

IZA DP No. 5475

**Did We Overestimate the Role of Social Preferences?  
The Case of Self-Selected Student Samples**

Armin Falk  
Stephan Meier  
Christian Zehnder

February 2011

# **Did We Overestimate the Role of Social Preferences? The Case of Self-Selected Student Samples**

**Armin Falk**

*University of Bonn  
and IZA*

**Stephan Meier**

*Columbia University  
and IZA*

**Christian Zehnder**

*University of Lausanne*

Discussion Paper No. 5475  
February 2011

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## ABSTRACT

### **Did We Overestimate the Role of Social Preferences? The Case of Self-Selected Student Samples<sup>\*</sup>**

Social preference research has received considerable attention among economists in recent years. However, the empirical foundation of social preferences is largely based on laboratory experiments with self-selected students as participants. This is potentially problematic as students participating in experiments may behave systematically different than non-participating students or non-students. In this paper we empirically investigate whether laboratory experiments with student samples misrepresent the importance of social preferences. Our first study shows that students who exhibit stronger prosocial inclinations in an unrelated field donation are not more likely to participate in experiments. This suggests that self-selection of more prosocial students into experiments is not a major issue. Our second study compares behavior of students and the general population in a trust experiment. We find very similar behavioral patterns for the two groups. If anything, the level of reciprocation seems higher among non-students suggesting that results from student samples might be seen as a lower bound for the importance of prosocial behavior.

JEL Classification: C90, D03

Keywords: methodology, selection, experiments, prosocial behavior

Corresponding author:

Stephan Meier  
Columbia University  
Graduate School of Business  
710 Uris Hall  
3022 Broadway  
New York, NY 10027  
USA  
E-mail: [sm3087@columbia.edu](mailto:sm3087@columbia.edu)

---

<sup>\*</sup> We thank Simon Gaechter, Lorenz Goette and Rafael Lalive for helpful comments.

# 1 Introduction

Social preferences such as concerns for distributional fairness and reciprocity have received considerable attention in recent economic research. In fact, in terms of citations and impact, social preference research is among the leading research programs in behavioral economics and has influenced the way many economists think about important economic and social phenomena (Fehr and Gächter, 2000; Cooper and Kagel, forthcoming).<sup>1</sup> A growing game theoretical literature has conceptualized the notion of fairness and social preferences in terms of inequity aversion (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000), fairness intention (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006), in addition to motives of reputation and image concerns (Benabou and Tirole, 2006; Seabright, 2009; see also Sobel, 2005). These models are helpful to understand the interaction of social preferences and economic incentives and have been applied to important economic questions such as contract theory (Fehr et al., 2007; Bartling and von Siemens, 2010), the role of fairness in the process of institution formation (Kosfeld et al., 2009), or the analysis of bargaining outcomes (Goeree and Holt, 2000).

Given their potentially far reaching implications in terms of economic predictions and policy recommendations, it is important to realize that the empirical foundation of social preferences is largely based on laboratory experiments with self-selected students as samples. A simple counting exercise reveals that in five field journals, which have persistently published lab experimental papers on social preferences, the vast majority of all subject pools consist of self-selected students. For the years 2004 to 2009 the

---

<sup>1</sup>Examples comprise tax policies and public economics in general (Bowles and Hwang, 2008; Ackert et al., 2009; Tyran and Sausgruber, 2006), contract enforcement (Fehr et al., 1993), wage formation and relational contracting (Brown et al., 2004; Falk et al., 2006), public goods provision (Andreoni, 1995; Fehr and Gächter, 2000; Masclet and Villeval, 2008), or organizational economics (MacLeod and Malcolmson, 1998; MacLeod, 2007; Ellingsen and Johannesson, 2008; Fehr et al., forthcoming).

overall percentage of lab papers on social preferences using student samples, including papers on reciprocity, fairness, altruism, envy and inequity aversion, is 89 percent.<sup>2</sup> We report data only for these five journals but the general statement applies to studies in other journals as well. In other words our knowledge concerning the nature of social preferences is almost exclusively based on self-selected student samples. This is a potential problem, because students participating in experiments might behave systematically different than non participating students or non-students. Student samples are obviously not representative of the general population. For example, they differ systematically from non-student populations in terms of social background, age or cognitive skills. Do they differ with respect to social preferences as well? If participating students behave more prosocially, we may have overestimated the potential of social preferences for the analysis of economic outcomes. In this case we should be more careful in plugging behavioral assumptions based on a very specific sample into models that derive implications for the general population.

In this paper we provide empirical evidence to test whether laboratory experiments with student samples lead to a systematic overestimation of the importance of social preferences. In particular we address two potential problems: First, experiments rely on volunteers, creating a problem of self-selection. This may bias outcomes in experiments if people who self-select into experimental studies are more prosocially inclined than people who do not participate in experiments. Second, most laboratory experiments are conducted with university undergraduates. While using students as subjects is very convenient, they are not representative of the general population. This is problematic if social preferences like trust and reciprocity are more prevalent among

---

<sup>2</sup>The journals and the respective percentages of studies that include student samples (for the years 2004 to 2009) are *Games and Economic Behavior* (83 percent), *Journal of Public Economics* (80 percent), *Economic Journal* (86 percent), *Journal of Economic Behavior & Organization* (100 percent), *Experimental Economics* (79 percent), total (85 percent), own calculations. Depending on classification (e.g., with regard to the definition of social preferences) these numbers vary and should therefore be taken for illustrative purposes only.

students than non-students. In a first study we therefore analyze whether participating students are more prosocial than non-participating students. The ideal data set to test for potential differences between participants and non-participants would provide information on prosocial preferences of *all* students while observing who participates in experiments and who does not. This type of data is usually not available simply because we know preferences typically only for participants in experiments. Moreover, if we know preferences from non-experimental data, e.g., survey studies, we do not observe decisions to participate in an experiment. In our first study we present results using a novel data set that combines preference measures for both participants and non-participants. In particular we use a naturally occurring donation decision as a measure of participants' and non-participants' prosocial inclination. Our results show that with respect to their prosocial inclination, students who participate in laboratory experiments are not significantly different from non-participating students. These findings resonate with recent results of a study by Cleave et al. (2010) who make use of a classroom experiment to study selection into experiments and also fail to find a bias regarding social preferences.

Our second study compares behavior of students and the general population. While social preferences of non-student subject pools have been measured before (see, e.g., Naef et al., 2008; Bellemare et al., 2008; Fehr and List, 2004), we are not aware of a study that provides a *ceteris-paribus* comparison between participants recruited out of a typical student pool and participants recruited out of a representative sample of the general population. Existing studies do either not use the same experimental procedures for the two groups or they do not recruit the non-students out of a representative sample.<sup>3</sup> In our study we use the same recruitment procedure, the same instructions,

---

<sup>3</sup>For example, in Bellemare and Kroeger (2007) non-students participate at home, while students come to the lab for the experiment. This difference in the decision environment might explain part of the difference between the two subject pools. The same argument applies to Carpenter et al. (2008) where students participate online and non-students mainly in a paper survey and List (2006) where

the same decision process and the same financial incentives for both our subject pools. We measure social preferences with the help of a variant of the trust game, one of the most important experimental paradigms to measure social preferences (Berg et al., 1995). We then compare first and second mover decisions in students with decisions of participants recruited from a sample of the general population. Our results show that first movers exhibit a very similar behavioral pattern. There is no significant difference in trusting behavior between students and non-students. In terms of reciprocation (second movers), higher investments are reciprocated with higher repayments in both groups. However, the level of reciprocation is significantly *lower* for students than for non-students. Thus if anything the level of reciprocation seems higher among non-students implying an even greater importance of reciprocity than assumed from student samples.

Our paper contributes to a recent methodological debate about the role of experimental economics in the social sciences (see, e.g., Levitt and List, 2007; Benz and Meier, 2008; Falk and Heckman, 2009; List, 2009; Croson and Gächter, 2010; Bardsley et al., 2010; Henrich et al., 2010; Gächter, 2010). Some of this work has raised serious concerns about the relevance of lab findings with regard to the role of social preferences. In particular, in addition to potential distortions caused by self-selected student samples, researchers have also expressed concerns that low financial stakes, short durations, and high degrees of scrutiny may bias the results of lab experiments. We think that this debate is useful. We also believe that most issues can - and actually should be - investigated empirically, calling for more not less experiments. This paper provides a step in this direction. Our results suggest that using self-selected student samples does not contribute to a systematic overestimation of the importance of social preferences. Of course, our results do not exclude that laboratory experiments 

---

sports card dealers participate in experiments at a trade show, while students interact in the lab.

overestimate the role of social preferences for other reasons. We see our paper as a starting point and hope that future research will also carefully investigate the empirical relevance of other potential sources for biases.

The rest of the paper is organized as follows. Section 2 reports design and results of the field study on selection of students into experiments. The question whether students and non-students have different prosocial inclinations is discussed in section 3. Both sections discuss the research design first and then summarize the results. Section 4 concludes.

## **2 Study 1: Do Prosocial Students Select into Experiments?**

### **2.1 Research Design**

This section analyzes whether there is a tendency for prosocially inclined undergraduate students to self-select into laboratory experiments. We study decisions of students to participate in experiments organized by the experimental economics laboratory of the University of Zurich. Our sample consists of 16,666 undergraduates who registered at the University of Zurich between the winter term 1998 and the summer term 2004 and for whom registration at the University of Zurich is the first enrollment at a University. For all those students, we know whether they participated at least once in an economics experiment. In total 1,783 students participated at least once, i.e., the participation rate is about 11 percent.

To measure the extent of all students' prosocial inclinations we use a naturally occurring prosocial decision at the University of Zurich as a proxy. Each semester,



every student who renews his university enrollment has to decide whether he or she wants to contribute a given amount (up to CHF 12; CHF 1  $\sim$  USD 0.95) to two social funds which provide charitable services (financial support for foreign students (CHF 5) and free loans for needy students (CHF 7), for further details, see Frey and Meier, 2004a,b). The level of possible donations (CHF 0 to CHF 12) is thus very similar to stake sizes typically used in lab settings.

We believe that these donation decisions provide us with a good proxy for measuring prosocial inclinations. First, our measure does not rely on self-reported survey responses but on actual decisions. Second, donation decisions are made in private and never made public.<sup>4</sup> Third, students are unaware that their behavior is analyzed in a research study. Fourth, Benz and Meier (2008) show that students who decide to contribute money to the two funds contribute substantially more of their endowment in a modified dictator game in the lab than students who don't contribute to the two funds. Fifth, and most importantly, all students at the university have to decide about contributions to the charitable funds. This means that our measure is not subject to any selection issue.

We construct two measures of prosocial behavior based on donation decisions. Our first measure considers only the very first donation decision of a student, i.e., the donation decision right before the first semester starts. This has the advantage that we measure prosocial inclinations before students have taken any courses at the University, before they have been exposed to any lab recruitment efforts and before they have participated in any experiment. We can therefore rule out the possibility of reversed causality as participation in experiments cannot have influenced the decision to contribute to the funds. We construct a dummy variable *Contributed in first decision*

---

<sup>4</sup>Students make their donation decision directly on the enrollment form which is sent to their homes a couple of weeks before the semester begins.

(=1), which is 1 if a student contributed to at least one of the two funds in his or her first donation decision, and 0 otherwise. It turns out that about 75 percent of the students contribute to at least one fund in their first decision. The second measure exploits information on all decisions taken by students. For each individual, we calculate how often she or he contributes to at least one of the funds, relative to the total number of decisions. The variable, *Individual contribution rate*, varies between 0 (never contributed to at least one fund) to 1 (contributed in all possible decisions to at least one fund). The mean is 0.76, indicating that students, on average, contribute in 76 percent of the decisions to at least one of the funds.<sup>5</sup>

## 2.2 Results

Panel A in Table 1 reveals that participants differ in various dimensions from non-participants. For example, participants are younger when registering, more likely to study law or medicine and less likely to study arts. These differences indicate the relevance of self-selection of particular groups of students. In Panel B of Table 1 we investigate whether this selection is driven by differences in prosocial inclinations. The panel provides descriptive statistics of contributions to the two funds for participants and non-participants. The summary statistic does not show any significant difference in prosocial behavior between participants and non-participants. In their first decision, the same proportion of participants and of non-participants contributed to at least one of the two funds (75 percent). When we look at all possible decisions of a student, it turns out that participants contribute on average in 77 percent of all possible decisions, while non-participants' contribution rate is 76 percent. This difference is not statistically significant.

---

<sup>5</sup>The two measures are positively correlated (Spearman's Rho= 0.71,  $p < 0.001$ ).

[Table 1 about here]

The fact that individual contribution rates are very similar is also apparent when comparing the distribution of contribution rates (see Figure 1). This figure displays a strikingly similar distribution of donation rates between the two samples. A non-parametric Kolmogorov-Smirnov test does not reject the null hypothesis that the samples are drawn from the same distribution ( $p = 0.64$ ). Thus, the raw data analysis does not reveal any significant difference in prosocial inclinations of participants and non-participants.

[Figure 1 about here]

Table 2 reports Probit estimations, where the dependent variable is an indicator variable for the decision to participate in experiments and the independent variable is either our first measure of prosocial inclination (columns (1) and (2)) or our second measure (columns (3) and (4)). We report marginal effects in brackets. Column (1) shows that students who contribute to at least one of the funds in their first decision are not significantly more likely to participate in an experiment in the following semesters than those who don't. The marginal effect is positive but small. An individual who contributed in his or her first semester is only 0.1 percentage points more likely to participate in an experiment compared to an individual who did not contribute. Column (2) controls for potential cohort effects. Including cohort dummies does not substantially affect the result.

[Table 2 about here]

Columns (3) and (4) report regressions using the *Individual contribution rate* as a proxy for prosocial inclinations. This proxy is potentially influenced by students' experience at the University including their participation in experimental studies. Whether

we include cohort dummies or not (columns (3) and (4)), the results are very similar to the ones obtained from using only the first decision: Individuals who contribute more often to at least one of the two charitable funds are not significantly more likely to participate in experiments. The marginal effects indicate that the participation rate of students who contribute in all possible cases (contribution rate = 1) is about 1 percentage point higher than the participation rate of students who never contributed to the two funds (contribution rate = 0). This means that for an increase in the contribution rate of one standard deviation (s.d.=0.34), the participation rate increases by only 0.3 percentage points. In sum, our results do not support the hypothesis that students with stronger prosocial inclinations self-select into economic experiments.

## **3 Study 2: Do Students Behave More Prosocially?**

### **3.1 Research Design**

In the previous section we presented data indicating that students who behave more prosocially in a natural setup are not more likely to participate in experiments than other students. This suggests that within the group of students the bias due to self-selection on social preferences is likely to be small. However, it is still possible that students in general tend to behave differently than other groups of the population. Compared to the general population students are different, e.g., with respect to their age, socioeconomic, educational and cultural background, which might correlate with the extent of prosocial behavior. For example, different age cohorts have been shown to exhibit different tendencies of prosocial behavior (e.g., Sutter and Kocher, 2007; Gächter and Herrmann, 2010). If students have different moral codes and social norms, their behavior in the laboratory may not be a good indicator for other people's

behavior outside the laboratory.

To test whether students behave more prosocially than non-students, we conduct identical trust experiments using two distinct subject pools. Thus we only vary the subject pool and no details of the decision context. Therefore differences in prosocial behavior can only be caused by differences between the two subject pools. All participants in the experiments live in Zurich. However, while one group of our participants was recruited from the student pool at the University of Zurich, the other group was recruited from a representative sample of the population of the city of Zurich. To recruit the students the university administration provided us with a random sample of 1000 addresses of undergraduate students of the University of Zurich, i.e., the same subject pool that the experimental economics laboratory of the University of Zurich typically uses to conduct experiments. For the recruitment of the participants from the general population the Statistical Office of the City of Zurich provided us with a sample of 4000 addresses of citizens. The procedure with which the sample was drawn ensures that it is representative for the city population with respect to gender, age and foreigner status.

To measure social preferences, we use a variant of the trust game, a simple sequential two-person game introduced by Berg et al. (1995). At the beginning both subjects receive an endowment of CHF 20. The first mover then decides how much of his endowment to transfer to the second mover. The transfer can be any amount in steps of 2 CHF, i.e., 0, 2, 4, . . . , or 20 CHF. The chosen transfer is tripled by the experimenter and passed to the second mover. Contingent upon the first mover's transfer the second mover decides on a back transfer to the first mover. This back transfer can be any integer amount between 0 and 80 CHF. Incomes are determined as follows: the first mover earns his endowment minus his own transfer plus the back transfer of the second mover. The second mover gets his endowment plus three times the first mover's transfer

minus the back transfer.

In addition, first movers were asked to indicate their expectation about the back transfer of their second mover given their own transfer decision. A measure of participants' beliefs allows us to distinguish between different motivations behind first mover transfers and whether these motives vary across subject pools. We can identify, e.g., whether high investments are made in the expectation of high back transfers or simply reflect a desire to be altruistic.

In order to elicit second movers' willingness to reciprocate, we used the contingent response method. This means that each second mover, before knowing the actual first mover's investment, made a back transfer decision for each of the 11 possible investments (0, 2, . . . , 20) of the first mover. The advantage of the contingent response method is that it allows us to measure each second mover's willingness to reciprocate independently from the transfer which he actually received. This is important, because it enables us to make a clean comparison of the level of reciprocity, even if first movers behave differently between subject pools. For example, if students would transfer less than non-students, back transfers of students would be lower, even if their reciprocal inclination was the same (or even more pronounced).

For logistical reasons the experiment was conducted via mail correspondence. All potential participants received a mailing including a cover letter, detailed instructions, a decision sheet and a questionnaire. The cover letter informed subjects about the possibility to take part in a paid experiment, conducted by the University of Zurich.<sup>6</sup> Subjects returned the completed decision sheets and questionnaires to the experimenters, using a pre-stamped envelope. The instructions explained the rules and procedures of the experiment in detail. There was no difference in the instructions for students and

---

<sup>6</sup>In order to enhance the credibility that we would actually pay subjects we added the remark that the Legal Service of the University guarantees that the study is run exactly according to the rules stated in the instructions.

non-students. Both groups of participants were told that they were randomly matched with another anonymous person who lives in Zurich. The subjects had to complete the questionnaire and the decision sheet. First movers made their transfer decision<sup>7</sup> and second movers filled out a contingent response table for the back transfers. We also made clear to subjects that the study was run in accordance with the data protection legislation of the city of Zurich. In particular, we stated that all data will be used only for scientific purposes and not given to any third parties. Moreover, we guaranteed that data would be stored in anonymous form and that any information specific to persons would be destroyed immediately after the data collection was completed.

The questionnaire contained items on socioeconomic characteristics and individual attributes like gender, age, marital status, profession, nationality and number of siblings, questions about the social environment of the participants, their political orientation, and items on their social network such as how many phone calls they had made during the last week, their number of friends in the residential neighborhood and the number of their club memberships. Table 3 provides descriptive statistics for these variables for students and non-students. Not surprisingly, the table reveals that students and non-students differ significantly in many dimensions. Non-students are on average older, are more likely to be married, have a lower education, have more friends in their neighborhood and are more likely to be right-wingers. In addition, the table indicates that the fraction of female participants is higher in the student sample than in the non-student sample.

[Table 3 about here.]

---

<sup>7</sup>First movers could condition their transfer decision on the 12 residential district of their second mover. Whether and how non-student first movers discriminate between people who live in different districts of Zurich is investigated in detail in Falk and Zehnder (2010). In this paper we consciously abstract from this feature of our experiment. In the following all calculations are based on the average transfer of a first mover across all possible residential districts of second movers.

Table 3 also reveals that the response rate of students was somewhat higher than the response rate of the more general subject pool. Roughly 300 of the 1000 contacted students took part in the study (30 percent), while about 1000 of the 4000 contacted citizens of the City participated (25 percent). For each subject pool separately, we randomly formed pairs among all participants who had sent back the completed decisions sheets.<sup>8</sup> Using the transfer decision of the first mover we then checked the corresponding back transfer of the second mover and calculated the profits of the first mover and the second mover. In a second mailing all participants were informed about the outcome of the experiment, i.e., the investment and back transfer decisions and the resulting payoffs for both players. The second mailing also contained the cash payments in a sealed envelope.

### 3.2 Results

We first examine trusting behavior of first movers. Transfer decisions of first movers are only slightly different in the two subject pools. Non-students transfer on average 13.17 CHF to their second mover, while students transfer 13.47 CHF. In column (1) of Table 4 we investigate whether the observed difference of 0.30 CHF is statistically significant. For this purpose we regress the first mover decisions on a student dummy, which is one if the decision comes from a student and zero otherwise. The regression analysis reveals that the difference is not statistically significant ( $p = 0.63$ ).<sup>9</sup> However, even if first mover transfers do not differ between subject pools, it is possible that students and non-students have different motives for their decisions. In column 2 of Table 4 we

---

<sup>8</sup>As there were a few more first movers than second movers, some second movers were matched twice. The payoff of these players was determined by the decisions associated with the first match.

<sup>9</sup>First movers can only choose transfers between 0 and 20. 224 out of the 652 have chosen the highest possible level of 20. To account for censoring, we also estimate a Tobit regression of first mover decisions on the student dummy. The results confirm that there is no statistically significant difference between students and non-students regarding first mover transfers ( $p = 0.66$ ).



therefore regress first movers' back transfer expectation on the student dummy and the level of the transfer. The fact that the student dummy is not statistically significant indicates that the decisions of students and non-students are not driven by different beliefs about the behavior of second movers. Thus, our results suggest that neither beliefs about the trustworthiness and reciprocity of their fellow citizens nor transfer decisions in the experiment differ between students and non-students.

[Table 4 about here.]

We now turn to second movers' behavior, which informs us about reciprocal inclinations. In Table 5 we show average second mover back transfers conditional on first mover transfer. For every possible first mover transfer students make lower average repayments than non-students. All differences are statistically significant (see Table 5 for the corresponding p-values). Averaging over all backtransfers, students transfer back 15 percent less than non-students.

[Table 5 about here.]

The fact that students transfer back less than non-students does not imply that they generally react less sensitive to first movers' transfers. In fact Figure 2 illustrates that the slope between first mover transfer and second mover back transfer is very similar. Put differently, students' and non-students' reciprocation pattern is very similar; the only difference being that students reciprocate on a lower absolute level. Results in column (3) in Table 4 confirm this. It shows an OLS regression with second mover's back transfer as the dependent variable. We regress back transfers on a student dummy, the first mover transfer and the interaction effect between student dummy and first mover transfer. The coefficient of the student dummy is negative and significant, i.e., students transfer back significantly less than non-students. However,

the interaction effect is close to zero indicating that students and non-students exhibit a similar reciprocal inclination as suggested by Figure 2. All results are robust if we use Tobit estimates to account for censoring.

[Figure 2 about here.]

## 4 Concluding remarks

This paper empirically tests whether laboratory experiments with students systematically misrepresent the importance of social preferences. Such an empirical test is critical as experimental methods become increasingly important in economics and experimental results, especially those on social preferences, often challenge traditional views and policy implications. The convention to mainly use students might affect the validity of laboratory experiments in two important ways: among the student population the more prosocial individuals may volunteer to participate and students' prosocial inclination may be more pronounced than that of the general population.

The results of our two studies show that, first, students who exhibit more prosocial behavior in an unrelated field donation are not more likely to participate in experiments. This suggests that there is no significant bias due to self-selection of excessively prosocial students into experiments. Second, student participants and non-student subjects show very similar behavioral patterns in our trust experiment. Although students make significantly lower second mover repayments, they exhibit a similar investment behavior and hold similar beliefs about second mover behavior as participants recruited from the general population. In addition, we find that students and non-students exhibit a similar reciprocal inclination: the increase in repayments in response to higher first mover investments is very similar between subjects pools. The results of our sec-

ond study resonate with earlier evidence suggesting that other subject pools tend to exhibit even more prosocial behavior than students (see e.g., Fehr and List, 2004; Egas and Riedl, 2008; Dohmen et al., 2008; Burks et al., 2009; Belot et al., 2010). Together, our two studies suggest that relying on self-selected students samples does not systematically overestimate the importance of social preferences. However, we think it is important that future studies also investigate the empirical relevance of other aspects of lab studies that might favor prosocial behavior (e.g., high degrees of scrutiny, short durations, or the salience of payoff comparisons).

It would be interesting to also learn more about selection patterns of non-student samples. In our second study we analyze prosocial behavior of two samples that volunteered to participate in the experiment. In light of our first study it seems plausible to assume the absence of important selection effects with respect to social preferences even for non-student samples, but we cannot rule out such a possibility with our data. Even if sorting takes place among non-students, however, our results suggest that whenever research is based on volunteers, recruiting subjects from the general population does not lead to dramatically different results than recruiting from a student subject pool. In particular, results obtained from student samples are not systematically biased towards overestimating the prevalence of prosocial inclinations.

## References

- ACKERT, L. F., A. B. GILLETTE, J. MARTINEZ-VAZQUEZ, AND M. RIDER (2009): “Risk Tolerance, Self-Interest, and Social Preferences,” Georgia State University Experimental Economics Center Working Paper Series 2009-04.
- ANDREONI, J. (1995): “Warm-Glow versus Cold-Prickle: The Effects of Positive and Negative Framing on Cooperation in Experiments,” *The Quarterly Journal of Economics*, 110, 1–21.
- BARDSLEY, N., R. CUBITT, G. LOOMES, P. MOFFATT, C. STARMER, AND R. SUGDEN (2010): *Experimental economics: Rethinking the rules*, Princeton University Press.
- BARTLING, B. AND F. VON SIEMENS (2010): “The Intensity of Incentives in Firms and Markets: Moral Hazard with Envious Agents,” *Labour Economics*, 17, 598–607.
- BELLEMARE, C. AND S. KROEGER (2007): “On representative social capital,” *European Economic Review*, 51, 183–202.
- BELLEMARE, C., S. KRÖGER, AND A. VAN SOEST (2008): “Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities,” *Econometrica*, 76, 815–839.
- BELOT, M., R. DUCH, AND L. MILLER (2010): “Who should be called to the lab? A comprehensive comparison of students and non-students in classic experimental games,” Discussion Papers 2010001, University of Oxford, Nuffield College.
- BENABOU, R. AND J. TIROLE (2006): “Incentives and Prosocial behavior,” *American Economic Review*, 96, 1652–1678.
- BENZ, M. AND S. MEIER (2008): “Do People Behave in Experiments as in the Field? Evidence from Donations,” *Experimental Economics*, 11, 268–281.

- BERG, J., J. DICKHAUT, AND K. MCCABE (1995): “Trust, Reciprocity, and Social History,” *Games and Economic Behavior*, 10, 122–142.
- BOLTON, G. E. AND A. OCKENFELS (2000): “ERC: A Theory of Equity, Reciprocity, and Competition,” *American Economic Review*, 90, 166–193.
- BOWLES, S. AND S.-H. HWANG (2008): “Social Preferences and Public Economics: Mechanism Design when Social Preferences Depend on Incentives,” Department of Economics University of Siena 530, Department of Economics, University of Siena.
- BROWN, M., A. FALK, AND E. FEHR (2004): “Relational Contracts and the Nature of Market Interactions,” *Econometrica*, 72, 747–780.
- BURKS, S., J. CARPENTER, AND L. GOETTE (2009): “Performance pay and worker cooperation: Evidence from an artefactual field experiment,” *Journal of Economic Behavior & Organization*, 70, 458–469.
- CARPENTER, J., C. CONNOLLY, AND C. MYERS (2008): “Altruistic behavior in a representative dictator experiment,” *Experimental Economics*, 11, 282–298.
- CLEAVE, B. L., N. NIKIFORAKIS, AND R. SLONIM (2010): “Is There Selection Bias in Laboratory Experiments?” Department of Economics - Working Papers Series 1106, The University of Melbourne.
- COOPER, D. AND J. KAGEL (forthcoming): *The Handbook of Experimental Economics, Vol. 2*, chap. Other Regarding Preferences: A Selective Survey of Experimental Results.
- CROSON, R. AND S. GAECHTER (2010): “The science of experimental economics,” *Journal of Economic Behavior & Organization*, 73, 122–131.
- DOHMEN, T., A. FALK, D. HUFFMAN, AND U. SUNDE (2008): “Representative Trust And Reciprocity: Prevalence And Determinants,” *Economic Inquiry*, 46, 84–90.

- DUFWENBERG, M. AND G. KIRCHSTEIGER (2004): “A theory of sequential reciprocity,” *Games and Economic Behavior*, 47, 268–298.
- EGAS, M. AND A. RIEDL (2008): “The economics of altruistic punishment and the maintenance of cooperation,” *Proceedings of the Royal Society - Biological Sciences*, 871–878.
- ELLINGSEN, T. AND M. JOHANNESSON (2008): “Pride and Prejudice: The Human Side of Incentive Theory,” *American Economic Review*, 98, 990–1008.
- FALK, A., E. FEHR, AND C. ZEHNDER (2006): “Fairness Perceptions and Reservation Wages - The Behavioral Effects of Minimum Wage Laws,” *Quarterly Journal of Economics*, 121, 1347–1381.
- FALK, A. AND U. FISCHBACHER (2006): “A theory of reciprocity,” *Games and Economic Behavior*, 54, 293–315.
- FALK, A. AND J. HECKMAN (2009): “Lab Experiments Are a Major Source of Knowledge in the Social Sciences,” *Science*, 326, 535–538.
- FALK, A. AND C. ZEHNDER (2010): “Social Capital at the City Level: A Case Study,” Working paper, University of Bonn and University of Lausanne.
- FEHR, E. AND S. GACHTER (2000): “Cooperation and Punishment in Public Goods Experiments,” *American Economic Review*, 90, 980–994.
- FEHR, E., O. HART, AND C. ZEHNDER (forthcoming): “Contracts as Reference Points - Experimental Evidence,” *American Economic Review*.
- FEHR, E., G. KIRCHSTEIGER, AND A. RIEDL (1993): “Does Fairness Prevent Market Clearing? An Experimental Investigation,” *The Quarterly Journal of Economics*, 108, 437–59.
- FEHR, E., A. KLEIN, AND K. M. SCHMIDT (2007): “Fairness and Contract Design,” *Econometrica*, 75, 121–154.

- FEHR, E. AND J. A. LIST (2004): “The Hidden Costs and Returns of Incentives - Trust and Trustworthiness among CEOs,” *Journal of the European Economic Association*, 2, 743–71.
- FEHR, E. AND K. M. SCHMIDT (1999): “A Theory Of Fairness, Competition, And Cooperation,” *The Quarterly Journal of Economics*, 114, 817–868.
- FREY, B. S. AND S. MEIER (2004a): “Pro-Social Behavior in a Natural Setting,” *Journal of Economic Behavior and Organization*, 54, 65–88.
- (2004b): “Social Comparison and Pro-Social Behavior: Testing Conditional Cooperation in a Field Experiment,” *American Economic Review*, 94, 1717–22.
- GAECHTER, S. (2010): “(Dis)advantages of student subjects: What is your research question?” *Behavioral and Brain Sciences*, 33, 92–93.
- GAECHTER, S. AND B. HERRMANN (2010): “The limits of self-governance when cooperators get punished: Experimental evidence from urban and rural Russia,” *European Economic Review*, forthcoming.
- GOEREE, J. K. AND C. A. HOLT (2000): “Asymmetric inequality aversion and noisy behavior in alternating-offer bargaining games,” *European Economic Review*, 44, 1079–1089.
- HENRICH, J., S. HEINE, AND A. NORENZAYAN (2010): “The weirdest people in the world,” *Behavioral and Brain Sciences*, 33, 61–83.
- KOSFELD, M., A. OKADA, AND A. RIEDL (2009): “Institution Formation in Public Goods Games,” *American Economic Review*, 99, 1335–55.
- LEVITT, S. D. AND J. A. LIST (2007): “What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?” *Journal of Economic Perspectives*, 21, 153–174.
- LIST, J. A. (2006): “The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions,” *Journal of Political Economy*, 114, 1–37.

- (2009): “Social Preferences: Some Thoughts from the Field,” *Annual Review of Economics*, 1, 1:21.1–21.20.
- MACLEOD, W. B. (2007): “Can Contract Theory Explain Social Preferences?” *American Economic Review*, 97, 187–192.
- MACLEOD, W. B. AND J. M. MALCOMSON (1998): “Motivation and Markets,” *American Economic Review*, 88, 388–411.
- MASCLET, D. AND M.-C. VILLEVAL (2008): “Punishment, inequality, and welfare: a public good experiment,” *Social Choice and Welfare*, 31, 475–502.
- NAEF, M., E. FEHR, U. FISCHBACHER, J. SCHUPP, AND G. WAGNER (2008): “Decomposing Trust: Explaining National and Ethical Trust Differences,” Working paper, Institute for Empirical Research in Economics, University of Zurich.
- RABIN, M. (1993): “Incorporating Fairness into Game Theory and Economics,” *American Economic Review*, 83, 1281–1302.
- SEABRIGHT, P. B. (2009): “Continuous Preferences and Discontinuous Choices: How Altruists Respond to Incentives,” *The B.E. Journal of Theoretical Economics*, 9.
- SOBEL, J. (2005): “Interdependent Preferences and Reciprocity,” *Journal of Economic Literature*, 43, 392–436.
- SUTTER, M. AND M. KOCHER (2007): “Trust and trustworthiness across different age groups,” *Games and Economic Behavior*, 59, 364–382.
- TYRAN, J.-R. AND R. SAUSGRUBER (2006): “A little fairness may induce a lot of redistribution in democracy,” *European Economic Review*, 50, 469–485.



Figure 1: Distributions of Individual Contribution Rates in Study 1

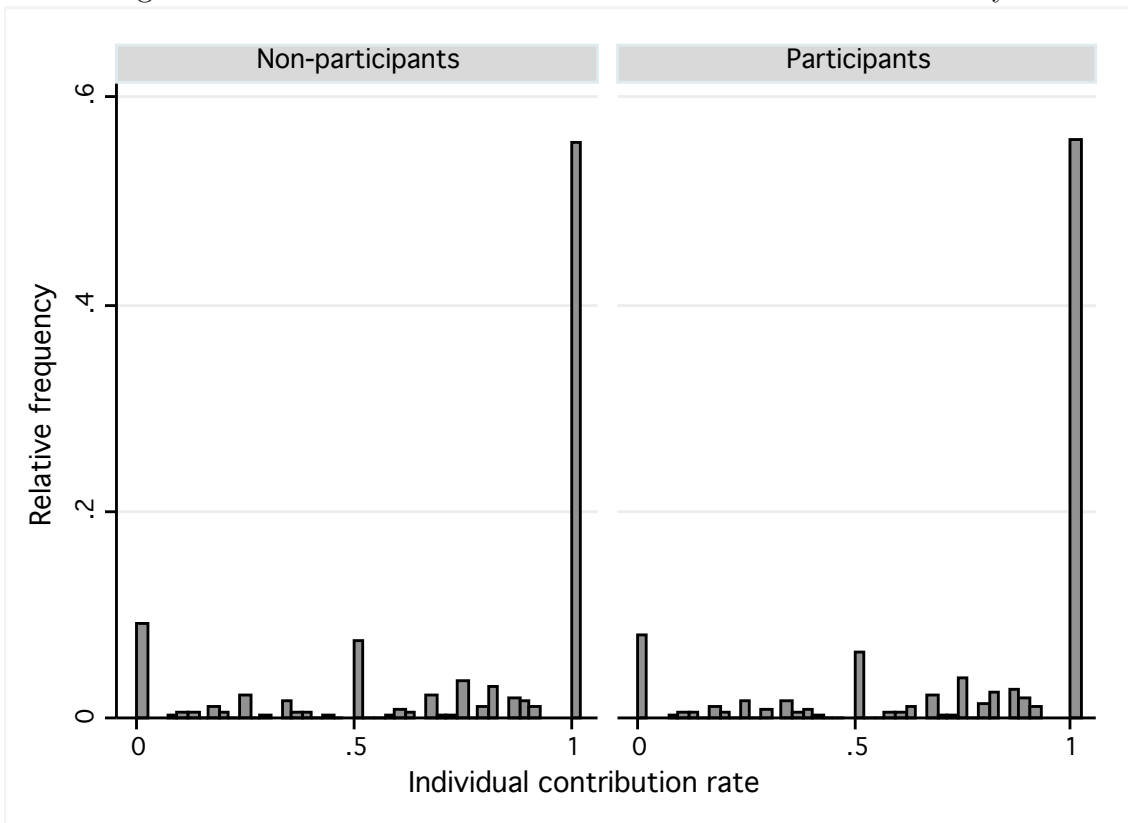


Figure 2: Back transfers of Students and Non-Students in Study 2

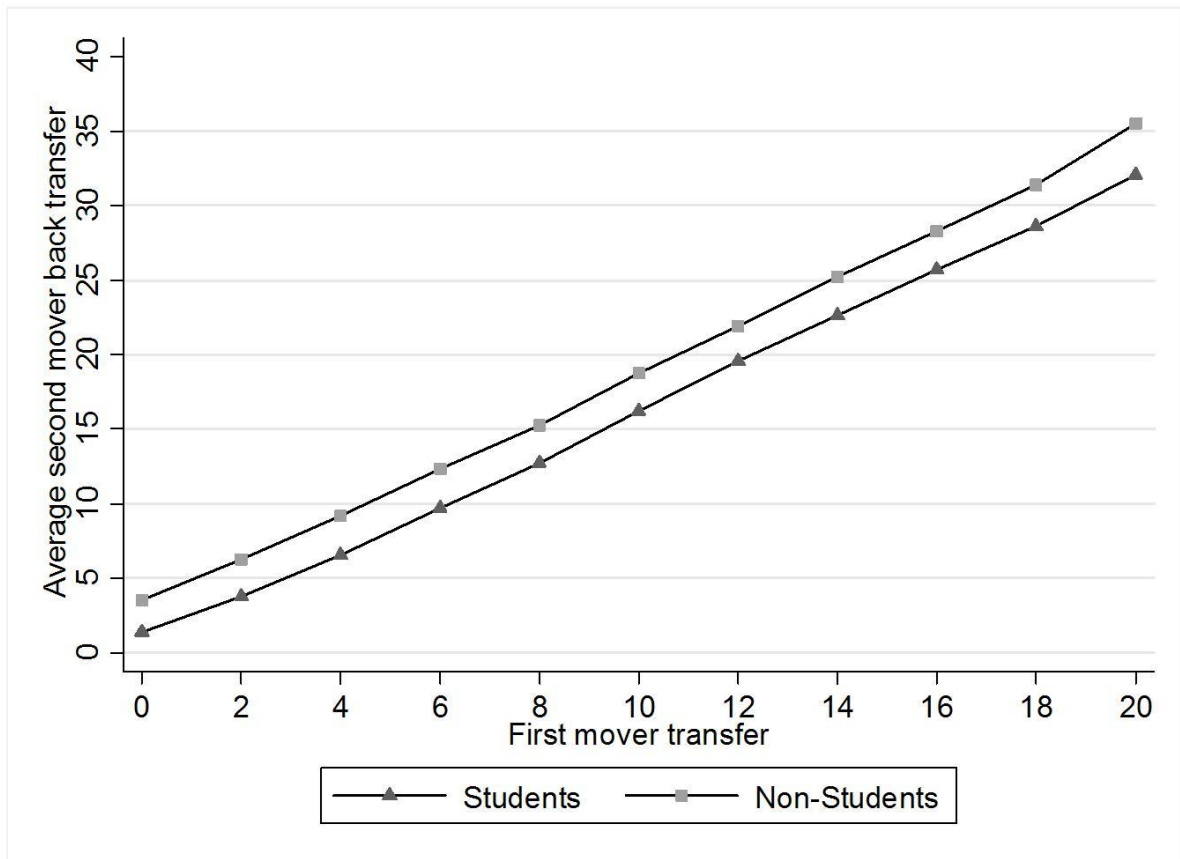


Table 1: Summary Statistics of Study 1

Variable	Non-participants		Participants		$t$ -test/ $\chi^2$ -test <sup>1</sup>
	Mean	s.d.	Mean	s.d.	
<b>Panel A: Observable characteristics</b>					
Age at registration	21.94	4.21	21.07	2.87	$p < 0.01$
No. of semesters	5.34	3.26	5.97	3.15	$p < 0.01$
Gender (Women=1)	0.57	0.50	0.53	0.50	$p < 0.01$
Nationality (Foreigner=1)	0.08	0.27	0.07	0.25	$p < 0.05$
Computer science	0.04	0.18	0.03	0.16	$p = 0.21$
Economics & Business	0.13	0.32	0.14	0.34	$p < 0.05$
Theology	0.01	0.08	0.003	0.05	$p < 0.05$
Law	0.16	0.36	0.25	0.42	$p < 0.01$
Medicine	0.07	0.26	0.18	0.38	$p < 0.01$
Veterinary medicine	0.03	0.16	0.03	0.16	$p = 0.64$
Arts faculty	0.47	0.49	0.33	0.46	$p < 0.01$
Natural science	0.10	0.30	0.05	0.21	$p < 0.01$
<b>Panel B: prosocial behavior</b>					
Contributed in first decision (=1) <sup>2</sup>	0.75	0.43	0.75	0.43	$p = 0.80$
Individual contribution rate <sup>3</sup>	0.76	0.34	0.77	0.33	$p = 0.20$
No. of observations	14,884		1,783		

*Note:*

<sup>1</sup>  $\chi^2$ -tests for categorical variables and  $t$ -tests otherwise.

<sup>2</sup> Contributed to at least one fund in the first decision (before attending any courses).

<sup>3</sup> Proportion of decisions (out of her total), she contributed to at least one fund.

Table 2: Participating in Experiments Depending on Prosocial Behavior in the Field

Variable	(1)	(2)	(3)	(4)
Contributed in first decision (=1)	0.008 (0.030) [0.001]	0.011 (0.030) [0.002]		
Individual contribution rate			0.050 (0.038) [0.009]	0.053 (0.040) [0.010]
Cohort dummies	No	Yes	No	Yes
Constant	-1.248** (0.026)	-1.238** (0.038)	-1.281** (0.032)	-1.269** (0.044)
No. of observations	16,666	16,666	16,666	16,666
Pseudo R squared	0.000	0.008	0.000	0.008

*Note:*

Dependent Variable: Participating at least once in an laboratory experiment (= 1). Probit regressions. Standard errors in parentheses. Marginal effects in brackets. Cohort dummies control for the semester/year in which students registered. Level of significance: \*\*  $p < 0.01$ , \*  $0.01 < p < 0.05$ .

Table 3: Summary Statistics

	Non-Students			Students			t-test/ $\chi^2$ -test <sup>5</sup>
	Mean	Std. Dev.	# Obs.	Mean	Std. Dev.	# Obs.	
Female	0.37	0.48	998	0.55	0.50	295	p < 0.01
Age	45.09	16.18	979	26.72	5.65	291	p < 0.01
Only Child	0.10	0.31	1001	0.13	0.34	295	p = 0.25
Foreigner	0.09	0.29	998	0.11	0.31	295	p = 0.47
Married	0.60	0.49	1001	0.09	0.29	295	p < 0.01
High Education <sup>1</sup>	0.56	0.50	888	1.00	0.00	295	p < 0.01
Club Memberships	1.32	1.95	997	1.05	1.39	295	p = 0.03
Friends <sup>2</sup>	7.76	16.54	992	3.92	5.51	291	p < 0.01
Phone Calls <sup>3</sup>	10.05	9.38	986	11.75	12.10	290	p = 0.01
Leftwinger <sup>4</sup>	0.53	0.50	985	0.58	0.49	290	p = 0.10
Rightwinger <sup>4</sup>	0.22	0.42	985	0.14	0.34	290	p < 0.01

*Note:*

<sup>1</sup> Dummy variable that takes on the value 1 if the person has obtained the general qualification for entrance to university or technical college.

<sup>2</sup> Number of friends in current residential neighborhood

<sup>3</sup> Number of private phone calls within the last week

<sup>4</sup> Dummy variable that indicates whether the person conceives herself as a leftwinger (rightwinger); originally measured on a five-point scale from 1 (leftwing) to 5 (rightwing). The dummy takes on the value 1 if the person has ticked "1" or "2" ("4" or "5") and 0 otherwise.

<sup>5</sup>  $\chi^2$ -tests for categorical variables and *t*-tests otherwise.

Table 4: Regression Analysis of First Mover and Second Mover Behavior

Dependent Variable	(1) FM Transfer	(2) FM Belief	(3) SM Back Transfer
Student	0.299 [0.611]	0.562 [0.765]	-2.297** [0.483]
First Mover Transfer		1.497** [0.046]	1.597** [0.036]
Student x FM Transfer			-0.056 [0.067]
Constant	13.169** [0.287]	-2.620** [0.422]	2.907** [0.285]
# Observations	652	652	7076
R squared	0.0004	0.5856	0.488

*Note:*

Column (1): OLS-estimations with robust standard errors in brackets. The dependent variable is the average transfer for each first mover (see Footnote 7 in the text for details). Column (2): OLS-estimations with robust standard errors in brackets. The dependent variable is the average expected back transfer for each first mover. Column (3): OLS-estimations with robust standard errors clustered on individual in brackets. The dependent variable consists of second mover repayments. As repayment decisions are elicited with the contingent response method, we have eleven observations per second mover (one for each possible first mover transfer). Level of significance: \*\*  $p < 0.01$ , \*  $0.01 < p < 0.05$

Table 5: Second Mover Back Transfers Conditional on First Mover Transfer

FM Transfer	SM Back Transfers			WMW-Test <sup>1</sup>
	Non-Students	Students	Difference	
0	3.52	1.25	2.27	p < 0,01
2	6.20	3.72	2.48	p < 0,01
4	9.10	6.51	2.59	p < 0,01
6	12.27	9.59	2.68	p < 0,01
8	15.27	12.45	2.82	p < 0,01
10	18.71	16.05	2.66	p < 0,01
12	22.00	19.16	2.84	p < 0,01
14	25.29	22.10	3.20	p = 0,01
16	28.34	25.31	3.03	p = 0,02
18	31.36	28.42	2.94	p = 0,07
20	35.55	31.67	3.88	p = 0,05

*Note:*

<sup>1</sup> Non-parametric Wilcoxon-Mann-Whitney Test.