Journal of Economic Psychology 59 (2017) 87–98

Contents lists available at ScienceDirect

Journal of Economic Psychology

journal homepage: www.elsevier.com/locate/joep

Spillovers from targeting of incentives: Exploring responses to being excluded

Francisco Alpízar [a], Anna Nordén [b,c,*], Alexander Pfaff [d], Juan Robalino [a]

[a] Environment for Development Center for Central America, CATIE, 7170 Turrialba, Costa Rica
[b] Department of Economics, University of Gothenburg, Vasagatan 1, PO Box 640, SE 405 30 Gothenburg, Sweden
[c] Department of Physical Geography and Ecosystem Science, Lund University, Sweden
[d] Duke University, Sanford School of Public Policy, Durham, NC 27708, USA

A R T I C L E I N F O

Article history:
Received 10 March 2015
Received in revised form 30 January 2017
Accepted 3 February 2017
Available online 14 February 2017

JEL classification:
C91
D03
H4

PsycINFO classification:
2260
2360
3040

Keywords:
Monetary incentives
Targeting
Spillovers
Economic experiments
Behavioral economics

A B S T R A C T

A growing set of policies involve transfers conditioned upon socially desired actions, such as attending school or conserving forest. However, given a desire to maximize the impact of limited funds by avoiding transfers that do not change behavior, typically some potential recipients are excluded on the basis of their characteristics, their actions or at random. This paper uses a laboratory experiment to study the behavior of individuals excluded on different bases from a new incentive that encourages real monetary donations to a public environmental conservation program. We show that the donations from the individuals who were excluded based on prior high contributions fell significantly. Yet the rationale used for exclusion mattered, in that none of the other selection criteria used as the basis for exclusion resulted in negative effects on contributions.

1. Introduction

Public policies often use incentives to improve public health and education or to lower poverty and environmental degradation by promoting behaviors more valuable to society than to the individual decision maker.[1] Yet in many cases, given limited resources, incentives are not offered to all potential recipients. One reason for exclusion is that authorities want to avoid

* Corresponding author at: Centre for Environment and Sustainability (GMV), Chalmers University of Technology and University of Gothenburg, Aschebergsgatan 44, 3rd floor, 412 96 Gothenburg, Sweden.

E-mail addresses: falpizar@catie.ac.cr (F. Alpízar), anna.norden@chalmers.se (A. Nordén), alex.pfaff@duke.edu (A. Pfaff), robalino@catie.ac.cr (J. Robalino).

[1] Some policies focus on outcomes for individuals, e.g., "conditional cash transfers" based on vaccination or school attendance (Fizsbein & Schady, 2009; Miller & Babiarz, 2013). Varied terms describe them including "pay for performance" (Miller & Babiarz, 2013) and "performance-based payments" (Pattanayak, Wunder, & Ferraro, 2010; Pfaff, Robalino, Sanchez-Azofeifa, Andam, & Ferraro, 2009).

http://dx.doi.org/10.1016/j.joep.2017.02.007
0167-4870/

transfers that do not change behavior, e.g. paying a mother to bring her kids to regular health checks when she planned to do so even in the absence of the incentive. Little if anything has been said about the effects of excluding individuals from such targeted incentives.

This paper uses a laboratory experiment to study individuals *excluded* from receiving a new incentive. We explore whether exclusion affects prosocial behavior and whether the rationale for exclusion matters.[2] When an incentive targets those who have not previously acted prosocially in an effort to induce additional prosocial behavior, we hypothesize that those who did act prosocially in the past – and thus were excluded from the incentive for that very reason – may deem their exclusion to be unfair and react negatively.[3] Such responses could offset gains obtained from establishing the incentive.

Potential reactions of this type generally are not considered within the design of incentives. One reason is that they are assumed away in economic theory since, for the excluded individuals, relevant prices and incomes are unchanged. Yet a lack of changes in relevant prices and income does not rule out negative reactions based upon, for example, concerns about a policy's fairness. If those excluded from access to an incentive may in fact feel that their exclusion was unfair then, as a result, they may intentionally shift behavior in ways that run counter to the incentive's aims.

Envy, spite, and inequity aversion (Bolton & Ockenfels, 2000; Dur & Glazer, 2007; Fehr & Schmidt, 1999; Goel & Thakor, 2005; Pillutla & Murnighan, 1996; Straub & Murnighan, 1995) also could explain negative reactions to exclusion. Rabin (1993) finds that we treat nicely those who treat us fairly but treat poorly those who do not treat us well – in line with documentation of reciprocity (for theory see Falk & Fischbacher, 2006). Preferences for fairness are suggested by relatively equal resource divisions in dictator games and by costly punishment of those who propose unequal resource allocations (Dawes, Fowler, Johnson, McElreath, & Smirnov, 2007; Fehr & Schmidt, 2006).

Yet shifts in prosocial behavior after exclusion could, instead, be due to changes in the social value of prosocial behavior. If behavior is driven by a desire to be perceived by others as altruistic, then a subsidy to donations could spoil their signaling value (Ariely, Bracha, & Meier, 2009; also see Akerlof, 1980; Andreoni & Bernheim, 2009; Bénabou & Tirole, 2006; Ellingsen & Johannesson, 2008; Hollander, 1990). We note that this would be independent of the selection rule employed. Generally, there are many possibilities for what exactly drives responses to exclusion. In this paper, we explore varied responses to exclusion without learning all of the underlying reasons for them.

To that end, we implemented a laboratory experiment where Costa Rican university students made three rounds of real monetary donations to a public environmental conservation program.[4] The experiment began with the well-known game in which one player acting as dictator is given money to allocate between herself and another player, the receiver (see Forsythe et al., 1994; Hoffman, McCabe, & Vernon, 1996; Kahneman, Knetsch, & Thaler, 1986). In our experiment, the recipient was not present in the room but instead was a conservation program widely perceived to pursue social goals (we follow Carpenter, Connolly, & Myers, 2008; Eckel & Grossman, 2003 in using a program as the recipient).

Our experiment compares treatments involving selection and exclusion to an incentive, with a control treatment. In selection treatments, subjects received money endowments in each of three rounds, and the following procedures were followed. Round 1 was a simple standard dictator game. Round 2 then introduced a regulator, randomly selected from the participants, who chose a selection rule that determined who would receive the incentives in Round 3. The regulator received a payoff that rose with the amount of donations (without the regulator actually doing anything). The presence of a regulator who chooses the selection rule, and enjoys donations, allows excluded actors who feel they are treated unfairly to "punish" the regulator by reducing their donations. Importantly, the subsidy is unexpectedly introduced and paid only in the last round. This is done to rule out any dynamic and strategic behavior that could result from expectations about future payments.

To determine who would get the subsidy, the regulator could use one of three selection rules: *additionality* – selecting those who made 'low' (i.e., below a threshold) Round 2 contributions to the public good; *reward* – selecting those who made high (above threshold) Round 2 contributions; and *random* – selecting the individuals with access to the incentive based solely upon a lottery.[5]

In the control treatment, no incentive was ever introduced. Subjects played three rounds with the following procedures. Round 1 was again just a simple standard dictator game. Rounds 2 and 3 then added a regulator but with no subsidy, i.e., no choice actually being made by the regulator. We compare the outcomes of this control treatment with those of the selection treatments using a difference-in-difference approach implemented in regression analysis. We do this both for absolute and relative changes in contributions. Thus, we examine the changes in contributions for each individual and test whether the observed individual change is partly due to our treatments. The regression analysis allows us to control for socioeconomic characteristics and "subject type", thereby allowing the estimation of the effects of exclusion and inclusion without potential biases due to, for instance, characteristics of the individuals or the 'experimenter effect' (see, e.g., Hoffman et al., 1996; Zizzo, 2010).

We provide empirical evidence that stakeholders *excluded* from monetary incentives may act less prosocially than before the incentive was introduced – even without changed prices or incomes for those excluded. Equally important, this exclu-

[2] We use the term "prosocial" to describe types of behavior that would not be predicted by the narrow definition of a purely selfish *homo economicus* since they provide some benefit to others at some form of cost to oneself.

[3] Studies in psychology and economics show that even those receiving a monetary incentive for an action considered prosocial may exhibit less pro-social behavior (via "motivational crowding out"), compared with when the action is driven by intrinsic motivation (Deci, Ryan, & Koestner, 1999; Frey, 1993, 1994; Gneezy & Rustichini, 2000; James, 2005).

[4] The program, Bosque Vivo, helps to conserve forest ecosystems in Costa Rica, the location for our experiments.

[5] These labels were not used with the participants in order not to generate any signals about expected behavior.

sion impact depended upon the selection rule that was used for targeting the incentive, i.e., to determine exclusion. Targeting the incentive to those who had acted less prosocially in previous rounds had two significant impacts: it *increased* the donations in this group, as intended; but it also *reduced* donations by those excluded, i.e., the higher initial contributors.

In contrast, other targeting rules yielded no negative spillovers from the incentive. Rewarding past prosocial behavior had no such effect, nor did random selection for the incentive. The reward rule, which steers incentives towards those with high prior prosocial contributions in previous rounds, also yielded no increase in contributions by selected subjects.

These results are relevant for policy design, as exclusion from incentives is common in policy. Programs have limited funds, large pools of potential recipients and often, in addition, mandates to maximize the impacts of those funds by paying recipients only if that will change key behaviors. Conditional cash transfer (CCT) programs (see Fizsbein & Schady, 2009) are a classic example in which, due to budget constraints and political pressure, agencies target those who on their own would not invest in education or vaccination. For them, the incentive could improve the outcomes. Another common policy now, in the environmental policy field, is payment for ecosystem services (PES). Again, budget constraints and political pressure can lead agencies to opt for selection rules favoring those likely to deforest while excluding those who have a history of protecting their forest.

Both examples describe the use of ad hoc versions of our additionality rule: selecting those who made below threshold contributions to the public good in the past. Still, most practitioners in charge of CCTs and PES programs in the developing world are reluctant to implement such rule, and use variations of either our reward rule or random selection within a prioritized population.

The remainder of the paper is as follows. Section 2 describes our experimental design, i.e., the modified dictator game, as well as the subject pool. Section 3 presents our findings. Section 4 concludes the paper and considers the implications for the design of targeted incentives policies.

2. Experimental design

In this section, we describe the basic experimental structure, the treatments and the analytical procedures we employ to test for effects of excluding subjects from an incentive to donate. As seen in Fig. 1, our experiment has selection treatments, where individuals receive or are excluded from an incentive based on a selection rule; and a control treatment, where no incentive is provided.

Our experiment is based on dictator games. The dictator gets 10 tokens, each worth 1000 colones,[6] to allocate between herself and a public program for forest conservation, i.e., a public good (using a public recipient follows Carpenter et al., 2008; Eckel & Grossman, 2003). At the end of each session, all contributions to the program were actually made anonymously, on site, via the internet, in order to make the credibility of the experimental structure as clear as possible.

Subjects allocated their endowments in each of three rounds. After each round, they were informed about their payoffs. At the beginning of the experiment, they were informed that one of those three rounds would be randomly selected and payments would be made based on behaviors in that round. Hence their experimental endowment was kept constant throughout the three rounds.

For each treatment, the first round was a dictator game. Before the second round, regulators were randomly selected, with one regulator per nine dictators.[7] The regulator remained anonymous to the dictators and, once chosen, remained in their roles. The regulator's payoffs depended on the total contributions made to the conservation program by all of the nine dictators with whom she was associated.[8] Regulators were paid from funds additional to the dictators' endowments, keeping the funds potentially going to the program constant across rounds. Regulators' payoffs give dictators a mechanism to "punish" or "reward" any regulator, via decreases or increases in their donations.

At the beginning of Round 2, all subjects learned about the presence of a regulator. They also learned that the regulator played no role in Round 2. Yet while senders' earnings are not affected by the introduction of a regulator, the regulator's earnings are a function of the senders' choices. With social preferences, this could affect contributions. A comparison of Round 1 and Round 2 allows us to test whether the mere presence of a paid regulator, even without a role that can affect senders' earnings, influenced the dictators' behavior. At the beginning of the third, and last, round, the regulator is then invited to choose a selection rule that determined which subset of dictators receive the subsidy. By design, the regulator has information about his potential expected payment in Round 2 (albeit not the distribution of contributions itself) before choosing the selection rule that would be used in Round 3. In this way, the regulator should make a value judgment regarding what rule she likes more,[9] and what rule is, according to her, the best given her expected payoff in Round 2. The other sub-

[6] The exchange rate at the time of the experiment was 500 colones/USD. Each dictator got 20 USD, a substantial amount for a Costa Rican university student (about five lunches at the cafeteria). High stakes were used to increase saliency, though Kocher, Martinsson, and Visser (2008) did not find significant stake effects in a study of contribution to public goods.

[7] Exceptions were made in five of the sessions due to less than thirty subjects participating.

[8] Specifically, the regulator's payoff equaled the average of all donations given by the dictators she regulated.

[9] Thus, the rules chosen likely reflect the regulators' preferences: 41% chose the additionality rule, 41% chose the reward rule, and 18% chose the random rule. Regulators who contributed more than five tokens in the first round, i.e., before being chosen as regulator, tended to prefer the reward rule (p-value = 0.07; chi-square test). In other words, regulators who give a lot to start with tend to prefer the reward rule. This shows that regulator preferences played a role in the choice of selection rule, precisely as we had intended.

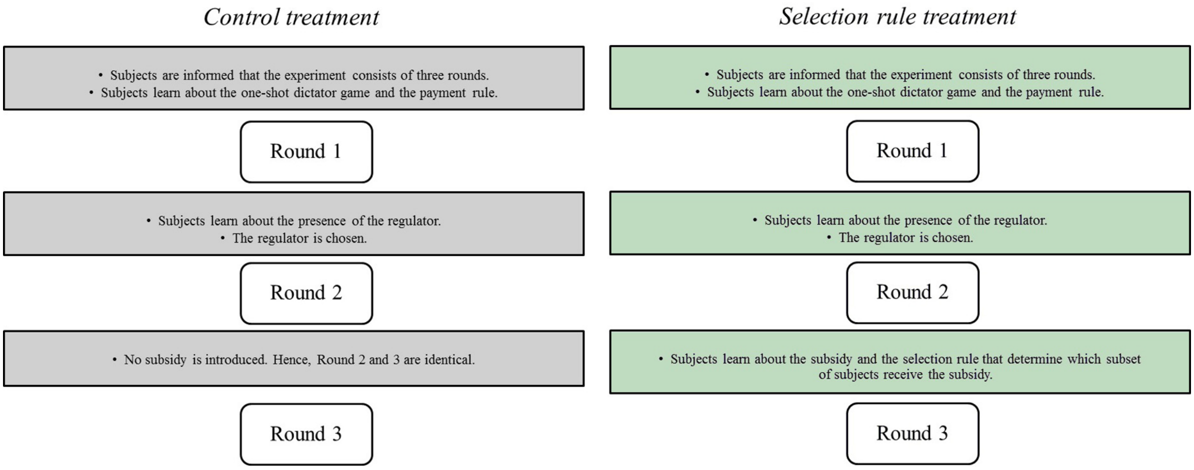

Fig. 1. Experimental design.

jects are aware of this, we should say, such that they could react to the selection rule using the knowledge that the regulator chose the rule with a reason.

Importantly, the subjects in selection treatments learned about the subsidy to encourage donations only after Rounds 1 and 2, prior to the last round. This rules out strategic behavior by the dictators, as they knew Round 3 was to be the last round (our experiment is not designed to explore dynamic strategic behavior). Individuals in the control play Round 3 without anybody receiving a subsidy.

2.1. Selection rules

The main goal of this paper is to examine how those excluded from the subsidy react. Three selection rules were implemented in the selection rule treatments. The first, which here – but not in the script – we call the *additionality rule*, selected subjects who gave two or fewer tokens (≤2000 colones) in Round 2 for the subsidy. All others were excluded. Those selected faced an incentive equal to 50% of their donation in Round 3, paid to them after they donated.[10] To ensure that donations went to the conservation program, we funded the subsidy with funds above and beyond the endowments. As we wanted subjects to link the subsidy to their donation, i.e., not see it as matching funds (List & Lucking-Reiley, 2002), the subsidy refund was handed back to the dictators (unlike in a matching setup). This may bring us closer to the operational feel or framing of incentive programs (see Eckel & Grossman, 2003 concerning the importance of framing).

In dictator games with similar subjects, average giving is around 20% of endowments (see Camerer, 2003 for a review), hence our threshold of 2000. An *additionality rule* targeting those who did not contribute much in the absence of focused incentives is a standard idea for programs that are aiming specifically to raise contributions above the giving that would occur without any incentive, i.e., are aiming for "additionality" (see Angelsen, 2008; Rawlings & Rubio, 2005).

Under the second rule, here called the *reward rule*, the subjects who contributed more than five tokens (i.e., 6000–10,000 colones) to the conservation program in Round 2 were selected for the subsidy. The rest of the subjects were excluded. Thus, those who contributed over 50% of their Round 2 endowments were rewarded with the same incentive as was just described above.

Both the *additionality* and the *reward* rule use past behavior to decide whom to exclude. The third rule, here called the *random rule*, selected subjects using a random draw with a 50% chance of each subject being selected for the incentive. This de-links selection from prior behavior, which allows us to test whether selection and exclusion *per se* have the same effect on behavior as selection based on past behavior. The rationale for the selection or exclusion may well matter.

Finally, we also implemented Round 3 sessions with a regulator but without any subsidy. Thus, nobody was selected and, consequently, also nobody was excluded. These sessions will be used as the *controls* for comparing their Round 3 minus Round 2 changes to selection treatments.

Participants knew not only whether they were selected or rejected for the subsidy before making their decisions in Round 3 but also why, i.e., the basis for selection or exclusion. To eliminate strategic behavior, subjects were again reminded that Round 3 was to be the last round.

Again, both the additionality and the reward rule are based on behavior in Round 2. Thus, if we had continued the experiment for more rounds, an agent excluded in Round 3 due to having contributed a lot in Round 2 might be tempted to lower

[10] Thus, if the dictators' payoff from giving G is $(10 - G)$ without incentives, now it is effectively $(10 - G/2)$. Given the timing or mechanics of giving, perhaps the payoff would be perceived as the equivalent $(10 - G + P)$, where $P = G/2$.

contributions in Round 3 to qualify for the subsidy in Round 4. Our design eliminates such strategic incentives and, in practice, most CCTs and PES programs define a cut-off date and actually check past behaviors, so as to forbid strategic responses to the selection rule. Given this, we believe our experiment fits well with real practice.

Finally, our subjects had no information on the share of dictators excluded from the subsidy. That is likely to be a realistic feature. It also helps to focus our experiment on the selection rule itself and less on relative comparisons across people, e.g., on how many others suffer exclusion, a piece of information that in itself might change the perception of the fairness of the decision.

In the control treatment, no subsidies are offered. This allows us to compare reactions by those excluded from the subsidy, for each selection rule, with change in behavior over time by those not excluded from any subsidy, controlling for contributions in Rounds 1 and 2 (Fig. 1).

2.2. Implementation procedure

We conducted a paper-and-pen experiment with students at the University of Costa Rica.[11] The students were paid 2500 colones (about 5 USD) simply for participating in the experiment. We recruited the subjects by distributing flyers and then had them sign up on a participation list for each session on a first-come, first-served basis. Each session had room for up to 30 subjects. All subjects had to show a university ID in order to ensure that they had not participated before.

The experiment was conducted with 438 students in total (392 dictators and 46 regulators). We conducted 16 sessions, all at the University of Costa Rica. Each lasted about 1 h and involved 15–30 students. Instructions were given orally, using PowerPoint to make them clear and easy to follow.[12] Before a session started, the students were informed it would last about 1 h. They were also asked for their informed consent and were given the option to leave. The decisions were confidential and the rooms were prepared so that all of the subjects were seated separately. At the end of the experiment, all of the subjects were asked to complete a questionnaire.

2.3. Analytical procedure

To start, we explore whether the mere introduction of a regulator affected donations. We ran two sessions repeating Round 1 without a regulator[13] then compared the difference between rounds with the difference that arose between Round 2 with a regulator and Round 1 without a regulator. We find no significant effect upon these differences from introducing a regulator.[14] Still, we focus on the change between Rounds 2 and 3, which have in common the regulator within the treatment.

We run OLS regressions for absolute (relative) changes in individual contributions between Rounds 3 and 2, for each of these selection rules, always using as a baseline for comparison the absolute (relative) difference in contributions between Rounds 3 and 2 in the control treatment. We are, *de facto*, using a difference-in-differences (DiD) approach. We do this for the selected and for the excluded individuals, separately, in order to explore the effect of the selection rules on both groups. Regressions allow robustness tests using socio-economic variables plus control for initial choices. The latter is especially important if one is concerned about systematic differences in the behavior in the baseline that might remain in the data, despite having chosen the subjects randomly for treatments or control. We use "contributions in Round 1" as an explanatory variable to control for "subject type", irrespective of whether the subject is assigned to treatment or control. This variable is indeed significant, i.e., past behaviors are good predictors of behavior in subsequent rounds. Yet, most importantly, this does not change the results. Our methodological approach is more refined than a comparison of means, particularly in the context of the large number of zero contributions that is expected in the experiment.

3. Results

Table 1 presents descriptive statistics for our sample. Total average contribution in Round 1 is 2.53 tokens for the standard dictator game with a public conservation program as the receiver. At 2530 colones (about 5 USD) that is about 25% of the endowment, in line with previous studies. In Round 3, regulators were free to choose the selection rule. They were informed about their own payoff in Round 2, i.e., the mean contribution of their group in Round 2. By design, our study is about excluded subjects reacting to the rule used to exclude them, so a priori we see no issue arising from the regulator strategically choosing a selection rule. Still, we tested whether there is a systematic pattern in rule selection resulting from the regulator's available information, i.e. whether those sessions in which the different rules were implemented differ systematically in terms of the contributions in Round 2. Comparing mean contributions in Round 2 between groups where different rules were used yields the following results in a Mann-Whitney test: additionality vs. other rules, p-value = 0.55; reward vs. other rules, p-value = 0.34; random rule vs. other rules, p-value = 0.64. Hence, we find no significant effect for any of the three rules, thereby ruling out strategic behavior. This is not surprising, as how to act strategic with respect to the choice of selection rules is rather unclear. The reward rule was chosen by the regulator in 31% of the cases, the addition-

[11] The protocol can be requested to the corresponding author.

[12] The original script, which is in Spanish, as well as the English translation of the script are available upon request.

[13] We find no differences in average contributions between Round 1 and 2 (p-value = 0.30; Wilcoxon test).

[14] Mann-Whitney test, p-value = 0.80.

Table 1
Descriptive statistics of the sample.

	Mean values (std. dev.)
Contribution in Round 1	2.53 tokens (2.13)
Share, control treatment	25% (0.44)
Share, additionality rule	31% (0.46)
Share, reward rule	31% (0.46)
Share, random rule	13% (0.33)
Male	65% (0.48)
Major in Economics & Business Administration	43% (0.50)
Major in Educational Studies	29% (0.45)
Major in Natural Science	20% (0.40)
Major in Social Science	5% (0.22)
Major in Law	3% (0.17)
Total years at the university	2.6 years (1.47)
Familiar with the conservation program	10% (0.30)

ality rule in 31% and the random rule in 13%, given that the control scenario was 25% of the observations by our design. Again by design, at no point did any of the dictators know the shares of subjects who were selected or excluded for an incentive. The average number of participants excluded from the subsidy was 42% for the additionality rule, with 87% for the reward and 52% for the random rules, respectively.

Table 2 presents average contributions in Rounds 2 and 3 for subjects in the control treatment and selection treatments. The controls are divided into subsamples satisfying the cut-off points within our selection rules. For each selection rule, we show contributions by selected and excluded individuals. The difference in the average contributions between the excluded and selected is due directly to the nature of each selection rule. For example, for the additionality rule, subjects are selected or excluded in Round 3 precisely because they contributed a little or a lot in Round 2.

We also show the absolute and relative differences between Rounds 3 and 2 for each treatment. It is only for the additionality rule that changes have the opposite sign for selected versus excluded. The same computations for the controls, even for those giving similar amounts in Round 2, yield substantially smaller differences. This might be driven by the fact that excluded individuals in both the additionality and the reward rule have more latitude to decrease contributions as their average contributions in Round 2 are higher compared to the control subgroups with the same cut-off points.

One could also argue that those who give a lot to start are systematically different than those who give little to start. Since two selection rules (i.e., additionality and reward) are based on those initial differences, we want to control for previous behavior to better isolate the effect of the rules themselves. We do this in several ways. We run separate regressions, for the additionality and for the reward rule, using only the subsamples of the controls which satisfy the cut-off points for the selection rules. We use absolute and relative differences in contributions as our dependent variable. We use Round 1 to control for the initial behavior because the Round 1 behavior seems to reveal something about the "contributor's type". We run separate regressions for excluded and selected.

For the Random rule, a small sample size seems to cause the relatively big differences between selected and excluded individuals in average contributions in Round 2, given that the excluded have a much larger share of zero contributors (39% > 17%). A larger sample most likely would not feature such different averages. Hence, conclusions should be drawn with caution for this rule.

3.1. Reactions of excluded individuals

Table 3 shows results for absolute differences in contributions between Round 3 and 2, for those excluded from the incentive, in four separate regressions. Table 4 shows the results for the relative difference in contributions. For the additionality rule, we observe a significant overall reduction of 0.84 tokens (840 colones) and a decrease in contributions of 18 percentage points compared to controls (column (1) in Tables 3 and 4).[15] This negative spillover occurs in the absence of any change in income or price. Interestingly, this type of spillover appeared to depend on the rationale for exclusion, since no significant spillovers are found for the exclusion via the reward or the random rule (columns (1) and (3) in Tables 3 and 4). This result shows that not receiving a monetary incentive for actions considered prosocial can, in fact, lead to less prosocial behavior.

The results do not change even when we controlled for student demographic characteristics (age, gender, economics major, year at the university, and if they knew about the conservation program that would be benefitted). The effect of being excluded by the additionality rule is still negative and statistically significant (column (2) in Tables 3 and 4) and the effect of being excluded by the reward rule is still not statistically different from zero (columns (3) and (4) Tables 3 and 4).

[15] A Mann-Whitney test confirms the significance of this result.

Table 2
Mean contributions in Rounds 2 and 3, and mean absolute and relative changes in contributions between these rounds by excluded and selected subjects.

	Control (all observations)	Controls for additionality rule (cut-off 2)		Additionality rule		Control for reward rule (cut-off 6)		Reward rule		Random rule	
	(No incentive)	(No incentive; Round2 > 2 tokens)	(No incentive; Round2 ≤ 2 tokens)	Excluded (Round2 > 2 tokens)	Selected (Round2 ≤ 2 tokens)	(No incentive; Round2 < 6 tokens)	(No incentive; Round2 ≥ 6 tokens)	Excluded (Round2 < 6 tokens)	Selected (Round2 ≥ 6 tokens)	Excluded (Lottery)	Selected (Lottery)
#obs.	99	24	75	50	70	97	2	107	16	26	24
Round 2	1.58 tokens	4.25 tokens	0.72 tokens	5.20 tokens	0.87 tokens	1.45 tokens	7.50 tokens	2.40 tokens	6.81 tokens	1.69 tokens	3.13 tokens
Round 3	1.70 tokens	4.21 tokens	0.89 tokens	4.70 tokens	1.77 tokens	1.54 tokens	9.50 tokens	2.62 tokens	7.31 tokens	2.00 tokens	4.46 tokens
Rd 3 − Rd 2	+0.12 tokens	−0.04 tokens	+0.17 tokens	−0.50 tokens	+0.90 tokens	+0.08 tokens	+2.00 tokens	+0.22 tokens	+0.50 tokens	+0.31 tokens	+1.33 tokens
$\frac{\text{Rd 3}-\text{Rd 2}}{\text{Rd 2}}$	+9.6%	+0.05%	+12.7%	−7.8%	+71.4%	+9.2%	+30.6%	+13.2%	+8.3%	+20.3%	+59.5%

Table 3
Results of OLS regressions for those excluded; dependent variable is the absolute difference in contributions between Rounds 3 and 2.

	(1)	(2)	(3)	(4)
Additionality rule	−0.84*** (0.32)	−0.77** (0.34)	–	–
Reward rule	–	–	0.22 (0.18)	0.19 (0.18)
Random rule	−0.07 (0.60)	−0.04 (0.61)	0.39 (0.29)	0.28 (0.30)
Contributions in Round 1	0.18** (0.09)	0.16* (0.09)	0.09** (0.04)	0.08* (0.05)
Student characteristics	No	Yes	No	Yes
Sample	Round 2 > 2	Round 2 > 2	Round 2 < 6	Round 2 < 6
#obs.	134	133	264	259
R^2	0.07	0.06	0.03	0.04

Standard errors in parentheses.
*** Significant at 1%.
** Significant at 5%.
* Significant at 10%.

Table 4
Results of OLS regressions for those excluded; dependent variable is the relative difference in contributions between Rounds 3 and 2.

	(1)	(2)	(3)	(4)
Additionality rule	−0.18** (0.08)	−0.17** (0.08)	–	–
Reward rule	–	–	0.07 (0.10)	0.06 (0.10)
Random rule	−0.03 (0.14)	−0.02 (0.15)	0.16 (0.16)	0.12 (0.16)
Contributions in Round 1	0.05** (0.02)	0.04* (0.02)	0.02 (0.02)	0.02 (0.02)
Student characteristics	No	Yes	No	Yes
Sample	Round 2 > 2	Round 2 > 2	Round 2 < 6	Round 2 < 6
#obs.	134	133	264	259
R^2	0.06	0.06	0.01	0.03

Standard errors in parentheses.
*** Significant at 1%.
** Significant at 5%.
* Significant at 10%.

3.2. Reactions of selected individuals

Table 5's OLS regressions shows whether the absolute differences in contributions between Round 3 and 2 for selected individuals result from the introduction of our selection rules, while Table 6 does the same for relative changes in contributions. We also run separate regressions, for the additionality and for the reward rule, with subsamples of the controls satisfying the cut-off points for the selection rules. Once again we define "contribution type" as per behavior in Round 1.

Columns (1) in Tables 5 and 6 show that receiving the incentive when selected on the basis of low contributions significantly increased the average contribution – by 0.74 tokens (740 colones) or 61 percentage points under additionality rule and 0.93 tokens (930 colones) or 44 percentage points under the random rule. The reward rule, though, yielded insignificant effects: individuals already contributing high amounts did not increase their contributions when receiving the incentive (see column (3) in Tables 5 and 6). Even when those high initial contributors get the incentive following selection under the random rule, their giving did not increase (column (3) in Tables 5 and 6).[16] The same results were obtained with student demographic controls (columns (2) and (4) in Tables 5 and 6).

3.3. Net effects of targeting

Our results suggest positive treatment effects for selected subjects under two selection rules: random assignment of the subsidy; and the additionality rule targeting the prior low contributors. Only one of those excludes subjects based on prior behavior. That rule – the additionality rule – is the only one showing negative spillovers: excluding the high prior donors reduced their donations.

[16] A Mann-Whitney test confirms the significance of this result.

Table 5
Results of OLS regressions for those selected, dependent variable is the absolute difference in contributions between Rounds 3 and 2.

	(1)	(2)	(3)	(4)
Additionality	0.74*** (0.22)	0.85*** (0.24)	–	–
Reward	–	–	−0.62 (0.61)	−0.75 (0.62)
Random	0.93*** (0.33)	0.96*** (0.34)	0.32 (0.43)	0.28 (0.44)
Contributions in Round 1	0.13** (0.06)	0.12** (0.06)	0.05 (0.10)	0.01 (0.10)
Student characteristics	No	Yes	No	Yes
#obs.	Round 2 < 3	Round 2 < 3	Round 2 > 5	Round 2 > 5
	185	182	112	111
R^2	0.11	0.13	0.02	0.07

Standard errors in parentheses.
*** Significant at 1%.
** Significant at 5%.
* Significant at 10%.

Table 6
Results of OLS regressions for those selected, dependent variable is the relative difference in contributions between Rounds 3 and 2.

	(1)	(2)	(3)	(4)
Additionality	0.61*** (0.15)	0.69*** (0.16)	–	–
Reward	–	–	−0.45 (0.41)	−0.53 (0.41)
Random	0.44* (0.23)	0.50** (0.23)	−0.04 (0.28)	−0.03 (0.29)
Contributions in Round 1	0.03 (0.04)	0.03 (0.04)	−0.05 (0.06)	−0.06 (0.07)
Student characteristics	No	Yes	No	Yes
#obs.	Round 2 < 3	Round 2 < 3	Round 2 > 5	Round 2 > 5
	185	182	112	111
R^2	0.09	0.12	0.04	0.10

Standard errors in parentheses.
*** Significant at 1%.
** Significant at 5%.
* Significant at 10%.

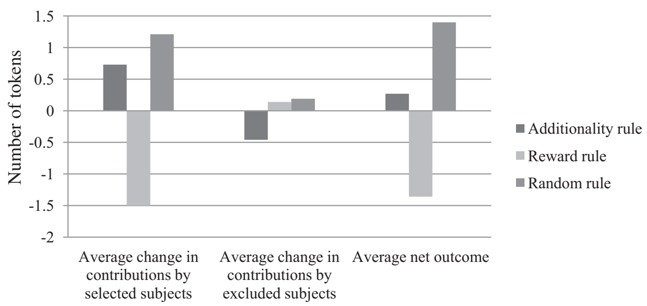

Fig. 2. Average net outcome comparing selection and exclusion effects. Average net outcome when comparing one selected and one excluded individual is not significant under the additionality rule yet is significantly negative for the reward rule but not significant for the random rule (Mann-Whitney tests p-value = 0.18, 0.01, and 0.12, respectively).

Fig. 2 (based on average change in contributions and the controls that satisfy the cut-off points for selection in Table 2) communicates the average change in contributions per person who is selected, the average change in contributions per person who is excluded, and net effect for each of the selection rules. The losses for the excluded counter the gains for the selected under additionality, as the average net rise of 0.27 tokens is not significant. This result shows that, in order to judge

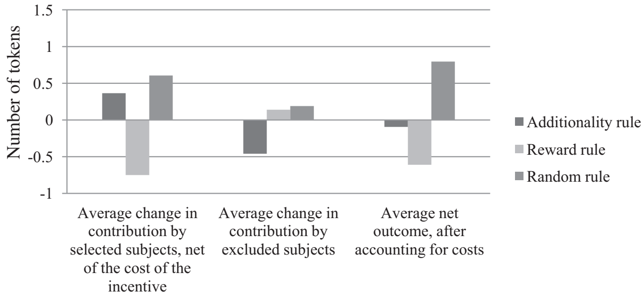

Fig. 3. Average net outcome comparing selection and exclusion effects including costs. Average net outcome when comparing one selected and one excluded individual and subtracting the cost of the incentive from contribution changes is not significant under the additionality yet significantly negative for the reward rules and significantly positive for the random rule (Mann-Whitney tests p-value = 0.36, 0.01, and 0.03, respectively).

total impacts of the rules typically used to assign incentives intended to promote prosocial behavior, scientists and policy makers really need to look at the reactions of both the selected and the excluded individuals.

There was no significant effect of exclusion for the other selection rules. Thus, the average net effect was negative for the reward rule and insignificantly positive for the random selection rule. Comparing net effects, randomizing performed better than additionality, yet the additionality rule performed better than the reward rule (p-value = 0.01 and 0.01, respectively; Mann-Whitney test).

From a program's point of view, the cost of paying the incentives also needs to be considered in correctly measuring the net gains of incentives. Fig. 3 repeats Fig. 2 with the average cost of the subsidy now subtracted from average contribution per person selected. The average cost of the subsidy is measured in tokens and calculated by dividing the average change in contributions of selected subjects by two. Now a net effect for the additionality rule is an insignificant loss of -0.10 tokens, while the net effect under the reward rule is significantly negative and we find a significant overall gain for the random rule.

The net effect of any selection rule could depend on the share selected or excluded. If almost all subjects are included, a few disgruntled individuals will not significantly reduce net gains. Yet if the share excluded is large, our results could imply that negative spillovers significantly reduce net gains from an additionality (low prior contributions) rule. Since our subjects were not informed about the share excluded but in actual policy settings they might be, and that, might affect spillovers, we cannot say what might happen if more (or fewer) people were excluded by an actual policy.

4. Conclusions

We provide evidence that stakeholders *excluded* from monetary incentives may choose to act less prosocially than before the incentive was introduced. This unintended effect of exclusion, or negative spillover from the incentive to the behavior of those not receiving it, which occurs even without any change in prices or income for the excluded, depended on the selection rule used for the incentive. Targeting an incentive to those who had acted less prosocially *increased* donations in that group, but it *reduced* the donations by those excluded, i.e., the highest prior contributors.

Neither rewarding past prosocial behavior nor randomly selecting subjects for incentives yielded negative responses among excluded individuals. The reward rule, which steers incentives to those with high prior contributions, yielded no increase in the contributions by those selected.

Our results suggest a tradeoff when targeting for additionality. A standard recommendation from an efficiency standpoint has been to give the incentive to those who can potentially increase contributions more as a consequence of the incentive. While we found that this indeed increased the contributions by those selected, here we provide evidence that there is a downside in terms of alienating those excluded when their exclusion is due to their high prosocial behavior.

Such results have important implications for the design of incentives and public policies. For instance, many programs have limited resources, such that those who receive incentives must be selected from a large pool of potential recipients, which implies the need to target or to exclude (noting the criteria of who is eligible and who is not may be well known). Further, when making decisions about the allocation of these resources, actors must consider the payoff from their use, i.e., whether they will be able to argue the uses they chose paid off in terms of shifts in behavior.

Previous literature has mostly focused on evaluating how different types of incentives increase contributions from those who receive the incentive (e.g. Eckel & Grossman, 2003, 2008; Karlan & List, 2006; List & Lucking-Reiley, 2002). As do we, they find that receiving rewards increase contributions. A large amount of research has also focused on the negative effects of eliminating incentives after having received them (Deci et al., 1999; Frey & Jegen, 2001; Meier, 2007). However, little has been said about the potential negative effects of rewards upon those individuals who do not receive incentives while knowing that others get them. As noted above, this situation of selection arises in a number of incentive policies and programs.

Our contribution is a focus upon those left out, with a test of whether the reason behind their exclusion matters. We show that negative effects arise in individuals that are left out from subsidies even if they had never received them. Depending on the rationale used for excluding individuals from the subsidy, the negative effects can be large enough to cancel out the gains of the subsidy.

The aggregate magnitude of such a negative effect likely depends on the share of participants excluded. Yet if the share of excluded individuals were very large, on the other hand the perception of being unfairly treated might be diluted, which could then result in less negative reaction. Further research is needed to explore this effect.

Another issue not explored in our paper is the 'social distance' between the agents involved. It is plausible that a larger distance magnifies or reduces negative reactions upon being excluded. This we also had to leave to future research. Given that we already varied the exclusion rationale, adding distance would have made our research design and data collection much more complex.

Finally, future studies certainly could explore alternative selection rules, as well as hybrid selection rules that following our results should try to maximize gains from selected participants without spoiling the behavior of those who did not initially need the incentive program to act prosocially. Our results show that our random selection rule, e.g., which was not linked to any past behavior, significantly increased contributions by those selected without alienating those excluded from the incentive. Other ways of delinking selection from past behavior could be explored as well.

Acknowledgments

We are grateful for valuable comments from two anonymous referees, Paul Ferraro, Peter Martinsson, Louis Preonas, Olof Stenman-Johansson, Mattias Sutter and participants in presentations at the AERE conference June 2012, the EAERE conference June 2013, and the EEA meeting August 2013. Funding from the Tinker Foundation, Inc., U.S for this project is gratefully acknowledged, as are funds provided by Paul Ferraro and financial support from Sida (Swedish development cooperation), Sweden and funding from CRED, the NSF DMUU center at Columbia University, U.S. All remaining errors are our own.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.joep.2017.02.007.

References

Akerlof, G. A. (1980). A theory of social custom, of which unemployment may be one consequence. *Quarterly Journal of Economics, 94*, 749–775.

Andreoni, J., & Bernheim, B. D. (2009). Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica, 77*, 1607–1636.

Angelsen, A. (Ed.). (2008). *Moving ahead with REDD: Issues, options and implications*. Bogor, Indonesia: CIFOR.

Ariely, D., Bracha, A., & Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review, 99*, 544–555.

Bénabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review, 96*, 1652–1678.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review, 90*, 166–193.

Camerer, C. (2003). *Behavioral game theory*. New York: Princeton University Press.

Carpenter, J., Connolly, C., & Myers, C. K. (2008). Altruistic behavior in a representative dictator experiment. *Experimental Economics, 11*, 282–298.

Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R., & Smirnov, O. (2007). Egalitarian motives in humans. *Nature, 446*, 794–796.

Deci, E., Ryan, R., & Koestner, R. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*, 627–668.

Dur, R., & Glazer, A. (2007). Optimal contracts when a worker envies his boss. *Journal of Law, Economics & Organization, 24*, 120–137.

Eckel, C. C., & Grossman, J. P. (2003). Rebate versus matching: Does how we subsidize charitable contributions matter? *Journal of Public Economics, 87*, 681–701.

Eckel, C. C., & Grossman, P. J. (2008). Subsidizing charitable contributions: A natural field experiment comparing matching and rebate subsidies. *Experimental Economics, 11*(3), 234–252.

Ellingsen, T., & Johannesson, M. (2008). Pride and prejudice: The human side of incentive theory. *American Economic Review, 93*, 990–1008.

Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior, 54*, 293–315.

Fehr, E., & Schmidt, K. (1999). Theory of fairness, competition, and cooperation. *Quarterly Journal of Economics, 114*, 817–868.

Fehr, E., & Schmidt, K. M. (2006). The economics of fairness, reciprocity and altruism – Experimental evidence and new theories. In *Handbook on the economics of giving, reciprocity and altruism*. Elsevier.

Fizsbein, A., & Schady, N. (2009). *Conditional cash transfers: Reducing present and future poverty* World Bank policy research report. Washington, DC: World Bank.

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior, 6*, 347–396.

Frey, B. S. (1993). Motivation as a limit to pricing. *Journal of Economic Psychology, 14*(4), 635–664.

Frey, B. (1994). How intrinsic motivation is crowded out and in. *Rationality and Society, 6*, 334.

Frey, B., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys, 15*, 589–611.

Gneezy, U., & Rustichini, A. (2000). A fine is a price. *Journal of Legal Studies, 29*(1).

Goel, A. M., & Thakor, A. (2005). *Optimal contracts when agents envy each other*. Unpublished.

Hoffman, E., McCabe, K., & Vernon, L. S. (1996). Social distance and other-regarding behavior in dictator games. *American Economic Review, 86*, 653–660.

Hollander, H. (1990). A social exchange approach to voluntary cooperation. *American Economic Review, 80*, 1157–1167.

James, H. S. (2005). Why did you do that? An economic examination of the effect of extrinsic compensation on intrinsic motivation and performance. *Journal of Economic Psychology, 26*(4), 549–566.

Kahneman, D., Knetsch, J. L., & Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *American Economic Review, 76*, 728–741.

Karlan, D., & List, J. A. (2006). *Does price matter in charitable giving? Evidence from a large-scale natural field experiment (No. w12338)*. National Bureau of Economic Research.

Kocher, M. G., Martinsson, P., & Visser, M. (2008). Does stake size matter for cooperation and punishment? *Economics Letters, 99*, 508–511.

List, J., & Lucking-Reiley, D. (2002). The effects of seed money and refunds on charitable giving: Experimental evidence from a university capital campaign. *Journal of Political Economy, 110*, 215–233.
Meier, S. (2007). Do subsidies increase charitable giving in the long run? Matching donations in a field experiment. *Journal of the European Economic Association, 5*(6), 1203–1222.
Miller, G., & Babiarz, K. S. (2013). *Pay-for-performance incentives in low- and middle-income country health programs (No. w18932)*. National Bureau of Economic Research.
Pattanayak, S. K., Wunder, S., & Ferraro, P. J. (2010). Show me the money: Do payments supply environmental services in developing countries? *Review of Environmental Economics and Policy, 4*, 254–274.
Pfaff, A., Robalino, J. A., Sanchez-Azofeifa, G. A., Andam, K., & Ferraro, P. (2009). Park location affects forest protection: Land characteristics cause differences in park impacts across Costa Rica. *B.E. Journal of Economic Analysis & Policy, 9*(2) (Contributions). Article 5. Available at: <http://www.bepress.com/bejeap/vol9/iss2/art5>.
Pillutla, M., & Murnighan, K. (1996). Unfairness, anger, and spite: Emotional rejections of ultimatum offers. *Organizational Behavior and Human Decision Processes, 68*, 208–224.
Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review, 83*(5), 1281–1302.
Rawlings, L. B., & Rubio, G. M. (2005). Evaluating the impact of conditional cash transfer programs. *World Bank Research Observer, 20*, 29–55.
Straub, P., & Murnighan, K. (1995). An experimental investigation of ultimatum games: Information, fairness, expectations and lowest acceptable offers. *Journal of Economic Behavior and Organization, 27*, 345–364.
Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics, 13*, 75–98.