Exp Econ (2008) 11: 282–298
DOI 10.1007/s10683-007-9193-x

Altruistic behavior in a representative dictator experiment

Jeffrey Carpenter · Cristina Connolly · Caitlin Knowles Myers

Published online: 15 February 2008

Abstract We conduct a *representative* dictator game in which students and random members of the community choose both what charity to support and how much to donate to the charity. We find systematic differences between the choices of students and community members. Community members are much more likely to write in their own charity, community members donate significantly more ($17), on average, and community members are much more likely (32%) to donate the entire $100 endowment. Based on this evidence, it does not appear that student behavior is very representative in the context of the charitable donations and the dictator game.

Keywords Altruism · Dictator game · Field experiment · Representativeness

JEL Classification C72 · C93 · D64

1 Introduction

Economists conducting behavioral research have often struggled with the extent to which convenience samples of college students are representative of larger, more

Electronic supplementary material The online version of this article (http://dx.doi.org/10.1007/s10683-007-9193-x) contains supplementary material, which is available to authorized users.

J. Carpenter (✉) · C. Connolly
Department of Economics, Middlebury College, 601 Warner, Middlebury, VT 05753, USA
e-mail: jpc@middlebury.edu

C. Connolly
e-mail: caconnol@middlebury.edu

C.K. Myers
Department of Economics, Middlebury College, 502E Warner, Middlebury, VT 05753, USA
e-mail: cmyers@middlebury.edu

general, populations. It could be, for example, that behavior is correlated with demographics but that researchers fail to identify such relationships because of the lack of variation in the student population. Another possible issue is that students simply may have little experience with many of the sorts of decision-making environments that interest economists. This worry has recently led a few economists to conduct experiments in the field with participants who either have more experience in the decision-making environment under investigation or are more representative of the broader population of decision-makers.[1] In general, these experiments have been well-received.

However, making sure that the participants in an experiment are representative of the populations to which one wishes to generalize the results is only one of two criteria set out by Egon Brunswik when he formulated the notion of a *representative design* in psychology (Brunswik 1956). Economists have been less interested in incorporating the second criterion into their experiments (Hogarth 2005). Brunswik's second criterion is that the situations faced by decision-makers are also representative of their environments.[2]

A good example of a non-representative design is the standard double blind dictator game (Forsythe et al. 1994) played by college students. This experiment has been used extensively to study and measure altruism. In the dictator game two people, who typically do not know each other, are charged with splitting a monetary pie. The dictator has full control over how the pie will be split—she can allocate any amount to the recipient and keep the rest. The amount given to the recipient is taken as a measure of altruism. Clearly, the fact that college students are often the participants in the dictator game means that the results may not be very representative; however, there are now a growing number of field studies in which other populations have been sampled from.

Carpenter et al. (2005a) for example show that although typical college students are sensitive to the fact that the recipient cannot veto the dictator's allocation, a more generally representative slice of participants who work in a warehouse facility in Kansas are not. In fact, the distribution of warehouse worker offers in the dictator game is indistinguishable from the distribution generated by an ultimatum game in which the recipient can reject the offer of the other player resulting in both players receiving \$0. In a much more ambitious project, Henrich et al. (2006) report on dictator games conducted in 15 diverse populations around the globe. Unlike the standard \$10 game played with students where allocations are often close to uniformly distributed on the interval [\$0, \$5], there is a lot of variation in the Henrich et al. data and the variation tends to correlate with the willingness of members of a population to enforce distribution norms.

The other problem with the standard dictator game is that the decision task is relatively unnatural and therefore may not be representative.[3] While people rou-

[1] Harrison and List (2004) refer to these as "artefactual" field experiments.

[2] Although our experiment is best categorized as artefactual by the Harrison and List (2004) system of classification, we do not think that if fits neatly in this category for this reason.

[3] We consider the notion of *representativeness* to be related to, but not the same as, *external validity*. For us experimental behavior is externally valid if it correlates with behaviors measured in more naturally

tinely give money to strangers who are obviously less well off in face-to-face interactions and they give to established charities with even more regularity, how often are you asked to give money to an anonymous stranger without any context or means of assessing exactly how deserving the stranger is (Camerer and Thaler 1995; Camerer 2003)? Just as dictator game research has been conducted with more representative populations, there have been a few studies that have examined more representative protocols. In the first, Eckel and Grossman (1996) used a convenience sample of students, but changed the protocol so that the recipient was the American Red Cross. This manipulation had the expected result of increasing average allocations. In another experiment, Fong (2005), again using student dictators, recruited people who qualify for welfare to play the role of the recipient in another modified dictator experiment. Her treatments varied the degree to which the recipients appeared willing to "pull themselves up by their bootstraps" and she found some evidence that recipient industriousness mattered.

Our contribution is to conduct a dictator experiment that is representative in both domains. We start by designing, what we consider to be, an even more representative version of the Eckel and Grossman (1996) dictator game. In this game, participants were allowed to choose among thirteen reputable charities to take the role of the recipient instead of being forced to give to the American Red Cross. By changing the game in this way, we can be more sure that the participants felt some attachment to whatever charity they picked. In fact, the game is clearly more representative because the participants could also write in whatever charity they wanted if they did not want to pick one of the thirteen on the list. We then conducted this experiment with two populations: a sample of students from Middlebury College, a residential liberal arts college in Vermont, and a sample of Vermonters drawn from the broader population.

The related issue of "generalizability" is discussed at length in List (2006). For our purposes, the important insight in this paper is that one must accept several assumptions when asserting that data gathered from a convenience sample of students will be a good indicator of behavior in other populations. One obvious concern is selection. Although one might invite participants from different subject pools using exactly the same procedures, it can easily be the case that the key determinants of participation vary from one group to the next. While List (2006) considers the implications of overlapping distributions of behavior, a related concern that is particularly important for our study is that the distributions of many of the important independent variables may not overlap nicely making it hard to assess the marginal effect of these factors. In the end, however, List (2006) shares our concern that the representativeness of the environment in which decisions are made is crucial for generalizability.

Our experiment was embedded in a broader survey so that we could also contribute by collecting the data necessary to test hypotheses about why student behavior might not be representative. Specifically, we collected information on a number of demographic characteristics (e.g., age, sex and education) that have been shown to matter in similar settings (as in List 2004 or Botelho et al. 2005), we collected standard

occurring contexts, regardless of the representativeness of the design. For example, if giving in a context-free student dictator experiment predicts real charitable donations, it is externally valid despite the design of the experiment not being particularly representative (Benz and Meier 2005).

survey measures of altruism to test the construct validity of our protocol, and we collected data that may shed light on the source and acquisition of altruistic traits.

To preview our main results, we find that student behavior is not representative of the behavior of members of the broader community. Students tended to be significantly less likely to write in their own choice of a charity and given this choice, they allocated significantly less as dictators. The robust demographic determinants of allocations include age, student status and sex. We also find evidence that the representative dictator game correlates with other measures of altruism and that altruistic traits are acquired (at least partially) from mothers and friends. These results echo the importance of subject pool differences stressed in Levitt and List (2007).

Like List (2007) and Bardsley (2008), we think that our results suggest the importance of institutions in distribution games and that by varying the institutional rules of the game we learn more about the distribution of social preferences in the lab and the field. Not only do institutions constrain the action space of decision-makers (e.g., you could never take money), the range of permissible actions may provide clues about what norms are appropriate. The fact that choosing the recipient seems to matter in both of our subject pools (although perhaps more so in the community sample) is just another institutional variation that helps predict behavior. That said, our point is slightly sharper: while varying institutions will help us identify preferences, varying them in a direction that might make the resulting experiment more representative, may expedite the project.

We proceed by describing our representative dictator protocol in the next section. We then provide an overview of our data, examine the choice of charity and the choice of how much to allocate to charity. We conclude by discussing how our results dovetail with other similar findings.

2 Designing a representative dictator experiment

Taking Eckel and Grossman (1996) as our starting point, we created a $100 dictator game in which participants first chose among the following 13 charities to take the role of the recipient: the American Cancer Society (ACS), the American Diabetes Association (ADA), Amnesty International (AI), Doctors without Borders (DwoB), the Humane Society (HS), Habitat for Humanity (HforH), the Nature Conservancy (NC), The United Nations Children's Fund (UNICEF), the United Service Organizations (USO), the United Way (UW), the Vermont Land Trust (VLT) or Vermont Public Radio (VPR). If the participant did not like any of the 13 choices she could write in her own choice on a fourteenth line.

Once the participant chose a charity, she was then asked to divide $100 between the charity and herself. The participants were told that after all the responses were collected we would pick 10% randomly and implement the allocation decision. In expectation, our game matched the $10 stakes used in many previous dictator experiments and Carpenter et al. (2005b) show that changing the stakes from $10 to $100 has no effect on the distribution of allocations. When all the responses were received we wrote checks directly to each of the charities for the total amounts that had been donated and, to keep the responses as anonymous as possible, we sent unnamed VISA

gift cards to the dictators in the amounts that they had chosen to keep for themselves. The exact wording of the instructions for the experiment appears in the Appendix (electronic supplementary material).

To learn more about how altruistic traits might be formed, we asked each participant three additional questions. We asked them how much they thought a random participant would allocate to charity, how much they thought that their mothers would allocate, and how much they thought that their best friends would allocate. The motivation behind the second and third of the questions was to test to what degree altruistic traits are transmitted from mother to child or from friend to friend. Of course the skeptic might think that the answers to these two questions might be confounded by what psychologists refer to as *projection bias* whereby participants project their own motives and views onto the anticipated behavior of others. For this reason we included the first question to control for projection bias. If the second two questions predict allocations, controlling for the first, we feel comfortable interpreting the second two as avenues of trait transmission. However, it should be noted that in the case of friend giving, it is unclear in which direction transmission works. The participant may be influenced by his or her friend, but the friend is likely similarly influenced by the participant. Moreover, it may be the case that people seek out friends with similar levels of altruism.

The student data were collected in the spring of 2006 via an online survey developed at Middlebury College and emailed to a list of students who had previously expressed some interest in participating in experiments. The community data were collected in the summer of 2006. We purchased a sample of 2000 addresses in the state of Vermont. The sample was drawn randomly on all but one criterion, sex. Because these data are a component of a larger project on altruism and volunteerism in which we are also collecting data from volunteer firefighters and because firefighters are predominately male, we over-sampled males in the community survey. Firefighters are not included in the community sample in this analysis. To get as many responses as possible, the community members could choose to complete the experiment and survey online, like the students, or they could return a paper version of the protocol that had been mailed to them with a stamped return envelope and a cover letter.[4] Whether responding by mail or online, participants used an alphanumeric response code that helped to maintain a sense of anonymity while allowing us to prevent multiple responses from a single participant.

Four hundred and ten community members and one hundred and fourteen students responded. The community response rate (21%) is particularly good considering surveyors are often impressed with a 10% response. Although there is a lot of demographic variation in our community data we can compare our sample to weighted Current Population Survey (CPS) data from December 2006 to assess how representative our responses are. Both the U.S. and Vermont are composed of 49% males and 51% females. Our community responses are distributed 67% male and 33% female for the reason mentioned above. While this appears to be a problem, we got back responses in exactly the proportion in which we sent surveys out indicating that there was no male–female sample selection bias.

[4] 90% of the community participants responded to the paper version of the protocol.

Our Vermont community respondents seem to be older (the mean age is 50 compared to a mean of 38 in the state as a whole and 35 in the U.S.). One explanation for this is that our sample was restricted to respondents who were at least 18 years old. When we restrict the CPS in the same way, we find more comparable means: 46 years old in both the U.S. and Vermont. Lastly, only our community respondents were asked to report their weekly earnings (using the same questions used in the CPS). Weekly earnings are also comparable across the three groups. In the U.S. overall mean weekly earnings were $743, in Vermont they were $615 and our community sample mean is $758.[5] Based on a few obvious demographics characteristics, our community sample appears to be similar to both the state of Vermont and the U.S. more generally.

3 An overview of the data

Table 1 describes the variables that are common to both the student and the community samples. Contrary to the standard, neutrally framed dictator game, our participants were quite generous. Overall, people allocated $68.12 (or 68 percent of the maximum) to the charity of their choice, on average. However, the average community member gave almost $20 more than a student. From the second row of Table 1

Table 1 Summary statistics

	Overall			Community		Student	
	N	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Allocation	522	68.119	31.937	72.302	31.581	54.694	29.376
Allocate all $100	522	0.404	0.491	0.480	0.500	0.161	0.369
Student (indicator)	524	0.237	0.425	–	–	–	–
Altruism (factor score)	519	0.000	0.857	−0.044	−0.857	0.142	−0.844
Age	519	44.019	17.958	50.789	14.701	21.752	5.010
Male (indicator)	518	0.629	0.483	0.676	0.469	0.475	0.501
Born in USA (indicator)	520	0.923	0.267	0.945	0.229	0.851	0.357
Born in VT (indicator)	524	0.353	0.478	0.105	0.308	0.430	0.500
Some College (indicator)[a]	523	–	–	0.291	0.455	0.000	0.000
College Degree (indicator)[a]	523	–	–	0.414	0.493	0.000	0.000
Graduate Degree (indicator)[a]	523	–	–	0.060	0.238	0.000	0.000
Itemize Deductions (indicator)	511	0.474	0.500	0.501	0.501	0.377	0.487
E(Random Participants Allocation)	516	45.713	24.751	48.125	24.787	38.089	23.128
E(Mother's Allocation)	511	63.970	35.736	62.697	36.447	67.944	33.246
E(Friend's Allocation)	517	56.868	34.501	59.863	35.361	47.379	29.819

[a]Education indicators measure educational attainment for community members only. Student is coded as a separate attainment category

[5]Neither of the earnings differences (our community sample versus Vermont or the U.S.) are statistically significant.

we can see that this difference is largely driven by the fact that community members were three times more likely to give away all $100. The amount given by students in our sample (55 percent of the maximum) is nearly twice the amount donated by students in Eckel and Grossman (1996). We suspect that the difference results from our participants being permitted to choose a recipient charity. Our participants are more likely to find a charity that they care about than those in Eckel and Grossman who are restricted to The Red Cross. In addition, participants in our experiment first choose a recipient and then choose how much money to donate. The act of first selecting a charity "to receive your donation" may prompt them to be more generous.

In row four of Table 1 we report the factor score as a summary of the ten altruism and empathy statements that the participants responded to in the survey. The statements, which are listed in the appendix, were taken from the NEO Personality Inventory for altruism (Costa and mcCrae 1992) and Barchard (2004). In general, these statements have high Cronbach alpha scores indicating that the items tend to be correlated and represent some latent characteristic—altruism. In our survey, the alpha score is 0.68 which is comparable to other implementations. By construction, the mean altruism factor score for the entire pool of respondents is zero, but it is interesting that the mean score is actually higher ($|t| = 2.11$, $p = 0.03$) for students.

We collected standard demographic information including age, sex, whether or not the respondent was born in the U.S., whether or not the respondent was born in Vermont, education level and whether or not the respondent itemizes deductions for tax purposes. There is little variation in educational attainment among students; nearly all students in our sample would be classified has having had "some college." Yet it seems likely that a student who has completed some college (but plans to finish) may be quite different than a community member who has completed some college and may or may not plan to finish the degree. For this reason, indicators of educational attainment apply only to community members. One could think of *student* as being part of a mutually exclusive and collectively exhaustive list of educational attainment categories so that each individual is either a student or has completed one of the listed level of educational attainment. It is hard to get reliable income data from students because it is not clear whether one should collect the student's income or the family's income. With that in mind, we chose to ask about itemization as a crude proxy of the participant's income.

In the last three rows of Table 1 we summarize the responses of our participants to how much they thought that a random participant, their mother and their best friend would allocate to charity. For both respondent groups people attributed the most altruism to their mothers and the least to a random participant. It is interesting that, on average, students attribute less altruism to random participants ($|t| = 3.99$, $p < 0.01$) and their friends ($|t| = 3.44$, $p < 0.01$) but relatively more to their mothers ($|t| = 1.42$, $p = 0.15$). Of particular interest is the students' pessimistic view of their friends which will be proven rational in the next section.

To get accurate population level estimates of the effects of the regressors listed in Table 1, it is important that there is some common support across the two populations. To examine the overlap in support, Fig. 1 provides histograms by respondent group. The lighter histogram in each figure projecting upwards represents the student data and the dark histograms projecting downwards are for the community members. In

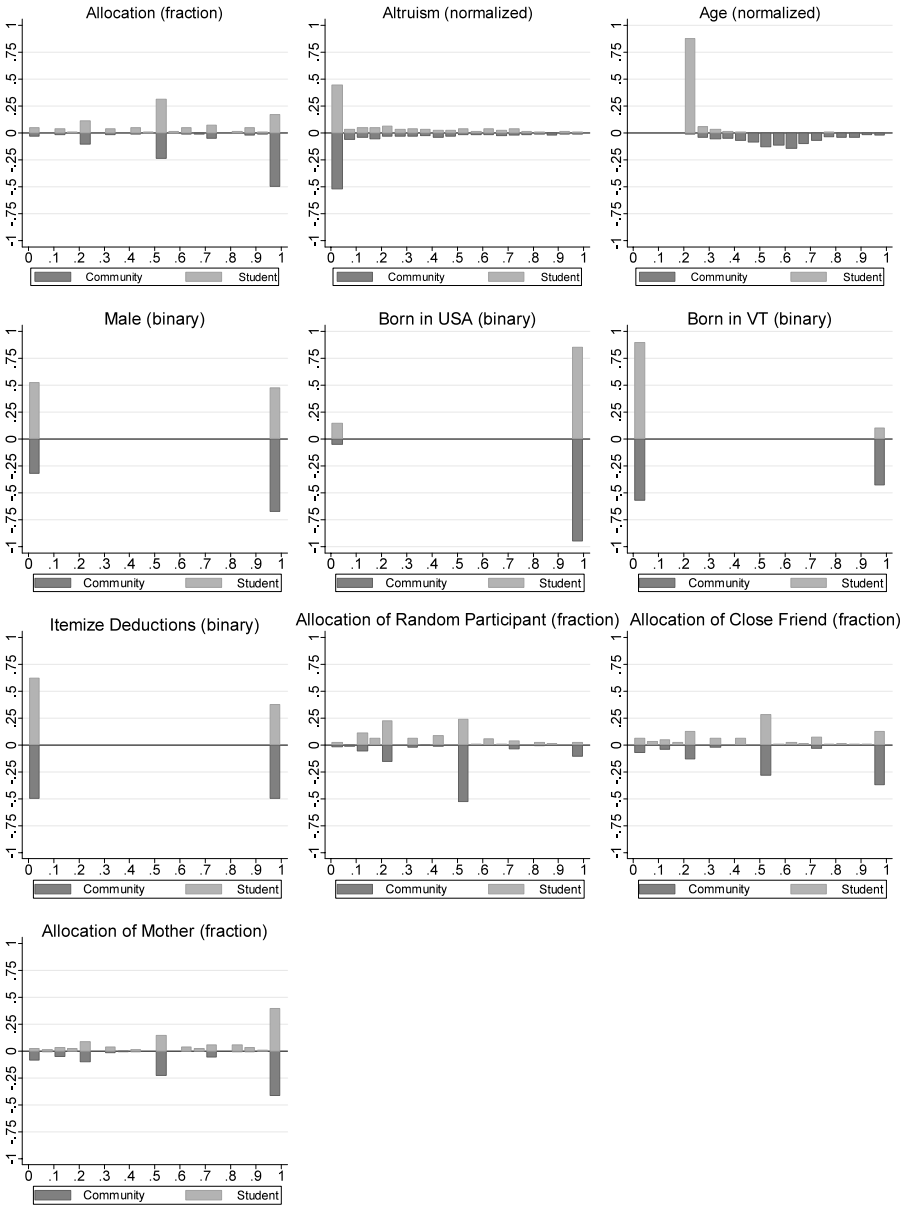

Fig. 1 Is there common support among the treatment groups?

general, there is considerable overlap in the distributions. In the upper left corner we see the two histograms of allocations to charity. The modal allocation is 50% of the money while the mode for the community members is to actually give it all away. Another thing to notice about Fig. 1 is the fact that while most of the students are

in their early 20s we were able to collect a few older students from the community implementation of the experiment.

4 Charity choice

There is a lot of variation in the choices that participants made about the charities to which money should be donated, suggesting the importance of choice. To get a better sense of these choices, Fig. 2a reports the frequencies with which each charity was chosen by participant group. Interestingly, although every charity has some support, community members chose to write in their own charities more than they chose any of the provided choices. At the same time, the students only chose to write in a charity twice. For the students the most popular charity was Doctors without Borders. These differences are both significant at the 1% level. Based on probit models of charity choice, the students were 19% less likely to write in a charity than to choose one of those listed. Student were also 21% more likely to pick Doctors without Borders than any other charity.

In Fig. 2b we report the total amount of money allocated to each of the charities. The important fact about this figure is that it looks very similar in shape to Fig. 2a. This suggests that the mean allocation does not vary that much from one charity to another. In other words there do not appear to be certain charities that were allocated more money on average than others. Indeed, the Kruskal–Wallis test indicates that the average donations (within a respondent group) come from a common population.

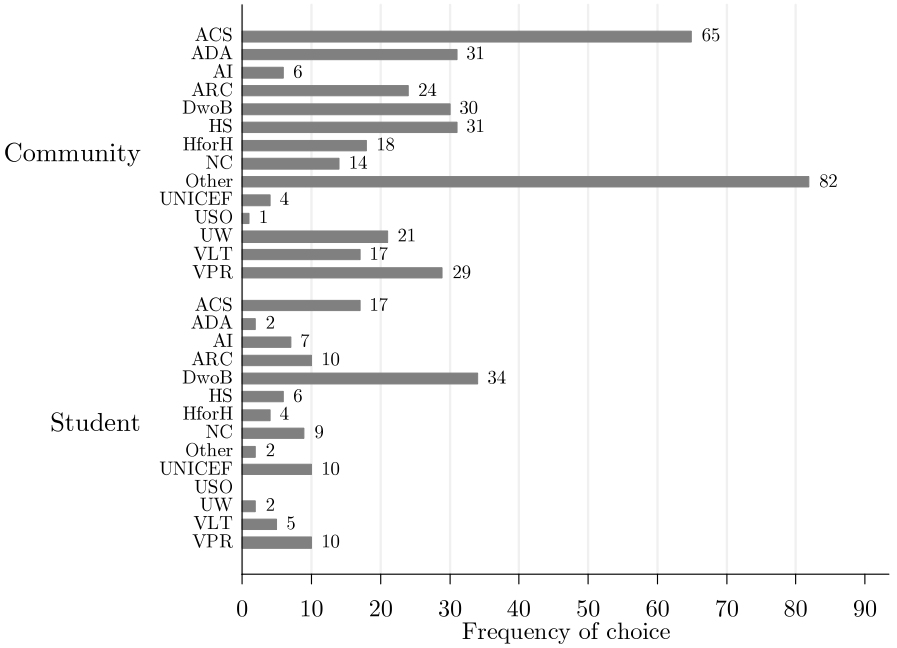

Fig. 2a Frequency each charity was chosen

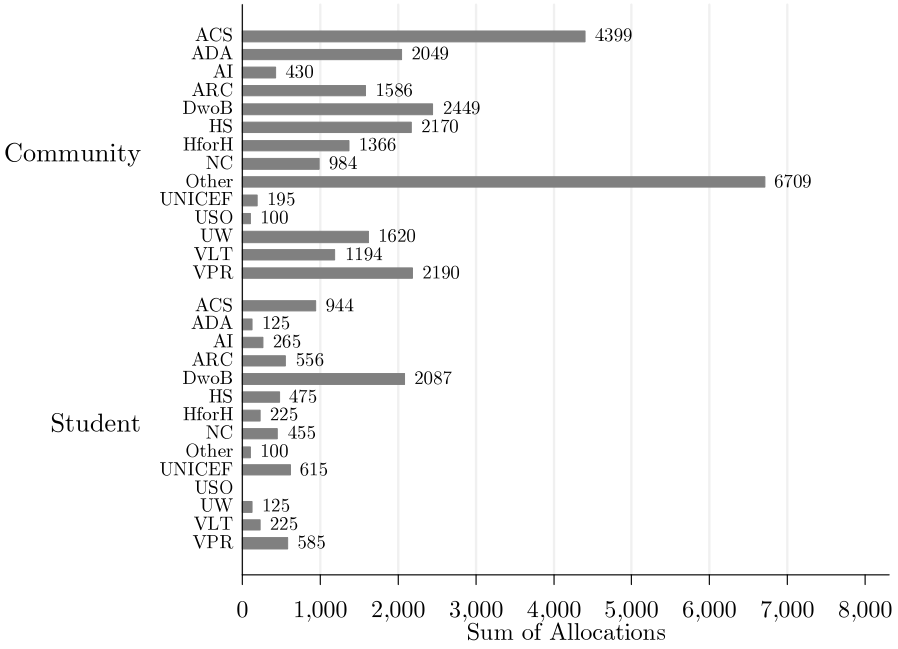

Fig. 2b Total allocated to each charity

For the community members the chi-squared statistic is 10.11 ($p = 0.61$) and for the students the chi-squared is 19.02 ($p = 0.12$). This result has implications for the next section in which we regress allocation decisions on the regressors reported in Table 1. One could imagine that we need to control for charity choice in these regressions, however adding charity fixed effect never affects the other coefficients appreciably.

5 Allocation choice

Students allocate significantly less to charity. Using a simple t-test, the difference in means ($17.61) is highly significant ($|t| = 5.51,\ p < 0.01$). This difference remains significant if we use the more conservative nonparametric Wilcoxon rank-sum test ($z = 5.62,\ p < 0.01$). If we focus on the fraction of people giving away all of the money, the proportions test confirms that students are much less likely to allocate all the money to charity ($z = 6.31,\ p < 0.01$).

In Table 2 we try to account for this large difference in allocations. Because responses are truncated at $0 and $100, we report coefficients from a Tobit model of giving. Accounting for truncation increases the point estimate of the difference in student and community member underlying propensity to give from $17 to $31 in column (1). In column (2) we test the construct validity of the representative dictator game. One's altruism factor score is positively associated with how much one gives in the representative dictator game ($p < 0.05$) suggesting that the game does measure

Table 2 Tobit models of the determinants of allocations

	(1)	(2)	(3)	(4)	(5)
Student (indicator)	−31.236***	−32.043***	−10.823	−12.848*	8.571
	(5.40)	(5.37)	(8.75)	(6.96)	(23.12)
Altruism (factor score)		5.443**	5.493*	3.722*	3.593
		(2.74)	(2.82)	(2.25)	(2.51)
Age			0.608***	0.527***	0.535***
			(0.18)	(0.15)	(0.15)
Male (indicator)			−0.02	−2.085	2.774
			(4.91)	(3.94)	(4.68)
Born in US A (indicator)			1.241	−4.969	−5.246
			(8.64)	(6.77)	(9.17)
Born in VT (indicator)			−3.016	−5.509	−6.103
			(5.34)	(4.26)	(4.56)
Some College (indicator)			2.875	−0.5	−0.435
			(7.15)	(5.75)	(5.74)
College Degree (indicator)			11.944*	1.358	0.719
			(7.05)	(5.67)	(5.71)
Graduate Degree (indicator)			23.984*	0.404	−0.912
			(13.25)	(10.85)	(10.89)
Itemize Deductions (indicator)			13.307***	3.999	4.677
			(4.79)	(3.85)	(4.52)
E(Random Participants Allocation)				0.332***	0.338***
				(0.09)	(0.10)
E(Mother's Allocation)				0.298***	0.326***
				(0.06)	(0.06)
E(Friend's Allocation)				0.556***	0.522***
				(0.06)	(0.07)
Student * Altruism					−0.47
					(5.83)
Student * Age					−0.634
					(0.75)
Student * Male					−16.161*
					(8.53)
Student * USA					4.078
					(14.12)
Student * VT					7.761
					(13.31)
Student * Itemize					−2.595
					(8.86)
Student * E(Random)					−0.058
					(0.23)
Student * E(Mother)					−0.106
					(0.13)

Table 2 (*Continued*)

	(1)	(2)	(3)	(4)	(5)
Student * E(Friend)					0.146
					(0.17)
Intercept	89.238***	89.180***	44.412***	0.253	−3.173
	(2.91)	(2.89)	(13.16)	(10.96)	(12.99)
Pseudo R-squared	0.009	0.011	0.024	0.086	0.088
N	522	517	501	488	488

Note: Double censored Tobit: standard errors in parentheses: *indicates significant 10%, **5%, ***1%

altruism. In column (3) we add the demographics and see that doing so cuts the difference between students and community members by two-thirds. Our income proxy, *Itemize Deductions*, indicates that people with higher incomes allocate more to charity, as one might expect. Controlling for income, there seems to be a separate effect of education. Keeping in mind that *student* is part of a list of educational attainment indicators, we see that students give $11 less than those with a high school degree or less (the omitted education category), although the estimate is not statistically significant. Students don't look very different than those with some college. The more highly educated members of the sample also contribute more. The most robust association, however, appears to be between *Age* and allocations. Older people behave significantly ($p < 0.01$) more generously.[6]

In column (4) we add the expectation responses and find that all three are highly positively correlated with allocations. Projection appears to motivate responses to the *E(Random Participant's Allocation)* question because the coefficient is positive and highly significant. Controlling for projection, however, we see that there is a positive relationship between the allocations of mothers and friends and that of the participant. We expect that the coefficient on *E(Mother's allocation)* likely reflects the transmission (via some combination of nature and nurture) of altruistic attitudes from mother to child. However, as the adage goes, one can pick one's friends, but not one's family. For this reason, the coefficient on *E(Friend's allocation)* reflects a correlation between the allocations of friends that results partly from transmission of attitudes and partly from the tendency to seek out friends with similar attitudes. For this reason, it is not surprising that the relationship between the altruism of friends is stronger that of mothers.

In column (5) we allow the coefficients on each of our regressors to vary by respondent population. Age remains significant and there does not seem to be any interaction between age and student status. Further, the coefficients on the expectations questions remain the same indicating there is not much of a differential effect of expectations in the student population; however, there is one differential effect that is very conspicuous: compared to their counterparts in the community, male students

[6] Allowing age to enter quadratically adds nothing to the analysis. The quadratic term is never significant.

are dramatically less altruistic.[7] Given the coefficient on the student indicator is no longer significant, our survey has allowed us to explain exactly why students are not representative. Students appear less altruistic because they are young relative to the other members of the population and because male students are particularly ungenerous.

We augment the analysis of allocation choice in Table 3 by focusing on those respondents who chose to give away all $100. In the first four columns Table 3 reports the marginal effects from probit regressions of whether or not one gave away all the money on the same regressors as in Table 2.[8] Many of the same factors motivate giving it all away. As in Table 2, there is a large difference between students and community members. Students are 32% less likely to give away all the money and this difference drops only slightly with the addition of the altruism factor score, the demographics and the expectations. Although age seems to matter again, it is only when we add the interaction of *Male* and *Student* in column (5) that we explain most of the difference between students and community members—male students are differentially less likely to give away all the money.

There are a few other things to notice about Table 3. While the altruism factor score does predict allocation choices is does not predict giving away all the money. One's expectations of how much others will allocate to charity predict both the amount one donates and whether or not one donates all $100. As before the influence of one's best friend appears to be more salient than that of one's mother, although both matter. And finally, as in Table 2, the F-test of the joint significance of the coefficients on the student interactions terms (minus *Student * Male*) indicates the effects of the (other) regressors does not change much from one respondent group to the other ($F = 0.98$, $p = 0.45$).

We perform several (unreported) robustness checks aimed at gauging whether students and community members systematically differ in their levels of attachment to charitable organizations and whether this could have affected selection into the experiment or the amount donated. As mentioned previously, adding charity fixed-effects does not alter the essential results. It may also be the case that community members have stronger ties to local charities in particular, so we ran the analysis after omitting individuals who wrote in a charity selection. Standard errors increase, but the point estimates are largely unchanged. Finally, community members who chose to respond to the survey may have been motivated by ties to a listed organization to which they volunteer. Students, in contrast, had selected into the potential subject pool prior to knowing any details of this particular experiment. We suspect that this difference in recruitment is unlikely to cause selection bias because students, like community members, were still free to choose whether or not to participate. As a rough test, we use a measure of volunteer status, although we don't include it in the tables because the measure is somewhat problematic. The question about volunteering changed between the two surveys; students were asked if they had volunteered in the previous

[7] Indeed, the joint F-test of whether the interaction terms (excluding the Student * Male regressor) are different from zero suggests that they are not ($F = 0.28$, $p = 0.97$).

[8] Because Ai and Norton (2003) show that the marginal effect of the interaction terms in nonlinear models is not the interaction effect, we present the results of a linear probability model in column (5) of Table 3. Standard errors are robust.

Table 3 Probit models of the determinants nf allocating all $100

	(1)	(2)	(3)	(4)	(5)
Student (indicator)	−0.319***	−0.320***	−0.214***	−0.289***	0.189
	(0.04)	(0.04)	(0.08)	(0.07)	(0.17)
Altruism (factor score)		0.019	0.021	0.018	0.011
		(0.03)	(0.03)	(0.03)	(0.03)
Age			0.004**	0.004**	0.003*
			(0.00)	(0.00)	(0.00)
Male (indicator)			0.007	0.008	0.048
			(0.05)	(0.05)	(0.05)
Born in USA (indicator)			0.028	−0.016	0.025
			(0.09)	(0.11)	(0.11)
Born in VT (indicator)			−0.036	−0.076	−0.057
			(0.05)	(0.06)	(0.05)
Some College (indicator)			0.025	0.014	0.02
			(0.07)	(0.08)	(0.07)
College Degree (indicator)			0.117*	0.025	0.037
			(0.07)	(0.08)	(0.07)
Graduate Degree (indicator)			0.193	−0.024	−0.003
			(0.12)	(0.12)	(0.10)
Itemize Deductions (indicator)			0.125***	0.067	0.054
			(0.05)	(0.05)	(0.05)
E(Random Participants Allocation)				0.003***	0.002**
				(0.00)	(0.00)
E(Mother's Allocation)				0.003***	0.002***
				(0.00)	(0.00)
E(Friend's Allocation)				0.006***	0.005***
				(0.00)	(0.00)
Student * Altruism					−0.022
					(0.04)
Student * Age					−0.004
					(0.00)
Student * Male					−0.169**
					(0.07)
Student * USA					−0.114
					(0.13)
Student * VT					0
					(0.10)
Student * Itemize					−0.016
					(0.08)
Student * E(Random)					0.001
					(0.00)
Student * E(Mother)					−0.002
					(0.00)

Table 3 (*Continued*)

	(1)	(2)	(3)	(4)	(5)
Student * E(Friend)					−0.001
					(0.00)
Pseudo R-squared	0.061	0.062	0.12	0.312	0.331
N	522	517	501	488	488

Note: Marginal effect from Probits reported in columns (1)–(4); in (5) we report a linear probability model for ease of calculating marginal effect of interactions; standard errors in parentheses; *indicates significant at 10%, **5%, ***1%

week, while community members were asked if they had volunteered in the previous month. Using this (inconsistent) measure, approximately half of both the community members and students report regular volunteer activity. We run the analysis from column (5) of Tables 2 and 3 with the addition of an indicator of volunteer status as well as an interaction with student status. Again, the results do not change appreciably.

6 Discussion

We conduct a representative dictator game in which both students and random members of the community choose both what charity to support and how much to donate to the charity. We find systematic differences between the choices of students and community members. Community members are much more likely to write in their own charity, community members donate significantly more ($17), on average, and community members are much more likely (32%) to donate the entire $100 endowment. Based on this evidence, it does not appear that student behavior is very representative in the context of the charitable donations and the dictator game.

The survey that accompanied our experiment allows us to offer specific reasons why students are not representative. Among the determinants of allocation choices, the robust factors appear to be age and sex. Over all, students give less because young people are less generous and because there is a differential effect of being a male student. Male students are particularly selfish. This result echoes findings in a variety of similar contexts. For example, List (2004) finds that younger people tend to be less cooperative in a field public goods experiment, men tend to be less likely to give and give less in a fundraiser and men (and young people) tend to be less cooperative in a television game show version of the prisoner's dilemma. In their analysis of the behavior of a random draw of Dutch society in a trust experiment, Bellemare and Kroger (2005) show that females and older people are significantly more trusting and that students, females and people over the ages of 45 are significantly more reciprocal. Finally, in a bargaining context, Gueth et al. (2007) find that older participants and women are more averse to inequality indicating that their student data is also not very representative.

In the future, economists should worry more about the *overall* representativeness of their designs. To this point some researchers have begun to worry about whether

or not a convenience sample of students behaves like other specific populations (e.g. Burns 1985; Cardenas 2003 or Burks et al. 2005), and even in a couple of cases, whether student behavior is representative of a random sample of the local population (e.g., Fehr et al. 2003 or Dohmen et al. 2005). However, few experiments also worry about the representativeness of the underlying task. As we have seen making the decision more representative of the sort of decisions people normally make about philanthropy has generated very different behavior that what is typically seen in the context-free lab.

Acknowledgements We thank the NSF (SES-Career 0092953) for funding and John List and two thoughtful referees for comments.

References

Ai, C., & Norton, E. (2003). Interaction terms in logit and probit models. *Economics Letters*, *80*, 123–129.

Barchard, K. (2004). The nature and measurement of emotional intelligence abilities: basic dimensions and their relationships with other cognitive ability and personality variables. *Educational and Psychological Measurement*, *64*, 437–462.

Bardsley, N. (2008). Dictator game giving: altruism or artefact? *Experimental Economics*. doi:10.1007/s10683-007-9172-2.

Bellemare, C., & Kroger, S. (2005). *On representative social capital*. CIRPEE Working Paper 05-04.

Benz, M., & Meier, S. (2005). *Do people behave in experiments as in real life? Evidence from donations*. University of Zurich, Institute for Empirical Research in Economics Working Paper No. 248.

Botelho, A., Harrison, G., Hirsch, M., & Rutstrom, E. (2005). Bargaining behavior, demographics and nationality: what can the experimental evidence show? In J. Carpenter, G. Harrison, & J. List (Eds.), *Field experiments in economics*. Greenwich: JAI Press.

Brunswik, E. (1956). *Perception and the representative design of experiments*. Berkeley: University of California Press.

Burks, S., Carpenter, J., & Goette, L. (2005). *Performance pay and the erosion of worker cooperation: field experimental evidence*. Middlebury College Department of Economics Working Paper.

Burns, P. (1985). Experience and decision making: a comparison of students and businessmen in a simulated progressive auction. In V. Smith (Ed.), *Research in experimental economics*. Greenwich: JAI Press.

Camerer, C. (2003). *Behavioral game theory: experiments on strategic interaction*. Princeton: Princeton University Press.

Camerer, C., & Thaler, R. (1995). Anomalies: ultimatums, dictators and manners. *Journal of Economic Perspectives*, *9*, 209–219.

Cardenas, J. C. (2003). *Bringing the lab to the field: more than changing subjects*. Javeriana University Department of Economics Working Paper.

Carpenter, J., Burks, S., & Verhoogen, E. (2005a). Comparing students to workers: the effects of social framing on behavior in distribution games. In J. Carpenter, G. Harrison, & J. List (Eds.), *Field experiments in economics. Research in experimental economics*. Greenwich: JAI Press.

Carpenter, J., Verhoogen, E., & Burks, S. (2005b). The effect of stakes in distribution experiments. *Economics Letters*, *86*, 393–398.

Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa: Psychological Assessment Resources.

Dohmen, T., Falk, A., Huffman, D., Wagner, G., Sunde, U., & Schupp, J. (2005). *Individual risk attitudes: new evidence from a large, representative, experimentally-validated survey*. IZA discussion paper no. 1730.

Eckel, C., & Grossman, P. (1996). Altruism in anonymous dictator games. *Games and Economic Behavior*, *16*, 181–191.

Fehr, E., Fischbacher, U., Rosenbladt, B., Schupp, J., & Wagner, G. (2003). *A nation-wide laboratory examining trust and trustworthiness by integrating behavioral experiments into representative surveys*. University of Zurich, Institute for Empirical Research in Economics Working Paper #141.

Fong, C. (2005). *Empathic responsiveness: evidence from a randomized experiment on giving to welfare recipients*. Carnegie Mellon Department of Social and Decision Sciences Working Paper.

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

Gueth, W., Schmidt, C., & Sutter, M. (2007). Bargaining outside the lab: a newspaper experiment of a three-person ultimatum game. *The Economic Journal*, *117*, 449–469.

Harrison, G., & List, J. (2004). Field experiments. *Journal of Economic Literature*, *42*, 1009–1055.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., et al. (2006). Costly punishment across human societies. *Science*, *312*, 1767–1770.

Hogarth, R. (2005). The challenge of representative design in psychology and economics. *Journal of Economic Methodology*, *12*, 253–263.

Levitt, S., & List, J. (2007). What do laboratory experiments measuring social preferences reveal about the real world?. *Journal of Economic Perspectives*, *21*, 153–174.

List, J. (2004). Young, selfish and male: field evidence of social preferences. *The Economic Journal*, *114*, 121–149.

List, J. (2006). Field experiments: a bridge between lab and naturally occurring data. *Advances in Economic Analysis and Policy*, *6*, Article 8.

List, J. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy*, *115*, 482–493.