THEY COME TO PLAY

SUPPLY EFFECTS IN AN ECONOMIC EXPERIMENT

Jeffrey Carpenter, Allison Liati, and Brian Vickery

ABSTRACT

Our experiment challenges the standard, social preference, interpretation of choices in the double blind dictator game played in the lab without any context. We present treatments formulated to minimize the social preference reasons to give and, despite this, the allocations are identical to our replication of the standard double blind game, implying that altruism might be the wrong interpretation of giving. Instead, we hypothesize that giving might be driven by participants coming to the lab ready 'to play'. The fact that there are strong correlations between participant responses to an attention deficit, hyperactivity disorder questionnaire and both the rate and level of giving provides direct support for the hypothesis that lab participants impulsively give money away. However, we also show that having players earn their endowments attenuates the bias.

KEY WORDS • demand effect • dictator game • experiment • impulsivity • social preference

Introduction

Psychologists have worried about 'demand effects' since the Hawthorne experiments in the 1920s and 1930s in which the investigators realized that much of the change in productivity of the young women at a Western Electric assembly plant was related to the simple fact that they were participating in an experiment (Roethlisberger and Dickson 1939). Since then demand effects in psychology have been associated with people under examination seeing themselves as research assistants who (mostly) want to help the experimenters. These junior research assistants try to figure out what the experiment is about and act in ways that confirm

 Vol 22(1): 83–102. http://rss.sagepub.com DOI: 10.1177/1043463109358486

hypotheses. In the Hawthorne study, it was obvious that the researchers were looking for the determinants of productivity and the participants did what they could to help even if they did not always guess correctly on the sign of the comparative static. In one illumination study, for example, the investigators recorded stable productivity even as the amount of light on the shop floor was reduced to 'an amount equal to that on an ordinary moonlit night' (Roethlisberger and Dickson 1939: 17).

Economists believe that they have attenuated this sort of demand effect by paying people in experiments based on the actions that they take. At a minimum, pay for performance in economic experiments increases the cost to participants for acting contrary to their material incentives. However, economists worry about a slightly different demand effect that may still bias behavior even after we make it costly to behave as junior research assistants. Economists worry that participants view their experience in the lab as a potentially repeated game and therefore may perceive that the probability of playing the next stage (and making more money) depends on how one plays in the current stage (Hoffman et al. 1996; to be written as HMS hereafter). If you are a 'nice' player today, you will be more likely to be invited back to play again. This sort of demand effect explains why economists began running 'double blind' experiments in which players knew that they could never be identified and so there was no point in investing in a reputation as a nice player.

Economists now appear confident that they have eliminated demand effects by paying for performance and by using double blind procedures. However, we have uncovered another bias that is particularly insidious because it is more likely to be present in experiments that, on first reading, seem clever and clean. An example using one of the most common economic experiments, the dictator game (DG), makes the point. In the typical DG one player, the dictator, is given an amount of money and told that she can send as much as she wants to another anonymous participant, the recipient, and keep the rest. Clearly there is no incentive for a selfish person to send anything, so giving money is often interpreted as a measure of social preference – specifically altruism – although giving has also been interpreted as demonstrating inequality aversion. These interpretations may have been premature because early versions of the DG were often tainted by demand effects. For example, the DGs reported on in Forsythe et al. (1994) were conducted single blind and in one treatment the payoffs were purposely hypothetical to test for the psychological notion of a demand effect. Indeed, both biases seem to be present because Forsythe et al. found significantly

more fairness in the hypothetical DG and HMS showed that offers are considerably smaller in the double blind DG. While offers fall when the stakes are real and the experiment is conducted double blind, approximately one-third of the people in the HMS experiment still allocated some money to the recipient. The question is whether the behavior of this residual third can be attributed to altruism or if something else is at work.

Put yourself in the role of the dictator in a double blind DG played for real stakes. You understand the instructions (it doesn't get much more straightforward), you understand that there is a one-for-one tradeoff between sending and keeping money, and you understand (and hopefully believe) that there is no way for the experimenter to identify what you have done. So why would you send any money? Maybe you do feel altruistic towards your partner, but why? You are both very likely to be college students which means you are both also likely to be relatively affluent and even if you did want to donate money to someone who might really need it, you can't do so for sure because your donation will be randomly assigned to someone in the other room. For that matter there is a good chance that you will end up 'donating' to someone who already has more than you. Alternatively, you might be inequality averse (à la Fehr and Schmidt 1999) and therefore you may want to rectify the fact that you were given money and the recipient was not. However, if you are that concerned with inequality, why would you not consider the inequality in personal wealth among all the participants in the session? If you use cues to assess where you are in the participant income distribution, you could easily find yourself in a situation in which you should give away all the money, but this almost never happens.

So why then do you send money? Our research suggests that people may send money in the DG just because they came to the experiment intent on 'playing'. That is, because they invested the time responding to recruiters and getting to the lab, they want to do something and find it very hard to do the equivalent of nothing, in this case to keep all the money. In broader terms, the choices may be made impulsively. Indeed, our data suggests that after employing the standard procedures to eliminate demand effects, the remaining third of our participants may allocate money to the recipient because they come to play not because they are altruistic.

How can we separate social preference reasons for giving from 'coming to play' reasons? After replicating the results of HMS, we conducted a second treatment which we call the *two-way dictator game* (TWDG) in which both players are given a $20 endowment and again asked to give randomly to someone in the other group. As in the

baseline, participants are randomly assigned to two separate rooms but, unlike the baseline, they are now *both* given envelopes with $20 inside. Simultaneously, participants in the two rooms decide how much to pass to a random person in the other room. Although each participant makes his or her choice privately, information is complete in the sense that players in both rooms know that everyone in the session will both make an allocation and randomly receive one. Think about giving in this game. Before, the myopic dictator might have focused on game wealth (and disregarded possible differences in overall wealth) and felt altruistic towards the recipient or may have been bothered by the inequality of the endowment. But, in the TWDG, inequality aversion suggests that you do not give because then you will just increase your chances of being worse off than the recipient, especially given the fact that randomization implies that there is a very small chance that you will give to the same person who can give to you.

The most plausible specifications of altruism, that posit some effect of deservingness (e.g. Fong 2001 or Fong 2005), also predict that no one gives when the initial endowment is perfectly egalitarian. Of course it may be impossible to eliminate all possible formulations of altruism in the DG. For example, if altruism is captured in the participant specific parameter α_i, then 'Mother Theresa' types with $u_i(x_i, x_j) = x_i + \alpha_i x_j$ and $\alpha_i > 1$ would still give as would warm glow types who derive personal satisfaction from giving with $u_i(x_i, \alpha_i)$. However, we can be reasonably confident that there are few Mother Theresas in the population because these people would give it all away, something that almost nobody does; so the real problem is the possibility of warm glow.

As it turns out, there is little empirical support for warm glow altruism. Bolton and Katok (1998) run two DGs. In the first condition the initial allocation (dictator, recipient) is (15, 5) and in the second it is (18, 2). Konow (2005) runs the standard (10, 0) and (10, 4). If dictators are motivated by warm glow (i.e. they enjoy the act of giving for its own sake), then they would give the same amount regardless of the recipient's endowment; however, they don't. In both experiments, the modal positive gifts are such that the final payoffs between the dictator and recipient are equalized. These results not only suggest that warm glow is not common, they also point to inequality aversion as being the dominant concern in the DG. Our TWDG therefore controls for all the standard deviations in the DG: demand effects should be minimized by using our double blind procedure and inequality aversion is mitigated when both players start with equal endowments. As a result, we conjecture that what is left is that people send money because they 'come to play'.

Although endowing both parties with $20 should allay most concerns about fairness motives in the TWDG, some may still worry about the fact that both parties make dictator choices. As unlikely as it is, players in the two groups might be matched in which case an optimistic and inequality averse dictator might send money because she expects to get some. Even if she does not expect to be matched with someone in the other group, there are complicated linkages that might motivate her to send money randomly if she thinks it will reduce the final inequality across the groups. To mitigate any remaining reciprocity and fairness concerns, we also ran a *one-way dictator game* (OWDG). The OWDG is identical to the TWDG in that both groups of participants are endowed with $20 but it is different because only one of the two groups can send money. Now dictators can be sure that they will receive no money from the other group.

One striking feature of our data is that when we compare the distribution of allocations in our treatments to our replication of HMS, we find no difference. Neither the central tendency of the distributions nor the incidence of giving is different. This suggests that the context-free double blind DG played between college students in the lab may not be much of a measure of social preference.[1] However, it might be a pretty good measure of impulsivity. To directly test this alternative, in our post-experiment survey we had our participants respond to a standard attention deficit, hyperactivity disorder (ADHD) questionnaire. Using factor analysis to summarize the 45-question scale, our measure of ADHD predicts both sending money and the amount sent at better than the 1% level of confidence. This result was surprising but it provides further evidence that people come to the lab ready to play and it is very hard for some of them to keep a grip on their impulses.

All is not lost, however. We conclude by showing that a less common experimental manipulation cuts the incidence of impulsivity in half. In a fourth treatment we have our participants stuff envelopes to earn their endowments. Similar to Cherry et al. (2002), we find that the rate of sending money falls to less than 20% and only one person sends more than a dollar when dictators earn their endowments.

The remainder of the paper provides the details of our research. In the next section we describe our experimental procedures. In the third section we discuss the distribution of behavior in our treatments compared to the data from HMS. In the fourth section we show that our results are robust to the inclusion of a number of demographic factors that could affect behavior, including our measure of ADHD. In the final section we discuss the implications of our research.

Experimental Design

Altogether, we conducted four double blind DGs with a total of 210 Middlebury College students. In each case participants were randomly sorted into two groups who made decisions simultaneously and independently. To connect our research to the literature, we began by replicating the double blind conditions of HMS. In this treatment, which we call our *replication*, we used instructions based on Forsythe et al. (1994) but updated the stakes to $20.[2] There were 46 participants (23 observations) in this treatment. In our second experiment, the TWDG, we tried to minimize the social preference reasons for giving money by allocating players in both groups a $20 endowment and by allowing players in both groups to pass money. This design is similar to Ben-Ner and Putterman (2004) with a few important differences: allocation decisions happened simultaneously, not sequentially and there was essentially no possibility of reciprocity because each participant's allocation was randomly given to a person in the other group. There were 48 participants and observations in this treatment.

To address any remaining social preference implications of allowing both groups to send money in the TWDG, we also conducted the OWDG in which only one of the two groups sent money although both groups started with $20 endowments. This is similar to treatment B in Cox (2004); however, Cox was interested in behavior in the trust game so any transfers were tripled by the experimenter. Obviously, the tripling provides an incentive to give for participants motivated by a preference for social efficiency (as in Engelmann and Strobel 2004) so our OWDG employs a standard non-subsidized transfer. We recruited 76 people to participate in the OWDG resulting in another 38 observations.

In the last treatment, *earned endowment*, 40 participants performed a real effort task to earn their endowments before they took part in the TWDG.[3] The production task was to fold, staple, and address flyers that would be sent via campus mail to other Middlebury students. To keep interactions among the participants to a minimum, production was done individually and in silence. To prevent income effects, each participant received a flat rate of $20 for half an hour of work. The production task was completely separate from the TWDG in that no mention of the dictator experiment was made until after the production task was finished. Further, the production task began after randomly splitting the participants into two groups to minimize any interactions among the potential dictators and recipients.

The protocol of the TWDG is representative. The instructions were handed out and read aloud while all the participants were in the same room. To answer the few questions that arose, the experimenters simply reiterated the appropriate passage from the text. Participants were then randomly sorted into two groups and the groups were taken to different large classrooms. One at a time, participants were given an opaque envelope containing $20 in cash and a slip with a participant number. They were then told to go to an adjoining room where they were to privately remove any money that they did not want to pass to someone in the other room along with the slip containing their participant number. To assure anonymity, they were told not to show their participant number to anyone, including the monitors. On the way back into the room, they deposited their envelopes into a large box. Before they returned to their seats, they were given a survey to fill out as the other participants made their choices. The last part of the survey was to fill out an ADHD scale borrowed from Young (2004).[4] When everyone was finished, the box containing the allocations was brought into the hall where the transfers were recorded (the participant numbers were written on the inside flap of the envelopes). After everyone was done with the survey, the boxes were brought to the opposite room and the participants were randomly assigned envelopes as they left. Unless a player revealed his or her participant number, there would be no way to match individuals with their behavior. In this sense each treatment was double blind anonymous.

Our hypothesis is that some allocations in the DG are made impulsively. We included the ADHD questionnaire because previous research suggests that being diagnosed with ADHD correlates with making impulsive choices (Malloy-Diniz et al. 2007). The full Young (2004) ADHD scale, which also has been shown to correlate with impulsivity (Young and Gudjonsson, 2005), is composed of 112 questions scored on a 5-point Likert scale varying from ‘never’ to ‘most of the time.’ We implemented the 45-question abridged version used to diagnose ADHD. This scale includes 25 items that relate to inattention, 11 items that relate to impulsivity, and 9 items relating to hyperactivity. As in Young (2004), we summarize the responses of our participants by calculating the factor loadings on the 45 items. In our implementation, the items that were assigned the highest scores included: Q3 Have you been inattentive, easily distracted, Q5 Have you had difficulty sustaining attention, Q10 Have you been reluctant to do tasks that require sustained mental effort, Q17 Have you had difficulty with quiet leisure activities, and Q24 Have you had difficulty concentrating? These five items also received relatively large scores in Young’s implementation.

Table 1. Summary statistics (by treatment)

	Overall	*Replication*	*Two-way*	*One-way*	*Earned endowment*
Amount sent	0.98	1.26	1.00	1.57	0.22
	(2.11)	(2.53)	(2.01)	(2.73)	(0.58)
Send?	0.32	0.35	0.37	0.39	0.17
	(0.47)	(0.49)	(0.49)	(0.49)	(0.38)
Grade point average	3.43	3.51	3.37	3.44	3.46
(GPA)	(0.29)	(0.20)	(0.34)	(0.29)	(0.26)
Combined SAT score	1389.41	1409.13	1355.84	1413.38	1395.60
	(87.99)	(91.70)	(85.89)	(74.15)	(91.21)
Female	0.39	0.48	0.40	0.42	0.33
	(0.49)	(0.51)	(0.49)	(0.50)	(0.47)
Number of siblings	1.33	1.17	1.37	1.26	1.42
	(1.02)	(0.83)	(1.02)	(0.98)	(1.17)
First born	0.60	0.61	0.50	0.68	0.62
	(0.49)	(0.50)	(0.50)	(0.47)	(0.49)
Age	20.57	20.30	20.52	20.50	20.85
	(1.16)	(1.02)	(1.20)	(1.22)	(1.10)
Economics major	0.57	0.48	0.58	0.60	0.58
	(0.50)	(0.51)	(0.50)	(0.49)	(0.50)
Work part time	0.55	0.56	0.58	0.55	0.50
	(0.50)	(0.51)	(0.50)	(0.50)	(0.51)
ADHD	0.00	0.08	0.07	–0.03	–0.10
	(0.97)	(0.96)	(0.94)	(0.90)	(1.08)

Note: mean, (standard deviation).

Data Overview

The data from our experiment is summarized in Table 1. Overall, the mean allocation was approximately a dollar (98 cents to be exact). However, the allocations ranged from nothing to $10 which was half the initial endowment. As another measure of behavior, 32% of our participants sent at least one dollar. The remaining rows in Table 1 provide summary statistics from our survey. The average participant was 20.57 years old, had a grade point average of 3.43 and 1.33 siblings. 39% of our participants were female, 60% were first born, 57% were economics majors, and 55% worked part time during the school year. As for the ADHD scale, we report the orthogonal varimax rotation resulting from a factor analysis of the 45 questions.[5]

We can also get a sense of any treatment differences from Table 1. Most of the demographics are very similar across treatments. In fact, only the treatment differences in combined SAT scores (specifically,

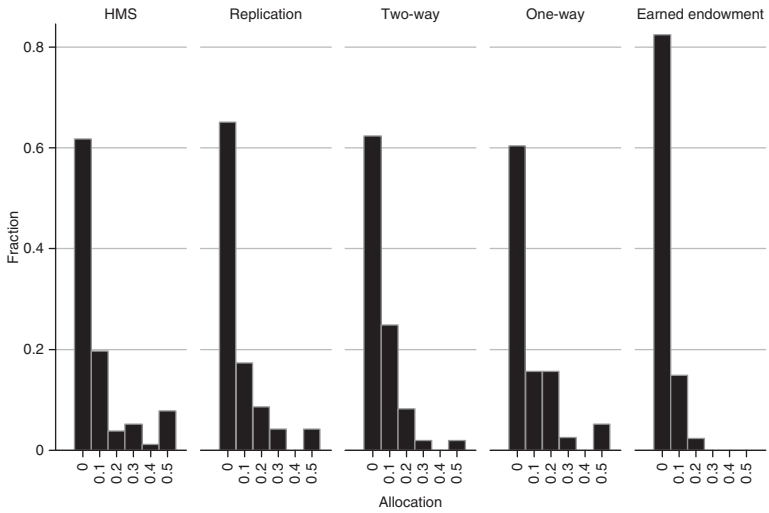

Figure 1. Dictator allocations by treatment

that the scores are lower in the TWDG) are significant at the 5% level. To help illustrate the behavioral differences in our experiment, Figure 1 presents histograms of the allocations by treatment. We report two statistics to be comprehensive in our analysis of any differences in the distribution of allocations. We report Wilcoxon rank sum statistics (z) to examine differences in the central tendency of the allocations and Fisher exact chi-squared statistics (χ^2) to test for differences in the incidences of allocating a positive amount. There are actually five histograms in Figure 1 because we also reproduce the HMS double blind data in the first cell on the left to demonstrate the extent to which we were able to replicate their results. In fact, our replication seen in the second cell from the left was nearly perfect ($z = -0.56$, $p = 0.57$; $\chi^2 = 0.08$, $p = 0.81$).

Interestingly, there also does not appear to be any significant difference between our double blind replication of HMS and our double blind TWDG ($z = 0.02$, $p = 0.98$; $\chi^2 = 0.05$, $p > 0.99$) or the OWDG ($z = 0.44$, $p = 0.66$; $\chi^2 = 0.13$, $p = 0.71$).[6] The implications of this result are considerable. Given we have minimized the social preference reasons for giving money in these two treatments, it appears that 'coming to play' may be a good explanation of allocations in *both* our treatments and the regular DG. For that matter, because we have replicated HMS so well, the allocations in their experiment might also have been chosen impulsively. Based on our results, one can make the case that once we have

removed any salient context (e.g. framing or the differential deservingness of the recipient) and controlled for both psychological and economic demand effects in the lab, what is left over might actually be the impulsive need to act when supplied a stimulus, instead of altruism.

The good news, however, is that having players earn their endowments is a viable way to remove much of the impulsive giving from the experiment. Comparing the last two cells of Figure 1, we see that having participants earn their endowments reduces the incidence of giving in the TWDG to 17% (from 38%) and removes all but one allocation larger than a dollar. Based on our two tests, earning one's endowment significantly reduces the giving seen in the other treatments ($z = -2.27$, $p = 0.02$; $\chi^2 = 6.47$, $p = 0.03$).

Data Analysis

Another way to make our analysis comprehensive is to control for a collection of other demographic factors that might affect allocation decisions and to examine both the level of giving and the incidence of giving. For example, Eckel and Grossman (1998) find that women in double blind DGs give away approximately twice as much money as men. With our lengthy exit survey we can account for the effects of dictator sex and all the other variables listed in Table 1, including our ADHD measure in which we have a particular interest because it allows us to directly link 'coming to play' with impulsivity.

Because allocations are bounded on both ends we use the Tobit regressor to test the robustness of our treatment differences in columns (1) to (3) of Table 2. Starting with column (1) we see that the differences between the treatments and our replication (the omitted treatment) and between the treatments and our earned endowment treatment are robust to the inclusion of the demographics, even though some of the demographics matter. The coefficients on the TWDG and the OWDG are small and insignificant confirming that there is no difference between behavior in our replication of HMS and the treatments. At the same time, the coefficient on the earned endowment treatment is large, negative and significant ($p = 0.03$) indicating that participants who earn their endowments tend to give away $3.08 less than in our replication. The important test, however, is whether the TWDG and the earned endowment point estimates are different. The chi-squared test suggests that they are ($p = 0.04$); participants give away significantly less money when they have to earn their endowments in the TWDG.[7]

Table 2. The determinants of sending and allocations

	Dependent variable is amount sent			*Dependent variable is 1 if amount sent > 0*		
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Two-way	–0.67	0.10	–0.36	–0.03	–0.02	–0.01
	(1.29)	(0.83)	(0.77)	(0.12)	(0.14)	(0.13)
One-way	0.47	1.06	0.58	0.003	0.001	0.02
	(1.27)	(0.81)	(0.76)	(0.12)	(0.14)	(0.15)
Earned endowment	–3.08**	–2.51***	–1.77**	–0.22*	–0.25*	–0.24**
	(1.38)	(0.92)	(0.80)	(0.10)	(0.11)	(0.11)
GPA	–6.31***	–1.55	–1.86**	–0.65***	–0.51***	–0.51***
	(1.63)	(1.02)	(0.91)	(0.17)	(0.19)	(0.21)
SAT score	0.42	0.67**	0.51*	0.05	0.14**	0.13**
(hundreds)	(0.49)	(0.32)	(0.28)	(0.05)	(0.06)	(0.07)
Female	1.79**	0.46	0.61	0.12	0.04	0.04
	(0.85)	(0.54)	(0.47)	(0.08)	(0.10)	(0.10)
Number of siblings	1.09***	0.44*	0.29	0.05	–0.001	0.01
	(0.41)	(0.27)	(0.24)	(0.04)	(0.04)	(0.04)
First born	1.27	0.24	0.26	0.15*	0.17*	0.16*
	(0.90)	(0.57)	(0.51)	(0.08)	(0.09)	(0.09)
Age	0.70*	0.83***	0.72***	0.08**	0.17***	0.18***
	(0.37)	(0.25)	(0.22)	(0.04)	(0.05)	(0.05)
Economics major	–0.58	–0.47	–0.35	–0.12	–0.15	–0.14
	(0.89)	(0.58)	(0.51)	(0.08)	(0.10)	(0.10)
Work part time	1.50*	0.11	–0.05	0.13*	–0.03	–0.02
	(0.85)	(0.55)	(0.49)	(0.08)	(0.09)	(0.09)
ADHD		2.65***	2.46***		0.47***	0.34*
		(0.29)	(0.59)		(0.11)	(0.22)
ADHD × Two-way			0.45			0.25
			(0.69)			(0.28)
ADHD × One-way			1.03			0.21
			(0.72)			(0.27)
ADHD × Earned			–1.16*			0.02
endowment			(0.71)			(0.26)
Constant	–3.27	–22.88***	–16.62**			
	(10.99)	(7.61)	(6.85)			
Observations	149	149	149	149	149	149
Pseudo R^2	0.12	0.31	0.35	0.21	0.48	0.50

Notes: *indicates significant at the 10% level, **5%, ***1%. Columns (1)–(3) are the results of Tobit regressions censored at zero. Columns (4)–(6) report the marginal effects from probit regressions with robust standard errors. The other interactions with ADHD are included in specifications (3) and (6) but only the treatment interactions are reported.

Column (1) also demonstrates that a standard deviation increase in grade point average reduces allocations by $1.83, older students give away more, an additional 70 cents per year, and having a part-time job increases allocations by $1.50. Lastly, we replicate Eckel and Grossman to some degree – our female participants gave away $1.79 more than our male participants. Considering that the average allocation was $0.98, the female–male difference is considerable; however, as we show next, the result is not robust.[8]

In column (2) of Table 2 we add our measure of ADHD to the analysis of dictator allocations. This addition appears to change a number of the previous point estimates. The coefficients on GPA, the Female indicator and the working indicator shrink substantially and are no longer significantly different from zero. Nevertheless, the treatment differences (or lack of differences) remain and allocations are highly correlated with our ADHD measure: a standard deviation increase in ADHD increases the dictator's allocation by $2.54. This result provides direct evidence that impulsive participants give away significantly more money in the DG.

We see in Figure 2, however, that the role of ADHD may depend on the treatment. To confirm that the reduction in allocations in the earned endowment treatment is due to the attenuation of impulsivity (perhaps because people have been doing something already), in column (3) we interact our measure of ADHD with the treatments. As one would

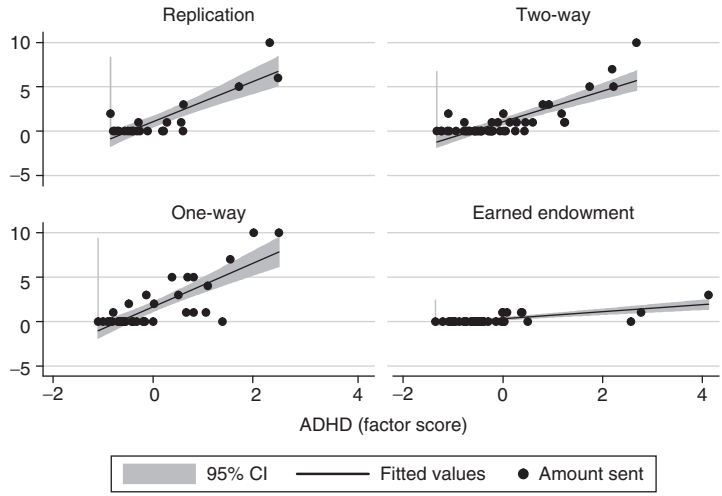

Figure 2. Dictator allocations and ADHD by treatment

expect, based on Figure 2, only the ADHD × Earned endowment coefficient is significant at the 10% level and negative indicating that those people who score high on the ADHD scale do tend to give away less money once they have earned it. The correct test, however, uses the TWDG as the baseline instead of the replication. Indeed, there is a significant difference between the ADHD × Earned endowment and the ADHD × Two-way coefficients ($\beta = 1.62$, $p < 0.01$) suggesting that limiting attention to the TWDG does not change the fact that earning one's endowment reduces the tendency to impulsively give money away.

All the results from our analysis of the amount allocated by the dictators remain when we use the probit estimator (with robust standard errors) to examine the incidence of giving instead. In columns (4) to (6) of Table 2 we see that participants in the treatments are no more or less likely to give compared to our replication of HMS but when they have to earn their endowments they are more than 20% less likely to give. The demographics results suggest that participants with higher GPAs are less likely to give, higher combined SAT scores are associated with more giving, first borns are more giving and older students are more likely to give. That said, the important result is that participants who measure high in ADHD do not just give more money away, they are also more likely to give and earning one's endowment reduces this impulse. A standard deviation increase in impulsivity predicts a 33% increase in the likelihood of giving.[9]

Discussion

We conducted an experiment that challenges the standard, social preference, interpretation of choices in the double blind DG conducted in the lab with students. In this sense our work is similar to the recent contributions by List (2007) and Bardsley (2008) who develop a different argument by demonstrating that changing the basic DG so that dictators can take in addition to give also results in lower transfers. The implication is that reference points and expectations seem to matter too. However, these papers do not include the treatments we examine here. Specifically, there are no earned endowment treatments; nor are there treatments in which endowments are equalized and the transfers are not inflated by the experimenter. In our simple set up, because all players were given $20 to start, neither inequality aversion nor deservingness-based altruism should motivate people to give. Despite this, allocations are identical to our replication of the standard double blind DG implying that altruism might be the wrong

interpretation of giving. Instead, we hypothesize that giving might be driven by participants coming to the lab ready 'to play.' The fact that there is a strong correlation between participant responses to an impulsivity questionnaire and both the rate and level of giving provides support for this hypothesis.

While we show that it is likely that 'coming to play' biases our interpretation of behavior in games like the DG, we also show that having participants earn their endowments reduces the amount of 'coming to play' substantially. That said, the exact mechanism by which earning one's endowment reduces impulsivity is not known; nor has our experiment exactly estimated the size of this bias. It may be that the initial production task or the more general assignment of a property right as in Hoffman and Spitzer (1985) or Hoffman et al. (1994) reduces the urge to act or it may be that earning the money makes one more focused on the implications of giving it away as in Spraggon and Oxoby (2008). As a practical matter, however, earned endowments should probably become more common.

Our data helps make a case for the importance of impulsive behavior in the DG; however, we also think that the degree to which impulsivity or 'coming to play' matters in the DG depends on the protocol, specifically the context in which the recipient is portrayed – is there a reason to give? Compared to the double blind DG conducted in the sterile lab between students, the protocol of Eckel and Grossman (1996) in which students make donations to the Red Cross seems more likely to create an environment in which altruism displaces impulsivity, at least to some degree. Similarly, the protocol of Fong (2005) in which the recipients were demonstrably poor or the fact that Benz and Meier (2005) find some correlation between contextualized lab giving and giving in the real world give us more confidence that they are measuring differences in social preference.

The problem of 'coming to play' surely affects behavior in other context-poor lab experiments. For example, the third party punishment version of the DG run by Fehr and Fischbacher (2004), in which there is a third participant whose job is to monitor and punish the dictator based on how much she gives to the second party, might overstate the likelihood of third party intervention because the only way the third party can participate is by punishing the dictator. Fehr and Fischbacher report a 60% rate of intervention but in a social dilemma setting in which third parties also made contributions to a public good, Carpenter and Matthews (2005) report an intervention rate of only 10%. In general, we think that people should be wary of experiments that seem too 'clean' if the cleanliness

comes from severe restrictions in the ways in which people participate. Another way of putting this is that economic experiments may be the only domain in which Say's law (the colloquial version) applies. Simply supplying choices to subjects in experiments may create a demand for behavior if that behavior is the only way to participate. Instead of demand effects, experimenters might start worrying about 'supply effects.'

Appendix – Sample instructions and the ADHD scale (borrowed from Young 2004)

Instructions

Payment and confidentiality

You have been asked to participate in an economics experiment. For your participation today you will receive an amount of money that will depend on the decisions participants make in the experiment. This amount will be paid to you in cash at the end of the experiment.

Today's experiment will be conducted according to a double blind protocol. This means that no one (including the other participants and the people conducting the experiment) will ever be able to match your decisions to you specifically. In this experiment each of you will be randomly paired with another person. You will not be told who this person is and the other person will not be told who you are, either during or after the experiment.

The experiment

After we finish reading the instructions together, you will be randomly split into two groups, group A and group B. The groups will be separated, and each of the participants will start with an envelope containing $20 in cash. [Earned endowment: This $20 is the money you earned during the first stage of the experiment.] When called on, a participant will go into the adjoining room alone to decide how much of this $20 (s)he wants to transfer to someone at random in the other group. Each of you will make this decision without knowing how much you will receive from the other group. When both groups are finished, you will receive an envelope containing the amount transferred to you from someone in the other group.

Your final payment will be:

$20 − amount you transfer + amount transferred to you.

Let's now go through the procedures in more detail. At the beginning of the experiment everyone will receive an opaque envelope containing

$20 in cash and a slip of paper that has your participant number written on it. To protect your anonymity, do not reveal this number to anyone (even the people running the experiment). When called on you will go to the adjoining room to privately make your choice of how much money to transfer to someone in the other group. Remove from the envelope the amount of money that you want to keep. Any money left in the envelope will be transferred to the other participant. After making your decision you will place your envelope in the box near the front of the room.

When you have made your decision you will be given a survey to fill out while the other participants make their decisions. Write your participant number on the top of the survey, fill out the survey, and then place it back in the opaque envelope that it came in. Your participant number will link your survey responses to your envelope, but nobody (including the people running the experiment) will ever be able to link you to your participant number.

After everyone has finished allocating money, the people running the experiment will take the two boxes (one from each group) that contain the opaque envelopes into the hall to record the amounts transferred. They will then bring the box from group A to group B and vice versa. One at a time, each participant will then go to the front of the room to randomly receive an envelope from the other group and turn in his or her survey.

Are there any questions?

ADHD scale

For the following statements, please circle the response that best fits. Never (1), Rarely (2), Sometimes (3), Often (4), Most of the time (5)

1. Have you failed to finish things you have started?
2. Have you been constantly fidgeting?
3. Have you been inattentive, easily distracted?
4. Have you made careless mistakes in your work or activities?
5. Have you had difficulty sustaining attention?
6. Have you not listened when people spoke to you directly?
7. Have you not understood instructions?
8. Have you had difficulty organizing tasks and activities?
9. Have you avoided tasks that require sustained mental effort?
10. Have you been reluctant to do tasks that require sustained mental effort?
11. Have you disliked doing tasks requiring sustained mental effort?
12. Have you lost items necessary for tasks or activities?
13. Have you been distracted by events happening around you?

14. Have you been forgetful in daily activities?
15. Have you fidgeted with your hands or squirmed in your seat?
16. Have you left your seat in situations where remaining seated is expected (e.g. in the cinema)?
17. Have you had difficulty with quiet leisure activities?
18. Have you felt 'on the go' or often acted as if driven by a motor?
19. Have you talked excessively?
20. Have you blurted out answers before questions have been completed?
21. Have you had difficulty waiting your turn?
22. Have you interrupted or intruded on others?
23. Have you daydreamed?
24. Have you had difficulty concentrating?
25. Have you engaged in attention-seeking behavior?
26. Have you lacked caution and acted in a dare-devilish way?
27. Have you felt as if you can't be bothered?
28. Have you felt bored and disinterested?
29. Have you been neat, tidy and organized?
30. Have you acted without thinking?
31. Have you behaved in a childish way?
32. Have you been able to sit comfortably through a movie or TV show?
33. Have you been able to remain seated at the dinner table?
34. Have you felt as though you were over-excited, as if you were 'flying' or going too fast?
35. Have you planned things ahead of time?
36. Have you had difficulty scheduling things on time?
37. Have you been late for appointments?
38. Have you jumped from task to task without finishing the first?
39. Have you put off difficult projects?
40. Have you procrastinated and complained?
41. Have you spoken out before thinking?
42. Have you made impetuous decisions?
43. Have you acted without thinking (e.g. impulse buying)?
44. Have you sought out excessive excitement and thrills?
45. Have you been able to focus on tasks and activities when necessary?

Acknowledgment

We thank Steve Burks, Colin Camerer, Peter Matthews and two referees for input on the design of our experiment and the National Science Foundation (CAREER 0092953) for financial support.

NOTES

1 Although there are no 'context-free' experiments, standard usage of this term refers to an experiment in which as many contextual cues as possible (e.g. deservingness) have been removed.
2 Sample instructions appear in the appendix.
3 In other words, the earned endowment treatment is identical to the TWDG treatment except for the fact that all participants earn their $20 endowments instead of being given them.
4 This scale also appears in the appendix.
5 In principle, the 45 questions of the ADHD scale are designed to measure a common unobservable characteristic – impulsivity. Faced with 45 measures of a common characteristic, it would be natural for an economist to create an index to summarize this data by simply adding the 45 responses for each respondent. However, it is not obvious that the linear combination that gives each question equal weight is the best summary. For this reason we used factor analysis to produce the linear combination of the questions that are the most highly correlated. The varimax rotation maximizes the variance of the squared loadings within factors to provide the most variation in the resulting index and the best interpretation.
6 The TWDG and OWDG treatments are also not significantly different ($z = 0.58$, $p = 0.56$; $\chi^2 = 0.03$, $p = 0.85$).
7 The coefficients on the OWDG and earned endowment indicators are also significantly different ($p < 0.01$).
8 Indeed the high coefficient on the Female indicator is due to using the Tobit regressor which accounts for the possibility of censoring. The intuition is as follows. The mean allocation conditional on giving for a woman was $3.67. For a man the same number is $2.56. So, part of the $1.79 reported in Table 2 can be accounted for by this difference of $1.11. However, $1.11 is still substantially lower than $1.79. The rest of the difference comes from the fact that women are more likely to give: 41% of women gave while only 30% of men did.
9 It appears from column (6), however, that earning one's endowment only focuses the attention of high ADHD players on how much they give away not on whether they should give.

REFERENCES

Bardsley, N. 2008. 'Dictator Game Giving: Altruism or artefact?'. *Experimental Economics* 11: 122–33.

Ben-Ner, A. and Putterman, L. 2004. 'Reciprocity in a Two Part Dictator Game'. *Journal of Economic Behavior & Organization* 53: 333–52.

Benz, M. and 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*.

Bolton, G. and Katok, E. 1998. 'An Experimental Test of the Crowding Out Hypothesis: The Nature of Beneficent Behavior'. *Journal of Economic Behavior and Organization* 37: 315–31.

Carpenter, J. and Matthews, P. 2005. 'Norm Enforcement: Anger, Indignation, or Reciprocity'. *Department of Economics, Middlebury College, Working Paper 0503*.

Cherry, T., Frykblom, P. and Shogren, J. 2002. 'Hardnose the Dictator'. *American Economic Review* 92: 1218–21.

Cox, J. 2004. 'How to Identify Trust and Reciprocity'. *Games and Economic Behavior* 46: 260–81.

Eckel, C. and Grossman, P. 1996. 'Altruism in Anonymous Dictator Games'. *Games and Economic Behavior* 16: 181–91.

Eckel, C. and Grossman, P. 1998. 'Are Women Less Selfish Than Men? Evidence from Dictator Experiments'. *Economic Journal* 108: 726–35.

Engelmann, D. and Strobel, M. 2004. 'Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments'. *American Economic Review* 94: 857–69.

Fehr, E. and Fischbacher, U. 2004. 'Third Party Punishment and Social Norms'. *Evolution and Human Behavior* 25: 63–87.

Fehr, E. and Schmidt, K. 1999. 'A Theory of Fairness, Competition, and Cooperation'. *Quarterly Journal of Economics* 114: 769–816.

Fong, C. 2001. 'Social Preferences, Self-Interest, and the Demand for Redistribution'. *Journal of Public Economics* 82: 225.

Fong, C. 2005. 'Evidence from an Experiment on Charity to Welfare Recipients: Reciprocity, Altruism and the Empathic Responsiveness Hypothesis'. *The Economic Journal* 117: 1008–24.

Forsythe, R., Horowitz, J., Savin, N. and Sefton, M. 1994. 'Fairness in Simple Bargaining Experiments'. *Games and Economic Behavior* 6: 347–69.

Hoffman, E., McCabe, K., Shachat, J. and Smith, V., 1994. 'Preferences, Property Rights, and Anonymity in Bargaining Games'. *Games and Economic Behavior* 7: 346–80.

Hoffman, E., McCabe, K. and Smith, V. 1996. 'Social Distance and Other-Regarding Behavior in Dictator Games'. *American Economic Review* 86: 653–60.

Hoffman, E. and Spitzer, M. 1985. 'Entitlements, Rights, and Fairness: An Experimental Examination of Subjects' Concepts of Distributive Justice'. *Journal of Legal Studies* 14: 259–97.

Konow, J. 2005. 'Mixed Feelings: Theory and Evidence of Warm Glow and Altruism'. Department of Economics, *Loyola Marymount University working paper*.

Malloy-Diniz, L., Fuentes, D., Leite, W. Borges, Correa, H. and Bechara, A. 2007. 'Impulsive Behavior in Adults with Attention Deficit/Hyperactivity Disorder: Characterization of Attentional, Motor and Cognitive Impulsiveness'. *Journal of the International Neuropsychological Society* 13: 693–8.

List, J. 2007. 'On the Interpretation of Giving in Dictator Games'. *Journal of Political Economy* 115: 482–93.

Roethlisberger, F. and Dickson, W. 1939. *Management and the Worker.* Cambridge, MA: Harvard University Press.

Spraggon, J. and Oxoby, R. 2008. 'Mine and Yours: Property Rights in Dictator Games'. *Journal of Economic Behavior and Organization* 65: 703–13.

Young, S. 2004. 'The YAQ-S and YAQ-I: The Development of Self and Informant Questionnaires Reporting on Current Adult ADHD Symptomatology, Comorbid and Associated Problems'. *Personality and Individual Differences* 36: 1211–23.

Young, S. and Gudjonsson, G. 2005. 'Neuropsychological Correlates of the YAQ-S and YAQ-I Self- and Informant-Reported ADHD Symptomatology, Emotional and Social Problems and Delinquent Behavior'. *British Journal of Clinical Psychology* 44: 47–57.

JEFFREY CARPENTER is an Associate Professor of Economics at Middlebury College and a research associate at the German Institute for the Study of Labor (IZA).
ADDRESS: 601 Warner Hall, Department of Economics, Middlebury College, Middlebury, Vermont, 05753 [email: jpc@middlebury.edu].

ALLISON LIATI is an economics and psychology student at Middlebury College and a research assistant on this project.
ADDRESS: 601 Warner Hall, Department of Economics, Middlebury College, Middlebury, Vermont, 05753 [email: aliati@middlebury.edu].

BRIAN VICKERY is an economics student at Middlebury College and a research assistant on this project.
ADDRESS: 601 Warner Hall, Department of Economics, Middlebury College, Middlebury, Vermont, 05753 [email: bvicker@middlebury.edu].