

Ek, Claes

**Working Paper**

## Some Causes are More Equal than Others? Behavioral Spillovers in Charitable Giving

Working Paper, No. 2015:29

**Provided in Cooperation with:**

Department of Economics, School of Economics and Management, Lund University

*Suggested Citation:* Ek, Claes (2015) : Some Causes are More Equal than Others? Behavioral Spillovers in Charitable Giving, Working Paper, No. 2015:29, Lund University, School of Economics and Management, Department of Economics, Lund

This Version is available at:

<https://hdl.handle.net/10419/260167>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

Working Paper 2015:29

Department of Economics  
School of Economics and Management

# Some Causes are More Equal than Others? Behavioral Spillovers in Charitable Giving

Claes Ek

October 2015



**LUND**  
UNIVERSITY

# Some Causes are More Equal than Others? Behavioral Spillovers in Charitable Giving

Claes Ek

Department of Economics, Lund University

October 2, 2015

## **Abstract**

People can often contribute to prosocial causes by several means; for instance, environmentally friendly activities include sorting household waste, buying organic products, and donating to NGOs. Policy to encourage prosocial behavior is sometimes directed only towards a particular activity, however, and such policies may give rise to ‘behavioral spillovers’, affecting efforts on other prosocial activities. We examine such spillovers in the lab. In a version of the dictator game, experimental subjects could donate to two different real-world charities, and to simulate activity-specific policy, the relative productivity of the charities varied. We hypothesize, first, that an increase in the productivity of one charity will ‘crowd out’ contributions to the other charity. Second, we introduce several treatments to test whether crowding occurs even across (possibly very) dissimilar alternatives. Crowding-out occurs significantly in all cases, but the effect is systematically weaker, the more dissimilar are the charity alternatives. In our most dissimilar treatment, it is only half as large as when alternatives are very similar.

*Keywords:* charitable giving, dictator game, public goods, prosocial behavior

*JEL classification:* C91, D03, H41

# 1 Introduction

People who set out to champion some prosocial cause — be it with respect to their fellow man, their community, the environment, or some abstract ideal — are likely to find that there is more than one way to make a contribution. This is perhaps especially apparent with regard to the environment, for which people can (and do) recycle household waste, buy organic products, adopt vegetarian diets, use public transport instead of their cars, and so on. Yet even when there are several activities to choose from, public policy sometimes focuses on just one. For example, it is quite common for local governments to promote household recycling in various ways, such as by information campaigns, building drop-off sites, and through economic incentives. Behavioral economists have by now devoted a great deal of research towards how such policies affect the activity being targeted (for monetary incentives, see e.g. Bénabou and Tirole, 2006; information campaigns, see Nyborg, 2011). Less well understood is how such policies spill over to contributions made through *other* activities. This paper contributes to filling that gap by presenting experimental evidence on productivity-driven spillovers across alternative prosocial activities.

Behavioral spillovers could, in principle, take either sign: policy directed towards one activity could crowd-in or crowd-out effort on other activities (Dolan and Galizzi, 2015). As an example, encouraging households to recycle their waste could drive an increase or a decrease in sales of organic products. Existing field studies on policy-driven behavioral spillovers, while few in number, have tended to find evidence of crowding-out. Jacobsen et al. (2012) studied a green-electricity program in Memphis, Tennessee, finding that among households that participated at a minimal level (paying the smallest possible increment to fund alternative energy), consumption of (non-green) electricity increased after they had joined the program. Similarly, Tiefenbeck et al. (2013) found that a campaign to conserve water at a housing complex in Massachusetts, while leading to decreased water use, also drove increased electricity consumption. Crowding-out, then, undermines a given policy, and policy impacts will be overestimated by any cost-benefit analysis that ignores spillovers.

Psychologically, negative spillovers tend to be explained in terms of ‘moral licensing’ (Cain et al., 2005; Sachdeva et al., 2009; Blanken et al., 2015), which is the tendency of people who have just behaved prosocially to feel justified in subsequently relaxing their moral constraints. Consequently, if a policy is successful in increasing effort on the activity being targeted, effort on other activities may drop as a result. The effect can also run in the opposite direction (‘moral cleansing’). For example, in a recent series of experiments (Ploner and Regner, 2013), subjects could donate to charity by participating in a real-effort task; this was followed by a dictator game. Those who contributed relatively little in the initial game tended to be especially generous in the dictator game.<sup>1</sup>

A related conceptualization of policy spillovers is that the number of ‘good deeds’ people are capable of within a given domain is either close to being fixed (so policies simply cause reallocation within that fixed ‘budget’), or can increase as a result of policy. The present experiment, which was conducted in Lund, Sweden, allowed for both possibilities. We ran a

---

<sup>1</sup>Some evidence exists on an opposing ‘moral-consistency’ effect suggesting that one prosocial act can spur another. Gneezy et al. (2012) argue that if prosocial behavior is sufficiently costly, people who nevertheless engage in it are led to conclude that they are a prosocial ‘type’ and will subsequently try to act in accordance. Findings in the same spirit are presented by Brown et al. (2012), who found that the more people donated to the 2004 Indian Ocean tsunami, the more likely they were to donate *more*, rather than less, to charity in the future.

version of the dictator game where subjects could donate money to two real-world charities by first earning ‘points’ in a real-effort task. The production-and-allocation decision was repeated over multiple rounds where, to mimic the effect of activity-specific policy, the number of donated Swedish crowns (SEK) per point of each charity was systematically varied. In this setting, negative spillovers follow from standard relative-price effects, but positive spillovers can also arise if, as a result of increasing charity-specific exchange rates, subjects earn more points overall and allocate more to both charities.

Our experimental prediction is one of incomplete crowding-out across alternatives: as the productivity (exchange rate) of one charity rises, fewer points are spent on the other, but total donations rise. This hypothesis is derived from a multi-activity generalization (Ek, 2015) of a version of the public-goods model originally developed by Brekke et al. (2003). Theoretically, crowding-out arises through a process of ‘self-image management’ which is conceptually similar to moral-licensing accounts. Our hypothesis is clearly validated by the experimental data.

The main novelty of our experiment, however, derives not from the demonstration of negative spillovers but rather from an effort to also chart their *reach*. That is, do significant spillovers arise even across (very) dissimilar activities? The policy relevance of answering this question should be obvious. As for predictions, we suspect that spillovers are less pronounced for dissimilar alternatives. Behavioral economists have long argued that much of human decision making is characterized by ‘mental accounting’ (Tversky and Kahneman, 1981; Read et al., 1999; Thaler, 1999), one aspect of which involves dividing expenditures into finite budgets or categories (for an interpretation of mental accounts as psychological categories, see Henderson and Peterson, 1992). A defining characteristic of such budgets is that they are not fungible (Thaler, 1985; Heath and Soll, 1996; Thaler, 1999): people are reluctant to move resources from one account to cover expenses that belong to another account, such as from food purchases to entertainment. For this reason, price shifts may have greater impact within its relevant account than across accounts. In our case, if a person has charities A and B to choose from, and the productivity of A rises, she will be less prone to shift resources from B to A if these charities contribute towards different mental accounts than if they contribute to the same one.<sup>2</sup>

Hence, hypothesizing that the way people categorize alternatives may explain the magnitude of subsequent spillovers, we run our dictator-game variant in four between-subject treatments, isolating two dimensions along which charities may differ: (i) geographical scope, and (ii) the cause involved. In all treatments, one charity was UNICEF, a global organization concerned with the welfare of children. The other charity differed across treatments. In our baseline treatment, UNICEF was paired with a ‘Global Action Fund’ run by Save the Children in the US. Since both charities address children’s needs globally, they are similar along both dimensions. By contrast, another treatment paired UNICEF with the local office of the Swedish Society for Nature Conservation (SSNC), the largest Swedish environmental NGO. This pair is dissimilar along both dimensions. Finally, to isolate the effect of each dimension (local/global and cause involved), two intermediate treatments were included. In one, UNICEF was paired with the local office of Save the Children (local/global dimension); in

---

<sup>2</sup>Note that while economists have tended to consider mental accounting a bias, this type of behavior may be entirely consistent with the standard model of consumer choice, with mental accounts simply giving structure to the person’s substitution patterns. For example, it may well be rational (though, we suspect, rare) to view the distinct physical processes underlying climate change and eutrophication as represented by different public goods (and hence put them in different accounts) rather than by some monolithic ‘environmental good’.

the other, with WWF International, a major global environmental NGO (cause dimension). Our results show that while crowding-out occurs significantly in all treatments, the effect is systematically weaker, the more dissimilar are the charity alternatives. In the dissimilar treatment, it is only half as large as in the similar treatment.<sup>3</sup>

The structure of this paper is as follows. Section 2 briefly sketches our theoretical framework and presents the resulting predictions; Section 3 describes our experimental design as well as the particularities of each treatment. Section 4 presents our main results, devoting particular effort to arguing that observed treatment effects are in fact due to differences in charity-pair similarity rather than some confounder, such as the subjects’ relative preference for each charity. We also attempt, with mixed success, to link substitution patterns to psychological measures and other individual characteristics. Section 5 concludes the paper.

## 2 Theory and predictions

This section describes and motivates our experimental predictions. First, we sketch a simple theory of prosocial behavior and show that it predicts (incomplete) crowding-out. A similar version of the utility model analyzed here was considered in Ek (2015).

We consider a population of identical agents choosing the level of ‘effort’ to exert on two activities in order to contribute to a single public good  $G$ . Effort variables  $x_1, x_2$  are measured in time units; labor supply is fixed, so any time not spent contributing is devoted to leisure ( $L$ ) instead. It is assumed that  $G$  is large-scale enough that each agent approximates own impacts on the overall level of the public good by zero. However, agents are warm-glow altruists in the sense of Andreoni (1990), caring about the size of their own contribution. Each agent maximizes the additively separable function

$$U(x_1, x_2) = u(T - x_1 - x_2) + I(g(x_1, x_2, \theta_1, \theta_2)) \quad (1)$$

subject to  $0 \leq x_1 + x_2 \leq T$ , where  $T$  is total non-work time.  $u$  is the utility of leisure; it is concave and strictly increasing. Each agent’s personal public-good production  $g$  is strictly increasing in  $x_1$  and  $x_2$ . The ‘warm-glow’ component  $I$  is concave and strictly increasing in individual production  $g$ . Psychologically, warm glow has been interpreted as a favorable sense of social image or self-image as a generous person (Bénabou and Tirole, 2006), the desire to adhere to a moral norm (Brekke et al., 2003), or as the desire to help other contributors feel validated (Rotemberg, 2014). As the model includes no social interaction, we suggest that it is best interpreted in terms of self-image.

The fundamental trade-off in model (1) is standard: agents may either reduce effort and enjoy increased leisure, or increase efforts and, by extension, warm glow. However, the multidimensional production function implies an additional trade-off, between activities. The ideal contribution is common to both activities and measured in terms of production rather than effort, making this a consequentialist model where agents ask, ‘Am I making enough of a difference?’ rather than ‘Am I trying hard enough?’.

For all identical agents,  $\theta_1 > 0$  and  $\theta_2 > 0$  regulate the productivity of  $x_1$  and  $x_2$ , respectively. They are assumed to reflect activity-specific policies that may affect available

---

<sup>3</sup>In the simultaneous-decision framework we use in this paper, crowding-out becomes very similar to gross substitution across charities, with exchange rates as implicit prices. Another interpretation of our experiment, then, is that it maps subjects’ substitution behavior within the charity domain. Under this interpretation, we show that there are important systematic patterns to that behavior, one of which relates to the perceived similarity of charities.

technology as well as convenience and information, so long as these are relevant for how much is accomplished per unit of effort. For example, suppose the local government launches a campaign to facilitate household recycling by increasing the number of drop-off sites and by sending out leaflets on how to recycle. Both interventions may cause the relevant  $\theta$  to increase.<sup>4</sup> By contrast, ‘nudges’ such as highlighting average performance in the community will not.

To derive predictions for the effect of activity-specific policy, we will shift  $\theta_1$ , results for a shift in  $\theta_2$  being completely analogous. To keep things simple we will consider only the special case where each agent’s public-good production function is linear:  $g = \theta_1 x_1 + \theta_2 x_2$ .<sup>5</sup> Since utility-of-leisure  $u$  is determined by the sum of  $x_1$  and  $x_2$ , activities will then be perfect substitutes and it will be optimal to set  $x_1 = 0$  if  $\theta_1 < \theta_2$ , and  $x_2 = 0$  if  $\theta_1 > \theta_2$ . Note that when  $\theta_1 < \theta_2$ , marginal shifts in  $\theta_1$  does not affect the solution, including the optimal value of  $x_2$ . Together these facts imply that  $x_2$  is decreasing in  $\theta_1$ . This corresponds to negative-sign spillovers, so e.g. encouraging recycling would cause sales of organic groceries to drop.

The effect of  $\theta_1$  on  $x_1$ , however, is ambiguous. To see why, note first that when  $\theta_1 > \theta_2$ , at an interior solution for  $x_1$  the first-order condition is

$$-u'(T - x_1(\theta_1)) + \theta_1 I'(\theta_1 x_1(\theta_1)) = 0$$

Implicit differentiation with respect to  $\theta_1$  now yields

$$\frac{dx_1(\theta_1)}{d\theta_1} = -\frac{\theta_1 x_1 I'' + I'}{u'' + \theta_1^2 I''}$$

Here the numerator is indeterminate. Hence, although we would certainly expect  $x_1$  to increase as the threshold  $\theta_1 = \theta_2$  is passed,  $x_1$  may decrease beyond that point. It should be clear that the effect of  $\theta_1$  on the sum of efforts  $x_1 + x_2$  is similarly ambiguous. Intuitively, when productivity increases it becomes possible to do more with less effort, and this may be desirable since  $I$  measures self-ideal discrepancy only in terms of production.

Nevertheless, as we will now show, total contributions  $\theta_1 x_1 + \theta_2 x_2$  will tend to increase in line with  $\theta_1$ . Similarly to before, if  $\theta_1 \leq \theta_2$ , the optimal level of  $\theta_1 x_1 + \theta_2 x_2$  does not depend on  $\theta_1$ . When  $\theta_1 > \theta_2$ , since  $x_2 = 0$ , we can redefine (1) in terms only of  $\hat{x}_1 = \theta_1 x_1$ , then equal to total contributions; agents maximize

$$U(\hat{x}_1) = u(T - \hat{x}_1/\theta_1) + I(\hat{x}_1)$$

subject to  $0 \leq \hat{x}_1/\theta_1 \leq T$ . The first-order condition associated with an interior solution for  $\hat{x}_1$  is  $-(1/\theta_1)u' + I' = 0$ . Implicit differentiation now produces

$$\frac{d\hat{x}_1(\theta_1)}{d\theta_1} = \frac{u'' \hat{x}_1 - \theta_1 u'}{\theta_1^3 I'' + \theta_1 u''} > 0$$

---

<sup>4</sup>A requirement for policy to be truly activity-specific is that each parameter has no direct effect on the other activity’s productivity, so  $g''_{\theta_1 x_2} = g''_{\theta_2 x_1} = 0$ .

<sup>5</sup>As in Ek (2015), our predictions apply to any situation where  $g$  and  $I$  satisfy two conditions. First, the cross-partial derivative  $g''_{x_1 x_2}$  should not be too positive, ruling out situations that approximate a ‘weakest-link’ structure, e.g. an old-growth forest under pressure from unsustainable logging and pollution, with each representing a separate existential threat. Here the marginal benefit of addressing one threat will increase, the more action is taken on the other. Second,  $\partial^2 U / \partial x_1 \partial \theta_1 = I' g''_{x_1 \theta_1} + I'' g'_1 g'_{\theta_1}$  should be non-negative, corresponding to  $x_1$  and  $\theta_1$  being complements in a utility sense. Under these two conditions,  $x_1$  will also be increasing in  $\theta_1$ . Note that although sufficient, the conditions are not necessary for our predictions; indeed, the treatment in this section shows that  $x_2$  can be decreasing in  $\theta_1$  without  $x_1$  also being increasing.

assuming that we do not have  $I'' = u'' = 0$ . Thus, beyond the threshold  $\theta_1 = \theta_2$ , total contributions are increasing in  $\theta_1$ . In summary, we have the following predictions.

**Hypothesis 1.** When a particular prosocial activity is exposed to a directed positive productivity shock, agents decrease efforts on the other activity.

**Hypothesis 2.** Total contributions increase in response to an activity-specific positive productivity shock.

Finally, we predict that spillovers become attenuated for more dissimilar activities. As the model sketched above assumes that both activities contribute to the same public good, this hypothesis is not based on formal arguments, though as stressed in the introduction, it seems consistent with the tendency of people to engage in mental accounting (e.g. Thaler, 1999). While we have been unable to find any study within behavioral economics that directly manipulates similarity, results by Blackwell and McKee (2003) do suggest that if the alternatives are dissimilar in geographical scope, spillovers may be weak. In a variant of the public-good game, the authors let subjects allocate an endowment across three alternatives: a private account, a ‘group’ public good and a ‘global’ public good. Returns from the group PG were redistributed only to each participant’s particular group of four people, while returns from the global PG were redistributed to all subjects. Increasing the average per capita return on the global public good caused contributions to that good to rise, but did not decrease contributions to the group PG.

Also, while observing that behavioral spillovers sometimes arise even across dissimilar behaviors (Miller and Effron, 2010), psychologists have hypothesized that the effect may be stronger within a particular domain (such as environmental behavior) than across domains. For example, the recent meta-analysis by Blanken et al. (2015) test this hypothesis, though they find no evidence in favor of it. There are, however, at least three methodological differences between our experiment and the studies that they analyze. First, while our experiment involves simultaneous decision making, psychological studies are explicitly sequential, with an initial (licensing) stage followed by a subsequent decision stage. Second, the initial stage is often hypothetical, e.g. with subjects instructed to describe themselves using positive words. Third, even when the initial stage is choice-based, what is being varied is not the incentives to behave prosocially, but whether a prosocial choice (such as disagreeing with a racist statement) was actually made. Hence, the link between these studies and policy-driven spillovers tends to be rather indirect.

**Hypothesis 3.** Activity-specific positive productivity shocks drive less ‘crowding-out’, the less similar the activities are.

### 3 Experimental design

The experiment was conducted over nine afternoon sessions in a computer room at the Lund School of Economics and Management (LUSEM). Visually isolated cubicle-like spaces were created by putting up cardboard screens and roll-up banners. The experimental environment could fit 20 subjects, and was fully occupied in six of the sessions. Of the remaining three sessions, one had 19 participants, one had 14, and one had 7, with  $N = 160$  in total. Subjects



were initially recruited through email from the pool of Swedish-speaking students at LUSEM; for each session, additional participants were recruited on-the-fly on the LUSEM premises. Each session lasted approximately an hour, after which participants were paid a show-up fee of SEK 200 (about €20) in cash.

The experimental setting was a real effort-based variant of the dictator game. Subjects were paid only a show-up fee and were informed that they could generate, with no monetary cost to themselves, donations to two organizations (henceforth, ‘charities’): UNICEF Sweden, presented as ‘UNICEF’ to subjects; and one other (henceforth denoted ‘Other’). UNICEF is a global organization concerned with the welfare of children which pools all incoming regular donations internationally. The identity of the ‘Other’ charity differed across treatments; it will be further described below. Donating required that subjects first earn ‘points’, and the amount given to charity  $k \in \{UNICEF, Other\}$  was  $\theta_k x_k$ , where  $\theta_k > 0$  is here an exchange rate (SEK/point) and  $x_k \geq 0$  the number of points allocated towards  $k$ . To make donations credible, participants were informed that they would be e-mailed a receipt of the payment, and this was done within days of each session.

The decision to earn and allocate points was repeated over six rounds, with subjects informed that only the choices made in a single round, randomly selected at the end of the session, would translate into actual donations. Five of the rounds were oriented towards identifying crowding across activities; in each of these  $\theta_{UNICEF}$  was held constant at 10 SEK/point, while  $\theta_{Other}$  took on different values (4, 8, 10, 12, and 16 SEK/point).<sup>6</sup> The remaining sixth round had  $\theta_{UNICEF} = \theta_{Other} = 16$ , and was introduced to examine the effect of a general (as opposed to activity-specific) productivity shock. Being irrelevant for identifying spillover effects, however, this round will not be considered in this paper (with a few clearly marked exceptions). Although all subjects encountered the same exchange rates at some point during a session, the sequence of values of  $\theta_{UNICEF}$  and  $\theta_{Other}$  was randomized within and between subjects.

Each round consisted of three computer screens.<sup>7</sup> The first screen simply presented the name and current-round exchange rate of each charity. Subjects were given no further information whatsoever on the charities. The second screen confronted subjects with a slider-adjustment task (Gill and Prowse, 2012). With 48 sliders on the screen and an allotted time of 120 seconds, each slider set to position 50 out of a possible 100, i.e. in the very middle, earned them a single point.<sup>8</sup> After the set time elapsed, subjects progressed to the third screen, where they were shown the number of points they had just earned and were asked to allocate these in integer amounts across the two charities. There was no public or private feedback. The placement of the two charities relative to each other (left/right) was randomized across rounds and subjects and had no significant effect on either the total number of points earned or the allocation across alternatives.

Our four between-subject treatments are summarized in Table 1. The similarity of the charities was systematically varied across treatments by isolating two dimensions along which

---

<sup>6</sup>In the theory discussed in the previous section, having  $\theta_1 = \theta_2$  would imply that both charities are equally efficient. Since a gift of some fixed size may in fact be more productive if given to one real-world charity than if given to another, the exchange rates described here should be interpreted only as limited components of the ‘overall’ productivity parameters of the theoretical model. Note that explicitly allowing for underlying efficiency differences (e.g. by  $g = a\theta_1 x_1 + b\theta_2 x_2$ , with  $a, b > 0$ ) does not affect our predictions with respect to  $x_2$  or total donations  $\theta_1 x_1 + \theta_2 x_2 \neq g$ .

<sup>7</sup>The experiment was programmed using the software z-Tree (Fischbacher, 2007).

<sup>8</sup>Thus, referring back to equation (1), the experiment was designed to reflect a situation where  $T = 48$  and  $u$  represents the utility of not exerting effort in the slider-adjustment task.

charities may differ: (i) geographical scope, and (ii) the cause involved. In all treatments, one charity was UNICEF, but the alternative (Other) differed across treatments.

Table 1: *Treatment summary*

Treatment	N	No. sessions	Other (non-UNICEF) charity	Gender (% male)	Age (avg.)
GLOBAL-CHILD	40	2	Save the Children Global Action Fund	47.5	23.05
LOCAL-GREEN	40	2	SSNC Lund	52.5	22.95
LOCAL-CHILD	40	3	Save the Children Lund	42.5	23.65
GLOBAL-GREEN	40	2	WWF International	50	22.65

In the GLOBAL-CHILD treatment, UNICEF was paired with the ‘Global Action Fund’ run by Save the Children in the US. Although the fund’s website<sup>9</sup> states that donations benefit children “in the United States and around the world”, subjects were presented only with the name ‘Save the Children Global Action Fund’, so it is likely that the global dimension was salient. Since both charities address children’s needs globally, they are similar along both dimensions, and should be relatively close substitutes.

By contrast, the LOCAL-GREEN treatment had UNICEF paired with the Lund office of the Swedish Society for Nature Conservation (SSNC), which is dissimilar along both dimensions. SSNC is the largest Swedish environmental NGO, and its local branch engages in awareness raising through lectures, workshops, and excursions into the local environment, in addition to lobbying local policy makers.

Finally, to isolate the effect of each dimension (local/global and cause involved), two intermediate treatments were also included: LOCAL-CHILD and GLOBAL-GREEN. In the former, UNICEF was paired with the local Lund office of Save the Children (local/global dimension); in the latter, with WWF International, a major global environmental NGO (cause dimension). Admittedly, the distinctions are not entirely clear-cut; in particular, as there is no need for disaster relief in Sweden, the Lund office of Save the Children focuses on education and community work (especially with respect to child refugees), and so its ‘cause’ might be viewed as somewhat different from UNICEF. Since this pair is nevertheless likely to be seen as more similar than UNICEF versus the Lund office of SSNC, our hypotheses are little affected.

At the end of each experimental session, participants responded to a questionnaire. First, they reported demographic variables: sex, age, and study program. Most subjects were enrolled either in a Bachelor of Science in Business and Economics (67 subjects), Political Science Bachelor (24 subjects), or various Master programs at LUSEM (25 subjects). 16 subjects were not enrolled in any program, and six were students at the Faculty of Engineering at Lund University (LTH).

Second, we elicited various personality variables: a ‘satisficing’ measure (Turner et al., 2012), the Conscientiousness dimension of Big-Five personality theory (John and Srivastava, 1999); the three ‘personal values’ dimensions of Benevolence, Universalism, and Achievement (Schwartz, 2003); and cognitive-reflection skills (by the test presented in Frederick, 2005). Except for the cognitive-reflection test, these sections of the questionnaire were in English; the rest of the survey was in Swedish (translated versions of all questions specific to our study are available in Appendix A).<sup>10</sup>

<sup>9</sup><https://secure.savethechildren.org>

<sup>10</sup>As a robustness check, at a time between 10 days and a month from the experimental session, subjects were invited to re-take the satisficing, conscientiousness and personal values tests through an online tool. We will not use this data set, however, as participation rates were low and we were able to match online results to experimental data for only 78 of the 160 subjects.

Third, each subject graded (on 5-point scales) his or her knowledge of each charity, along with its perceived favorability and identity/self-image value. We also asked subjects to grade the overall similarity of each charity pair.<sup>11</sup>

Fourth, we asked for overall charity preferences, including the amount donated to any cause per month, the amount donated to each of the charities included in the experiment, and the frequency of giving to people begging on the street.

Fifth and finally, participants were explicitly instructed to sort all five charities that were included in any treatment into any number of categories; they were told to focus on the categories they found “most relevant”. We then constructed a dummy variable indicating whether or not each subject thought both charities in his/her particular session belonged to the same category.

Table 2: *Perceptions of charities*

<i>All subjects (N = 160)</i>						
Charity	Donors (no. subjects)	Favorability (average)	Identity (average)	Knowledge (average)	Similar to UNICEF (average)	Categorized with UNICEF (no. subjects)
UNICEF	5	4.43 (0.75)	2.71 (1.31)	3.66 (0.96)	-	-
STC (Global)	2	4.48 (0.70)	2.70 (1.33)	3.51 (0.90)	3.80 (1.20)	110
SSNC (Local)	0	3.53 (0.84)	2.06 (1.21)	1.94 (1.02)	1.84 (0.94)	4
STC (Local)	1	3.88 (0.87)	2.41 (1.38)	2.64 (1.25)	2.43 (1.11)	26
WWF (Global)	6	4.34 (0.78)	2.58 (1.41)	3.41 (1.19)	2.86 (1.30)	34
<i>Within relevant treatment (N = 40)</i>						
Charity	Donors (no. subjects)	Favorability (average)	Identity (average)	Knowledge (average)	Similar to UNICEF (average)	Categorized with UNICEF (no. subjects)
STC (Global)	0	4.53 (0.64)	2.65 (1.39)	3.58 (0.87)	4.1 (1.10)	28
SSNC (Local)	0	3.55 (0.93)	2.23 (1.10)	1.98 (0.97)	1.98 (1.17)	1
STC (Local)	0	3.83 (0.93)	2.40 (1.43)	2.78 (1.27)	2.50 (1.30)	7
WWF (Global)	3	4.40 (0.63)	2.53 (1.30)	3.50 (1.09)	3.23 (1.40)	6

Standard deviations in parentheses. ‘Within relevant treatment’ includes subjects in treatments where donating to the charity in question was possible. For the questions underlying this data set, see Appendix A.1 (number of donors, favorability, identity, knowledge, similarity with UNICEF) and Appendix A.3 (categorization).

Summary statistics on how each charity was perceived on average by subjects are given in Table 2. The table presents, from left to right, the number of subjects who reported donating to each charity, average ratings of the favorability and self-image/identity value of each charity as well as subjects’ knowledge of it, the average perceived degree of similarity between each charity and UNICEF, and finally the number of subjects who placed each charity in the same arbitrary category as UNICEF. Since subjects rated not only the charities included in their

<sup>11</sup>Subjects also reported the perceived effectiveness of each charity, as well as whether the charity was thought to be as effective as UNICEF. These variables were, however, highly correlated with the favorability and overall similarity variables, respectively, and will not be used in our analysis, though their wording is still reported in Appendix A.

own treatment, we have sorted the results by ‘all subjects’ as well as only ‘within relevant treatment’. The latter includes only answers by those subjects who had the opportunity to give to the charity in question, so e.g. the STC Global row of that part of the table includes only participants in the GLOBAL-CHILD treatment. Since UNICEF was available in all treatments, it is redundant here and not included. We see that results are similar in both parts of the table. Importantly, variation in the similarity item, as well as results from explicit categorization by subjects, confirm that our treatments were successful in manipulating the perceived similarity of the charity pairs.

## 4 Results

Our first priority in analyzing the experimental dataset is to check whether results are consistent with predictions. Before turning to formal testing, however, we inspect the data visually. Figure 1 provides an initial look at the spillover patterns within each treatment. Results seem encouraging; we see, first, that all lines slope downwards, indicating that there was crowding-out across all charity pairs. Indeed, since  $\theta_{UNICEF}$  was fixed at 10 SEK/point, subjects not only allocated less points to UNICEF as  $\theta_{Other}$  increased, but donated less money to that charity as well. Also, the spillover effect was much more pronounced in some treatments than in others, in a manner conforming with expectations. Our similar treatment, GLOBAL-CHILD, saw the most crowding across charities (steepest line); the least similar, LOCAL-GREEN, had the least crowding (flattest line); and our intermediate treatments exhibited moderate crowding. Because the LOCAL-CHILD and GLOBAL-GREEN lines are very close to each other, the local/global and cause dimensions seem equally important in driving spillovers.

It also seems that spillovers are asymmetric. When UNICEF is the more productive option in terms of SEK/point, subjects contribute a large share of points to this charity in all treatments, and major differences between treatments only arise when UNICEF is the less productive charity (to the right of the  $\theta_{Other} = 10$  mark). This suggests that UNICEF was seen as a default option by most subjects. While we can only speculate as to why, it seems plausible that it is partly because UNICEF is generally both well known and well liked (Table 2). It may also be simply because the productivity of UNICEF was fixed at 10 SEK/point, while  $\theta_{Other}$  varied.

Figure 2 reinterprets the data in relative terms, presenting the share of total points allocated to UNICEF, i.e.  $x_{UNICEF}/(x_{UNICEF} + x_{Other})$ . In terms of relative slope and height, there are hardly any differences between Figures 1 and 2. The reason is that, as shown in the left panel of Figure 3, the mean number of total points earned by subjects was roughly constant for different values of  $\theta_{Other}$  and for different treatments. The pattern is consistent with evidence reported by Imas (2014) that people’s prosocial efforts are unresponsive to the magnitude of the resulting benefits. As a result, total amounts donated ( $\theta_{UNICEF}x_{UNICEF} + \theta_{Other}x_{Other}$ ) clearly increased with  $\theta_{Other}$ , as shown in the right panel of Figure 3. Moreover, consistent with our theory, donations seem to have increased especially strongly when  $\theta_{Other} > \theta_{UNICEF} = 10$ .<sup>12</sup>

While suggestive, the above figures only provide information averaged at the treatment

---

<sup>12</sup>Figures 1-3 also imply that total amounts contributed to Other increased with  $\theta_{Other}$ . While not predicted by our model, this finding is consistent with a number of studies showing that contributions in dictator and public-good games depend positively on the price or productivity of giving (e.g. Andreoni and Miller, 2002; Blackwell and McKee, 2003).

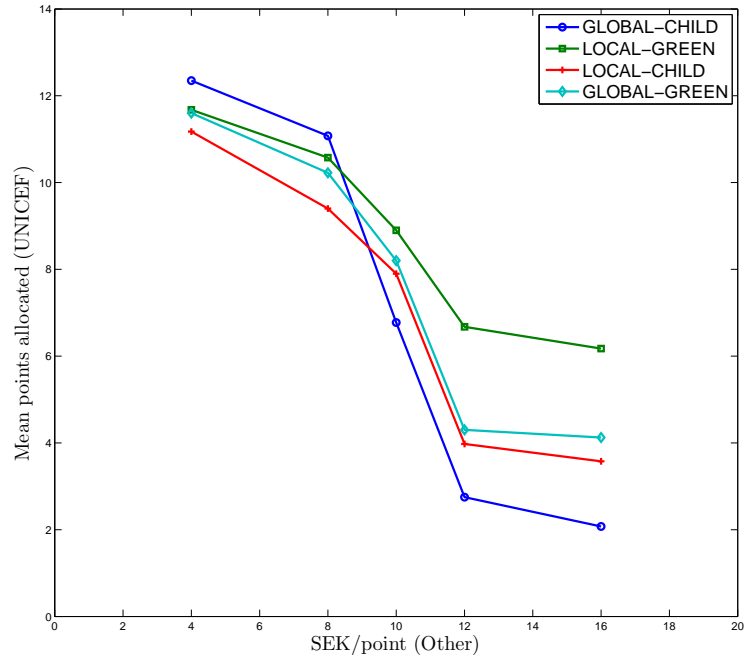


Figure 1: *Graphical illustration of spillovers and treatment effects. Allocation in absolute terms.*

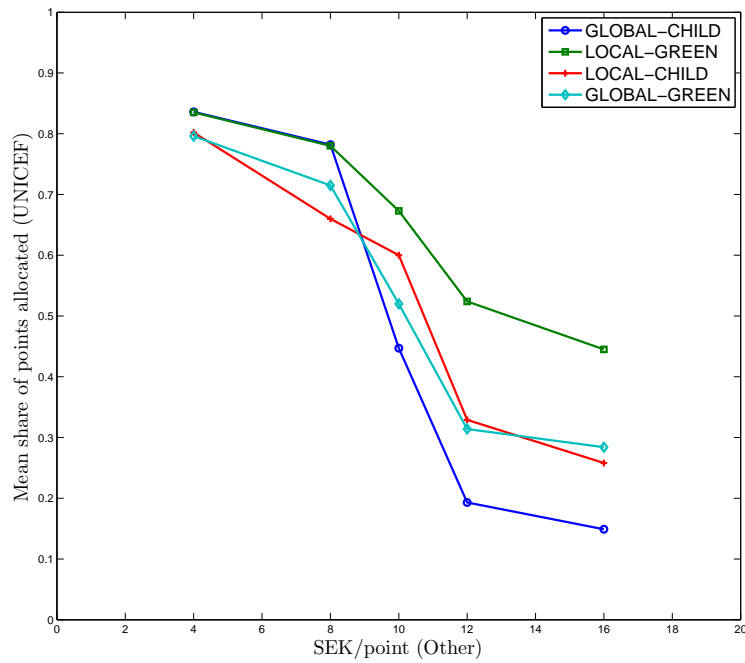


Figure 2: *Graphical illustration of spillovers and treatment effects. Allocation in relative terms.*

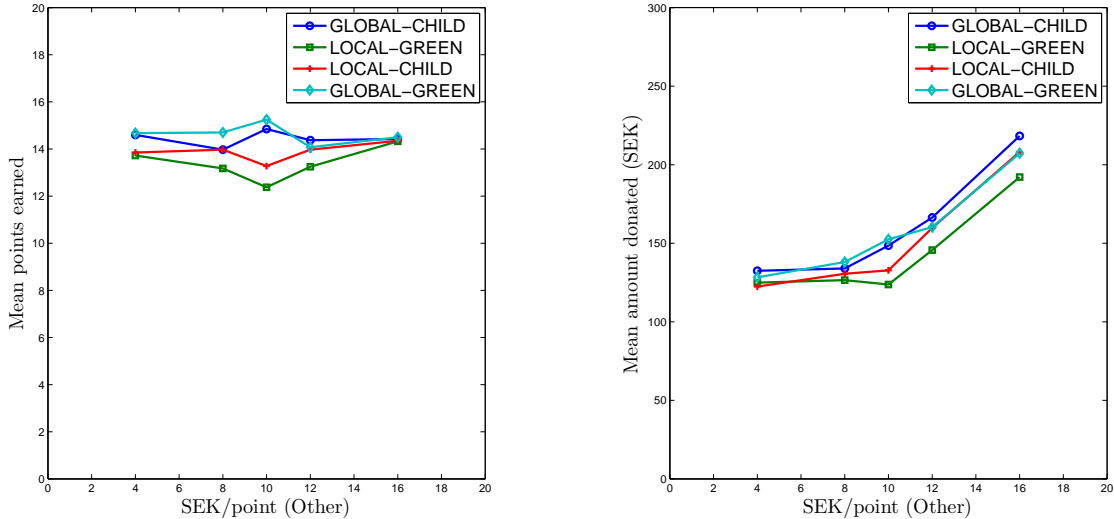


Figure 3: *Effect of activity-specific productivity on overall giving. Left: total efforts. Right: total amounts donated.*

level. To analyze the data using within-treatment variation, we perform various formal tests. First, we confirm the asymmetric pattern mentioned above by a series of nonparametric rank-sum tests. For each value of  $\theta_{Other}$ , we run three tests comparing the mean number of points allocated to UNICEF in GLOBAL-CHILD versus all other treatments. Table 3 gives differences in means and significance levels for all tests, confirming that differences between treatments with respect to contributions to UNICEF tend to emerge only when  $\theta_{Other} \geq \theta_{UNICEF}$ .<sup>13</sup>

Table 3: *Results of non-parametric rank-sum tests: points allocated to UNICEF*

	LOCAL-GREEN	LOCAL-CHILD	GLOBAL-GREEN
$\theta_{Other} = 4$	-0.675	-1.175	-0.75
$\theta_{Other} = 8$	-0.5	-1.675	-0.85
$\theta_{Other} = 10$	2.125*	1.125	1.425
$\theta_{Other} = 12$	3.925***	1.225	1.55
$\theta_{Other} = 16$	4.1***	1.5	2.05**

*Note:* table presents differences in means for  $x_{UNICEF}$  between GLOBAL-CHILD and each other treatment, conditional on  $\theta_{Other}$ . Also given are significance levels (\*\*\*)  $p < 0.01$ , (\*\*)  $p < 0.05$ , (\*)  $p < 0.1$ ) when testing for equality of means in  $x_{UNICEF}$ .

Table 3, in analyzing contribution levels, provides an indirect formal test for treatment effects. However, a more direct test looks for differences in the *slope* of contributions, as in Figure 1. This corresponds to running regressions of the form<sup>14</sup>

$$x_{UNICEF,i} = \beta_0 + \beta_1 \theta_{Other,i} + \beta'_2 \mathbf{X}_i + \beta'_3 \theta_{Other,i} \mathbf{X}_i + \beta_4 \mathbf{T}_i + \beta_5 \theta_{Other,i} \mathbf{T}_i + \epsilon_i \quad (2)$$

where  $\mathbf{X}_i$  is a vector of control variables, and  $\mathbf{T}_i$  is a vector of treatment dummies. In our

<sup>13</sup>We also tried using LOCAL-GREEN as our benchmark, comparing the mean  $x_{UNICEF}$  in this treatment with the means of each other treatment. This also confirms the asymmetry observed in Figures 1 and 2, and unlike GLOBAL-CHILD, LOCAL-GREEN is significantly different from LOCAL-CHILD for  $\theta_{Other} > 10$ .

<sup>14</sup>Robust standard errors are clustered at the subject level in all regressions reported in this paper.

regressions, we treat GLOBAL-CHILD as the baseline, so  $\mathbf{T}_i$  includes dummy variables for LOCAL-GREEN, LOCAL-CHILD, and GLOBAL-GREEN. Note that treatment effects are given by interaction terms  $\beta_5 \theta_{Other,i} \mathbf{T}_i$  in (2).

Results are given in Table 4.<sup>15</sup> Although regression equation (2) includes non-interacted terms for relevant treatment and (if applicable) control variables, the table reports results only on  $\theta_{Other}$  itself and all interaction terms. Column 1 presents outcomes from the simplest version of (2), with no control variables; hence, it closely corresponds to Figure 1. We see that in the similar GLOBAL-CHILD treatment (the non-interacted  $\theta_{Other}$  coefficient), there is clear evidence of crowding-out. Moreover, as expected, this effect is less pronounced in the other treatments, though significantly so only for LOCAL-GREEN, the least similar charity pair, where the effect’s magnitude is roughly halved. Column 2 shows that the treatment effect is robust to adding demographic controls (sex, age, study program) as well as each observation’s round number.

Table 4: *Treatment effects*

VARIABLES	Dependent variable in all regressions: $x_{UNICEF}$			
	OLS (1)	OLS (2)	Tobit (3)	Tobit (4)
$\theta_{Other}$	-0.979*** (0.144)	-0.518 (0.496)	-1.329*** (0.213)	-0.724 (0.696)
$\theta_{Other} \times \text{GLOBAL-CHILD}$		Reference category		
$\theta_{Other} \times \text{LOCAL-GREEN}$	0.469*** (0.177)	0.457*** (0.173)	0.728*** (0.244)	0.715*** (0.233)
$\theta_{Other} \times \text{LOCAL-CHILD}$	0.273 (0.177)	0.243 (0.169)	0.416* (0.249)	0.380 (0.236)
$\theta_{Other} \times \text{GLOBAL-GREEN}$	0.270 (0.183)	0.252 (0.176)	0.433* (0.251)	0.406* (0.239)
$\theta_{Other} \times \text{Round}$		-0.0455* (0.0270)		-0.0617* (0.0348)
Demographic variables	NO	YES	NO	YES
Observations	800	800	800	800
R-squared	0.244	0.278	0.0504 (pseudo- $R^2$ )	0.0561 (pseudo- $R^2$ )

Robust standard errors clustered at the subject level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Only coefficients on interaction terms reported. All interacted variables were included separately in each regression.

Out of all 800 observations, 185 have the boundary solution  $x_2 = 0$ , and we may suspect that some subjects, if able, would have been willing to set  $x_2$  to a negative number in order to further increase donations to their preferred option. It may therefore be of interest to compare OLS to a model which assumes that the distribution of  $x_2$  values is censored at zero. Hence, columns 3 and 4 of Table 4 repeat the analysis using a Tobit regression model. Although unsurprisingly the magnitude of the treatment effect is somewhat greater in the Tobit regressions (causing all treatment differences to be at least marginally significant), on the whole differences between the models are minor.

Next, we pool data across all treatments and regress total efforts only on  $\theta_{Other}$ . Results confirm that there is no significant relationship between charity-specific productivity and total

<sup>15</sup>We ran a number of variants of the OLS regressions summarized in Table 4 and beyond. First, we tried using LOCAL-GREEN rather than GLOBAL-CHILD as the baseline (Tables 4 and 5 only). Second, we included subject fixed effects. Third, we replaced  $x_{Other}$  as dependent variable with the share-of-points variable  $x_{UNICEF}/(x_{UNICEF} + x_{Other})$ . For all of these extensions, results were qualitatively (and indeed, often quantitatively) very similar to the OLS results presented in this paper.

efforts.<sup>16</sup> There is also little indication that the number of points earned reacted to changes in the maximum of  $\theta_{UNICEF}$  and  $\theta_{Other}$  ( $p = 0.251$ ).<sup>17</sup> Furthermore, in a substantial departure from the contribution patterns typically observed in endowment-based dictator games, the number of points earned was approximately normally distributed, and the number of points earned was below 5 in only 17 observations (out of 960), two of which had  $x_{UNICEF} + x_{Other} = 0$ . It thus appears that most subjects found the slider-adjustment task relatively easy to perform.

Finally, we confirm that total amounts donated ( $\theta_{UNICEF}x_{UNICEF} + \theta_{Other}x_{Other}$ ) increased with  $\theta_{Other}$  ( $\beta = 6.600$ ,  $p = 0.000$ ), as seen in the right panel of Figure 3. Indeed, although the increase seems more pronounced when  $\theta_{Other} < 10$ , regressing total donations on  $\theta_{Other}$  using only observations where  $\theta_{Other} \leq 10$  still yields a significantly positive slope coefficient ( $p = 0.002$ ).

#### 4.1 Robustness tests

All results presented so far are consistent with our predictions. We will now argue that treatment effects are in fact caused by the differing similarity of charity pairs and not by some confounding factor, such as differences in how well known and well regarded each ‘Other’ charity is compared to UNICEF. Returning to Table 2, although patterns are not clear-cut, we see e.g. that SSNC Lund was both the least well known and the least popular charity, so subjects may have given less to SSNC than to the Save the Children Global Action Fund because of the latter’s relative popularity rather than because of their varying similarity with UNICEF.

To settle the issue, we examine individual variation within each treatment. Figure 4 provides a look at the raw correlation between substitution patterns and perceived similarity, as well as some possible confounders. All panels are based on a simple pooled interaction regression of the form

$$x_{UNICEF,i} = \alpha + \beta\theta_{Other,i} + \gamma z_i + \delta\theta_{Other,i}z_i + \epsilon_i$$

for some interacted variable  $z_i$ . In the upper-left panel,  $z_i$  is the perceived similarity between UNICEF and Other, and the figure shows the marginal effect of  $\theta_{Other}$  on  $x_{UNICEF}$  (i.e. spillovers) for various values of this variable. Hence, the slope of the line represents  $\delta$ , the difference that  $z_i$  makes for the magnitude of spillovers. Clearly, greater perceived similarity is correlated with stronger crowding-out. This is confirmed by the underlying regression, where the estimated  $\delta < 0$  has  $p = 0.000$ .

The other three panels represent similar correlations for the *difference* in perceived favorability, identity/self-image value, and knowledge between UNICEF and Other (calculated as  $z_i = z_{UNICEF,i} - z_{Other,i}$ ). Although all suggest a positive slope — for instance, the more favorable UNICEF is perceived compared to Other, the less crowding-out — none are significantly increasing. In addition, of the four lines, the one associated with the similarity variable is the steepest: a one-point change in similarity has a greater effect on spillovers than a one-point change in any of the three difference variables. This suggests that of the four variables, similarity is the most important for explaining spillover patterns.

<sup>16</sup>Overall effort was not constant over time, as regressing total points on the ‘round’ variable produces a highly significant ( $p=0.000$ ) coefficient of 0.674.

<sup>17</sup>If the general-productivity shock round (where  $\theta_{UNICEF} = \theta_{Other} = 16$ ) is included in the sample, the coefficient is positive and marginally significant ( $p = 0.106$ ).



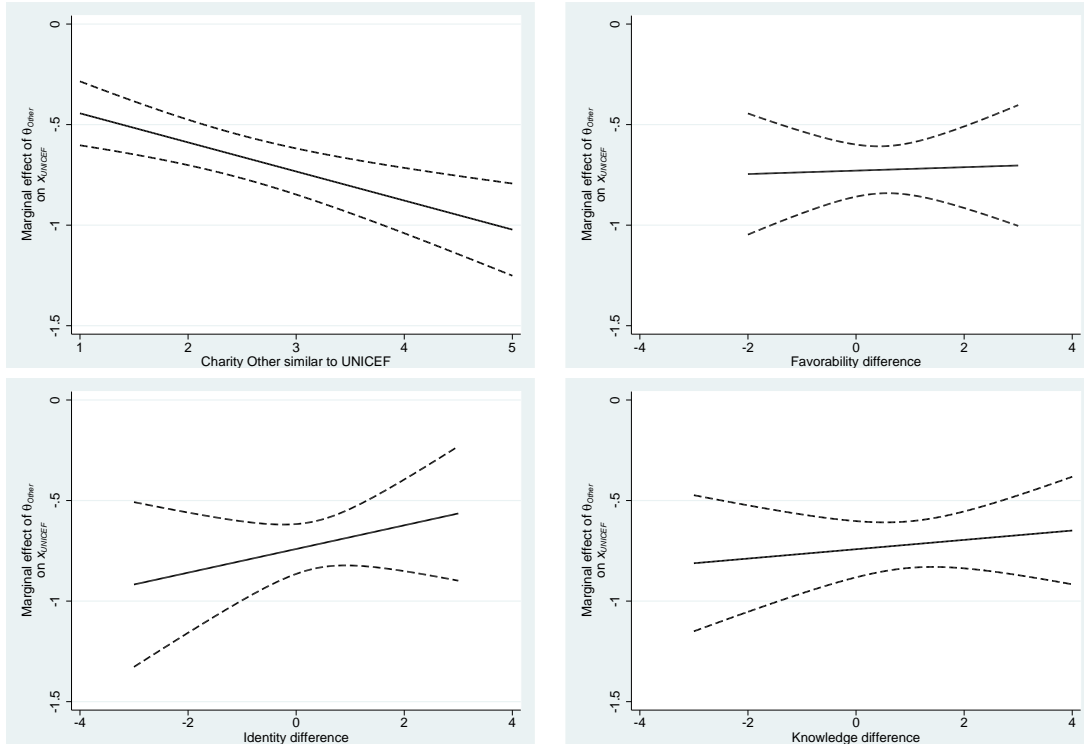


Figure 4: *The impact of charity-specific preference and belief variables on spillovers. Dashed lines give 95% confidence intervals.*

A more careful test involves adding these four variables as controls in regression (2). These results are presented in Table 5. Column 1 adds differenced variables for perceived favorability, identity/self-image value, and knowledge; clearly, this has only a minor impact on treatment effects, suggesting they are not driven by such differences among charities. Indeed, if anything, including these variables makes the treatment effect more prominent compared to Table 4. By contrast, adding also the similarity variable (column 2) does in fact render the treatment effect insignificant, while the similarity variable itself is highly significant and has the expected sign.<sup>18</sup> As in Table 4, when we repeat the analysis using the Tobit model (columns 3 and 4), results are broadly similar. This confirms that variation in substitution patterns across treatments does seem to reflect the perceived similarity of the charities, rather than e.g. their relative popularity among subjects.

## 4.2 Exploring individual characteristics

This section exploits individual heterogeneity in an attempt to answer the question, ‘Who exhibits strong spillovers?’ As we do not have particular hypotheses in mind, the analysis is exploratory. We begin with checking whether the sequence of  $\theta_{Other}$  values makes a difference for subject behavior. In particular, we may be concerned that subjects ‘anchor’ their beliefs

<sup>18</sup>In a variant of these regressions, we included the favorability, identity/self-image, and knowledge variables for Other instead of first taking differences. The pattern with respect to treatment effects and the similarity variable was very similar, with one exception, namely that identity/self-image was significant at the 5 % level both when similarity was included and when it was not.

Table 5: *Robustness checks*

VARIABLES	Dependent variable in all regressions: $x_{UNICEF}$			
	OLS (1)	OLS (2)	Tobit (3)	Tobit (4)
$\theta_{Other}$	-0.388 (0.511)	0.261 (0.586)	-0.535 (0.710)	0.291 (0.797)
$\theta_{Other} \times \text{GLOBAL-CHILD}$			Reference category	
$\theta_{Other} \times \text{LOCAL-GREEN}$	0.620*** (0.195)	0.324 (0.207)	0.882*** (0.255)	0.502* (0.265)
$\theta_{Other} \times \text{LOCAL-CHILD}$	0.331* (0.170)	0.0973 (0.184)	0.471** (0.235)	0.174 (0.252)
$\theta_{Other} \times \text{GLOBAL-GREEN}$	0.266 (0.177)	0.159 (0.179)	0.396* (0.237)	0.267 (0.237)
$\theta_{Other} \times \text{Round}$	-0.0498* (0.0270)	-0.0452* (0.0265)	-0.0682* (0.0348)	-0.0616* (0.0341)
$\theta_{Other} \times \text{Favorability difference}$	-0.0858 (0.0659)	-0.0830 (0.0623)	-0.0931 (0.0844)	-0.0941 (0.0793)
$\theta_{Other} \times \text{Identity difference}$	0.0801 (0.0746)	0.0231 (0.0810)	0.145 (0.101)	0.0662 (0.108)
$\theta_{Other} \times \text{Knowledge difference}$	-0.0687 (0.0545)	-0.0488 (0.0527)	-0.0989 (0.0689)	-0.0709 (0.0663)
$\theta_{Other} \times \text{UNICEF similar to Other}$		-0.139*** (0.0496)		-0.182*** (0.0657)
Demographic variables	YES	YES	YES	YES
Observations	800	800	800	800
R-squared	0.343	0.354	0.0723 (pseudo- $R^2$ )	0.0754 (pseudo- $R^2$ )

Robust standard errors clustered at the subject level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Only coefficients on interaction terms reported. All interacted variables were included separately in each regression.

about reasonable contribution levels based on first-period exchange rates. To examine this possibility, we generate dummy variables for each possible value of first-period  $\theta_{Other}$ . As shown in Table 6, there were no differences in either the level of effort (column 1) or the magnitude of spillovers (column 2).<sup>19</sup> A Tobit regression corresponding to column 2 (not reported) yields qualitatively identical results.

Next, we ask whether subjects who earned more total points were more likely to exhibit crowding-out. As an initial test, in the regression

$$x_{UNICEF,i} = \alpha + \beta\theta_{Other,i} + \gamma(x_{UNICEF,i} + x_{Other,i}) + \delta\theta_{Other,i}(x_{UNICEF,i} + x_{Other,i}) + \epsilon_i \quad (3)$$

the interaction term is negative and highly significant ( $p = 0.000$ ), suggesting that people who earned more points substituted more. The question is tricky, however, because we would expect subjects who earned many points to shift more points simply because on average they have more points to shift. Thus, this result probably does not tell us very much about the personal characteristic of tending to earn many points. A better test looks at the *share* of points given to UNICEF; that is, we replace the dependent variable in (3) with  $x_{UNICEF}/(x_{UNICEF} + x_{Other})$ .<sup>20</sup> The two observations where  $x_{UNICEF} + x_{Other} = 0$  were

<sup>19</sup>Note that in Table 6, we have dropped entirely those subjects who had  $\theta_{UNICEF} = \theta_{Other} = 16$  in the first round; including them does not make any difference. Neither does interacting the maximum of exchange rates in the first round, or the minimum, regardless of whether such subjects are included or not (there is a single exception: when these subjects are not included, those who faced a maximum exchange rate of 16 in the first round exhibit less crowding-out, with  $p = 0.071$ , than those who faced a maximum exchange rate of 10; given the sheer number of regressions involved, we find this less than persuasive).

<sup>20</sup>As model (1) is not expressed in share terms, strictly speaking this analysis does not relate to our hypotheses.

Table 6: *Anchoring effects*

VARIABLES	Dep. variable: $x_{UNICEF} + x_{Other}$ (1)	Dep. variable: $x_{UNICEF}$ (2)
$\theta_{Other}$		-0.668*** (0.173)
$I(\theta_{Other} = 4 \text{ in round 1})$		Reference category
$I(\theta_{Other} = 8 \text{ in round 1})$	-0.430 (0.996)	0.297 (2.408)
$I(\theta_{Other} = 10 \text{ in round 1})$	1.413 (1.519)	3.680 (2.787)
$I(\theta_{Other} = 12 \text{ in round 1})$	0.874 (1.264)	0.0420 (2.846)
$I(\theta_{Other} = 16 \text{ in round 1})$	-0.320 (0.991)	0.0683 (2.403)
$\theta_{Other} \times I(\theta_1 = 4 \text{ in round 1})$		Reference category
$\theta_{Other} \times I(\theta_1 = 8 \text{ in round 1})$		-0.0597 (0.207)
$\theta_{Other} \times I(\theta_1 = 10 \text{ in round 1})$		-0.310 (0.229)
$\theta_{Other} \times I(\theta_1 = 12 \text{ in round 1})$		0.0694 (0.229)
$\theta_{Other} \times I(\theta_1 = 16 \text{ in round 1})$		0.1000 (0.201)
Constant	13.64*** (0.858)	13.85*** (2.063)
Observations	695	695
R-squared	0.018	0.219

Robust standard errors clustered at the subject level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

not included in this regression. Results, shown in Figure 5, suggest there is little correlation between total efforts and crowding-out in share terms; the interaction coefficient in the underlying regression has  $p = 0.210$ .

Finally, we turn to the personal characteristics elicited in our questionnaire. Table 7 summarizes a number of OLS regressions of  $x_{UNICEF}$  on  $\theta_{Other}$  and a number of interacted (and, as in Table 4, uninteracted and unreported) variables.<sup>21</sup> For columns 1 through 10, separate regressions were run on single variables or groups of variables, while for columns 11 and 12 all variables were included in the same regression. In column 11, it is likely that for several variables, the modifying effect on spillovers is itself confounded by total points earned, for the same reason as in our discussion of (3). For example, judging from column 11, we might think that the 15 subjects who show to the world that they support a charity (the Giving identity variable) are less likely to substitute points, perhaps because of a tension between defining oneself as a donor to some particular charity and maximizing donations across all charities. However, these subjects also earned on average almost two points less than their peers, a significant difference ( $p = 0.064$ ); and adding total points to the regression causes the significance of the Giving identity variable to disappear.

There are some clear patterns. First, Master students substitute significantly more, possibly consistent with an ‘indoctrination effect’ on economics students (Frank et al., 1993; Bauman and Rose, 2011). It is conceivable that Master students, having been longer exposed to economic theory, have also become more responsive to relative price shocks. Alternatively, the pattern could be understood as a type of experimental demand effect, with Master stu-

<sup>21</sup>We also ran the corresponding Tobit regressions, with very similar results.

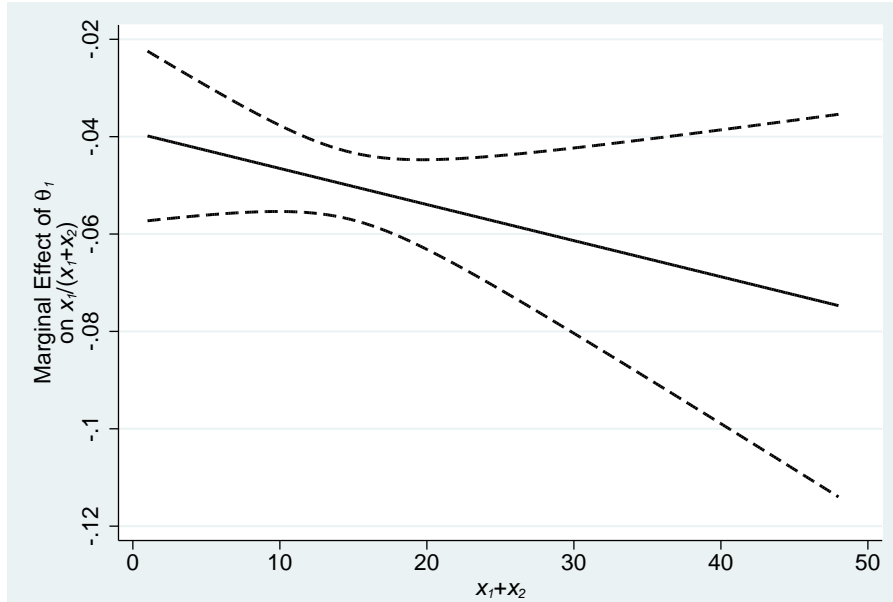


Figure 5: *Total efforts and spillovers in share terms*

dents more likely to suspect that the experimenter is looking for, or expecting, substitution behavior. However, although subjects were recruited either from courses at the Lund School of Economics and Management or on the LUSEM premises, our questionnaire did not ask Master students to specify what program they were enrolled in. Thus, it is possible that results are driven not by economics majors, but rather by e.g. business majors.

Second, Table 7 indicates that subjects who often give money to beggars are also more likely to exhibit strong negative spillovers. However, this pattern turns out to be dependent on a single outlier. This subject reported giving money to beggars very frequently (20 times per month), earned a large number of points in the experiment, and moreover exhibited strong negative spillovers (in both absolute and relative terms). Dropping the five observations associated with this individual renders results in Table 7 insignificant.<sup>22</sup>

Third and finally, of the psychological measures, two are significant: self-reported ‘satisficers’ and subjects with good cognitive-reflection skills both exhibit greater crowding-out. While these two types may seem like opposites, they may behave similarly for different reasons. Both, it seems, arrive at the conclusion that most of the points should be allocated to the charity with the largest exchange rate, because more money to charity is a good thing. However, the groups have different allocation styles. To show this, we define a variable for the number of rounds (out of five) in which a given subject chose a boundary solution ( $x_1 = 0$  or  $x_2 = 0$ ). Then, we regress this variable on all the personal characteristics included in column 12 of Table 7. In this regression, we exclude the outlier mentioned in the previous paragraph as well as the two subjects who had  $x_1 = x_2 = 0$  in one period. Results show that subjects with high cognitive-reflection scores were significantly more likely to choose a

<sup>22</sup>As an additional test, we ran a regression identical to column 11 in Table 7, except that the dependent variable was the share (rather than absolute amount) of points allocated to UNICEF. Results were very similar to those reported here. The only qualitative differences were that the satisficing variable was marginally significant at  $p = 0.101$ , and that giving to beggars was insignificant even without excluding outlier observations.

Table 7: *Explaining spillovers by individual characteristics*

VARIABLES	Dependent variable in all regressions: $x_{UNICEF}$											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\theta_{Other}$	-0.456** (0.177)	-0.300 (0.426)	-0.616*** (0.0867)	0.392 (0.582)	-0.749 (0.475)	-0.236 (0.526)	-0.533*** (0.0974)	-0.725*** (0.0624)	-0.665*** (0.0711)	-0.740*** (0.0631)	1.104 (0.900)	1.410 (0.871)
$\theta_{Other} \times Sex$	-0.182 (0.119)										-0.0869 (0.121)	0.0343 (0.108)
$\theta_{Other} \times Age$		-0.0185 (0.0184)									0.00744 (0.0202)	0.0189 (0.0216)
$\theta_{Other} \times Business/Econ. Bachelor$			Reference category								Reference category	
$\theta_{Other} \times Pol. Sci. Bachelor$			-0.150 (0.142)								0.0158 (0.158)	-0.0972 (0.136)
$\theta_{Other} \times Master$			-0.475** (0.202)								-0.495** (0.200)	-0.395*** (0.150)
$\theta_{Other} \times No\ study\ program$			-0.0718 (0.195)								-0.0748 (0.209)	0.00952 (0.209)
$\theta_{Other} \times LTH$			-0.359 (0.471)								-0.239 (0.438)	-0.156 (0.347)
$\theta_{Other} \times Other\ program$			0.0352 (0.176)								0.0984 (0.189)	0.0824 (0.163)
$\theta_{Other} \times Satisficing$				-0.0274* (0.0146)							-0.0290* (0.0148)	-0.0246* (0.0129)
$\theta_{Other} \times Conscientiousness$					0.000693 (0.0138)						-0.000517 (0.0155)	-0.00598 (0.0128)
$\theta_{Other} \times Universalism$						-0.0176 (0.0154)					-0.0211 (0.0148)	-0.0192 (0.0145)
$\theta_{Other} \times Achievement$						0.00819 (0.0168)					0.0202 (0.0163)	0.0120 (0.0152)
$\theta_{Other} \times Benevolence$						-0.0143 (0.0246)					-0.0137 (0.0249)	0.00532 (0.0225)
$\theta_{Other} \times Cognitive\ reflection\ score$							-0.105** (0.0436)				-0.0966** (0.0462)	-0.119*** (0.0450)
$\theta_{Other} \times Monthly\ giving$								-1.77e-05 (0.000258)			-8.58e-05 (0.000287)	-0.000206 (0.000265)
$\theta_{Other} \times Monthly\ giving\ (beggars)$									-0.0353* (0.0191)		-0.0439** (0.0202)	-0.0275** (0.0138)
$\theta_{Other} \times Giving\ identity$										0.148 (0.194)	0.366** (0.168)	0.148 (0.168)
$\theta_{Other} \times (x_{UNICEF} + x_{Other})$												-0.0655*** (0.0154)
Constant	11.03*** (2.112)	9.278* (5.140)	13.44*** (1.043)	-0.237 (7.761)	13.93** (6.247)	10.55 (6.453)	13.06*** (1.176)	15.10*** (0.774)	14.32*** (0.885)	15.17*** (0.780)	-7.683 (10.69)	-12.92 (9.303)
Observations	800	800	800	800	800	800	800	800	800	800	800	800
R-squared	0.227	0.221	0.246	0.228	0.220	0.225	0.225	0.227	0.224	0.223	0.291	0.497

Robust standard errors clustered at the subject level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$   
 Only coefficients on interaction terms reported. All interacted variables were also included separately in each regression.

boundary solution ( $\beta = 0.322$ ,  $p = 0.000$ ), but satisficers were not ( $p = 0.852$ ).<sup>23</sup> In one interpretation, while allocating *most* points to the charity with the highest exchange rate is probably a salient rule of thumb, the stronger rule of allocating *all* points to this charity (and hence, none to the other) is more restrictive, may appear aversive at first glance, and is therefore more likely to be followed by people who consider their choices carefully.

In conclusion, perceived similarity of the alternatives as well as some (but not many) individual characteristics appear to be important for explaining spillovers. This is underlined by a final regression (not reported) combining the final columns of Table 7 and Table 4, without including the outlier participant just mentioned. In this regression, apart from a few treatment dummies and the total-points variable (which was negative and highly significant), three interaction coefficients were significant at the 10% level, namely: perceived similarity of UNICEF and Other, the Master dummy, and cognitive-reflection scores.

## 5 Concluding remarks

Does policy targeting a particular prosocial activity spill over onto other activities? This paper has argued, on the basis of experimental data, that it may. In particular, we have presented evidence of an ‘incomplete crowding-out effect’ within the charitable-giving domain: spillovers attenuate, but do not reverse, the policy’s impact. Furthermore, we have presented data indicating that spillovers decrease in magnitude as activities grow more distant in terms of similarity. Hence, policy affecting an environmental activity may spill over onto other environmental acts but, perhaps, not onto e.g. charitable giving to the poor. Whether this is actually the case is, of course, an empirical question which merits investigation.

In this context, it is noteworthy that the similarity item included in our questionnaire was highly predictive of negative spillovers both within and across treatments. There may be scope for using questions of this type in ex-ante policy evaluation. Ideally, of course, spillover effects should be estimated by well-identified empirical analysis rather than by survey. Nevertheless, once more empirical evidence on spillovers has accumulated, it is worth checking whether similarity ratings are highly correlated with results; if they are, survey items may eventually prove useful as a cheap shortcut for policy analysts.

Our experiment differs from most, if not all, previous studies on behavioral spillovers and moral licensing in that it involves simultaneous rather than sequential decision making. However, our interpretation of the moral-licensing effect is that it reflects the more basic process of managing one’s self-image as a sufficiently moral person. Experimental manipulations (such as the amount initially donated in Ploner and Regner, 2013) work because they cause shifts in the perceived degree of moral self-worth, which in turn affects behavior in other morally charged situations. If our interpretation is correct, the pattern should generalize from sequential to simultaneous decision making, and can inform predictions for either setting. Certainly the theory underlying this paper captures much of the flavor of self-image management. Hence, we suggest that our novel methodology can be profitably used in future research on policy spillovers.

In this paper, we have focused on the direct effect of activity-specific productivity. This

---

<sup>23</sup>On the whole, the results of this regression mirror those of column 12 in Table 7, with Master students also more likely to choose a boundary solution ( $p = 0.000$ ). Some additional patterns arise, however. At the 5 % level, men were significantly more likely to allocate at the boundary, while subjects who had high values of  $x_1 + x_2$ , high Conscientiousness scores, and a strong identity for giving were less likely to do so.

is not to argue that relative productivity is the only factor underlying choices on how to contribute; our assumption of no social interaction, in particular, has been made for simplicity rather than realism. In addition, policy may impact behavior indirectly through means other than productivity, including by signaling to agents in the economy that the public good targeted is more important than previously thought. The main contribution of this paper has been to provide evidence on one of the relevant factors at work in determining the sign and magnitude of spillovers.

## References

- J. Andreoni. Impure Altruism and Donations to Public Goods: A Theory of Warm-glow Giving. *The Economic Journal*, 100(401), 1990.
- J. Andreoni and J. Miller. Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, 70(2), 2002.
- Y. Bauman and E. Rose. Selection or indoctrination: Why do economics students donate less than the rest? *Journal of Economic Behavior & Organization*, 79(3), 2011.
- C. Blackwell and M. McKee. Only for my own neighborhood? Preferences and voluntary provision of local and global public goods. *Journal of Economic Behavior & Organization*, 52(1), 2003.
- I. Blanken, N. van de Ven, and M. Zeelenberg. A Meta-Analytic Review of Moral Licensing. *Personality and Social Psychology Bulletin*, 41(4), 2015.
- R. Bénabou and J. Tirole. Incentives and Prosocial Behavior. *American Economic Review*, 95(5), 2006.
- K.A. Brekke, S. Kverndokk, and K. Nyborg. An economic model of moral motivation. *Journal of Public Economics*, 87(9), 2003.
- S. Brown, M.N. Harris, and K. Taylor. Modelling charitable donations to an unexpected natural disaster: Evidence from the U.S. Panel Study of Income Dynamics. *Journal of Economic Behavior & Organization*, 84(1), 2012.
- D.M. Cain, G. Loewenstein, and D.A. Moore. The Dirt on Coming Clean: Perverse Effects of Disclosing Conflicts of Interest. *The Journal of Legal Studies*, 34(1), 2005.
- P. Dolan and M.M. Galizzi. Like ripples on a pond: Behavioral spillovers and their implications for research and policy. *Journal of Economic Psychology*, 47, 2015.
- C. Ek. A Multi-Activity Approach to Public-Good Contributions, 2015. Unpublished working paper.
- U. Fischbacher. z-tree: Zurich Toolbox for Ready-Made Economic Experiments. *Experimental Economics*, 10(2), 2007.
- R.H. Frank, T. Gilovich, and D.T. Regan. Does Studying Economics Inhibit Cooperation? *Journal of Economic Perspectives*, 7(2), 1993.

- S. Frederick. Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 19(4), 2005.
- D. Gill and V. Prowse. A Structural Analysis of Disappointment Aversion in a Real Effort Competition. 2012.
- A. Gneezy, A. Imas, A. Brown, L.D. Nelson, and M.I. Norton. Paying to Be Nice: Consistency and Costly Prosocial Behavior. *Management Science*, 58(1), 2012.
- C. Heath and J.B. Soll. Mental Budgeting and Consumer Decisions. *Journal of Consumer Research*, 23(1), 1996.
- P.W. Henderson and R.A. Peterson. Mental Accounting and Categorization. *Organizational Behavior and Human Decision Processes*, 51(1), 1992.
- A. Imas. Working for the ‘warm glow’: On the benefits and limits of prosocial incentives. *Journal of Public Economics*, 114, 2014.
- G.D. Jacobsen, M.J. Kotchen, and M.P. Vandenbergh. The behavioral response to voluntary provision of an environmental public good: Evidence from residential electricity demand. *European Economic Review*, 56(5), 2012.
- O.P. John and S. Srivastava. The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives. In L.A. Pervin and O.P. John, editors, *Handbook of personality: Theory and research*. New York: Guilford Press, 2nd edition, 1999.
- D.T. Miller and D.A. Effron. Psychological License: When it is Needed and How it Functions. 43, 2010.
- K. Nyborg. I don’t want to hear about it: Rational ignorance among duty-oriented consumers. *Journal of Economic Behavior & Organization*, 79(3), 2011.
- M. Ploner and T. Regner. Self-image and moral balancing: An experimental analysis. *Journal of Economic Behavior & Organization*, 93, 2013.
- D. Read, G. Loewenstein, and M. Rabin. Choice Bracketing. *Journal of Risk and Uncertainty*, 19(1-3), 1999.
- J.J. Rotemberg. Charitable giving when altruism and similarity are linked. *Journal of Public Economics*, 114, 2014.
- S. Sachdeva, R. Iliev, and D.L. Medin. Sinning Saints and Saintly Sinners. *Psychological Science*, 20(4), 2009.
- S.H. Schwartz. A proposal for measuring value orientations across nations, 2003. Chapter 7 in *Questionnaire development report of the European Social Survey*. Retrieved from [http://www.europeansocialsurvey.org/docs/methodology/core\\_ess\\_questionnaire/ESS\\_core\\_questionnaire\\_human\\_values.pdf](http://www.europeansocialsurvey.org/docs/methodology/core_ess_questionnaire/ESS_core_questionnaire_human_values.pdf).
- R. Thaler. Mental Accounting and Consumer Choice. *Marketing Science*, 4(3), 1985.
- R.H. Thaler. Mental Accounting Matters. *Journal of Behavioral Decision Making*, 12(3), 1999.



- V. Tiefenbeck, T. Staake, K. Roth, and O. Sachs. For better or for worse? empirical evidence of moral licensing in a behavioral energy conservation campaign. *Energy Policy*, 57, 2013.
- B.M. Turner, H.B. Rim, N.E. Betz, and T.E. Nygren. The Maximization Inventory. *Judgment and Decision Making*, 7(1), 2012.
- A. Tversky and D. Kahneman. The Framing of Decisions and the Psychology of Choice. *Science*, 211(4481), 1981.

# Appendices

## A Survey questions

The following questions were posed in Swedish and have been translated.

### A.1 Charity-specific preferences and beliefs

1. *Do you regularly give money to X?* (regular giving, Charity X)

*To what extent do you agree with the following statements?*

2. *X does valuable work* (favorability of Charity X)
3. *Money given to X has a large effect* (effect of Charity X)
4. *I would be willing to wear a symbol of X on my jacket* (identity/self-image value of Charity X)
5. *I know roughly what work is done by X* (knowledge of Charity X)<sup>24</sup>
6. *When it comes to donations to charity, I think that gifts to UNICEF is roughly the same thing as gifts to X* (similarity of Charity X to UNICEF, overall)
7. *A Swedish crown given to UNICEF has as large an effect as a Swedish crown given to X* (marginal product of Charity X relative to UNICEF)

### A.2 Overall giving preferences

1. *Roughly how much (in SEK) do you donate to various charitable organizations each month?* (monthly giving)
3. *How often (in number of occasions) do you give money to beggars in a month?* (monthly giving to beggars)
4. *Do you show to the world that you support a charity? This might be for example by sticking a symbol on your jacket/bag or to sometimes wear a T-shirt with a logo.* (overall giving identity)

### A.3 Categorization

*Consider the following charities:*

*Save the Children Global Action Fund  
The local office of SSNC in Lund*

---

<sup>24</sup>For the Save the Children Global Action Fund, the wording was slightly modified as ‘...*what work is done based on X*’.

*WWF International*  
*The local office of Save the Children in Lund*  
*BRIS in Lund*  
*UNICEF*

*You are now to sort these organizations/alternatives into categories. Use the categories that you find most relevant.*