CEDEX

CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

The University of Nottingham

Discussion Paper No. 2012-14

J. Anderson, S. V. Burks,
J. Carpenter, L. Götte,
K. Maurer, D. Nosenzo,
R. Potter, K. Rocha and
A. Rustichini

May 2012

Self-Selection and Variations in the Laboratory Measurement of Other-Regarding Preferences across Subject-Pools: Evidence from One College Student and Two Adult Samples

CeDEx Discussion Paper Series
ISSN 1749 - 3293

CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

The Centre for Decision Research and Experimental Economics was founded in 2000, and is based in the School of Economics at the University of Nottingham.

The focus for the Centre is research into individual and strategic decision-making using a combination of theoretical and experimental methods. On the theory side, members of the Centre investigate individual choice under uncertainty, cooperative and non-cooperative game theory, as well as theories of psychology, bounded rationality and evolutionary game theory. Members of the Centre have applied experimental methods in the fields of public economics, individual choice under risk and uncertainty, strategic interaction, and the performance of auctions, markets and other economic institutions. Much of the Centre's research involves collaborative projects with researchers from other departments in the UK and overseas.

Please visit http://www.nottingham.ac.uk/cedex for more information about the Centre or contact

Sue MacCormick
Centre for Decision Research and Experimental Economics
School of Economics
University of Nottingham
University Park
Nottingham
NG7 2RD
Tel: +44 (0)115 95 15469
Fax: +44 (0) 115 95 14159
sue.maccormick@nottingham.ac.uk

The full list of CeDEx Discussion Papers is available at

http://www.nottingham.ac.uk/cedex/publications/discussion-papers/index.aspx

Self-Selection and Variations in the Laboratory Measurement of Other-Regarding Preferences across Subject Pools: Evidence from One College Student and Two Adult Samples

Jon Anderson[1], Stephen V. Burks[2], Jeffrey Carpenter[3],
Lorenz Götte[4], Karsten Maurer[5], Daniele Nosenzo[6 *],
Ruth Potter[7], Kim Rocha[7], Aldo Rustichini[8]

May 19, 2012
[forthcoming, Experimental Economics]

Abstract: We measure the other-regarding behavior in samples from three related populations in the upper Midwest of the United States: college students, non-student adults from the community surrounding the college, and adult trainee truckers in a residential training program. The use of typical experimental economics recruitment procedures made the first two groups substantially self-selected. Because the context reduced the opportunity cost of participating dramatically, 91% of the adult trainees solicited participated, leaving little scope for self-selection in this sample. We find no differences in the elicited other-regarding preferences between the self-selected adults and the adult trainees, suggesting that selection is unlikely to bias inferences about the prevalence of other-regarding preferences among non-student adult subjects. Our data also reject the more specific hypothesis that approval-seeking subjects are the ones most likely to select into experiments. Finally, we observe a large difference between self-selected college students and self-selected adults: the students appear considerably less pro-social.

Keywords: methodology; selection bias; laboratory experiment; field experiment; other-regarding behavior, social preferences, prisoner's dilemma, truckload, trucker.

JEL: C90, D03.

[*]Corresponding author. E-mail: daniele.nosenzo@nottingham.ac.uk; Tel: 0044(0)1158467492; Fax: 0044(0)1159514159
[1] Division of Science and Mathematics, University of Minnesota, Morris (600 East 4th Street, Morris, MN 56267, USA)
[2] Division of Social Science, University of Minnesota, Morris (600 East 4th Street, Morris, MN 56267, USA), IZA (P.O. Box 7240 D-53072, Bonn, Germany), and CeDEx, University of Nottingham (University Park, Nottingham, NG7 2RD, UK)
[3] Department of Economics, Middlebury College (Warner Hall, 303 College Street, Middlebury, VT 05753, USA), and IZA (P.O. Box 7240 D-53072, Bonn, Germany)
[4] Faculty of Business and Economics, University of Lausanne (Internef CH-1015 Lausanne, Switzerland) and IZA (P.O. Box 7240 D-53072, Bonn, Germany)
[5] Department of Statistics, Iowa State University (Snedecor Hall Ames, IA 50012, USA)
[6] School of Economics and CeDEx, University of Nottingham (University Park, Nottingham, NG7 2RD, UK)
[7] Division of Social Science, University of Minnesota, Morris (600 East 4th Street, Morris, MN 56267, USA)
[8] Department of Economics, University of Minnesota, Twin Cities (4-101 Hanson Hall, 1925 4th Street S Minneapolis, MN 55455, USA) and Faculty of Economics, University of Cambridge (Sidgwick Avenue,Cambridge,CB3 9DD, UK)

1. Introduction

A considerable body of evidence has now accumulated from economic experiments that many individuals exhibit "other-regarding preferences": not only do they care about their personal material payoffs from social and economic interactions, but they also care about the payoffs of other agents with whom they interact. The designs of such experiments control, at least to a good first approximation, for potentially confounding reasons - such as repeated interactions or reputation effects - that could lead to what appears to be other-regarding behavior, but is really sophisticated self-interest. For example, in the context of voluntary cooperation games only about a third of the participants in experiments typically behave in accordance with own monetary-payoff maximization. The majority of individuals seem instead motivated by other-regarding considerations. For example, more than half of the participants in public goods game experiments are found to be "conditionally cooperative" – they are willing to forgo material gain and cooperate if others cooperate as well (see, e.g., Fischbacher et al. 2001; Herrmann and Thoni 2009; Kocher et al. 2008). Analogous evidence of the importance of other-regarding preferences has been documented by experimental studies using dictator games, bargaining games, trust games, and gift-exchange games (for reviews see, e.g., Camerer 2003; Fehr and Schmidt 2006).

However, most of the economic experiments providing evidence for the importance of other-regarding motives have been conducted using samples of undergraduate college students who self-selected into participation in the studies. Generalizations from studies using self-selected college student samples could be problematic for two reasons. First, experimental studies relying on self-selected samples may overestimate the importance of other-regarding preferences if the process by which participants self-select into experiments is correlated with their preferences. For example, as suggested by Levitt and List (2007, p. 166) "… *volunteers … who have social preferences or who readily cooperate with the experimenter and seek social approval might be those who are most likely to participate in the experiment*." If this were the case, the pervasiveness of social and other-regarding behaviors documented in economic experiments could substantially reflect the endogenous process by which the experimental participants were selected rather than the underlying propensities of the population. Second, college students clearly differ in many ways from the general population (e.g. in terms of age,

education, social class and experience with markets and economic environments), and in principle it is possible that they may also differ in the strength of their other-regarding concerns.

In this paper we address these concerns by examining how other-regarding preferences measured in a laboratory experiment vary across three different samples of experimental subjects. One sample consists of undergraduate students who self-select into the laboratory experiment. The two other samples consist of participants recruited among the non-student adult population. An important difference between these two non-student samples is in the procedures used to recruit participants: in one case the recruitment procedures were similar to those used for recruiting undergraduates, and participants could self-select into the experiment. In the other case, the recruitment procedures allowed for very little self-selection of participants.

As described in detail in Section 2, we measure other-regarding preferences using a sequential social dilemma game in which players choose between an uncooperative action that leaves earnings unaffected, and cooperative actions that are costly for the player, but benefit their partner and increase total earnings. In the experiment decisions were elicited using the strategy method and subjects played both in the role of first-mover and in the role of second-mover. We use decisions in the role of second-mover to classify subjects in three main categories: Free-Riders, who do not display other-regarding concerns and choose the own-material-payoff maximizing actions, Conditional Cooperators, who behave cooperatively only if the first-mover behaves cooperatively, and Unconditional Cooperators, who behave cooperatively regardless of how the first-mover behaves. The latter two types both exhibit other-regarding concerns in the sense that they choose actions that are inconsistent with own-material-payoff maximization. Finally, our data also include a measure of subjects' need for social approval (the "Unlikely Virtues Scale", developed by Patrick et al., 2002), which we use to examine whether, as suggested by Levitt and List (2007), approval-seeking is positively related to decisions to self-select into experimental studies.

We report our results in Section 3. To examine whether other-regarding preferences are more widespread among self-selected participants than among non-self-selected participants we compare the distribution of cooperation types across the two samples of adult non-students. We find that self-selection does not distort the measuring of other-regarding preferences: the proportions of Free Riders, Conditional Cooperators, and Unconditional Cooperators do not differ significantly between the two groups. We also do not find any difference in the need for

social approval of self-selected and non-self-selected adult participants. To examine whether there are differences in the extent to which students and non-students engage in other-regarding behaviors we compare the sample of self-selected college students and the sample of self-selected adults. We find that the share of individuals exhibiting other-regarding concerns is remarkably smaller among college students, even after controlling for observable differences in socio-demographic characteristics between the two subject pools. Our finding that the impact of self-selection on measurements of other-regarding preferences in our two adult samples is negligible is in line with the results of two recent studies that also examine the issue of self-selection in economic experiments among college student subjects (Cleave et al., 2011; Falk et al., forthcoming 2012). The finding that college students are less other-regarding than non-students is also in line with the existing literature comparing student and non-student samples across experimental games. We review and discuss these related literatures in Section 4, and briefly summarize our conclusions in Section 5.

2. Experimental Design and Procedures

2.1. Subject pools

The data used in this paper were collected from 1,261 subjects who took part in the "Truckers & Turnover Project," an extensive experimental study run in two locations over two years and comprised of several decision tasks and questionnaires (Burks et al. 2008). Participants in the experiment belonged to one of three different samples, which differ in whether subjects were undergraduate college students or not, and/or in the procedures used to recruit them.

One-hundred subjects were students at the University of Minnesota, Morris (UMM). They were recruited by e-mail through the opt-out student list at UMM. An initial invitation e-mail was sent out asking for those interested in participating as paid volunteers in experiments to respond. Responders were then contacted via e-mail with information about specific session times and potential earnings, and allocated to sessions based on availability. These recruitment procedures are similar to those typically used for economic experiments. In particular, note that these participants self-selected into the experiment. We thus refer to this sample of subjects as *Self-Selected Students*.

Ninety-two subjects were recruited from the non-student population living in the vicinity of Morris, which is a town of 5,000 in a rural area. Recruitment was done by placing posters on

business bulletin boards in Morris. Posters contained information about the experiment, potential earnings, and possible session times, and had pre-paid mail-back postcards for those interested in participating. Responders were then contacted by telephone to arrange session allocation. Thus, the recruitment procedures used for this subject pool were similar to those used for college students. In particular, these ninety-two subjects also self-selected into the experiment, as the subjects in the student sample did. We thus refer to this sample as *Self-Selected Non-Students*.

The remaining one-thousand and sixty-nine subjects were also recruited from a population of subjects who were not enrolled in university or college. These subjects were trainee truck drivers at a driver training school in the U.S. Midwest operated by a large trucking firm which provides basic training to its new-to-the-industry employees. The researchers had the cooperation of the trucking firm that runs the school, and the data collection was designed to make the opportunity cost of participation particularly low. At the beginning of the class day, one of the authors (Burks) approached the trainees and conducted an informed consent process, in which he explained the goals and procedures of the experimental study to potential subjects. While trainees were informed that participation in the study was voluntary and that those who wished not to participate were excused, the relatively low opportunity cost of participating and the credible guarantee of confidentiality from the University[1] resulted in a very high participation rate: 91% of those offered the opportunity chose to join the study.[2] Thus, there is very little self-selection into the experiment by this group of subjects. We refer to the trainees sample as *Non-Self-Selected Trainee Truckers*.[3]

---

[1] Specifically, subjects were informed that the data was going to the University and not the firm (their new employer), and the role of the University's Institutional Review Board (IRB) in enforcing the promise of individual confidentiality was explained.

[2] With the cooperation of the training school, the study was run on Saturdays that came in the middle of a two-week residential basic training program. Lunch was provided and the buses to and from the trainee's lodgings arrived at an early hour and left at the end of the afternoon. Only a half day of training activity was scheduled, so trainees were split into two groups and in the morning one did training while the other took part in the study, with the reverse in the afternoon. Those not participating in the study with their group did not have extra training available and had to spend the time in a break room.

[3] In Section 3.1 we assess the potential implications for our results of the fact that 9.2% of trainee drivers did not take part. Here we would like to remark that our strategy of running an experiment with trainee truckers as a method to gather data on a relatively non-self-selected sample is similar to the use of classroom experiments with college students by related studies that address the selection issue (e.g., Cleave *et al.*(2011) or Eckel and Grossman (2000)). Running an experiment which is not announced in advance during the course of a regularly scheduled class is meant to minimize the potential for selection, although participation remains voluntary, and a non-random selection of students may be absent, as well. Recruiting an adult sample to voluntary participation is in general likely to make it harder, not easier, to achieve a sample with low self-selection relative to student samples. Students may be more prone to comply with requests made by a relevant authority figure (the professor/experimenter). Even so there is normally still some self-selection (for example, 2% of potential subjects declined to participate in the classroom

All subjects were exposed to the same experimental protocol (see Burks *et al.* (2008) for more details). At the beginning of each session subjects were guided through a consent form that explained the conditions for participation in the study. The experiment was set up as two two-hour-long blocks that subjects spent doing tasks with the researchers, either on computers or with paper and pencil, with a short break in between.[4] The part of the experimental design used in the current study is described in detail in the next sub-section. At the beginning of each two-hour-long block subjects received a fixed payment of $10 for their participation, and could earn additional money in the course of the experiment depending on their performance.[5] Sessions were run with groups ranging from 20 to 40 subjects at a time. Four sessions were conducted with Self-Selected Students, three sessions with Self-Selected Non-Students, and forty-six sessions with Non-Self-Selected Trainee Truckers.

2.2. Experimental measurements

Our measurement of subjects' other-regarding preferences is based on the decisions they made in the following social dilemma game. At the outset of the "Two-Person Sending Task" two players, Person 1 and Person 2, are each allocated $5. Person 1 moves first and chooses an amount $s_1 \in \{\$0, \$5\}$ to send to Person 2. Person 2 learns Person 1's decision and then chooses an amount $s_2 \in \{\$0, \$1, \$2, \$3, \$4, \$5\}$ to send back to Person 1. Any amount sent by either player is doubled by the experimenter, and this is common knowledge.[6] After Person 2's decision, the game ends. Instructions for the Two-Person Sending Task are available in Appendix A. Payoffs were shown to subjects using a payoff table (see Appendix B).

In the experiment subjects played the game exactly once, and were asked to make decisions in both roles knowing that the final assignment to roles would be randomly determined at the end of the experiment. On a first screen subjects were asked to make a decision in the role of Person 1, and on a second screen a decision in the role of Person 2. Person 2's decisions were elicited using the strategy method, i.e. subjects had to specify the amount they intended to transfer to Person 1 both for the case where Person 1 had sent $0 and for the case where Person 1 had sent

---

experiments of Cleave *et al.* (2011)). There is also a potential cost that is not likely to be as high with adults: student participants in classroom experiments, exposed to an authority figure as the experimenter, may have an increased potential for experimenter demand effects.

[4] The full set of activities thus took four and a half hours. The computerized tasks were programmed and implemented with the software z-Tree (Fischbacher 2007).

[5] The fixed payments were doubled for Self-Selected Non-Students because on average they faced relatively higher opportunity costs to participation, since they had to come to campus from the surrounding town.

[6] Thus, a technically correct label for this game is a sequential and strategic form of the prisoner's dilemma.

$5. Thus, subjects in the experiment were asked to make three decisions in total: one decision in the role of Person 1 and two decisions in the role of Person 2. Once all decisions had been made, subjects were anonymously and randomly matched with another participant in the room, were randomly assigned a role, and were shown their payoffs according to the decisions they had made in that role. On average, subjects earned $8.32 from the social dilemma game, with a minimum of $0 and a maximum of $16.[7]

The sequential social dilemma game described above was the first task that subjects performed in the experiment. Of the subsequent tasks that subjects had to complete one is of particular interest for the purposes of this study. After their choices in the social dilemma game, subjects were asked to fill out the brief form of the Multidimensional Personality Questionnaire (MPQ) developed by Patrick *et al.* (2002). This is a personality profile test consisting of eleven different scales representing primary trait dimensions, and one 13-item scale (the "Unlikely Virtues Scale") which provides a stand-alone index of social desirability.[8] Scores in the version of the Unlikely Virtues Scale administered to subjects can range from 13 to 52 and high scores result when subjects over-report uncommon "good behaviors" (e.g. they answer positively to questions such as "*Never in my whole life have I taken advantage of anyone*") and under-report common "bad behaviors" (e.g. they answer negatively to questions such as "*I have sometimes felt slightly hesitant about helping someone who asked me to*"). In our sample we find that the intercorrelations of the 13 items of the scale exceed the standard threshold (Chronbach's alpha is 0.73) and so we conclude that our implementation has resulted in capturing the desired latent trait. We will use the Unlikely Virtues Scale to examine whether there is any relation between approval-seeking (in the form of more socially desirable responding) and decisions to take part in economic experiments.

As part of the experimental design, subjects completed a questionnaire collecting basic socio-demographic information. Table 1 presents a summary of subjects' socio-demographic characteristics disaggregated by subject pool.[9]

---

[7] Before each decision screen subjects were also asked to predict the behavior of the other participants in the room, and received additional earnings for correct answers, which is why the highest earnings were $16.00 (see Burks *et al.* (2008)).

[8] The Unlikely Virtues Scale developed by Patrick *et al.* (2002), actually consists of 14 items. Due to a programming error, one item was not included in the questionnaire administered to participants in the experiment.

[9] Although the subject pools were not intended to be representative of the corresponding population, we report for comparison a summary of socio-demographic characteristics of the population of Morris (residence of two of the three subject pools) for the period 2005-2009: Age (median): 30.3 years; Female: 54.9% ; Non-White or Hispanic:

[Insert Table 1 about here]

Although there is a fair amount of overlap in most of the socio-demographic dimensions, there are also important differences across the three subject pools. In terms of age, as expected, Self-Selected Students are on average younger than the two non-student groups, and although both adult groups exhibit a wide age range, Non-Self-Selected Trainee Truckers are younger on average than Self-Selected Non-Students. The three subject pools also differ in terms of gender composition: both non-student groups are predominantly composed of male subjects, while the majority of college students are female. In terms of years spent in education, Non-Self-Selected Trainee Truckers are less educated than both Self-Selected Students and Self-Selected Non-Students. Other notable differences across groups are in terms of their racial composition (with Self-Selected Non-Students being less likely to be classified as "Non–White or Hispanic"), and in terms of disposable income (Self-Selected Students have higher incomes than both other groups, and Self-Selected Non-Students have higher incomes than Non-Self-Selected Trainee Truckers). In the data analysis presented in the next section we will use regression analysis to account for these differences across subject pools.

3. Results

We are mainly interested in addressing the following questions: 1) Is the pervasiveness of other-regarding motives overstated by measurements based on samples of self-selected participants?; and 2) Are other-regarding preferences as widespread among student subjects as among non-student subjects? To examine these questions, we start by classifying subjects in different "preference types" based on the cooperativeness of their choices in the role of second-mover in the social dilemma game described above. We then compare how the distributions of these types vary across the three subject pools.[10] To address the first question, we compare the two samples of non-student adults, which differ in whether subjects self-selected into the study or not. To address the second question we compare the two samples of self-selected participants, which

8.3%; Years of Education Completed for the population aged 25 or above (mean): 12.9 years; Marital Status of the population aged 15 or above: 33% married (source: 2005-2009 American Community Survey 5-Year Estimates).

[10] In Appendix C we follow an alternative approach to address our research questions, and directly compare the amounts transferred by second-movers across subject pools instead of using these amounts to classify subjects in different 'preference types' (we thank an anonymous referee for suggesting this alternative approach). The results of this alternative approach are qualitatively equivalent to those reported in Section 3.

differ in being drawn from college student versus non-student adult populations.[11] The section concludes by examining the differences across samples in subjects' need for social approval as measured by the Unlikely Virtues Scale.

3.1. Other-regarding preferences across subject pools

We measure subjects' other-regarding preferences using decisions in the role of second-mover in the social dilemma game described in Section 2. The use of the strategy method allows us to observe two decisions from each participant in the role of second-mover: one for the case where the first-mover behaves uncooperatively and sends \$0, and one for the case where the first-mover is cooperative and sends \$5. This allows us to classify subjects into three well-defined types depending on how cooperatively they respond to the first-mover's actions: "Free Riders", "Conditional Cooperators" and "Unconditional Cooperators".[12] These categories tie in with those discussed in the social dilemma games literature, which distinguish between subjects who never cooperate regardless of what others do ("Free Riders" or "Defectors"), subjects who are willing to cooperate as long as others are also willing to do so ("Conditional Cooperators"), and subjects who are prepared to cooperate even if others defect ("Unconditional Cooperators" or "Altruists").[13]

In the context of our social dilemma game, we classify as Free-Riders those subjects who behave uncooperatively and choose the payoff-maximizing action (return \$0) irrespective of the amount sent by the first-mover. Subjects who choose the most cooperative action available (send back \$5) if the first-mover sends \$5, but behave uncooperatively and send back \$0 otherwise are classified as "Conditional Cooperators". Finally, subjects who always choose the most

[11] The experiment also delivers data on unconditional cooperation decisions by subjects in the role of first-mover. Compared to decisions as second-mover, it is more difficult to infer other-regarding motives from first-movers' choices since these may also reflect considerations about the profitability of cooperating, false-consensus effects, etc. (see, e.g., (Gächter et al. forthcoming 2012). For this reason, in the main text we focus on decisions in the role of second-mover, and only briefly discuss here first-mover's behavior. In the role of first-mover, 74% of Self-Selected Non-Students, chose to transfer \$5 to the second-mover. This is significantly more than the fraction of Self-Selected Students, choosing to do so (55%, $\chi^2(1) = 6.93, p = 0.008$). The share of Non-Self-Selected Trainee Truckers sending \$5 is 67%, which is not significantly different from that of Self-Selected Non-Students ($\chi^2(1) = 1.38, \mathrm{p} = 0.239$). Further analysis of the first mover behavior of Non-Self-Selected Trainee Truckers may be found in Burks *et al.*(2009)

[12] In order to have a well-defined classification of subjects' cooperativeness one needs to observe their behavior in both subgames. Observing second-movers' behavior in only one subgame may not be sufficient. For example, observing a second-mover who sends \$0 when the first-mover sends \$0 does not reveal whether she is a 'conditional cooperator' who defects when the first-mover defects, or whether she is instead motivated by material payoff maximization. The use of the strategy method solves this problem by allowing us to observe how a second-mover responds to both possible decisions of the first-mover.

[13] See, e.g., Camerer and Fehr (2006); Fehr and Gächter (2000); Fischbacher and Gächter (2010).

cooperative action available and send back \$5 irrespective of what the first-mover sends to them are classified as "Unconditional Cooperators.". While Free Riders do not display other-regarding concerns as they always choose the action that maximizes their own material payoff, note that both Conditional and Unconditional Cooperators exhibit other-regarding concerns, as they are willing to forgo material gain to increase the payoff of the other player.

This approach allows us to classify 61% of the Self-Selected Students, 53% of the Self-Selected Non-Students, and 62% of the Non-Self-Selected Trainee Truckers. To assign the remaining participants to a type category we calculate, for each subject, the Euclidean distance between his or her decisions and the decisions that each of the three types would make, and then assign the subject to the least distant type category.[14] We can thus classify all but 25 subjects (2 in Self-Selected Students, and 23 in Non-Self-Selected Trainee Truckers): these participants are classified separately as "Others."[15] Figure 1 shows the distribution of types across the different subject pools.

[Insert Figure 1 about here]

A first notable feature of Figure 1 is the similarity between the distributions of cooperation types across the two non-student adult subject pools. The shares of subjects that we classify as Free Riders, Conditional Cooperators and Unconditional Cooperators are remarkably similar across adult participants who self-selected into the experiment and those who did not self-select into the experiment. In fact, a comparison between these two groups reveals that the distribution of types does not differ significantly ($\chi^2(2) = 0.61, p = 0.737$). Thus, measurements based on samples of non-student volunteers who self-select into the experimental environments do not seem to overestimate the prevalence of other-regarding preferences. This is the case even when we correct for the 9.2% attrition rate in the Non-Self-Selected Trainee Truckers sample. Assuming the extreme case that the 109 trainees who did not participate are uniformly those who are *least* other-regarding and who would then be classified as Free Riders, we would still fail to

[14] Formally, if $x_{\$0}$ is the amount that a subject returns when the first-mover sends \$0 and $x_{\$5}$ is the amount returned when the first-mover sends \$5, we compute the distance of the subject's decisions from the Free Rider type as $D_{FR} = \sqrt{(x_{\$0} - 0)^2 + (x_{\$5} - 0)^2}$, from the Conditional Cooperator type as $D_{CC} = \sqrt{(x_{\$0} - 0)^2 + (x_{\$5} - 5)^2}$, and from the Unconditional Cooperator type as $D_{UC} = \sqrt{(x_{\$0} - 5)^2 + (x_{\$5} - 5)^2}$.

[15] All participants classified as Others cannot be classified because they are equally distant from a Free Rider and an Unconditional Cooperator. In the remainder of this sub-section we will focus on the three major cooperation types and ignore the 25 subjects classified as Others.

reject the null hypothesis that the distribution of other-regarding preferences types does not differ between the non-student groups ($\chi^2(2) = 4.585, p = 0.101$).

A second feature that emerges from Figure 1 is that the share of individuals exhibiting other-regarding concerns is clearly smaller among Self-Selected Students than among Self-Selected Non-Students. The share of college students who display some form of other-regarding behavior is 63%, while 79% of the non-students are classified either as a Conditional Cooperator or as an Unconditional Cooperator. In fact, we can reject the hypothesis that the two groups are sampled from the same population at the 1% level ($\chi^2(2) = 23.52, p < 0.001$).

As a robustness check for these results, we use regression analysis which allows us to control for observable differences across subject pools. We use a multinomial logit regression model where the dependent variable is a categorical variable describing whether a subject is classified as a Free Rider, a Conditional Cooperator or an Unconditional Cooperator. In Model I we only use dummy variables for the different subject pools as regressors (note that the reference category is the group of Self-Selected Non-Students). Model II expands Model I by adding the set of controls for socio-demographic characteristics listed in Table 1: age, gender, years of education, number of siblings, a dummy variable describing the subject's marital status, a dummy variable for racial characteristics, and a set of dummy variables for different income categories. To allow for potential nonlinearities we also include quadratic terms of the continuous explanatory variables (age, years of education, and number of siblings). The regression results are reported in Table 2.[16]

[Insert Table 2 about here]

Starting with the equations of Model I, we confirm that the distribution of types across non-student subject pools is not affected by whether participants did or did not self-select into the experiment. The odds of being classified as a Free Rider, a Conditional Cooperator, or an Unconditional Cooperator are not different between the group of Non-Self-Selected Trainee Truckers and the regression reference group, Self-Selected Non-Students, at any conventional

[16] The multinomial logit model relies on the assumption known as the 'independence of irrelevant alternatives' (IIA) whereby introducing or removing any category type from our classification should have the same proportional impact on the probability of the other categories. We tested the IIA assumption using the two tests presented by Long and Freese (2006), the Hausman test and the Small-Hsiao test. The results show no evidence that the IIA assumption has been violated (these tests results are available from the authors upon request).

significance level. These results hold also in Model II where we add controls for socio-demographic characteristics.

Model I also confirms that Self-Selected Students are significantly less likely to engage in one specific form of other-regarding behavior as compared to Self-Selected Non-Students. Equations Ib and Ic show that Self-Selected Students have substantially lower odds of being classified as Unconditional Cooperators. The odds of being an Unconditional Cooperator rather than a Free Rider are 92% lower for a college student than for a Self-Selected Non-Student. Similarly, being a college student decreases by about 89% the odds of being classified as an Unconditional Cooperator rather than a Conditional Cooperator. The effects are significant at the 1% level for both equations in Model I, and remain statistically significant in Model II after controlling for observable differences between subject pools (at the 1% level for equation IIb, at the 5% level for equation IIc). Interestingly, this difference across samples is unique to the Unconditional Cooperator behavioral category. When we consider the second type of other-regarding behavior that is possible in our experiment, we find that Self-Selected Students are not significantly less likely than Self-Selected Non-Students to be classified as a Conditional Cooperator rather than a Free Rider (see equations Ia and IIa).

Among the controls for socio-demographic characteristics included in Model II, age has a positive impact on the odds of being classified as an Unconditional Cooperator (see equations IIb and IIc). This finding is in line with results from other studies showing that older people tend to be more cooperative than younger people (e.g., Carpenter et al. 2005; List 2004). Interestingly, having more years of education appears to increase the odds of being classified as a Conditional Cooperator relative to any of the other two type categories (see equations IIa and IIc). In both equations the effects are significant at the 1% level.

3.2. Need for social approval across subject pools

Overall, our results on behavior in the social dilemma game suggest that adult volunteers who self-select into economic experiments and those who do not self-select do not differ significantly in their other-regarding inclinations. Research from social psychology, however, suggests that a dimension in which self-selected and non-self-selected volunteers might also differ is in their need for social approval (see, e.g., Rosenthal and Rosnow (1969)). Approval-seeking participants may be systematically more prone to cooperate with the experimenter (e.g. by behaving in accordance with the perceived experimental objectives, or with what is perceived

to constitute "appropriate" behavior), and this may also distort measurements collected through experiments.

To examine whether the need for social approval is higher among participants who self-select into the experiment than among non-self-selected participants Table 3 shows, disaggregated by subject pool, participants' scores in the Unlikely Virtues Scale, a stand-alone index of social desirability with higher scores indicating more socially desirable responding.[17]

[Insert Table 3 about here]

Table 3 reveals that self-selected participants did not respond in a more socially desirable manner to the Unlikely Virtues Scale questions than participants who did not self-select into the experiment. In fact, the group of Trainee Truckers, who did not self-select into the experiment, scored highest in the Unlikely Virtues Scale, although the difference is not large.[18] Table 3 also shows that Students' scores are lower than the scores of Self-Selected Non-Students. A two-sided Mann-Whitney-U-test reveals that the difference is statistically significant ($p < 0.001$).[19]

A Tobit regression controlling for subjects' socio-demographic characteristics produces the same results. The regression estimates show that, *ceteris paribus*, Non-Self-Selected Trainee Truckers score about 1 point higher than Self-Selected Non-Students in the Unlikely Virtues Scale, and the difference is statistically significant (p = 0.047). On the other hand, Students on average score about 2.94 points lower than Self-Selected Non-Students, and the difference is highly significant (p = 0.000).[20, 21]

[17] For 29 Non-Self-Selected Trainee Truckers there are missing data for some of the items composing the Unlikely Virtues Scale. To compute a score for these subjects we impute the neutral midpoint of the scale for those items whose answers are missing. Results do not change if we conduct the analysis excluding these 29 subjects.

[18] A two-sided Mann-Whitney-U-test shows that the difference is not statistically significant ($p = 0.102$).

[19] Although it is standard to simply sum the responses to summarize the Unlikely Virtues Scale, we also conducted a factor analysis. The analysis resulted in one eigenvalue above one and using the resulting factor scores we find similar results: two-sided Mann-Whitney-U-tests reveal that socially desirable responding is somewhat more prevalent among Non-Self-Selected Trainee Truckers than Self-Selected Non-Students (p = 0.094), and more prevalent among Self-Selected Non-Students than Self-Selected Students (p = 0.000) .

[20] The regression also shows that approval-seeking is positively correlated with age (p = 0.002) and with the dummy variable for Non-White or Hispanic subjects (p = 0.000). Full regressions results are available upon request.

[21] Another interesting question is whether there is a link between the need for social approval and other-regarding preferences. To address this question we re-ran the multinomial logit regression reported in Table 3 (Model II) adding the Unlikely Virtues Scale scores to the list of explanatory variables. We find that, if anything, social desirability slightly increases (by about 3%) the odds of being classified as a Free Rider rather than a Conditional Cooperator, and the effect is significant at the 10% level. None of the other comparisons is statistically significant.

4. Discussion of Main Findings in Relation to the Literature

Questions have been raised in the experimental economics literature about whether the incidence of other-regarding behavior observed using college student samples might be biased upwards, either because those with other-regarding tendencies differentially self-select into participation as experimental subjects, or because college students might be more generally other-regarding than adult non-student subjects (see Section 1). The two main results of our study are as follows. First, with regards to the self-selection issue, we do not find any significant difference in the distribution of other-regarding preferences across non-student subjects who self-selected and who did not self-select into the experiment. This suggests that self-selection effects among such subjects are not likely to have a significant impact on the measurement of other-regarding preferences. Second, when comparing (self-selected) student and non-student subjects, we find that the share of subjects who are motivated by other-regarding considerations is remarkably larger among non-students. In this section we discuss these findings in relation to the existing experimental literature.

Only a few studies have investigated empirically the possibility that more pro-socially inclined individuals self-select into economic experiments.[22] Eckel and Grossman (2000) use a dictator game to compare the behavior of "volunteers" (students recruited through prior announcements in graduate and undergraduate classes) and "pseudo-volunteers" (students recruited from a class to immediately participate in the experiment during class time in order to minimize potential self-selection issues). They find that volunteers are significantly less generous than pseudo-volunteers and that they behave in a less extreme manner, suggesting that they are more motivated by monetary incentives. However, as also noted by Eckel and Grossman (2000), these differences may reflect the fact that pseudo-volunteers were exposed to a more authoritative environment, which could have amplified potential experimenter demand effects (see also Zizzo (2010)). These concerns are minimized in our setting, since there was no close relation between the non-self-selected participants and the experimenters conducting the study, and the informed-consent process emphasized the strict confidentiality of individual data including specifically that it would never be available to their managers at the trucking firm. In a recent study, Cleave *et al.*(2011) recruited 1,173 students from an introductory microeconomics

[22] Another strand of the literature has examined the implications of selection for the elicitation of risk preferences, see, e.g., Harrison et al., (2009); von Gaudecker et al., (2011).

class into experiments carried out during class time in order to minimize self-selection, and measured their other-regarding preferences using a trust game. They later invited the 1,173 students to take part in a subsequent laboratory experiment, and examine whether those who decided to eventually participate in the experiment had different preferences than the original population. They do not find a selection bias based on other-regarding preferences. Finally, Falk *et al.*(forthcoming 2012) have examined whether students' decisions to take part in economic experiments organized by the experimental economics laboratory at the University of Zurich are related to their pro-social inclinations as measured by their contributions to two charitable funds providing financial support to foreign and needy students to which all students must decide about donating. They find that students who take part in experiments are not more pro-social than non-participant students.

The results from our study complement the evidence gathered in these studies with students, and extend it in three ways. Most importantly, we examine self-selection not among undergraduate students but among non-student adults. This holds particular relevance in light of the growing number of experimental studies relying on self-selected subject samples drawn from non-student adult populations (see, e.g. Carpenter and Seki 2011; Rustagi et al. 2010; Voors et al. 2011). Moreover, the fact that other-regarding behavior is more frequently observed in non-student than student samples (a finding which we corroborate in our paper), hints that any correlation between other-regarding preferences and the process by which participants self-select into experiments should be more clearly visible in non-student samples. In other words, by focusing on a sample in which pro-social behavior may be relatively more abundant, we provide a strong test of the conjecture that the proportion of subjects exhibiting pro-social behavior in experiments is driven by the endogenous process through which participants self-select into the studies. A second novel contribution of our study relative to the existing literature is that we also examine selection effects driven by a desire for social approval, a possibility that was suggested, for example, by Levitt and List (2007), but that has not been addressed by the complementary papers. Finally, in our study we simultaneously address both self-selection and the comparison of students and non-student adults using exactly the same experimental protocol.

Turning to the comparison between students and non-students, several other studies have compared undergraduate student samples with adult samples across a variety of games where pro-social inclinations may matter. A general result from these studies is that there seem to be

more pro-sociality among non-students than among college students. For example, three studies (Bellemare and Kroger 2007; Falk et al. forthcoming 2012; Fehr and List 2004) compare student and non-student samples in trust game experiments. Fehr and List (2004) use one-shot trust games both with and without a punishment option whereby first-movers can impose a fine on second-movers if they return less than the first-mover's desired payback. They conduct experiments in Costa Rica with undergraduate students and CEOs from the coffee mill sector. In both versions of the game, they find that CEOs transfer more money than students in the role of first-mover and pay back more money in the role of second-mover. Bellemare and Kröger (2007) use a standard one-shot trust game to compare a sample of undergraduate students with a representative sample of individuals drawn from the general population in the Netherlands. They find that the representative sample is more trusting and more trustworthy than the student sample. Falk et al. (forthcoming 2012) also use one-shot trust games in experiments conducted with students and members of the general population in Switzerland. Non-students are found to be more trustworthy than students, whereas trusting behavior does not differ across samples. Interestingly, in both Bellemare and Kröger (2007) and Falk et al. (forthcoming 2012) differences in trust and trustworthiness between samples are statistically insignificant once standard socio-demographic background characteristics (age, gender, education, etc.) are taken into account, suggesting that sample variations in observable characteristics explain most of the behavioral differences observed in the experiments. This differs from what we find in our social dilemma game, where differences in cooperativeness between students and non-students persist even when we control for a similar set of socio-demographic characteristics.

In the context of bargaining games, Carpenter *et al*. (2005) study student and non-student (warehouse worker) samples in the US using ultimatum game and dictator games. They observe the usual result that among students there is a large fraction of high offers in the ultimatum game (UG) which drops to a low fraction in the dictator game (DG), but find that almost all non-students make high offers in both games. This difference in DG giving is analogous to the present results about the difference between students and non-students in the proportion of Unconditional Cooperators. Güth et al. (2007) conduct a three-person UG with readers of a weekly news magazine in Germany, including 626 readers who reported themselves to be 'students' at the time of the experiment. Relative to the non-student fraction of the magazine

readers, students are less likely to offer the equal split and more likely to accept unequal offers.[23] Carpenter et al. (2008) compare students and non-students in a DG experiment with a charity in the role of the recipient. They find that non-students donate significantly more ($17) than students on average and are about 30% more likely to donate the entire $100 endowment. Hoffman and Morgan (2011) conduct a battery of 'social preference experiments' on business people and students, including a dictator game and a version of the trust game. They find that business people are more generous, more trusting and more trustworthy than students.

Turning to social dilemma game experiments, four studies (Belot et al. 2010; Carpenter and Seki 2011; Gächter et al. 2004; Stoop et al. 2009) have compared students and non-students samples in laboratory public good game (PGG) experiments. All studies find that non-students are significantly more cooperative than students. Gächter et al. (2004) collected data from 639 subjects (339 students and 300 non-students) in several cities and villages in Russia and Belarus, and compare their behavior in a one-shot PGG experiment. They find that non-student subjects contribute significantly more (on average 1.5 tokens out of a 20 tokens endowment) than students.[24] Belot et al. (2010) examine the behavior of student and non-student subjects in a 10-round 20-tokens-endowment repeated PGG in a laboratory experiment in the UK. They also find that non-students are more cooperative than students (average non-students' contributions start 1.2 tokens higher in the first-round and are still 0.6 tokens higher in the last-round).[25] Both Carpenter and Seki (2011) and Stoop et al. (2009) compare the behavior of students and fishermen in repeated PGG experiments, respectively in Japan and the Netherlands. Carpenter and Seki (2011) use a 10-round repeated PGG, with a 'social disapproval' stage introduced in the last 5 rounds whereby participants could send costly messages to their group signaling their dissatisfaction with the pattern of contributions. They find that professional fishermen contribute significantly more than students, both in the presence and absence of the social disapproval mechanism. Stoop et al. (2009) conduct a 6-round PGG laboratory experiment using students and recreational fishermen as subjects. They also find that fishermen are more cooperative than students, especially in the later periods of the game. Overall, the evidence from these four studies

[23] Interestingly, offers and acceptance rates of students and non-students do not differ if the non-student sample is restricted to a sub-group of participants in the same age groups as students.
[24] However, differences in contributions vanish once participants' socio-economic characteristics are accounted for.
[25] Belot et al. (2010) also compare the choices of students and non-students in other games where other-regarding preferences may be relevant (a dictator game and a trust game). They find that non-students are more other-regarding than students in these games as well.

across four different societies corroborates our finding that students are less cooperative than non-students.[26]

Most closely related to our study, Burks et al. (2009a) use a simplified version of the sequential prisoner's dilemma game used in the present paper to compare undergraduate students in Zürich to bicycle messengers in Zürich and San Francisco, and find that the latter are significantly more cooperative than the former. In particular, they also classify subjects according to their (conditional) cooperativeness and find that there are far fewer Unconditional Cooperators and far more Free Riders among students than among messengers.[27]

5. Conclusion

Taken together, the findings from our study and the related literature suggest that measurements of other-regarding preferences based on self-selected samples, whether of adults, as in the present study, or of college students, as in the existing literature, are not systematically biased upwards. The pro-social inclinations of subjects who self-selected into the experiments do not appear to be significantly different from the inclinations of those who had instead very little opportunity to self-select into the study. Because it is relatively difficult in practice to arrange a non-student adult sample which avoids the potential for self-selection bias (a problem avoided in our case because we were able to design the study so that nearly all of the potential subjects took part), providing some systematic evidence about the effects of self-selection on the laboratory measurement of other-regarding preferences for a subject pool of this type is the signal contribution of our paper. Further, a common picture emerging from a significant collection of experimental studies, including ours, is that the use of college student samples may lead to *underestimating* the pervasiveness of other-regarding preferences. There now appears to be substantial accumulated evidence suggesting that measurements obtained from undergraduate college students represent a lower bound on the extent to which individuals of advanced industrial societies exhibit other-regarding behaviors in behavioral economic laboratory experiments.[28]

[26]Also related is Cardenas (2005) who conducts common pool resources game experiments with students and villagers in Colombia, and finds that villagers are more cooperative than students.
[27] Burks et al. (2009a) label Unconditional Cooperators "Altruists" and Free Riders "Egoists."
[28] Where the populations of advanced industrial societies fall in the full range of behavior typical of humans as a species is an open question that our data do not address; see, for example, the discussion in Henrich *et al.*(2010).

Acknowledgments: We thank the editor and two anonymous referees for useful comments. We received helpful comments from Simon Gächter, John Galbraith, Herbert Gintis, Nikos Nikiforakis and participants at the 2011 International Meeting of the Economic Science Association in Chicago (IL). The Truckers and Turnover Project acknowledges financial and in-kind support from the cooperating firm, and financial support from the MacArthur Foundation's Research Network on the Nature and Origin of Norms and Preferences, the Sloan Foundation's Industry Studies Program, the Trucking Industry Program at Georgia Institute of Technology, the University of Nottingham, and the University of Minnesota, Morris. Götte acknowledges support from the Federal Reserve Bank of Boston, and Nosenzo from the Leverhulme Trust (ECF/2010/0636). The views expressed are those of the authors, and do not necessarily reflect those of the supporting entities.

REFERENCES

Bellemare, C., & Kroger, S. (2007). On Representative Social Capital. *European Economic Review, 51*(1), 183-202, doi:http://www.elsevier.com/wps/find/journaldescription.cws_home/505541/description#description.

Belot, M., Duch, R., & Miller, L. (2010). Who should be called to the lab? A comprehensive comparison of students and non-students in classic experimental games. *University of Oxford, Nuffield College Discussion Papers*(2010-001).

Burks, S., Carpenter, J., & Goette, L. (2009). Performance Pay and Worker Cooperation: Evidence from an Artefactual Field Experiment. *Journal of Economic Behavior and Organization, 70*(3), 458-469, doi:http://www.elsevier.com/wps/find/journaldescription.cws_home/505559/description#description.

Burks, S., Carpenter, J., Götte, L., Monaco, K., Porter, K., & Rustichini, A. (2008). Using Behavioral Economic Field Experiments at a Firm: the Context and Design of the Truckers and Turnover Project. In S. Bender, J. Lane, K. Shaw, F. Andersson, & T. von Wachter (Eds.), *The Analysis of Firms and Employees: Quantitative and Qualitative Approaches* (pp. 45-106). Chicago and London: NBER and University of Chicago

Burks, S., Carpenter, J., Götte, L., & Rustichini, A. (2009). Cognitive Skills Affect Economic Preferences, Social Awareness, and Job Attachment. *Proceedings of the National Academy of Science (USA), 106*(19), 7745-7750.

Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction* (Roundtable Series in Behaviorial Economics). Princeton, NJ: Princeton University Press.

Camerer, C. F., & Fehr, E. (2006). When does "economic man" dominate social behavior? [Review]. *Science, 311*(5757), 47-52, doi:10.1126/science.1110600.

Cardenas, J. C. (2005). Groups, Commons and Regulations: Experiments with Villagers and Students in Colombia. In B. Agarwal, & A. Vercelli (Eds.), *Psychology, Rationality and Economic Behaviour: Challenging Standard Assumptions*: Palgrave Macmillan.

Carpenter, J., Burks, S., & Verhoogen, E. (2005). Comparing Students to Workers: the Effects of Social Framing on Behavior in Distribution Games. In *Research in Experimental Economics* (Vol. 10, pp. 261-290, Vol. Field Experiments in Economics).

Carpenter, J., Connolly, C., & Myers, C. (2008). Altruistic Behavior in a Representative Dictator Experiment. *Experimental Economics, 11*(3), 282-298, doi:http://www.springerlink.com/link.asp?id=102888.

Carpenter, J., & Seki, E. (2011). Do Social Preferences Increase Productivity? Field Experimental Evidence from Fishermen in Toyama Bay. *Economic Inquiry, 49*(2), 612-630, doi:http://www.blackwell-synergy.com/loi/ecin/.

Cleave, B., Nikiforakis, N., & Slonim, R. (2011). Is There Selection Bias in Laboratory Experiments? *Institute for the Study of Labor Discussion Papers*(#5488), 1-34.

Eckel, C., & Grossman, P. (2000). Volunteers and Pseudo-volunteers: The Effect of Recruitment Method in Dictator Experiments. *Experimental Economics, 3*(2), 107-120, doi:http://www.springerlink.com/link.asp?id=102888.

Falk, A., Meier, S., & Zehnder, C. (forthcoming 2012). Do Lab Experiments Misrepresent Social Preferences? *Journal of the European Economic Association*.

Fehr, E. and S. Gächter (2000). "Cooperation and Punishment in Public Goods Experiments." *American Economic Review 90*(4): 980-994.

Fehr, E., & List, J. (2004). The Hidden Costs and Returns of Incentives--Trust and Trustworthiness among CEOs. *Journal of the European Economic Association, 2*(5), 743-771, doi:http://www.mitpressjournals.org/loi/jeea.

Fehr, E., & Schmidt, K. (2006). The Economics of Fairness, Reciprocity and Altruism - Experimental Evidence and New Theories. In S. Kolm, & J. M. Ythier (Eds.), *Handbook of the Economics of Giving, Reciprocity and Altruism* (Vol. Volume 1, pp. 615-691): Elsevier.

Fischbacher, U. (2007). Z-Tree: Zurich Toolbox for Ready-Made Economic Experiments. *Experimental Economics, 10*(2), 171-178, doi:http://www.springerlink.com/link.asp?id=102888.

Fischbacher, U., & Gächter, S. (2010). Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments. *American Economic Review, 100*(1), 541-556, doi:http://www.aeaweb.org/aer/.

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are People Conditionally Cooperative? Evidence from a Public Goods Experiment. *Economics Letters, 71*(3), 397-404, doi:http://www.elsevier.com/wps/find/journaldescription.cws_home/505574/description#description.

Gächter, S., Herrmann, B., & Thoni, C. (2004). Trust, Voluntary Cooperation, and Socio-economic Background: Survey and Experimental Evidence. *Journal of Economic Behavior and Organization, 55*(4), 505-531, doi:http://www.elsevier.com/wps/find/journaldescription.cws_home/505559/description#description.

Gächter, S., Nosenzo, D., Renner, E., & Sefton, M. (forthcoming 2012). Who Makes A Good Leader? Cooperativeness, Optimism And Leading-By-Example. *Economic Inquiry,* DOI: 10.1111/j.1465-7295.2010.00295.x.

Güth, W., Schmidt, C., & Sutter, M. (2007). Bargaining Outside the Lab--A Newspaper Experiment of a Three-Person Ultimatum Game. *Economic Journal, 117*(518), 449-469, doi:http://www.blackwellpublishing.com/journal.asp?ref=0013-0133.

Harrison, G., Lau, M., & Rutstrom, E. (2009). Risk Attitudes, Randomization to Treatment, and Self-Selection into Experiments. *Journal of Economic Behavior and Organization, 70*(3), 498-507, doi:http://www.elsevier.com/wps/find/journaldescription.cws_home/505559/description#description.

Henrich, J., Heine, S., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2/3), 1-75, doi:doi:10.1017/S0140525X0999152X.

Herrmann, B., & Thoni, C. (2009). Measuring Conditional Cooperation: A Replication Study in Russia. *Experimental Economics, 12*(1), 87-92, doi:http://www.springerlink.com/link.asp?id=102888.

Hoffman, M., & Morgan, J. (2011). Who's Naughty? Who's Nice? Social Preferences in Online Industries. *University of California, Berkeley*(unpublished manuscript).

Kocher, M., Cherry, T., Kroll, S., Netzer, R., & Sutter, M. (2008). Conditional Cooperation on Three Continents. *Economics Letters, 101*(3), 175-178, doi:http://www.elsevier.com/wps/find/journaldescription.cws_home/505574/description#description.

Levitt, S., & List, J. (2007). What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World? *Journal of Economic Perspectives, 21*(2), 153-174, doi:http://www.aeaweb.org/jep/.

List, J. (2004). Young, Selfish and Male: Field Evidence of Social Preferences. *The Economic Journal, 114*(January), 121-149.

Long, S. J., & Freese, J. (2006). *Regression Models for Categorical Dependent Variables Using Stata* (Second Edition ed.). College Station, Texas: Stata Press.

Patrick, C., Curtin, J., & Tellegen, A. (2002). Development and Validation of a Brief Form of the Multidimensional Personality Questionnaire. *Psychological Assessment, 14*(2), 150-164.

Rosenthal, R., & Rosnow, R. (1969). The Volunteer Subject. In R. Rosenthal, & R. L. Rosnow (Eds.), *Artifact in Behavioral Research* (pp. 61-118). New York and London: Academic Press.

Rustagi, D., Engel, S., & Kosfeld, M. (2010). Conditional Cooperation and Costly Monitoring Explain Success in Forest Commons Management. [Article]. *Science, 330*(6006), 961-965, doi:10.1126/science.1193649.

Stoop, J., Noussair, C. N., & van Soest, D. (2009). From the lab to the field: Cooperation among fishermen. *Munich Personal RePEc Archive*(#28924).

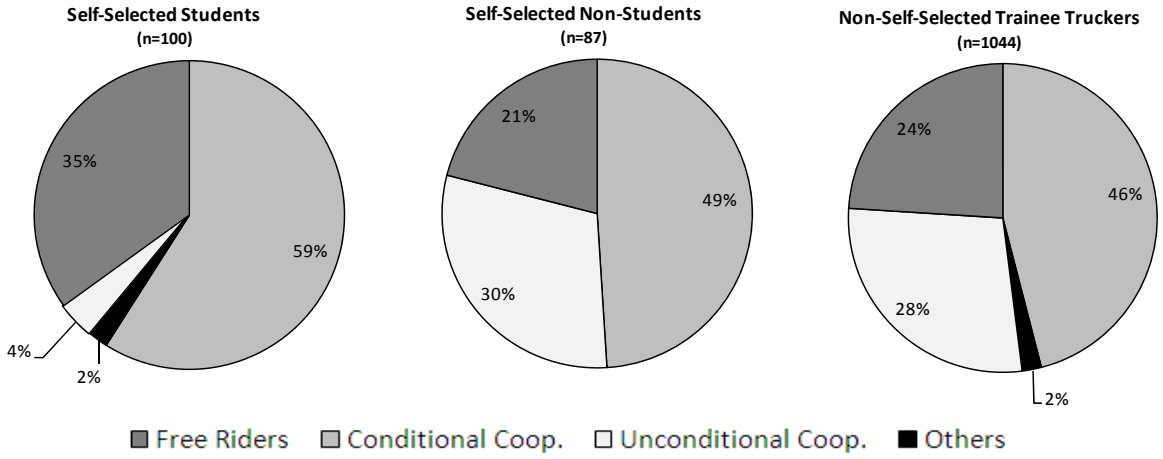
von Gaudecker, H.-M., van Soest, A., & Wengström, E. (2011). Experts in Experiments: How Selection Matters for Estimated Distributions of Risk Preferences. *Institute for the Study of Labor Discussion Papers*(#5575).

Voors, M., Bulte, E., Kontoleon, A., List, J., & Turley, T. (2011). Using Artefactual Field Experiments to Learn about the Incentives for Sustainable Forest Use in Developing Economies. *American Economic Review, 101*(3), 329-333, doi:http://www.aeaweb.org/aer/.

Zizzo, D. (2010). Experimenter demand effects in economic experiments. *Experimental Economics, 13*(1), 75-98.

[Figures]

Figure 1: Distribution of Other-Regarding Preferences across Subject Pools



[Tables]

Table 1: Socio-Demographic Characteristics, Disaggregated by Subject Pool

	Self-Selected Students (n=100)	**Self-Selected Non-Students** (n=87)	**Non-Self-Selected Trainee Truckers** (n=1044)
Age, *median (min. – max.)*	20.5 (18 - 41)	42.8 (21 - 66)	36.2 (21 - 69)
Female *(%)*	0.61	0.41	0.10
Non – White or Hispanic *(%)*	0.23	0.05	0.19
Number of Siblings, *mean (s.d.)*	2.0 (1.2)	2.9 (1.6)	2.9 (2.0)
Years of Education Completed, *mean (s.d.)*	14.2 (1.1)	14.3 (2.0)	13.0 (1.7)
Marital Status *(%)*			
Married or in marriage-type relationship	0.03	0.66	0.48
Single/Separated/Divorced/Widowed	0.97	0.34	0.52
Income Category *(%)*			
$0-$10,000	0.07	0.01	0.39
$10,000-$20,000	0.00	0.05	0.16
$20,000-$30,000	0.08	0.11	0.15
$30,000-$40,000	0.00	0.21	0.11
$40,000-$50,000	0.18	0.13	0.07
$50,000-$60,000	0.00	0.11	0.05
$60,000-$70,000	0.21	0.07	0.03
$70,000+	0.46	0.31	0.04

Numbers of subjects with complete questionnaire data. These restricted samples will be used for the data analysis in Sect. 3. The variable "Years of Education Completed" was derived by asking subjects to indicate the highest level of education they had completed at the time of the experiment. For Students we refine this measure by distinguishing between freshmen, sophomores, juniors and seniors using the self-reported number of accumulated college credits at the time of the experiment. For Students the variable "Income Category" was derived from the question "*Which range best fits the annual income of your parents (step-parents)?*". For Self-Selected Non-students "Income Category" was constructed by combining their answers to the questions: "*Not counting your earnings, which range best fits the annual income you and your household have from other sources?*" and "*Which range best describes the annual earnings you would normally expect from your usual jobs?*". For Non-Self-Selected Trainee Truckers, which were undertaking full-time training and were thus unemployed at the time of the experiment, only answers to the first question was used. Other variables are self-explanatory. The variable Number of Siblings was coded as missing for the 27 subjects who reported having more than 10 siblings.

Table 2: Multinomial Logit Regressions

	Cond. Coop. vs. Free Rider		*Uncond. Coop. vs. Free Rider*		*Uncond. Coop. vs. Cond. Coop.*	
	Ia	*IIa*	*Ib*	*IIb*	*Ic*	*IIc*
Self-Selected Students	-29.4	-32.7	-92.1***	-84.7***	-88.8***	-77.3**
	(.323)	(.362)	(.000)	(.006)	(.000)	(.019)
Non-Self-Selected Trainee Truckers	-20.0	15.0	-17.7	5.5	2.8	-8.2
	(.445)	(.677)	(.541)	(.884)	(.914)	(.783)
Age	-	-5.7	-	39.8***	-	48.2***
		(.532)		(.002)		(.000)
Age^2 / 100	-	3.9	-	-1.5	-	-5.2
		(.686)		(.880)		(.498)
Gender (1 if Female)	-	-16.0	-	-21.7	-	-6.8
		(.404)		(.334)		(.759)
Non – White or Hispanic	-	-14.5	-	-15.9	-	-1.7
		(.392)		(.424)		(.931)
Number of Siblings	-	-7.0	-	7.3	-	15.3
		(.423)		(.503)		(.125)
Number of Siblings^2 / 100	-	-4.9	-	-5.1	-	0.3
		(.572)		(.604)		(.977)
Years of Education Completed	-	36.2***	-	3.0	-	-24.3***
		(.000)		(.757)		(.000)
Years of Education Completed^2 / 100	-	8.8	-	3.5	-	-4.9
		(.302)		(.691)		(.486)
Marital Status (1 if Single/etc.)	-	-20.1	-	-4.1	-	20.0
		(.171)		(.814)		(.242)
Income Category						
$10,000-$20,000	-	-0.1	-	32.6	-	32.8
		(.996)		(.286)		(.207)
$20,000-$30,000	-	-34.4*	-	-13.7	-	31.4
		(.062)		(.556)		(.240)
$30,000-$40,000	-	-18.9	-	30.5	-	61.0*
		(.429)		(.347)		(.059)
$40,000-$50,000	-	85.0*	-	117.4**	-	17.5
		(.052)		(.026)		(.556)
$50,000-$60,000	-	-4.3	-	31.5	-	37.5
		(.902)		(.473)		(.312)
$60,000-$70,000	-	22.3	-	157.8**	-	110.7**
		(.610)		(.028)		(.035)
$70,000+	-	22.2	-	6.5	-	-12.9
		(.502)		(.876)		(.698)
N.	*1206*	*1206*				
Wald χ2	*20.36*	*93.50*	*(same as columns Ia and IIa)*		*(same as columns Ia and IIa)*	
Prob > χ2	*0.000*	*0.000*				
Pseudo R^2	*0.016*	*0.046*				

Multinomial logit regression with robust standard errors. Dependent variable is subject's preference type: whether a subject is classified as type *m* (listed first in the column heading) rather than type *n* (listed second in the column heading). Results are reported as percentage changes in the odds ratios, which multiply the odds ratio of the reference subject type, which is: Self-Selected Non-Student, Male, Married, Adult, White (Non-Hispanic), Income category $0-$10,000. Dummy variables are treated in the standard manner. For continuous variables (Age, Years of Education Completed, Number of Siblings and their quadratic terms) the Table reports changes after a standard deviation increase in the explanatory variable (standard deviations are: 11.5 for Age, 1.42 for Age^2/100, 1.69 for Years of Education Completed, 0.06 for Years of Education Completed^2/100, 1.96 for Number of Siblings, 0.07 for Number of Siblings^2/100). Continuous variables are centered at their mean (means are: 36.2 for Age, 13.2 for Years of Education Completed, and 2.8 for Number of Siblings), and quadratic terms are computed for the mean-centered variables. P-values are reported in parentheses. A constant is included in all models, but omitted from the Table output. Significance levels: * 10% ; ** 5%; *** 1%.

Table 3: Unlikely Virtues Scale Scores, Disaggregated by Subject Pool

	Self-Selected Students (n=100)	**Self-Selected Non-Students** (n=87)	**Non-Self-Selected Trainee Truckers** (n=1044)
Mean	29.8	33.6	34.3
Standard Deviation	4.05	3.35	4.32
Min. – Max.	18 - 43	25 - 42	21 - 52

APPENDIX A: INSTRUCTION SCRIPT FOR THE SOCIAL DILEMMA EXPERIMENT

What follows is the text of the script spoken by Burks to each group of subjects. Subjects also saw an abbreviated version of these instructions on their computer screens, while they were listening to the instructions. The table mentioned in this text is in Appendix B.

ACTIVITY ONE: TWO-PERSON SENDING DECISION.

This activity does not take very long to do, but it is the most complicated thing to explain that we will do all day. So please bear with me as I give you the details.

You are going to make the decisions in this task ONCE. Some things you do today are going to be repeated, but the decisions in this task are not among the things we will repeat.

The BASICS OF THIS ACTIVITY ARE VERY SIMPLE, and so let me start there.

In this activity there are two different roles, Person 1 and Person 2. When we figure out your payoff you will be either a Person 1, or a Person 2, but not both. Each Person 1 will be matched with a Person 2 here in this room, but neither of you will ever know which specific other driver trainee you have been matched with.

Whether you are a Person 1 or a Person 2, the basics are the same. You will have a new amount of five dollars put in your account at the beginning of this activity by us. You have to decide whether to keep this five dollars, or to send it to the other person you are paired with. If you KEEP the money, it is yours at the end of the activity. If you SEND the money, we will double it, so that the person you send it to gets twice what you sent. Likewise, the person you are matched with will be making a SIMILAR decision about their five dollars. If they keep it, then it is theirs at the end of the activity, but if they send it to you, you will get twice what they sent.

So, that is THE BASIC OUTLINE. NEXT WE WILL LOOK AT THE DETAILS.

If you are a Person 1 your decision is simple. You have to decide whether to send your five dollars to Person 2, or to keep it. If you keep it, it is yours at the end of the activity, but if you send it to Person 2, we will *double* it, so that Person 2 actually gets $10.

Now, Person 2 also gets to decide about sending money to Person 1. But there are a couple of special features to Person 2's decision.

The first special feature is that Person 2 doesn't just have a yes-no choice about sending the five dollars. Instead, Person 2 can send any exact dollar amount to Person 1. So Person 2 can send: $0, $1, $2, $3, $4, or $5. Just like before, however, whatever Person 2 keeps is theirs at the end of the activity, and whatever Person 2 sends will be *doubled* by the researchers.

Example: if Person 2 keeps $4 and sends $1 to Person 1, Person 1 will actually receive $2,
Example: if Person 2 keeps $2 and sends $3 to Person 1, Person 1 will actually receive $6.

The second special feature is that Person 2 gets to decide what to do under two different cases. The first case is how much they want to respond if Person 1 has not sent them money. This choice will be on the LEFT side of the Person 2 choice screen. The second case is how much they want to respond if Person 1 has sent them money. This choice will be on the RIGHT side of

the Person 2 choice screen. In both cases the rules are the same: Person 2 can choose how many dollars to keep and how many to send, and whatever is sent is doubled.

TABLE OF PAYOFFS IS HANDED OUT. (table is provided following end of script text)

PLEASE DO NOT WRITE ON THIS SHEET, AS WE WILL RE-USE IT.

Look at payoff handout sheet.

Let's look at the top table. As you can see from looking at the first column (from Line 1 to Line 6), the top table is for the case in which Person 1 decides to send Person 2 $0. If you look at the second column of the top table (from Line 1 to Line 6), you can see all of the possible choices Person 2 has about how to respond. Finally, in each line, if you follow the arrow to the right, you see two more columns that show what the two people, Person 1 and Person 2, make in earnings. So, Lines 1 through 6 show each response Person 2 can make to the decision of Person 1 to send $0, and the payoffs each of them receive.

Example: look at Line 1. From the first two columns, this is the situation in which Person 1 sends $0, and Person 2 responds by also sending back $0. Following the arrow to the right, you can see that Person 1 makes $5, because he keeps $5, and gets $0 from Person 2, and Person 2 also makes $5, because he keeps all of his initial $5, and also receives nothing from Person 1.

Example: look at Line 4. From the first two columns, this is the situation in which Person 1 sends $0, and Person 2 responds by sending back $3. Following the arrow to the right, you can see that Person 1 makes $11, because he keeps $5, and also gets $3 doubled to $6 from Person 2. But Person 2 makes $2, because he sent $3 of his initial $5, but received nothing back.

Example: look at Line 6. From the first two columns, this is the situation in which Person 1 sends $0, and Person 2 responds by sending back $5. Following the arrow to the right, you can see that Person 1 makes $15, because he keeps $5, and also gets $5 doubled to $10 from Person 2. But Person 2 makes $0, because he sent all of his initial $5, but received nothing back.

Now let's look at the bottom table. This repeats the same pattern as the top table, except that, as you can see from looking at the first column (from Line 7 to Line 12), it is for the case in which Person 1 sends $5 to Person 2. If you look at the second column of the bottom table (from Line 7 to Line 12), you can see all of the possible choices Person 2 has about how to respond. Finally, in each line, if you follow the arrow to the right, you see two more columns that show what the two people, Person 1 and Person 2, make in earnings. So, Lines 7 through 12 show each response Person 2 can make to the decision of Person 1 to send $5, and the payoffs each of them receive.

Example: look at Line 7. From the first two columns, this is the situation in which Person 1 sends $5, and Person 2 responds by sending back $0. Following the arrow to the right, you can see that Person 1 makes $0, because he sent all his initial $5, and gets $0 back from Person 2, but Person 2 makes $15, because he keeps all of his initial $5, and also gets $5 doubled to $10 from Person 1.

Example: look at Line 9. From the first two columns, this is the situation in which Person 1 sends $5, and Person 2 responds by sending back $2. Following the arrow to the right, you can see that Person 1 makes $4, because he sent all of his initial $5, and gets $2 doubled to $4 back

from Person 2. And Person 2 makes $13, because he kept $3 of his initial $5, also gets $5 doubled to $10 from Person 1

Example: look at Line 12. From the first two columns, this is the situation in which Person 1 sends $5, and Person 2 responds by sending back $5. Following the arrow to the right, you can see that Person 1 makes $10, because he sent all his $5, and gets back $5 doubled to $10 from Person 2. And Person 2 also makes $10, because he sent back all of his initial $5, also got $5 doubled to $10 from Person 1.

Any questions now?

OK, now for the next to last special feature. This is very important. We are going to randomly assign the roles of Person 1 and Person 2 at the end of the activity, not at the beginning. So, we are going to ask everyone to make a decision first as a Person 1, and then second, a decision as a Person 2.

Let me repeat that: you will first make a decision IN CASE YOU ARE A Person 1, whether you will send $5 or not. Then, on a new screen, you will also make a two-decision IN CASE YOU ARE A Person 2: how much to send back if you got $0 (LEFT SIDE of the screen), and how much to send back if you got $5 (doubled to $10) (RIGHT SIDE of the screen).

So the way the payoffs will work is that first you will be matched by the computer with someone else here in the room. For example, #11 over here might be matched with #23 over there. Of course, let me remind you that we will never tell you with whom you were matched. Since everyone made both a Person 1 choice and then a Person 2 choice, once you have been matched, the computer will in effect flip a coin—it will randomly make one of you Person 1 and the other Person 2. Then it will look at your choices and those of the person you were matched with, and calculate your payoffs.

Finally, here is the last special feature. Before each decision screen there is another question. We are not only going to ask you what you want to do, I'd like to know what you think everyone else here today will do. So, we will also ask you to guess how others will handle the decisions you are about to make.

So for instance, right before the Person 1 screen asks you whether you will send $5 to Person 2 or not, we will also ask you to guess what percent of the people here in the room will send $5 as Person 1. We will pay you $1 extra if your guess is close (+/- 5%) to what people actually do.

And, when you make your decision as Person 2, you will have to tell us how much you want to send to Person 1 both when Person 1 sent $0 to you, and also when Person 1 sent $5 to you. Right before you make this decision, we will also ask you the average amount in dollars you think people in this room will send in each of these cases. We will pay you $1 extra for each guess that is close (+/- $.25) to what people actually choose.

So, to recap, you will have FOUR different screens of choices. FIRST, your best guess about the % of those here today who will send $5 as Person 1. SECOND, your own decision in case the computer makes you a Person 1 for the payoffs. THIRD, your best guess about how much people here will send as Person 2, for the case when they got nothing (LEFT SIDE of the screen), and again for the case in which they received $5 (RIGHT SIDE of the screen). FOURTH, your own choice in case the computer makes you a Person 2 for the payoffs, of how much to send back to

Person 1 when Person 1 sent you $0 (LEFT SIDE of the screen), and when Person 1 sent you $5 (RIGHT SIDE of the screen).

Any questions now?

At the end of the activity, the computer will show you what your earnings are in total, including both from your guesses about others and from your choices as Person 1 or Person 2.

OK, let's look at the instructions on the first computer screen. When you are happy you understand them, please click "continue" or "OK" in the lower right-hand corner of the screen. That will take you to a waiting screen, and when everyone is there, we will move to the first of two practice question screens, to make sure you understand how the payoff table works.

Practice Screen 1: Person 1 sends $0, and Person 2 responds by sending back $1.
Line 2 (Person 1 gets $7, Person 2 gets $4)

Practice Screen 2: Person 1 sends $5, and Person 2 responds by sending back $3. Line 10 (Person 1 gets $6, and Person 2 gets $12)

APPENDIX B: TABLE OF PD GAME PAYOFFS GIVEN TO SUBJECTS

	Person 1 Sends	Person 2 Sends		Person 1 Makes	Person 2 Makes
Line 1	$0	$0	→	$5	$5
Line 2	$0	$1	→	$7	$4
Line 3	$0	$2	→	$9	$3
Line 4	$0	$3	→	$11	$2
Line 5	$0	$4	→	$13	$1
Line 6	$0	$5	→	$15	$0

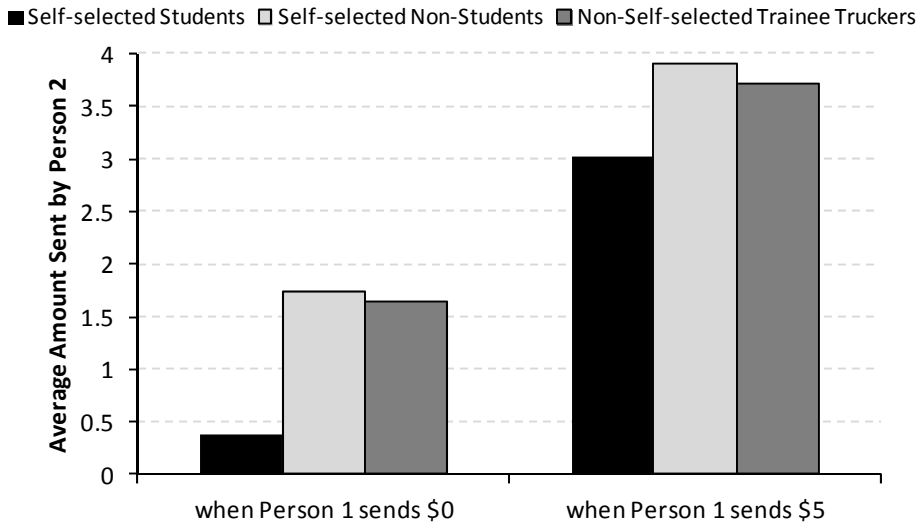
	Person 1 Sends	Person 2 Sends		Person 1 Makes	Person 2 Makes
Line 7	$5	$0	→	$0	$15
Line 8	$5	$1	→	$2	$14
Line 9	$5	$2	→	$4	$13
Line 10	$5	$3	→	$6	$12
Line 11	$5	$4	→	$8	$11
Line 12	$5	$5	→	$10	$10

APPENDIX C – ADDITIONAL STATISTICAL ANALYSIS

In this Appendix, we address our research questions following an alternative approach to analyze the data: instead of using the amounts transferred by Persons 2 to assign subjects to categories, as we did in Section 3 of the paper, we directly compare the amounts transferred across subject pools. The results of this alternative data analysis largely corroborate the results discussed in Section 3.

Figure C.1 below shows the average amount sent by subjects in the role of Person 2, disaggregated by subject pool. The Figure distinguishes between amounts that were transferred when Person 1 had sent \$0 and when Person 1 had sent \$5.

Figure C.1 – Average Amount Sent by Person 2, by Subject Pool



The amounts sent by the two non-student subject pools are, on average, similar: Self-selected Non-Students transferred an average of \$1.75 to Person 1 when Person 1 had sent \$0, and \$3.91 when Person 1 had sent \$5. Non-Self-selected Trainee Truckers transferred respectively \$1.64 and \$3.71 in these two situations. Using two-sided Wilcoxon rank sum tests, we do not find statistically significant differences in the amounts sent by Self-selected Non-Students and Non-Self-selected Trainee Truckers in either situation ($z = 0.926$, $p = 0.354$ for the situation where Person 1 sends \$0; $z = 0.892$, $p = 0.373$ for the situation where Person 1 sends \$0).

Figure C.1 also shows that Self-selected Students on average transfer lower amounts than Self-selected Non-Students, both when Person 1 sends \$0 (\$0.38 vs. \$1.75) and when Person 1 sends \$5 (\$3.02 vs. \$3.91). Using two-sided Wilcoxon rank sum tests, we reject the hypothesis that amounts transferred by Self-selected Students are the same as those of Self-selected Non-Students ($z = 5.858$, $p = 0.000$ when Person 1 sends \$0; $z = 3.052$, $p = 0.002$ when Person 1 sends \$5).

We further examine the differences in amounts sent across subject pools using two separate Tobit regressions (one for the situation where Person 1 sends \$0 and one for the situation where Person 1 sends \$5). In a first model, we just regress the amount sent in the role of Person 2 against dummy variables for the different subject pools. In a second model, we also use the set of control regressors used in Table 3 in Section 3 of the paper. Table C.1. reports the results of the regressions, which confirm our main findings: 1) whether participants did or did not self-select into the experiment does not affect the amount they transferred, 2) students transfer lower amounts than non-students.

Table C.1: Tobit Regressions

	When Person 1 sends $0		*When Person 1sends $5*	
	Ia	*IIa*	*Ib*	*IIb*
Self-Selected Students	-5.930***	-3.629***	-3.343***	-2.335*
	(.000)	(.002)	(.002)	(.060)
Non-Self-Selected Trainee Truckers	-0.428	-0.681	-0.793	0.003
	(.529)	(.388)	(.324)	(.997)
Age	-	0.094***	-	0.014
		(.000)		(.546)
Age^2 / 100	-	-0.098	-	-0.064
		(.533)		(.704)
Gender (1 if Female)	-	0.045	-	-0.766
		(.938)		(.217)
Non – White or Hispanic	-	0.447	-	-0.749
		(.378)		(.132)
Number of Siblings	-	0.213*	-	0.017
		(.096)		(.894)
Number of Siblings^2 / 100	-	1.854	-	-2.653
		(.607)		(.493)
Years of Education Completed	-	0.507***	-	0.324**
		(.000)		(.017)
Years of Education Completed^2 / 100	-	3.379	-	3.194
		(.274)		(.356)
Marital Status (1 if Single/etc.)	-	0.004	-	-0.600
		(.992)		(.194)
Income Category				
$10,000-$20,000	-	0.545	-	1.248*
		(.400)		(.083)
$20,000-$30,000	-	0.820	-	-0.362
		(.187)		(.582)
$30,000-$40,000	-	0.791	-	0.420
		(.278)		(.564)
$40,000-$50,000	-	0.918	-	2.180***
		(.260)		(.008)
$50,000-$60,000	-	1.102	-	0.056
		(.198)		(.954)
$60,000-$70,000	-	1.135	-	0.739
		(.224)		(.463)
$70,000+	-	0.011	-	0.599
		(.990)		(.496)
N.	*1206*	*1206*	*1206*	*1206*
Pseudo R^2	*0.004*	*0.029*	*0.004*	*0. 012*

Tobit regressions with robust standard errors. Dependent variable is amount sent by subjects in the role of Person 2. The reference subject type which is: Self-Selected Non-Student, Male, Married, Adult, White (Non-Hispanic), Income category $0-$10,000. Continuous variables are centered at their mean (means are: 36.2 for Age, 13.2 for Years of Education Completed, and 2.8 for Number of Siblings), and quadratic terms are computed for the mean-centered variables. P-values are reported in parentheses. A constant is included in all models, but omitted from the Table output. Significance levels: * 10% ; ** 5%; *** 1%.