ISSN 1755-5361

University of Essex

Department of Economics

Discussion Paper Series

No. 680 November 2009

Desert and Tangibility: Decomposing House Money Effects in a Charitable Giving Experiment

David Reinstein
Gerhard Riener

Desert and Tangibility: Decomposing House Money Effects in a Charitable Giving Experiment

David Reinstein and Gerhard Riener

December 11, 2009

Abstract

Several papers have documented that when subjects play with standard laboratory "endowments" they make less self-interested choices then when they use money they have either earned through a laboratory task or brought from outside the lab. In the context of a charitable giving experiment we decompose common "house money" effects into two components: the tangibility of cash in hand relative to money (or ecu's) promised on a computer screen, and the desert of earned money relative to random windfall gains. While both components are found to be significant in non-parametric tests, the former effect, which has been neglected in previous studies, has a stronger effect on total donations. These results have clear implications for experimental design, and also suggest that the availability of less tangible payment methods may increase charitable donations.[1]

1 Introduction

Several economists have found that when subjects play with standard laboratory "endowments" they make less self-interested choices then when they use money they have either "earned" through a laboratory task or brought from outside the lab (Cherry et al., 2002, Hoffman and Spitzer, 1985, Loomes and Burrows, 1994). This effect is typically interpreted as a result of Lockean desert effects (Rutstrom and Williams, 2000), fairness concerns (*a*

[1]Acknowledgements: We would like to thank the British Academy and the University of Jena for providing financial support for our research. We would also like to thank the many colleagues and seminar participants who have offered us valuable advice and comments.

Contact details: David is a lecturer (assistant professor) at the University of Essex. He can be contacted at drein@essex.ac.uk. Gerhard is a postdoctoral researcher at the University of Jena, his email is gerhard.riener@uni-jena.de.

This paper is ripe: please feel free to cite it!

la Rabin (1993)), or a different mental accounting over windfall gains (Sheffrin and Thaler, 1988, Thaler, 1985 and Thaler and Johnson, 1990); this is generally referred to as the *windfall effect*.

In the context of dictator games, the evidence for windfall effects is strong. Cherry et al. (2002) ran a series experiments in which dictators were "provisionally allocated" either $10 or $40 in cash in an envelope and asked to "propose" a division of it between themselves and another subject by leaving the amount they want the other subject to receive in the same envelope. In the baseline treatment the allocation was randomly determined, while in their *earnings* treatment the dictators had to arrive one hour earlier than the recipient subjects, and their endowment was based on performance in a cognitive task (solving GMAT questions), and this was common knowledge. Finally, their *double blind with earnings* treatment modified the *earnings* treatment to increase subject-experimenter anonymity.[2] Both their earnings and double-blind-earnings lead to significantly less generous dictator behavior; the latter treatment the dictators became almost entirely *hardnosed*, keeping nearly all of the money. Their results strongly suggest that both earnings ("legitimized" assets) and reputation-seeking concerns *vis-a-vis* the experimenter (which they term "strategic") significantly affect dictators' behavior.

Mittone and Ploner (2006) start with a replication of Cherry et al.'s double-blind-earnings treatment as their control group, and compare this with dictator behavior when potential *recipients* are allowed to take the quiz, but are not compensated for this. They find that dictators give significantly more, and are more likely to give something, when recipients have also taken the quiz, and this holds for all levels of earnings (success in GMAT questions). This suggests that fairness concerns are important, and that *relative* desert may be driving dictator decisions. Ruffle (1998) Oxoby and Spraggon (2008), and Cherry and Shogren (2008) find similar results on the importance of the receiver's effort.

In all of the experiments mentioned above the dictator subjects' "earnings" come from answering GMAT questions, some of which involve retailers' and consumers' decisions, dishonest job applicants, wealth, investments, money, and marketing. These may be triggering more self-interested behavior through a framing effect as in Vohs et al. (2006), rather than simply increased legitimacy of the dictators' own endowments. However, the estimated relationships between the dictators' gifts and the *recipient's* performance, hence the impact of fairness (relative desert) are robust to this critique.

[2]In the former treatments each dictator passes her envelope to the experimenter after making her choice. In the latter treatment, the dictators drop their envelopes into a box. Furthermore, in the latter treatment, the dictator has a two in 12 chance of getting an envelope filled with only blank paper and thus not having any money to pass at all – hence the experimenter can never infer with certainty that a particular dictator was "selfish," although the experimenter could still make a probabilistic judgment about subjects' behavior.

In contrast to the dictator environment, there is little evidence for windfall effects in voluntary contribution mechanism (henceforth VCM) experiments. Clark (2002) examines contribution rates in a linear VCM game. He finds no significant difference between contributions in the "own money" treatment, in which subjects are asked to bring $8 from outside the lab to purchase tokens, and the "house money" treatment, in which subjects are simply given the tokens.[3] However, as Clark's "own money" subjects are *also* allocated *house money* at the end of the experiment, they presumably have the same expected "windfall" earnings as the other subjects. Furthermore, Clark's subjects use tokens, and the earnings effect may be more salient when the rewards are more tangible.

Cherry et al. (2005) find that while heterogeneity in earnings decreased contributions, the origin of these earnings – whether derived randomly or through success in answering GMAT questions – did not have a significant effect. Kroll et al. (2007) and Spraggon and Oxoby (2009) allow subjects to choose a "strategy vector method" (see Fischbacher et al. (2001) and Keser and van Winden (2000)), and identify a surprising "inverse found money effect": "participants who earned their endowments and were matched with someone who did not were more unconditionally and conditionally cooperative."

Finally, there is some very recent evidence in the context of charitable giving itself. Carlsson et al. find windfall effects in a charitable giving experiment in both a laboratory and a field setting; subjects in both environments donate less when they have "earned" their pay by completing a survey. Soetevent (2009) compared treatments with different payment options (debit cards, cash, or both) in a door-to-door solicitation campaign, finding that smaller donors "drop out" when cash is disallowed, while participation increases when *only* cash (and not debit cards) is allowed. However, he does not decompose the effects of the greater anonymity of the cash payment, the "small coin nuisance," the lack of trust in debit card security, and the impact of the use of cash itself.

This literature (aside from Soetevent, 2009) has ignored a second component of the "house money" bias that limits the external validity of many laboratory results: people may treat money they are promised (or are given in the form of tokens) differently then cash they physically hold – we call this the *tangibility* effect.[4] We hypothesize three potential reasons why this may occur. First, psychology experiments (involving deception) demonstrate that subjects given "reminders of money" are both less helpful and less likely to ask for help in a variety of non-remunerated tasks (Vohs et al., 2006). Second, using cash may cause subjects

[3]Still, Harrison (2007), who reanalyzed Clark's data to deal with the potentially non-independent error structure, suggests that a house money effect is present. Still, the tangibility and earnings effects are not separable in this context

[4]The use of such payments is ubiquitous in experimental economics.

to more carefully consider the consumption they are sacrificing.[5] Finally, parting with cash may itself bring some disutility, perhaps through an attachment to this money similar to the "endowment effect" of Kahneman et al. (1991).[6]In the present paper we present the first salient, non-deceptive experimental evidence that the tangibility of the choice medium affects other-regarding decisions.

2 Experimental design[7]

We use a charitable giving experiment with a $2 \times 2 \times 2$ design to differentiate two components of the house money effect. Firstly, the treatments vary according to the extent to which subjects should see the money as *earned*; we compare giving behavior after compensation based on performance on a five minute task to behavior with a randomly allocated payment. The second dimension of variation involves the tangibility of the payment: we either give cash to the subjects *before* they decide how much to donate (and they physically place any donations they make into envelopes) or they allocate their donation from an "endowment" on the computer screen and they are paid cash at the end of the experiment. Thus, we separately test whether *earning the money* and having *cash in hand* affect giving behavior in the lab.[8]

Unlike many of the experiments previously mentioned, our subjects make decisions over donations to charitable foundations – institutions outside the laboratory. In line with Eckel and Grossman (1996), we see this as a more obvious and typical expression of other-regarding behavior than donations to a laboratory public good or towards another laboratory subject. Our environment also provides a more demanding test for tangibility and windfall effects. In the real world it is rare to be asked for a gift from a random non-needy stranger (or to receive such a gift); hence, it is not surprising that standard dictator games should be sensitive to framing effects. On the other hand charitable appeals and charitable giving are regularities, so subjects will have more experience with such decisions and their decisions should be less easily perturbed. While dictator giving to other subjects is highly sensitive to the level of social isolation, falling to very low levels in "double-blind" environments Hoffman

[5]Oberholzer-Gee and Eichenberger (1999) make a similar case, arguing that subjects do not fully consider the opportunity costs of the funds they give away in experiments. Also see Mazar et al. (2008) who find more cheating with exchangeable tokens then with cash.

[6]We leave distinguishing between these explanations for future research. Nonetheless all these hypotheses have similar implications for experimental design and for charitable fundraising.

[7]Pictures of key computer screens are given in the appendix.

[8]This third dimension of variation is in the choice set; we offer three charities instead of two in the *expanded choice set* treatment. This allows us to measure the sensitivity of our observed effect to variations in the choice set, and also pertains to our simultaneous work on "expenditure substitution" in charitable giving (for an earlier draft, see Reinstein (2008)).

et al. (1996)), charitable giving persists at significant levels even under highly anonymous conditions (Eckel and Grossman, 1996). Our setting may also better isolate the effect of asset legitimacy: intuition suggests that in the charitable giving context, subjects will focus less on their desert *relative* to the recipient(s) then they would in deciding how much to give to a fellow subject. Finally, our setting is relevant to an important sector of the economy.[9]

All treatments are assigned orthogonally; we have a (nearly) fully balanced design.[10] By construction, the distribution of initial allocations is the same for each treatment. Finally, all of our treatments involve the same strong level of anonymity.[11] The time spent in each treatment of our experiment is approximately the same, so subjects in each of our treatments should have the same earnings expectations.

The sessions were run at the Jena University Experimental Economics lab using the standard subject pool. In total 190 subjects participated in the experiments of which 54.2% were female.[12] The sessions were conducted in October 2008 (39 subjects), February 2009 (79 subjects), and September 2009 (72 subjects). While we ran each of the four payment regime treatments in a separate session, the participants were from the same subject pool and the times and dates of the experiment were stratified by treatment.[13] To avoid mixing payment types, we did not give subjects pre-experiment "show-up fee."

To guarantee anonymity, the lab was divided into an outer partition - which serves as a meeting room before the experiment and as a room for the administrators during the experiment - and an inner partition with computer terminals on which the subjects make decisions and answer questions. These were separated so that it was impossible to see the inner partition from the from the outer partition and vice-versa. For administrative purposes, a volunteer from the participants helped with the procedures whenever communication between the inner and the outer part of the lab was necessary. Furthermore, to ensure our credibility, this volunteer supervised the online donations made by the experimenters after

[9]Benz and Meier (2006) offer some support for the generalizability of lab charitable giving behavior.

[10]Because the treatments were run in separate sessions and their were some no-shows, the actual observations are very slightly off-balance, and the allocations are not precisely identically distributed by treatment, nor is the "choice set" treatment. However, these slight differences are controlled for in our multivariate regressions and in our balanced bootstrapped rank-sum tests. The lack of balance does not measurably affect any of our results. Our treatments are also not perfectly balanced over time. To test for session-specific effects, in the appendix we also report regressions with standard errors clustered by session, and controls controlling for time-of-day and date of session effects; our results are robust to all of these, and none of these are significant.

[11]See the protocol in the appendix for a full description of our careful procedure to insure subject-experimenter and subject-subject anonymity.

[12]We did not collect extensive demographics on our subjects in order to preserve subject-experimenter anonymity.

[13]Appendix D illustrates this balance, and shows our results are insensitive to the time and date of the session.

the other participants had been dismissed.

At the beginning of the experiment all subjects were assured that we would not be able to connect their name to the decisions they made. Next we asked for a volunteer to help us with administrative issues, mainly allocating the sealed envelopes with payouts at the end of the experiment.

The task Subjects in the performance treatments (PA and PC, described below) were told that their endowment would depend on their performance on a simple task. They were asked to add up five two-digit numbers[14] using only scratch paper and a pencil. The numbers were randomly drawn and presented to the subjects as in the example below:

12	77	34	55	62	__

The participants were given five minutes to solve as many tasks as possible. We argue that this task was sufficiently tedious to make subjects feel that they *earned* the money recieved. This task, although numerical, is less likely to cue self-interested "economic" thinking than the GMAT questions used in many previous studies.

The charitable giving stage This stage was a one-shot dictator game where subjects could donate none, some, or all of their endowment to any combination of the available charities in units of 50 Euro cents. All subjects were presented with *Brot für die Welt* (BfdW) – "Bread for the World", a German development aid agency and the World Wild Life Fund for Nature (WWF), a nature conservancy charity. For the *expanded choice set* treatment we also included *Deutsches Rotes Kreuz* (DRK) - the German Red Cross - which operates in similar areas as *BfdW*. Subjects were given information about each of the charities on the computer screen[15] and next had to decide how much (if anything) to donate to each available charity and enter this into the computer.[16] By using multiple charities we reduced the noise surrounding heterogeneous tastes for charities, and gained more useful data on a wider range of subjects.

[14]This task has been used in various occasions for testing competitiveness (e.g., (Niederle and Vesterlund, 2007)).

[15]Translations can be found in appendix B.

[16]The order of the presentation of the charities, both on the description screens and on the actual donation screen are stratified over subjects, in order to control (and test) for potential order effects.

Treatment 1: Performance / (on-screen) Account (PA) Subjects in the performance treatments were told that the probability of higher earnings increased in the number of tasks correctly completed, but we did not specify exactly how performance translated into payoffs.[17] After completing the task they were told how much this earned them. In *account* treatments they were allocated €5, €7.50 or €10 (shown on their computer screen) but were not yet given cash. They next made their donation decisions. At the end of the experiment they were (anonymously) given envelopes containing their earnings minus their total donations.

Treatment 2: Performance / Cash (PC) As in *PA*, subjects first completed the task first and learned how much they earned. However, unlike in the *account* treatments, subjects in *cash* treatments were paid in cash *before* they made giving decisions. After the task stage, the volunteer was prompted to come outside and bring the numbered envelopes containing the cash earnings into the inner part , where they hand each subject the envelope with his or her subject number on it. The payment envelopes were carefully assembled to look identical and have similar weights.[18] Subjects were instructed to inspect and count the money in private at their computer desks. Next, they made their donation decision(s) by entering these choices on the computer screen. Finally, subjects were asked to put the chosen contributions (in cash) into the donation envelope and seal it.

Treatment 3: Random / on screen Account (RA) In this treatment, subjects were allocated €5, €7.50 or €10 randomly on their computer screen. The donation stage followed, and payments were distributed as in *PC*.

Treatment 4: Random / Cash (RC) In the *RC* treatment the allocations were randomly determined (as in *RA*), and given to the subjects in identical envelopes as in *PC*. The donation and payments procedure also followed *PC*.

[17] We do not tell them that their pay was based on *relative* performance because we do not want them to compare themselves to other subjects in making their charitable contributions . This might lead them to believe that that subjects who earned more have a greater obligation to donate. In the treatments of October 2008, the first got €10, the second €7.50 and the rest of the subjects in the same session €5. In the sessions conducted in February, March, and September 2009, the participants who were in the upper tercile of solved tasks received €10, in the middle tercile €7.50 and in the lower tercile €5.

[18] We did this by using coins of different increments. To the extent that small coins are less desirable then bills this would lead to a bias *against* our finding of a tangibility effect. Since payments in performance and random treatments had the same distribution, this should not impact our "earnings effect" findings.

Table 1: Average proportion contributed by payment regimes

	Payment			
Allocation	*Random*	*Performance*	**Total**	**N**
Account	0.27	0.18	0.23	99
Cash	0.14	0.12	0.13	91
Total	0.21	0.15	0.18	
N	102	88		190

Wilcoxon rank sum tests		
P(Account> Cash)	0.57*	(0.06); [0.05]
P(Random> Performance)	0.58**	(0.05); [0.03]
P(Account/Random > Cash/Performance)	0.64***	(0.01); [0.00]
P(Account/Random > Cash/Random)	0.61*	(0.06); [0.05]
P(Account/Performance > Cash/Performance)	0.53	(0.59); [0.57]
P(Account/Performance > Cash/Random)	0.49	(0.86); [0.85]

p-values for simple rank sum tests in parentheses, *: p< 0.10, **: p<0.05, ***: p<0.01

In square brackets: p-values for bootstrapped rank sum tests, 1000 draws, balanced by all treatments and stake sizes.

3 Results

3.1 Summary Statistics

Table 1 compares the proportion of the endowments donated to any of the two (or three) charities, pooling across charity treatments. Subjects donated significantly less[19] when they were paid in cash then when their allocation was only shown on the computer screen (13% versus 23% of the total funds, pooling across all other treatments).[20]

Figure 1 shows the cumulative distribution of the share of earnings donated over the earnings and payment treatments. The distribution of contributions under on-screen entitlements (RA and PA) stochastically dominates the distribution under cash payments (RC and PC). Similarly, the distribution under random payments (RA and RC) stochastically dominates the distribution under performance-based earnings (PA and PC).[21]

[19] These differences are significant in Wilcoxon rank-sum tests, as well as in familiar parametric tests (available by request). Because of the aforementioned lack of balance (stemming from no-shows), we also report bootstrapped rank-sum tests in brackets, with each of the 1000 random draws (with replacement) exactly balanced by payment treatment, earnings treatment, choice set treatment, and stake size.

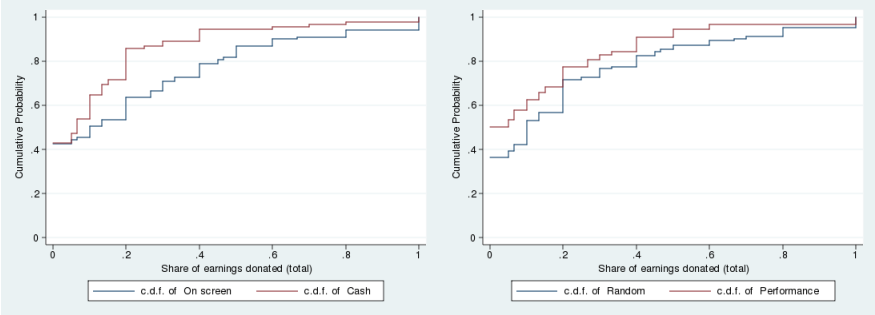
[20] This rate of giving is fairly consistent with results of previous experiments. E.g., in Eckel and Grossman (1996) subjects give 30% of their $10 cash allocation (they were also given a $5 show-up fee).

[21] On the other hand, as shown in table 2, the performance treatment lead to a significantly lower *propensity* to donate (50% versus 64%), while the cash treatment had no noticeable extensive margin effect. However, in Probit regressions (available by request) none of our treatments are significant.

Table 2: Number of subjects who donated by treatment

	Payment		**Allocation**	
Donated	*Random* N *(column %)*	*Performance* N *(col. %)*	*Account* N *(col. %)*	*Cash* N *(col. %)*
No	37 *(46%)*	44 *(50%)*	42 *(42%)*	39 *(43%)*
Yes	65 *(64%)*	44 *(50%)*	57 *(58%)*	52 *(57%)*
Total number	102	88	99	91
p-values of tests				
Pearson χ^2	0.06		0.95	
Fisher's exact	0.07		1.00	

Figure 1: Cumulative distribution functions of share of earnings donated



3.2 Multivariate Analysis

To control for observable (random) differences in treatment assignment and examine the heterogeneity of our observed effects in a standard framework, we regress total donations on controls for observable heterogeneity and treatment interactions.[22] These regressions again suggest that cash treatments significantly and consistently reduced generosity. Donations were also higher when subjects were paid according to their performance, but thiseffect was not statistically significant. The coefficients on higher stake sizes (€7.5 or €10) are small and not significant: subjects who earn more do not tend to donate more. In line with some previous work, (e.g., Eckel and Grossman (1998), List (2004)) women donated more than men. Total donations were not significantly different when a third charity was included. The interaction effects are not significant, although their positive sign and magnitude suggest that the treatments may not have an additive effect but may be "substitutes" – the summed coefficient representing the effect of cash and performance combined is very close to the coefficient on cash alone.

As we show in the appendix, the tangibility and windfall effects on donations to each charity are similar, and our results also hold for a fractional response Papke and Wooldridge (1996) regression of "share donated".

4 Conclusion

Our experiment is the first to document the tangibility effect; its magnitude appears at least as strong as the windfall effect (although the latter has a stronger effect at the extensive margin). Furthermore, by using a charitable giving context and a relatively neutral real-effort task, we add to the evidence that the legitimacy (*absolute* desert) of experimental subjects' *own* assets affect their other-regarding behavior. Our findings do not imply that experimenters should *always* use "tangible" cash. In the context of our experiment, we cannot say which contribution level is more externally valid. Whether the differences are because seeing money cues self-interest, because cash causes a more careful consideration of trade-offs, or because parting with cash is more painful, either frame (cash or endowment) may have external validity.[23] In the field many decisions are made without physical cash.

[22] We use both a standard OLS specification for familiarity. We use a Poisson specification, both because our data resembles count data (in increments of 50 cents) and because this specification deals with corner-solution (non-negative) data without being as sensitive to non-normality and heteroskedasticity as a standard Tobit regression Gourieroux et al. (1984); Arabmazar and Schmidt (1981). In the appendix we find similar results using a fractional regression specification. The cash and performance results are similar in zero-inflated Poisson regressions (available by request).

[23] On the other hand, cash is obviously better if it leads to greater experimenter credibility. But this is unlikely to have been a driver of our results, as all of our subjects had previously participated in economic

Table 3: Poisson and OLS regression on total donations

			Add. contr.		Gender contr.	
	(1)	(2)	(3)	(4)	(5)	(6)
	Psn.	OLS	Psn.	OLS	Psn.	OLS
Pay cash	-0.75*	-0.84*	-0.75*	-0.84*	-0.77**	-0.89**
	(0.30)	(0.33)	(0.30)	(0.34)	(0.29)	(0.34)
Pay by performance	-0.42	-0.54	-0.42	-0.54	-0.43	-0.58
	(0.31)	(0.39)	(0.31)	(0.40)	(0.30)	(0.41)
Cash × performance	0.31	0.44	0.32	0.44	0.41	0.56
	(0.59)	(0.49)	(0.59)	(0.50)	(0.60)	(0.52)
Third charity			0.25	0.26	0.24	0.26
			(0.23)	(0.25)	(0.22)	(0.24)
Stake: 7.5			-0.14	-0.13	-0.084	-0.084
			(0.27)	(0.29)	(0.26)	(0.28)
Stake: 10			0.077	0.093	0.064	0.100
			(0.29)	(0.32)	(0.27)	(0.31)
Female					0.49*	0.53*
					(0.23)	(0.25)
Combined coefficients						
Cash+perform+cash × perform	-0.76*	-0.94**	-0.76*	-0.94**	-0.73*	-0.92**
	(0.32)	(0.37)	(0.32)	(0.37)	(0.31)	(0.36)
Observations	190	190	190	190	190	190
R^2		0.048		0.056		0.079
Pseudo R^2	0.033		0.039		0.055	

Standard errors in parentheses, reported heteroskedasticity robust for OLS.

+ p<0.10, * p<0.05, ** p<0.01 for tests using heteroskedasticity-robust standard errors (for all columns)

All regressors are dichotomous (0,1) variables, dy/dx for discrete change of dummy variable reported .

However, researchers must be *aware* of this framing effect and take it into account. This distinction is important: economic experiments vary greatly along both dimensions, often simultaneously. For example, in Hoffman et al. (1996) the "single blind 2" treatment combines both a decreased social distance from the experimenter (relative to "single blind 1") and "a decision form for making the decision, instead of money". As noted in Hoffman et al. (1994), comparisons *between* experimental results must take into account differences in the decision medium.[24]

Our results may also be generalizable to real-world decision making, particularly over intangible "warm-glow" goods such as charitable donations. For example, rather than asking for cash, charitable organizations might do better to solicit donations in less tangible forms, such as through credit-cards or payroll deductions from future years' income.[25]

References

Arabmazar, A., Schmidt, P. (1981), Further evidence on the robustness of the Tobit estimator to heteroskedasticity, Journal of Econometrics, vol. 17(2):253–258.

Benz, M., Meier, S. (2006), Do People Behave in Experiments as in the Field?–Evidence from Donations, Institute for Empirical Research in Economics, University of Zurich, manuscript. Good connection between lab and field charitable giving.

Breman, A. (2006), Give more tomorrow: Evidence from a randomized field experiment, Unpublished manuscript.

Carlsson, F., He, H., Martinsson, P. (), Easy come, easy go-The role of windfall money in lab and field experiments, rapport nr: Working Papers in Economics 374.

Cherry, T., Frykblom, P., Shogren, J. (2002), Hardnose the Dictator, American Economic Review, vol. 92(4):1218–1221.

Cherry, T., Kroll, S., Shogren, J. (2005), The impact of endowment heterogeneity and origin on public good contributions: Evidence from the lab, Journal of Economic Behavior and Organization, vol. 57(3):357–365.

experiments.

[24] In comparing directly compare their dictator results to those of Forsythe et al. (1994), they note that "other aspects of the double blind procedures require experimental examination to identify what is driving the outcome; an envelope containing the cash might be an important factor." Our analysis confirms this speculation.

[25] See Breman (2006) for a similar suggestion, motivated by field experimental evidence.

Cherry, T., Shogren, J. (2008), Self-interest, sympathy and the origin of endowments, Economics Letters, vol. 101(1):69–72.

Clark, J. (2002), House Money Effects in Public Good Experiments, Experimental Economics, vol. 5(3):223–231.

Eckel, C., Grossman, P. (1996), Altruism in Anonymous Dictator Games, Games and Economic Behavior, vol. 16(2):181–191.

Eckel, C., Grossman, P. (1998), Are women less selfish than men?: Evidence from dictator experiments, The Economic Journal, vol. 108(448):726–735.

Fischbacher, U., Gächter, S., Fehr, E. (2001), Are people conditionally cooperative? Evidence from a public goods experiment, Economics Letters, vol. 71(3):397–404.

Forsythe, R., Horowitz, J., Savin, N., Sefton, M. (1994), Fairness in Simple Bargaining Games, Games and Economic Behavior, vol. 6(3):347–69.

Gourieroux, C., Monfort, A., Trognon, A. (1984), Pseudo maximum likelihood methods: applications to Poisson models, Econometrica, vol. 52(3):701–720.

Harrison, G. (2007), House money effects in public good experiments: Comment, Experimental Economics, vol. 10(4):429–437.

Hoffman, E., McCabe, K., Shachat, K., Smith, V. (1994), Preferences, property rights, and anonymity in bargaining games, Games and Economic Behavior, vol. 7(3):346–380.

Hoffman, E., McCabe, K., Smith, V. (1996), Social Distance and Other-Regarding Behavior in Dictator Games, American Economic Review, vol. 86:653–660.

Hoffman, E., Spitzer, M. (1985), Entitlements, rights, and fairness, Journal of Legal Studies, vol. 14(2):259–297.

Kahneman, D., Knetsch, J., Thaler, R. (1991), Anomalies: The endowment effect, loss aversion, and status quo bias, The Journal of Economic Perspectives, vol. 5(1):193–206.

Keser, C., van Winden, F. (2000), Conditional Cooperation and Voluntary Contributions to Public Goods, Scandinavian Journal of Economics, vol. 102(1):23–39.

Kroll, S., Cherry, T., Shogren, J. (2007), The impact of endowment heterogeneity and origin on contributions in best-shot public good games, Experimental Economics, vol. 10(4):411–428.

List, J. (2004), Young, Selfish and Male: Field evidence of social preferences*, The Economic Journal, vol. 114(492):121–149.

Loomes, G., Burrows, P. (1994), The Impact of Fairness on Bargaining Behaviour, Empirical Economics, vol. 19:201–221.

Mazar, N., Amir, O., Ariely, D. (2008), The Dishonesty of Honest People: A Theory of Self-Concept Maintenance, Journal of Marketing Research, vol. 45(6):633 – 644.

Mittone, L., Ploner, M. (2006), Is It Just Legitimacy of Endowments? An Experimental Analysis of Unilateral Giving, CEEL Working Paper N 02/2006.

Niederle, M., Vesterlund, L. (2007), Do Women Shy Away from Competition? Do Men Compete Too Much?, Quarterly Journal of Economics, vol. 122(3):1067–1101.

Oberholzer-Gee, F., Eichenberger, R. (1999), Focus Effects in Dictator Game Experiments.

Oxoby, R., Spraggon, J. (2008), Mine and yours: Property rights in dictator games, Journal of Economic Behavior and Organisation, vol. 65:703–713.

Papke, L., Wooldridge, J. (1996), Econometric methods for fractional response variables with an application to 401 (k) plan participation rates, Journal of Applied Econometrics, vol. 11(6):619–632.

Rabin, M. (1993), Incorporating Fairness into Game Theory and Economics, American Economic Review, vol. 83(5):1281–1302.

Reinstein, D. (2008), Substitution Between (and Motivations for) Charitable Contributions: An Experimental Study, Working paper.

Ruffle, B. (1998), More is better, but fair is fair: Tipping in dictator and ultimatum games, Games and Economic Behavior, vol. 23(2):247–265.

Rutstrom, E., Williams, M. (2000), Entitlements and fairness: an experimental study of distributive preferences, Journal of Economic Behavior and Organization, vol. 43(1):75–89.

Sheffrin, H., Thaler, R. (1988), The Behavioral Life-cycle Hypothesis, Economic Inquiry, vol. 26:609–643.

Soetevent, A.R. (2009), Payment Choice, Image Motivation and Contributions to Charity:Evidence from a Field Experiment.

Spraggon, J., Oxoby, R. (2009), An experimental investigation of endowment source heterogeneity in two-person public good games, Economics Letters, vol. 104(2):102–105.

Thaler, R. (1985), Mental Accounting and Consumer Choice, Marketing Science, vol. 4:199–214.

Thaler, R., Johnson, E. (1990), Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice, Management Science, vol. 36:643–660.

Vohs, K., Mead, N., Goode, M. (2006), The Psychological Consequences of Money, Science, vol. 314:1154–1156.

A Protocol – Not for publication[26]

Lab Setup

There must be two parts to the laboratory. One, the "Inside" is where the subjects sit at their desks/computers, and the other, the "Outside" is where we the experimenters, meet the subjects at the beginning and end. The Inside must not be viewable from the Outside, and this must be obvious. We set up subject computers, and the relevant handouts and numbers on desks Inside, but the server computer and the "N" z-leaf must be set up Outside. We need to have access to a printer Outside.

Set number of subjects in Background and Global. Sort clients and pre-fill envelopes with receipts and money and build three stacks.

Timing

1. Participants meet Outside
2. Give short description of what will happen [Briefing]
3. Ask for a volunteer. If there are more, select them by drawing balls from the Urn. Brief the volunteer
4. Participants draw a number from the box and are advised not to let us see it but to look at the number "Inside" the lab facility, and report to the desk with that number on it and follow the instructions. [Instructions performance]

[26]Note to editor and reviewers: this entire appendix, or some part of it, might be more suitable to put online, for space considerations. To emphasize this, we write "Not for publication" in each part of the appendix, although, if space permits, it might be helpful to have some parts of this in the publication.

5. [PERFORMANCE TREATMENTS ONLY] Start the WORK Stage

6. [CASH TREATMENT ONLY] Experimenter screen Prompts Subject number and payment. Look up subject number and computer number in subject table and put post it notes with subject number on the prepared envelopes [ENVELOPE]

   (a) Take pre-filled envelopes and put in: If subject shock=1: 3 donation envelopes If subject shock=0: 2 donation envelopes (Brot and WWF) and put subject number on post-it on the big brown envelope Prompt the volunteer to come out. Hand over the box with envelopes and instruct the volunteer to distribute the envelopes to the tables 10.[on screen] Instruct subjects to open the envelopes, count the money and enter the amount

7. DONATION PHASE – Donation

8. [CASH TREATMENT ONLY] Screen that reports payments tells subjects to put the amount they promised to donate into the appropriate plain white envelopes.

9. After Questionnaire

   (a) [CASH TREATMENTS ONLY] Subjects are instructed to collect their belongings and get up from their desks. They put the SPENDEN envelopes into the SPENDEN box and the BELEG into the BELEG Box. Volunteer makes sure that all are ready the subjects are to come "Outside" to meet us. Volunteer opens the "SPENDEN" [Donations] box and adds the actual donations. Volunteer observes that experimenter donated the correct amount of money (online, using credit card) and signs to this effect. We pay volunteer, and volunteer signs receipt of this. Volunteer also signs Volunteer Witness Form saying that "I witnessed that experimenters made [AMOUNTS HERE]€ payments to charities. This payment equaled the total of the actual subject contributions."

   (b) [ENTITLEMENT TREATMENTS ONLY] Subjects are instructed to collect their belongings and get up from their desks. They the BELEG into the BELEG Box. Volunteer makes sure that all are ready the subjects are to come "Outside" to meet us. Experimenters add up the donations recorded by ztree and show them to volunteer. Volunteer observes that experimenter donated the correct amount of money (online, using credit card) and signs to this effect. We pay volunteer, and volunteer signs receipt of this. Volunteer also signs Volunteer Witness Form saying that "I witnessed that experimenters made [AMOUNTS HERE]€ payments to charities."

Records and book-keeping

Receipt forms (and Volunteer Witness Form) and records of donations (email response from charity, credit card record, letters from charity when they arrive) will be sent to the graduate school administrator at the Economics Department, University of Essex, UK or the University of Jena for reimbursement. Subjects are told that they can contact the administrator [name provided to the subjects] if they are still skeptical and want to verify the (total, per session) donations made.

B Description of charities – Not for publication

Brot für die Welt (Bread for the World)

Brot für die Welt is a development organisation by the church founded in 1959 in Berlin. It is supported by all the country's Protestant and independent churches. The management of the organisation "Bread for the World" is located at the Diakonisches Werk der EKD eV, which is the legal entity of action. The annual fund-raising starts on the first Advent, the beginning of the liturgical year. Every action is under a particular theme, which will indicate specifically funded projects. Most development projects are assigned to various program topics. In 2007 "Brot für die Welt" mainly promoted measures to ensure food security and access to basic services such as education and health. Other supported areas are peacekeeping and democracy promotion, and the fight against HIV / AIDS. As of 2005, they have received over 1.6 billion Euro in donations for aid projects in Africa, Asia, Latin America and for several years in Eastern Europe. In 2006 Bread for the World received donations amounting to 51.5 million euros.

WWF

The WWF, the World Wide Fund For Nature, is one of the largest international nature conservancy organisation worldwide. It was founded in Switzerland in 1961 as World Wildlife Fund. . The WWF wants to halt the worldwide destruction of nature and create a future where humanity lives in harmony with nature. The WWF stands up for: conserving ecological diversity, the sustainable use of natural resources, and the reduction of pollution and harmful consumer behavior. Over the years the areas of expertise have grown from pure species preservation: Now general topics of protecting the environemnt and climate change are on the agenda of the WWF.

Deutsches Rotes Kreuz (German Red Cross)

The German Red Cross is committed to life, health, welfare, protection, peaceful coexistence and the dignity of all people. All people in need have the same entitlement to assistance, without regard to nationality, race, religion, sex, social status or political conviction. The DRC offers help solely on the degree of need and the urgency of the assistance. The voluntary assistance is used to restore the powers of self-help for people in need. The DRC will offer all services that are necessary to fulfill our mandate. They should meet the highest standards and quality requirements. In fulfillment of our own objectives , the DRC cooperates with all institutions and organizations in state and society that can be helpful and / or have similar objectives. However, we are preserving our independence. We respond to competition from others by improving the quality of our assistance, but also its economic viability.

C Screenshots of experimental stages – Not for publication

Figure 2: The real effort task

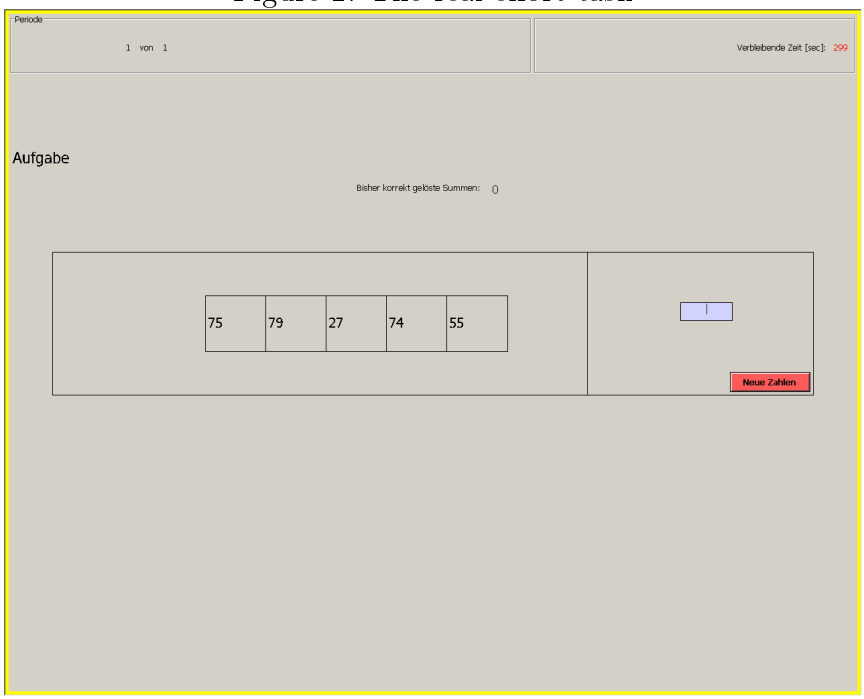
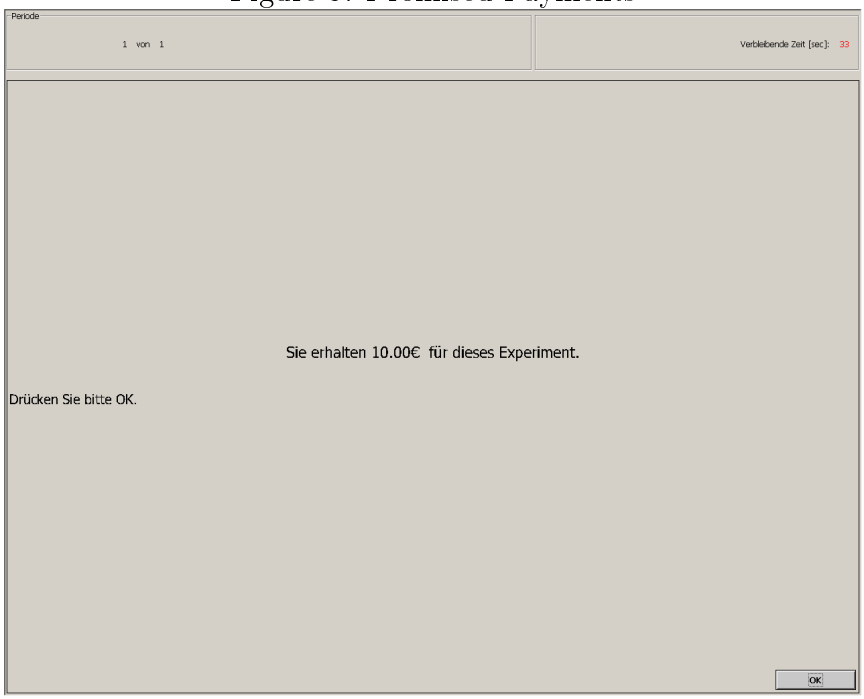


Figure 3: Promised Payments

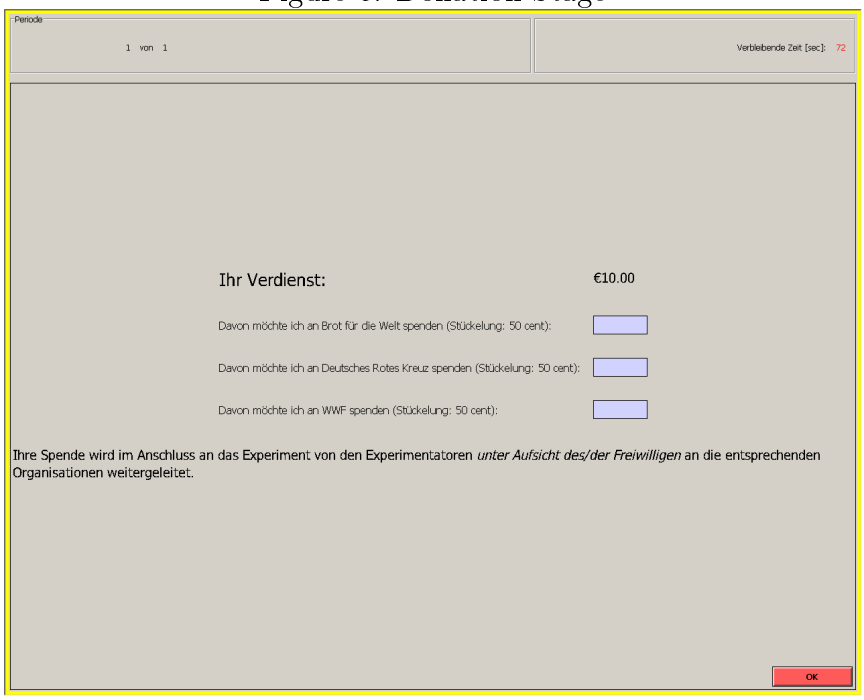


You obtain 10.00€ for this experiment. Please press OK.

Figure 4: Cash Payments

You obtain 10.00€ for this experiment.
The volunteer will now go to the outer part of the lab to get the envelopes and distribute them. Please remain seated in the meanwhile and do not talk to your neighbors.
As soon as you receive the money, please count it.
Press OK after you have counted the money and signed the receipt.

Figure 5: Donation Stage



Your earnings: €10.00

[Donation Decisions]

Your donation will be transferred under the supervision of the volunteer to the respective organizations.

D Robustness checks – Not for publication

Table 4 presents the results of the regression on the proportion of income donated, using Papke-Wooldridge estimator for fractional response variables. The coefficient on cash payments is still significant and negative, while the coefficient on performance pay loses significance when adding additional controls, but the coefficient itself does not change.

In table 5 we run the regressions from 3 split by charity. We find a similar pattern for the cash treatments over all charities, although the coefficients on cash are are only significantly negative for WWF and DRK.

Table 4: Ratio of income donated (Papke-Wooldridge estimator)

	(1) Base	(2) Exp. Controls	(3) Add. controls
Share of earnings donated (total)			
Pay cash	-0.81**	-0.75**	-0.81**
	(0.33)	(0.33)	(0.32)
Pay by performance	-0.57*	-0.50	-0.54
	(0.34)	(0.35)	(0.36)
Cash × performance	0.36	0.26	0.37
	(0.50)	(0.51)	(0.52)
Third charity		0.25	0.24
		(0.24)	(0.24)
Stake: 7.5		-0.67**	-0.62**
		(0.30)	(0.30)
Stake: 10		-0.79***	-0.81***
		(0.30)	(0.29)
Female			0.57**
			(0.26)
Constant	-0.98***	-0.74***	-1.05***
	(0.21)	(0.26)	(0.29)
Combined Coefficients			
Cash+perform+cash × perform	-1.02	-0.98	-0.98
	(0.34)	(0.33)	(0.33)
Observations	190	190	190

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 5: Poisson regressions by charity

	(1) Brot f.d. Welt	(2) WWF	(3) DRK
main			
Pay cash	-0.44	-0.72**	-1.39*
	(0.27)	(0.28)	(0.59)
Pay by performance	-0.41	-0.24	-0.76
	(0.28)	(0.24)	(0.52)
Cash × performance	0.22	0.036	1.74*
	(0.44)	(0.43)	(0.83)
Third charity	-0.11	-0.049	
	(0.21)	(0.19)	
Stake: 7.5	-0.19	0.11	-0.63
	(0.26)	(0.23)	(0.55)
Stake: 10	0.088	-0.11	0.38
	(0.24)	(0.25)	(0.39)
Female	0.73**	0.12	0.62
	(0.23)	(0.20)	(0.41)
Constant	-0.77**	-0.24	-0.99*
	(0.27)	(0.23)	(0.50)
Combined Coefficients			
Cash+perform+cash × perform	-0.62	-0.93	-0.41
	(0.30)	(0.30)	(0.45)
Observations	190	190	94
R^2			
Pseudo R^2	0.050	0.039	0.105

Standard errors in parentheses

+ p<0.10, * p<0.05, ** p<0.01

Marginal effects reported. Constant dropped.

Session and time of day effects

As the table below illustrates, our treatments are also not perfectly balanced over time:

Table 6: Schedule of sessions and treatments

Date (m/d/yr)	Time	Subjects in treatments			
		Account/Random	Cash/Random	Account/Performance	Cash/Performance
10/27/08	9:50	10			
	11:49	10			
10/28/08	9:51			10	
	11:43				9
02/25/09	9:30	18			
	11:56			15	
02/26/09	11:25		18		
03/02/09	10:42				18
	12:08		10		
10/30/09	10:11		18		
	11:33				18
	13:07	18			
	14:38			18	

To test for session-specific effects, we report regressions below standard errors clustered by session, and controls for time-of-day and time-of-year effects; our results are robust to all of these. We divide our session times into three categories: 9.30-10:30 am, 10:31-12 noon, and afternoon (12:01-14:38pm). The regressions below control for all of these "time dummies", and they are not jointly significant. We also divide our sessions into four "sets": those run in October of 2008, those run in February and March of 2009, and those run in October 2009. Again, these dummies are not jointly significant in any of the regressions below.

Selection on performance differences

It is conceivable that those who do better on the task earn more, and these people might be less generous on average. This "selection" might cause us to falsely attribute this to a desert effect – when we compare the high earners to those with high randomly-assigned endowments, the former would tend to give less. As evidence against this, we find the same effect across all stake sizes (results available by request). As payments when players get a tie score are randomly assigned, we can also control for the absolute level of performance (regression tables available by request). Adding a control for the "number of correctly solved sums" to the regressions in table 3 barely alters any of coefficients, and the coefficient on this control variable is tiny, insignificant, and tightly bounded around zero (e.g., if we add this variable to the first column of table 3 its coefficient has a 95% confidence interval of -0.08, 0.05).

This result supports Cherry et al. (2005), who write "the selection of high and low endowment dictators in the earned and windfall treatments differ (exam score versus random),

Table 7: Poisson and OLS regression on total donations

				Gender contr.
	(1)	(2)	(3)	(4)
	Psn.	OLS	Psn.	OLS
Pay cash	-0.69**	-.78**	-0.70**	-0.82**
	(0.16)	(0.18)	(0.15)	(0.18)
Pay by performance	-0.42	-0.54	-0.43+	-0.59+
	(0.26)	(0.31)	(0.22)	(0.29)
Cash × performance	0.26	0.38	0.33	0.48
	(0.34)	(0.29)	(0.33)	(0.29)
Third charity			0.23	0.26
			(0.22)	(0.25)
Stake: 7.5			-0.06	-0.05
			(0.31)	(0.34)
Stake: 10			0.10	0.14
			(0.27)	(0.29)
Female			0.51*	0.55+
			(0.24)	(0.27)
Time dummies <Chi-sq>/[F-test]	<0.23>	[0.12]	<0.54>	[0.26]
{P-value of test}	{0.89}	{0.89}	{0.76}	{0.77}
"Set" dummies <Chi-sq>/[F-test]	<0.85>	[0.48]	<1.41>	[0.59]
{P-value of test}	{0.65}	{0.63}	{0.50}	{0.57}
Observations	190	190	190	190
R^2		0.048		0.079
Pseudo R^2	0.034		0.057	

Standard errors in parentheses, reported clustered by session for OLS.

+ p<0.10, * p<0.05, ** p<0.01 for tests using standard errors clustered by session *for all columns.*

All regressors are dichotomous (0,1) variables, dy/dx for discrete change of dummy variable reported .

which may raise questions of sample selection, but previous research using this selection method has found this is not a significant concern."