment in public goods experiments', *American Economic Review*, vol. 90(4), pp. 980–94.

Fehr, E. and Schmidt, K. (1999). 'A theory of fairness, competition, and cooperation', *Quarterly Journal of Economics*, vol. 114, pp. 817–68.

Fehr, E., Kirchler, E., Weichbold, A. and Gaechter, S. (1998). 'When social norms overpower competition: gift exchange in experimental labor markets', *Journal of Labor Economics*, vol. 16(2), pp. 324–51.

Feldstein, M. and Taylor, A. (1976). 'The income tax and charitable contributions', *Econometrica*, vol. 44, pp. 1201–22.

Gertner, R. (1993). 'Game shows and economic behavior: risk-taking on card sharks', *Quarterly Journal of Economics*, vol. 108(2), pp. 507–21.

Hannan, L., Kagel, J. and Moser, D. (1999). 'Partial gift exchange in experimental labor markets: impact of subject population differences, productivity differences, and effort requests on behavior', University of Pittsburgh, Working Paper.

Harbaugh, W. T. and Krause, K. (2000). 'Children's contributions in public good experiments: the development of altruistic and free-riding behaviors', *Economic Inquiry*, vol. 38(1), pp. 95–109.

Isaac, R. M., Walker, J. M. and Williams, A. W. (1994). 'Group size and the voluntary provision of public goods: experimental evidence utilizing large groups', *Journal of Public Economics*, vol. 54(1), pp. 1–36.

Keeler, J., James, W. and Abdel-Ghany, M. (1985). 'The relative size of windfall income and the permanent income hypothesis', *Journal of Business Economics and Statistics*, vol. 3(3), pp. 209–15.

List, J. A. (2001). 'Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auctions for sportscards', *American Economic Review*, vol. 91, pp. 1498–507.

List, J. A. (2002). 'Friend or foe: a natural experiment of the prisoner's dilemma', University of Maryland, Working Paper.

List, J. A. and Lucking-Reiley, D. (2002). 'Seed money matters: experimental evidence from a university capital fundraising campaign', *Journal of Political Economy*, vol. 110, pp. 215–33.

Marwell, G. and Ames, R. E. (1981). 'Economists free ride, does anyone else?', *Journal of Public Economics*, vol. 15(3), pp. 295–310.

Metrick, A. (1995). 'A natural experiment in jeopardy!', *American Economic Review*, vol. 85(1), pp. 240–53.

Mixer, J. R. (1993). *Principles of Successful Fundraising: Useful Foundations for Successful Practice*, San Francisco: Jossey-Bass.

Reece, W. S. (1979). 'Charitable contributions: new evidence on household behavior', *American Economic Review*, vol. 69, pp. 142–51.

Samuelson, P. (1954). 'The pure theory of public expenditures', *Review of Economics and Statistics*, vol. 36, pp. 387–9.

Sobel, J. (2001). 'Interdependent preferences and reciprocity', University of California, San Diego, Working Paper.

Sugden, R. (1984). 'Reciprocity: the supply of public goods through voluntary contributions', ECONOMIC JOURNAL, vol. 94, pp. 772–87.

Tenorio, R. and Cason, T. (2002). 'To spin or not to spin? Natural and laboratory experiments from ''The Price is Right''', ECONOMIC JOURNAL, vol. 112, pp. 170–95.