

**Dieses Dokument ist eine Zweitveröffentlichung (Verlagsversion) /
This is a self-archiving document (published version):**

Robert Neumann

The framing of charitable giving: A field experiment at bottle refund machines in Germany

Erstveröffentlichung in / First published in:

Rationality and Society. 2019, 31(1), S. 98 – 126 {Zugriff am: 19.08.2019}. SAGE journals. ISSN 1461-7358.

DOI: <https://doi.org/10.1177/1043463118820894>

Diese Version ist verfügbar / This version is available on:

<https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-353902>

„Dieser Beitrag ist mit Zustimmung des Rechteinhabers aufgrund einer (DFGgeförderten) Allianz- bzw. Nationallizenz frei zugänglich.“

This publication is openly accessible with the permission of the copyright owner. The permission is granted within a nationwide license, supported by the German Research Foundation (abbr. in German DFG).

www.nationallizenzen.de/

The framing of charitable giving: A field experiment at bottle refund machines in Germany

Rationality and Society

2019, Vol. 31(1) 98–126

© The Author(s) 2018

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/1043463118820894

journals.sagepub.com/home/rss**Robert Neumann**

Technische Universität Dresden, Germany

Abstract

This article investigates the decision of consumers at bottle refund machines to either reclaim their bottle deposit or to donate the refund to a non-profit organization. The study documents the unique pre-intervention data on donating behaviour and introduces a field experiment to increase donation levels. The design comprised the strategic framing of the situation by highlighting different cues about the normative, descriptive and local expectations of charitable giving as well as cues about the warm glow of donating money. The experiment took place in 20 supermarkets in Germany and lasted for 12 months. By varying the experimental design and using different modelling approaches, the study arrives at the conclusion that individuals largely act consistent with the assumption having self-regarding preferences that are stable and difficult to change. Hence, our pre-test and post-intervention data stand in sharp contrast to results from lab experiments.

Keywords

Charitable giving, descriptive norms, difference-in-difference estimation, field experiment, multi-level modelling, social norms, warm glow

Introduction

Donating financial resources to charities has puzzled the social sciences for a long time. According to the narrow version of rational choice theory

Corresponding author:

Robert Neumann, Technische Universität Dresden, 01062 Dresden, Germany.

Email: robert.neumann@tu-dresden.de

(Opp, 1999), pure self-interest would lead individuals to refrain from donating money due to the lack of immediate benefit. Of course, given that American citizens gave roughly US\$281 billion to charities or non-profit organization in 2016¹ or donations of the German population to charities were estimated at 5.3 billion Euros in 2016 (Deutscher Spendenrat, 2017), this model has not been accurate in predicting donations to charities. Therefore, several theoretical suggestions have been introduced to explain individual decision making that takes into account the well-being of others. One prominent explanation highlights the internal reward individual donors experience from giving, metaphorically described as a *warm glow* of giving (Andreoni, 1990).

In their literature review, Bekkers and Wiepking (2011) offer several additional micro- and macro-level factors that may drive charitable giving. First, they identify macroconstellations that explain donation decisions, especially the strategies of asking for donations through solicitations and by describing the gains in reputation when money is donated. Second, next to an awareness of need and motives like altruism (Andreoni, 1989), they claim that the standard costs and benefits framework requires an extension by including several psychological factors that trigger charitable giving. For instance, efficacy (the belief that donating makes a difference), pro-social value orientations or the preservation of a positive self-image provide internal incentives to give resources to charities (Grossman and Van Der Weele, 2017).

This highlights the standard sociological take on the explanation of such individual behaviour that underscores the importance of norms as driving forces behind individual decisions to donate money, especially in the presence of internal or external sanctions (Elster, 1989; Hechter and Opp, 2001). According to this view, internal sanctions impede the violation of a social norm regarding the appropriateness of certain behaviour by inducing a feeling of guilt, shame or dissonance. External sanctioning would reflect the (costly) sanctioning of norm violations by formal institutions (law) or informal arrangements (Ostrom, 2000).

Recently, several studies have used field experiments to investigate which contextual characteristics may lead to an increase in the violation of social norms (Keizer et al., 2008; Keuschnigg and Wolbring, 2015) and under which conditions the violation of social norms is sanctioned (Balafoutas and Nikiforakis, 2012; Berger and Hevenstone, 2016; Przepiorka and Berger, 2016). These studies have focused on actions that equate norm violations with anti-social behaviour. Investigations of the impact of norms on pro-social behaviour remain restricted to lab experiments using dictator or ultimatum games (Kimbrough and Vostroknutov, 2016; Krupka and Weber, 2013). So far, there are only a few field

experiments that study the impact of norms in pro-social behaviour, for instance on pro-environmental behaviour (Cialdini, 2007; Cialdini et al., 2006; Goldstein et al., 2008). Evidence from field experiments on the effects of norms on charitable giving and of informational cues about warm glow behaviour on donation decisions remain rare or are so far missing.² This article tries to fill this research gap.

We will investigate the donating behaviour of individuals and the effect of norms and potential experiences of a warm glow on charitable giving within a unique setting: the decision at bottle refund machines in German supermarkets, where customers can either cash out their refund or donate it to a non-profit organization.

The contribution of our study is twofold. First, we examine the unique data set of 34 pre-intervention months on monthly aggregate donation and refund levels. A mere descriptive look at the data on donating behaviour of individual customers clearly paints a picture that results found in lab experiments during dictator games do not translate to real-life behaviour (Levitt and List, 2007; List, 2009).

Second, we use a field experiment in an attempt to increase monthly donation levels by framing the decision context (Tversky and Kahneman, 1981). We were given the opportunity to frame the decision of donating the refund with cues that highlight the expectations why to donate money to a non-profit organization. We follow the advice of others who have argued that to understand the role norms play within a rational choice framework, emphasis is required on the investigation of why individuals deviate from the standard model in a specific situation (Bicchieri, 2006: 19). To investigate this question for the context of pro-social behaviour in general and charitable giving in particular, it is crucial to vary or manipulate the expectations individuals face when having to decide between a self or other regarding course of action. Through the presentation of messages with respect to social or injunctive norms, descriptive norms, local norms as well as cues on warm glow behaviour, the field experiment investigates under which conditions customers are more likely to donate their refund.

One study closely related to our research area investigated the introduction of a donation button at refund machines in Sweden (Knutsson et al., 2013). The authors theorize that next to the warm glow experienced due to donating, individuals who are more likely to act benevolent may do so because of social pressure. The authors assume that due to this pressure, some donations that take place might be rather involuntary and customers might be willing to incur a cost to avoid being asked for a donation (see also Andreoni et al., 2017). They test their argument on data from a natural field experiment where they analyse the behaviour of customers after a donation option was introduced in markets of a Swedish supermarket chain. They

find that the level of refund revenue drops after the instalment of the donation option and conclude that customers switched locations to avoid the pressure of the donation request.

Our experiment is also related to several studies that have investigated how individuals can be influenced to start donating to charity or to increase donation levels. For instance, Frey and Meier (2004) investigated whether university students increase their donations if they were told how much students were willing to donate in the past. They find that the willingness to contribute to a fund increases if the students are informed about high (in contrast to low) donation levels in the past. Shang and Croson (2009) investigated the effect of varying information about donation levels of other individuals during a radio campaign. They are able to show that providing information about donation levels of others leads to an increase in amounts donated.

Another stream of research investigates whether it is possible to nudge individuals and steer them towards pro-social behaviour by re-framing the choice situation (Thaler and Sunstein, 2008). For instance, Altmann et al. (2014) investigated how default donation levels on online donation websites influence the decision to donate and whether total donation levels could be increased. They speculate that higher defaults may serve as information about recommended actions and individuals may prefer to follow the empirical expectations of others. Recently, researchers have shown that when individuals face the decision between a donation to a single recipient or having to choose from a list of potential recipients, both the donation levels and the donation frequency will increase if the latter option is presented (Schulz et al., 2016). Another strategy under ongoing scrutiny is the approach of matching donations and how this will influence individual donations. If an organization matches donations by individual donors with equal amounts (called linear matching) or lower matching rates, the frequency of donations by individuals usually increases, but average levels of per capita donations decrease (Huck and Rasul, 2011; Karlan et al., 2011; Karlan and List, 2007). The latter phenomenon reflects a crowding out effect (see also Meier, 2007) that may be avoided if matched donations are directed at other fund raising project (Adena and Huck, 2017). Together, these studies reflect a research agenda to investigate how field experiments are used to increase donation levels successfully.

Theory

The present investigation follows upon research that has challenged the assumption that individuals do not have stable preferences and that these preferences can be changed by reframing the context of the decision (DellaVigna, 2009: 318, 347). For instance, Tversky and Kahneman (1981)

famously showed that individuals facing two outcomes are less risk averse when facing a monetary loss (loss frame) in comparison to a gain of equal size (gain frame). Numerous lab experiments have documented the existence of framing effects, for example, that participants act more cooperative if the game they play carries a name that sounds more pro-social ('Community Game' vs 'Wall Street Game', see Liberman et al., 2004). Neumann and Mehlkop (2018) have shown with panel data that environmental choices with equal monetary consequences vary with regard to the framing of the costs of a decision. It mattered whether the choice option for a costly green alternative was framed as a forgone saving or a loss of profit. Framing effects also hold for non-economic decisions, with Nelson et al. (1997) showing different acceptance levels for a Ku Klux Klan rally, depending on whether it was framed as a democratic act of freedom of speech or a violation of public order.

The focus of this study is on the framing of decision for charitable giving. It tries to answer the question whether it is possible to increase donation levels through strategic framing of the situation. Several mechanisms have been proposed that attempt to explain what drives individuals to donate money without direct material payoffs. Here, we focus on the framing of the decision context with regard to two determinants that we wanted to test during our field experiment: norms and warm glow behaviour. Concurrent theoretical explanations for charitable giving did not apply to our study because they could not be tested or because we assume the motives were not at work in our setting.³

Norms

Bicchieri (2006: 8, 42) introduces a useful theoretical framework to identify some key features of norms and distinctions between different types of norms. Generally speaking, norms both prescribe certain actions as socially desirable and proscribe certain behaviour due to the presence of formal or informal sanctioning mechanism. But these general mechanisms will only partially account for norm following behaviour. First, individuals need to share the belief that a norm exists in situation *S*. Bicchieri (2006: 2–3, 11) denotes this awareness or contingency. In our application, we can assume the awareness of a norm of helping is given when customers decide to donate money. To either donate a non-profit organization or to claim the refund due to material self-interest reflects a situation *S* in which everybody should be aware about which one of the choices will reflect a socially approved decision of helping.

Furthermore, the conditions for revealing such norm following behaviour are (a) that there exist a set of empirical expectations about the behaviour of others and that (b) a normative expectation exists that corresponds to a belief

about the right way to act in a social situation (Bicchieri, 2006: 13–15). Normative expectations entail either the expectation by others how an actor should behave in a situation or the expectation about how an actor ought to behave, given the possibility to sanction norm violation. This corresponds to the conviction by others that next to mere behavioural regularities, such ‘oughtness norms’ (Hechter and Opp, 2001: 13) or ‘injunctive norms’ (Cialdini et al., 2006) exhibit their influence on individual decision making primarily through the anticipation of external (by a third party or the law) and internal punishment (shame, guilt). Hence, *social norms* are characterized by the awareness of its existence along with empirical and normative expectations. Both shape the conditions under which a ‘conditional preference’ (Bicchieri, 2006: 11) for following or breaking a social norm is revealed.

In contrast to social norms, *descriptive norms* merely rely on the existence of cues about what actions are perceived as common, appropriate or frequent in similar situations (Cialdini et al., 2006: 4). Here, the perception of the existence of a descriptive norm serves as an information signal as to how most individuals behave in these situations, without any particular moral obligation (Hechter and Opp, 2001). Cialdini and colleagues have investigated these mechanisms in several publications. Their studies on littering in public areas (Cialdini et al., 1990) and environmental behaviour (Goldstein et al., 2008) have shown that providing information about norms as ‘guidelines for appropriate’ or as ‘rules for accepted and expected behaviour’ in social situations greatly enhances norm following behaviour by individuals.

An extension of this perspective was put forward by Goldstein et al. (2008) who state that empirical information about the behaviour of others will trigger norm following behaviour if the expectations reflect the *locality* of a decision situation. For instance, highlighting the behaviour of other hotel guests in contrast to the behaviour of hotel guests who stayed in the exact same room increased the re-usage rates of towels in the latter condition that highlighted the ‘immediate situational circumstances’ (Goldstein et al., 2008: 472).

Testing the impact of social and descriptive norms requires ‘*manipulating expectations*’ (Bicchieri, 2010: 298, emphasis in the original). For our experiment, the distinction between normative and empirical expectations will be addressed by the different framing strategies. In addition, we test whether cues about descriptive norms exhibit stronger effects on the decisions of individuals if we highlight the immediate situational circumstances of the decision in the sense of a local norm.

Warm glow

One of the most prominent economic explanations for the question why individual actors either contribute to a public good or give money to

charities is that individuals derive a private utility from giving. James Andreoni (1990) coined the term ‘warm glow giving’ to describe that individuals may experience a warm glow, a metaphor for the internal reward they receive from giving to strangers without immediate payoff. The importance of this explanatory mechanism is highlighted by the remarks of Andreoni who claims that ‘experimental data is overwhelming in its support of warm-glow’ (Andreoni, 2006: 1226; Andreoni and Miller, 2002; Palfrey and Prisbrey, 1997).

Empirically, several investigations have highlighted that donors report the sentiment of ‘feeling good’ as a driving factor behind donation decisions (Wunderink, 2000). Neuropsychological studies link charitable giving to areas related to reward activation (overview in Fehr and Camerer, 2007; Harbaugh et al., 2007; Moll et al., 2006). Related research suggests that the effect of the warm glow on certain outcomes may depend on the framing of the decision context. Andreoni (1995) shows that individuals will increase contributions to a public good if the consequence of its provision are framed positively (positive vs negative externalities). He concludes that the differences may be explained by the existence of a warm glow that is (only) triggered by a positive framing of the decision context.

Interestingly, there seems to be lack of evidence that shows the impact of cues with respect to warm glow behaviour in realistic settings.⁴ Therefore, our experiment will apply cues about this internal reward of donating money to a charity.

The experiment

We study charitable giving in the field by analysing donations from customers of a local supermarket chain in the city of Dresden, Germany. The context of the experiment covers the decisions customers face when they return empty but refundable bottles or cans to the supermarket. Since 1 May 2006, all retailers in Germany are obliged by EU regulations to take back bottle types they sell in their stores: there is a bottle deposit on close to all types of bottles, ranging from 8 cents (bottle of beer) to 30 cents (PETs, cans) and entire cases (from 1.50 to 3.10 Euros) that is paid by the customers when the items are purchased. In general, customers are not obliged to return the bottles to the store, but dumping them in the trash results in a direct monetary loss that is usually avoided by most customers. Hence, customers who return their bottles have to decide whether they want to cash out the pre-paid refund or donate it to a German division of the well-known non-profit organization Order of Malta (Malteser Hilfsdienst).⁵ To donate their refund, customers enter the bottles or cases in a refund machine (see Figure 1) and then are required to press the ‘donate’ button twice, otherwise the refund receipt will be printed (see also Knutsson et al., 2013 for the same

procedure). Twenty out of 34 markets of the supermarket chain KONSUM Dresden provide the opportunity to donate the refund.

Prior to the start of our intervention, most of the machines carried a sticker with the logo of the organization and a photo of an older lady taken care of by a nurse, accompanied by a sentence to donate the first couple of bottles to the organization (for details, see the online appendix). With the start of the intervention, these stickers were replaced by our messages (shown in Figure S1 in the Supplemental material) if the markets were assigned to a treatment group. The treatment messages had the same size and were positioned in the same spot across all treated markets. The untreated control group received a sticker that carried the phrase ‘Please donate the first bottles to Malteser’, but did not show the picture of the older lady.

The treatments

Following the four separate theoretical approaches, we customized four different labels that only varied in the written message, while all other design elements were held constant (font size, colour, label and its position of the supermarket chain and non-profit organization). First, cues about social norms are presented by addressing both the content of the socially approved behaviour (‘helping those in need’) together with a message that addresses the normative feature of helping as an obligation. In the present example, an increase in donation levels and donation frequency is expected due to the desire for norm following behaviour to avoid informal sanctions, like feelings of guilt, as well as due to the normative belief about the behaviour of others and their belief about their own behaviour (Bicchieri, 2006):⁶

Social Norm Treatment Please donate your refund – we share an obligation to help those in need. Thank you for your donation. (p. 16)

In contrast, we restrict the information about the existence of a descriptive norm to the expectation about the behaviour of other individuals, without any moral imperative:

Descriptive Norm Treatment Many of our customers regularly donate their refunds. Thank you for your donation.

In addition, we will test whether the observation about the importance of situational circumstances for norm adherence translate to settings of consumer choices. Hence, we formulate:



Figure 1. Refund machine, labels were placed at eyesight level as shown in the photo.

Local Norm Treatment Many of our customers from this store regularly donate their refunds. Thank you for your donation.

In the case of the warm glow condition, we relied on the formulation of survey items that were used in the past to study the tendency for warm glow behaviour (see Liebe et al., 2011):

Warm Glow Treatment Please donate your refund – think of the good feeling in helping others. Thank you for your donation. (p. 116)

Experimental design

We follow the advice by Shadish et al. (2002) and apply and combine many design elements to test the effect of the interventions and to identify the causal mechanisms. We use the following approaches:

1. Untreated matched controls group design;
2. Switching replication design;
3. Removed and repeated treatments design;
4. Different outcome measurements (multiple substantive post-tests).

First, the Untreated Matched Controls Design reflects the standard experimental procedure that compares the outcome between treated and untreated units. With four treatments to be introduced across 20 markets, we matched markets into four groups. The matching was performed by overall levels of monthly refund revenue and monthly donations. This stratified matching on pre-test data assorts five markets into four groups (high revenue/high donations, high revenue/low donations, low revenue/high donations and low revenue/low donation). This procedure enables us to assign the four treatments randomly within each of the four groups, while one market within each group serves as the control group. To study the lasting effect of the interventions, this design of the first phase lasted for 4 months.

In the second phase of our experiment, we applied the Switching Replications Design that withdraws the treatments from the 16 markets treated during the first phase after the 4 months and randomly assigns the treatments to the remaining four markets that served as controls. This second phase of the experiment lasted for only 1 month.

The third phase covered the Removed and Repeated Treatment approach and reintroduced the interventions at markets that served as the treatment groups in the first phase. This phase lasted for another 2 months. Afterwards, we removed all treatments from all markets for 2 months and reintroduced treatments randomly within the matched groups. As the treatments were assigned randomly within the matched groups, some markets received the same treatment of the first intervention period a second time where others received a different one and others served as controls. Table 1 summarizes the sequence of the experimental procedures for one of the matched groups with five markets.

Finally, we check the results for robustness using different outcome measures. The main outcome measure used is the market-specific monthly donation, aggregated for all customers in a particular market. It was not

Table I. Schematic exemplary depiction of the experimental sequence within one group of matched markets M^A-M^E.

Month	1	2	3	4	5	6	7	8	9	10	11	12
Market	R	NR	NR	R		NR	NR			R	R	R
M ^A	○	X ₁	○	X ₁	○	X ₁	○	○	○	X ₄	○	○
M ^B	○	X ₂	○	X ₂	○	X ₂	○	○	○	○	X ₁	○
M ^C	○	X ₃	○	X ₃	○	X ₃	○	○	○	X ₃	○	X ₄
M ^D	○	X ₄	○	X ₄	○	X ₄	○	○	○	X ₂	○	X ₂
M ^E	○	○	○	○	X ₁	○	○	○	○	X ₃	○	X ₃
	Matched controls				Switched treatments	Repeated treatments		Removed treatments		Switched/repeated treatments		

The four treatments X₁-X₄ are assigned randomly (R) or non-randomly if repeated/switched by design (NR) during the 12 months with observations ○ (scheme and notation following Shadish et al. (2002)).

possible to use process data on every single decision (refund vs donate); therefore, we have to rely on aggregate data. Because precise pre-test data were available for this outcome measure, this is our first-best variable for the investigation of the interventions. In addition, we are able to control for the monthly number of donors, the ratio of donors to total customers and the amount of donation per capita. From a theoretical perspective, the treatments aim at increasing the rate of donating the refund, thus the monthly number of donors and the ratio between donors to overall customers actually would serve as the better measure for the effectiveness of our interventions. Unfortunately, due to technical constraints in the process of providing the data, the additional measures only came available during the first experimental period.⁷

Empirical strategies

To capture the effects of the interventions during the first experimental periods, we estimate difference-in-difference models (DID; see Angrist and Pischke, 2008; Card and Krueger, 1994). This allows us to control for time trends in donating behaviour, given that the general trend in donations is similar across all markets (parallel trends assumption, see Abadie, 2005). The standard DID estimation of an observed outcome Y_{imt} (where i are the observed donation in market m) uses a dummy variable $POST$ that indicates whether the data belong to a pre- or post-intervention period and a dummy variable $TREAT$ that takes on the value 1 if the observed outcome in market m was subject to the intervention (else 0). Finally, the DID_{mt} represents the interaction between these dummy variables, where the coefficient δ reflects the DID estimator for the observation of the treatment group after the intervention

$$Y_{imt} = \alpha + \lambda_t (POST_t) + \gamma_m (TREAT_m) + \delta DID_{mt} + \epsilon_{imt}$$

This basic model can be extended to capture multiple treatments and multiple time periods by introducing dummy variables for the additional interventions and post-treatment periods (Wooldridge, 2015). This will result in four treatment-specific DID coefficients that will provide the foundation for the interpretation whether the intervention worked or not.

One drawback of the DID model is that we are not able to use it for the second, third and fourth experimental phase, because we are not able to distinguish pre- and post-interventions as sooner or later all markets are subject to at least one intervention. Hence, this would likely raise criticism that we control for post-treatment observations. Therefore, we use a second identification strategy that accounts for both the longitudinal and

the hierarchical structure of our entire data set. We observe donation levels y_{imt} (Level 1) within markets m (Level 2) in month t and estimate a varying-intercept regression model that checks for the different treatment effects. We also control for aforementioned size effects of the markets by including three dummy variables that reflect whether a market belonged to one of the $k = 4$ groups that were created based on the average amount of refund and donation revenue (REV). Thus, we compute $k - 1$ Dummy variables REV_k . Hence, the effect of the interventions given the simple matching procedure will be tested via cross-level interactions, where

$$GROUP = \sum_1^{T-1} \sum_1^{K-1} \beta * TREAT * REV_k$$

$$Y_{imt} = \beta_1 + \beta_2 (TREAT_{im}) + \sum_{k=1}^3 \beta_{2+k} (REV_k) + GROUP + \zeta_m + \epsilon_{im}$$

Thus, we estimate a model for all periods and all treatments simultaneously and check for whether the effects of a treatment depend on its introduction within the matched groups of markets that are more similar.

Empirical results

Pre-test data

We first explore the data before the first intervention took place. We have data on 34 pre-test months across 20 stores. Markets with fewer observations either did not yet exist (e.g. market #14 ‘Meissner Straße’) or did not have a refund machine with the option to donate the refund, which would be installed later. Although we have monthly donation levels for close to all 34 pre-test months across all markets, again the caveat regarding our pre-test data is the level of aggregation of several indicators that provide information about the characteristics of markets. First, overall levels of refund (in Euros), total number of refund customers and donors were only available as annual data. Second, due to technical reasons of downloading data from the machines, some monthly information was reported missing and had to be replaced with averages derived from the available data.⁸

Figure 2 shows the overall levels of monthly donations in Euros for all 20 markets across all 34 pre-test months. Despite several spikes and the fact that all markets show considerable within variation, we assume that the ‘common trends’ assumption for conducting the difference-in-difference estimation procedure holds. The overall mean monthly donation during the 34 pre-test months was 25.58 Euros per market. Descriptive statistics for the available pre-test months is shown in Table 2.

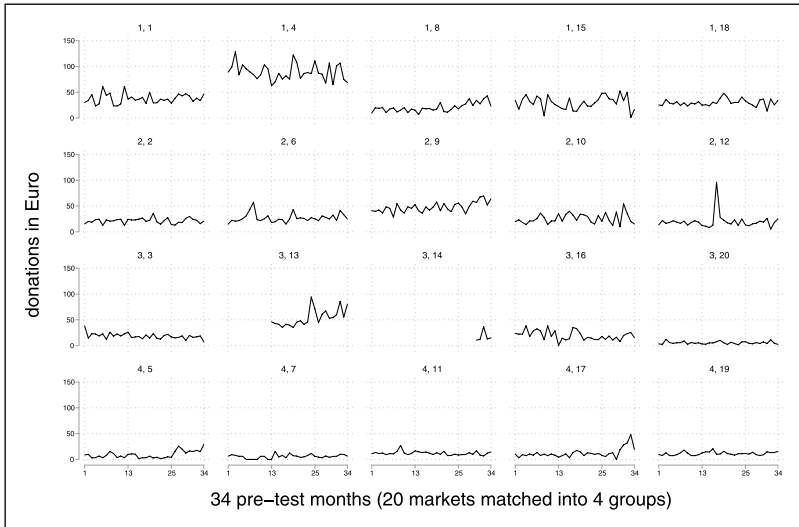


Figure 2. Total monthly donation levels for 34 pre-test months across 20 Markets (in Euro).

There is considerable variation in the donation levels across markets, with market number 4 showing the highest average donation levels across the 34 pre-test months with $y_4=88.94$ Euros and market 20 the lowest levels of average donations ($y_{20}=5.41$ Euros). Using the information from the annual data from 2013 and 2014 on refund customers and donors, we are able to (approximately) assess that during the 34 pre-test months, an average of only 1.11% of all customers donated their refund to charity. Figure 3 illustrates this discrepancy between total number of donors and number of customers for the year 2015 until the end of October when the intervention started.

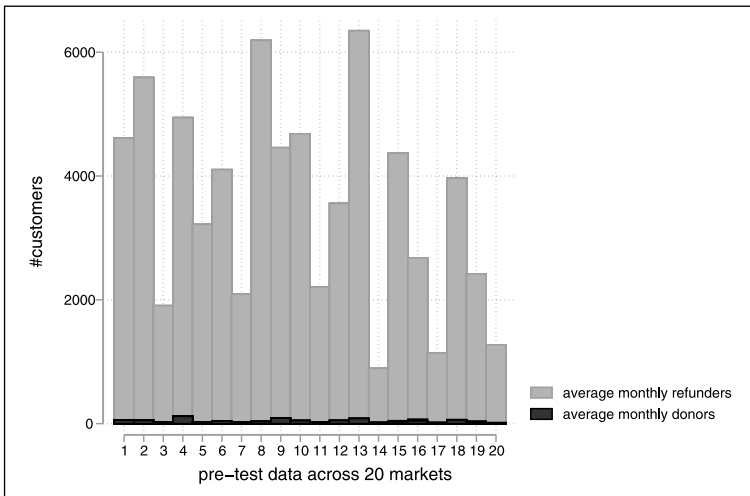
Summarizing briefly, customers are endowed with a small amount of their own money and are asked to allocate their resources by either sending money to charity or keeping it to themselves. This real-life example for a dictator game from lab experiments illustrates that results from the lab about the generosity or benevolence of individuals seem to be greatly exaggerated.⁹

Multivariate models

Before turning to the multivariate analyses, we present some descriptive statistics of the sequential steps of the experiment. Figure 4 shows the total amount of donations for all 20 markets in comparison to the last 10 months of the pre-test data. The coloured bars indicate the different phases of the

Table 2. Summary statistics of monthly refund and donation levels for 34 pre-test months.

	Number of data points	Mean	SD	Min	Max
Monthly donation (in Euros)	639	25.58	21.38	0.00	128.74
Number of donors	620	46.90	28.26	12.00	146.00
Monthly p.c. donation (in Euros)	603	0.55	0.31	0.00	3.16
Ratio donors/total customers	620	1.11	0.45	0.02	2.29

**Figure 3.** Total number of monthly refunders and donors for the last 10 pre-test months, plotted for the 20 markets separately. The average ratio of donors to all customers during all 34 pre-test month was estimated at 1.11%.

experiment. We observe an increase in the first month (November 2015) and a substantial spike in aggregated donations for December 2015. Afterwards, donation levels regress towards the mean level of donations during the pre-test period (indicated by the dotted line). The re-introduction of treatments in April and in August 2016 triggered an increase on total donations as well. It is noteworthy that the total of 835.23 Euros in December 2015 and 800.62 Euros in September 2016 represented the 2 months with the highest amounts of monthly donations ever recorded since the introduction of the donate option.

We now turn to the econometric analyses to test the effectiveness of the different interventions in comparison to the pre-test data. In a first step, we test for the impact of our treatments during the first experimental phase.

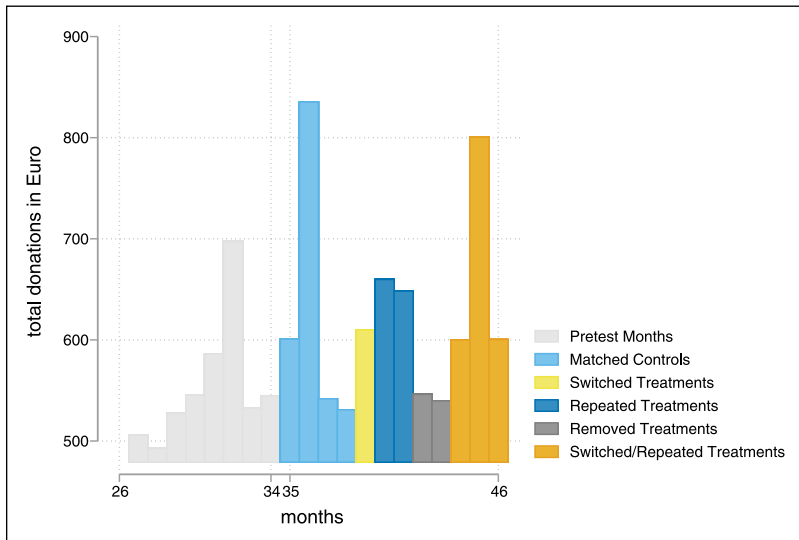


Figure 4. Sequencing of the experimental designs and the total monthly donations generated across all markets (in Euro).

Results are shown in Figure 5. The first take away is that the only intervention that appears to have a small impact on donation levels is the warm glow condition. Markets that were assigned with the cue on the emotional benefit of giving to strangers generated on average 10.61 Euros more in donations. This finding also holds if one uses the other outcome measure of per capita donations (see Figure 8 in Appendix 1). Second, if we check for the robustness of our results using the number of donors as the outcome measure, we find slightly increasing number of donors for the local norm condition, but no effect for the message that addresses warm glow behaviour. We also observe that the month December turned out to be a month with substantial increasing donation levels, which may be attributed to the Christmas festivities around this time of year. Note that we do not find an increase in total number of donors and we do not observe a similar ‘Christmas effects’ in the years 2013 or 2014.

Finally, monthly donation levels of the markets in the comparison group remain surprisingly similar to the ones in the four treatment groups. This may not come as a surprise because the comparison group was not characterized by a complete lack of informative cues (‘Please donate the first bottles to Malteser’).

As mentioned previously, we do not consider the DID estimation to test the treatment effects for the other experimental phases. Instead, we rely on

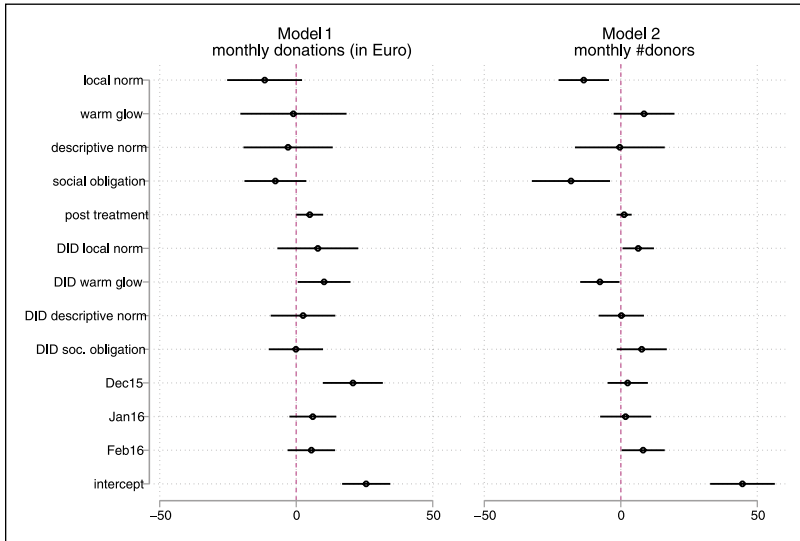


Figure 5. Difference-in-difference estimation results (with clustered standard errors) for the first experimental sequence of the matched control group design. Model 1 covers the outcome variable “monthly donations in Euro”; Model 2 comprises the results on the outcome variable “monthly number of donors”.

multilevel modelling to both explain the variation between and within markets. A null-model with only varying intercepts already replicates the conclusions drawn from Figure 2. With an intraclass correlation coefficient (ICC)=0.80, the larger part of the variance in pre-test donation levels can be attributed to the variation between markets instead of variation within. Furthermore, we check whether the variation in donation levels during the experimental period can be attributed to the market or the treatments, by estimating a three-level varying-intercept model for which donation levels vary within markets (Level 3) and within treatments (Level 2). Results suggest that most of the variation in the donation levels can be attributed to the market level in contrast to level of the treatments. Both results (not shown) underscore the necessity to match more similar markets together to improve the identification of the treatment effects.

In a next step, we focus on the two-level model (observations within markets) and restrict the estimation to the time periods covering the pre-test data and the first phase of the matched control design. Figure 6 illustrates the finding that the matched groups indeed vary as expected. Markets with initially lower levels of donation continue their trend after the intervention. Again, we find that cues on warm glow behaviour produce the highest

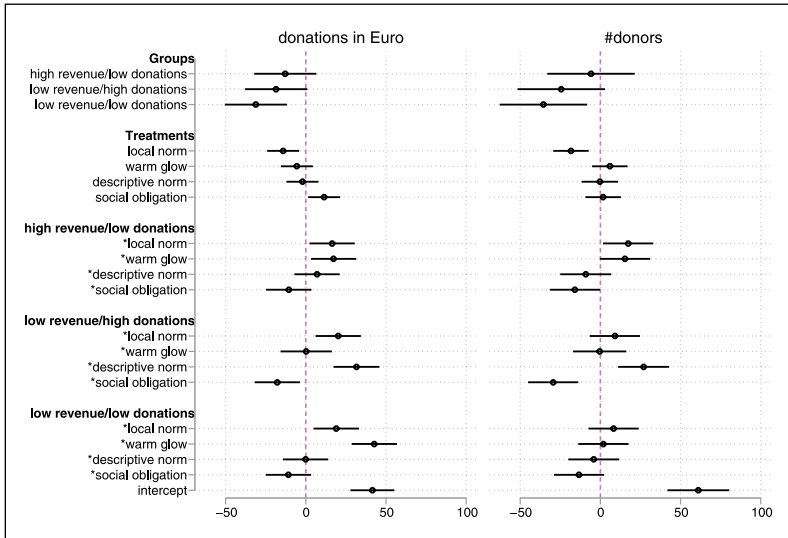


Figure 6. Estimation results from two varying-intercept models on donation levels and on number of donors for the first four months, including interaction effects between the treatments and dummies as indicators for the matched groups.

increases in donation levels, here especially within the group of markets that was characterized by lower levels of refund revenue and low donations levels. Second, we find that the cue about the locality of empirical expectations increases average donation levels within three of our matched groups of market, though the increases are rather small. Both results replicate the finding of the DID estimation. The results for the framing of the situation as a social obligation and as a descriptive norm are rather inconclusive.

We then extend our analyses to include all 12 months of the experiment. Figure 7 shows that if any intervention leads to higher donation levels, we perceive the local norm condition as the one with the most frequent impact on donation levels. While the intervention regarding the social obligation fails to increase donation levels and the message of a descriptive norm does so only within one of the matched groups (low revenue/high donations), the cues about the behaviour of other customers in a particular market (the local norm) generate consistent but small increases in donation levels across groups. Again, the cue on warm glow behaviour reveals its impact especially within small markets (by revenue) that showed a priori the lowest average donation levels.

Finally, we apply different outcome measures (number of monthly donors, the ratio between donors to total number of customers as well as the

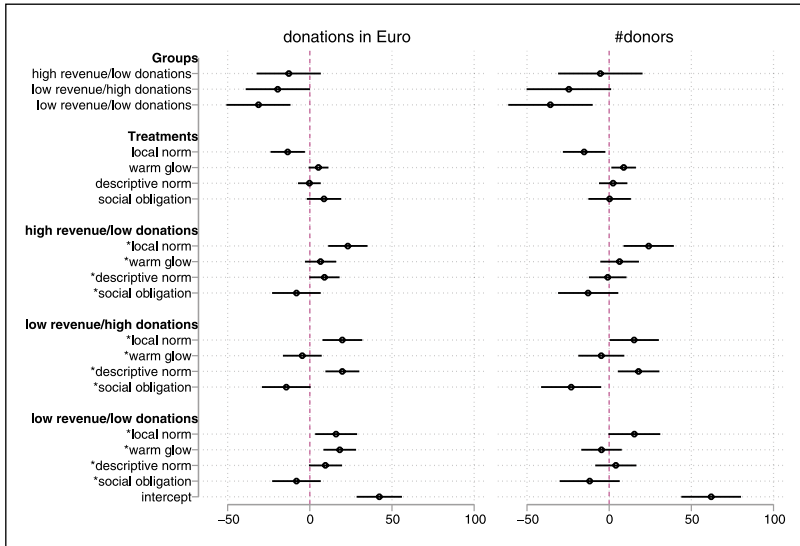


Figure 7. Estimation results from two varying-intercept models on donation levels and on monthly number of donors over the entire experimental period.

amount of donations per capita) that serve as a robustness check for our results. Summarizing the results for the number of donors as the outcome variable of interest (see Appendix 1), we find that the effects of the cue about the benefit of warm glow behaviour are more or less indistinguishable from zero across all groups. All other treatments do not generate consistent and homogeneous impact on the number individuals who opt to donate. In contrast, the ratio of donors to total customers increased when machines were assigned to the warm glow treatment group, while the other treatments fail to generate any lasting increases (see Figures 8 and 9 in Appendix 1). The only result that stands out is the sharp increase of per capita donations for the warm glow treatment in markets where the lowest overall donation levels were registered in the past. Putting it blatantly, the findings seem to underscore a behaviour that can best be described as the revealed behaviour of self-regarding altruists: supermarkets that were characterized with the lowest pre-test donation levels scored best in a condition that highlighted the self-regarding, non-monetary reward of giving to charity.

Summarizing the results, the experiments generated two takeaways. First, messages that frame donations as warm glow behaviour and the description of the local behaviour of other customers were the only cues that increased donation levels across the different experimental designs and experimental phases. Neither the framing of normative expectations nor that

of empirical expectations increased average donation levels by consumers in a consistent way. Second, despite the apparent success of the experiment – the experiment generated 5 out of the 6 months with the highest total donation levels ever recorded – the overall ratio between donors and customers who cashed out their refund remained always below 3%. We will consider some possible explanation for this finding in the concluding chapter.

Discussion

This study presents the results of a field experiment in Germany that investigated whether strategic information cues induce consumers to donate their bottle refund to a charitable organization. The cues conveyed messages about the normative appropriateness and the empirical expectations regarding the behaviour of other consumers as well as the self-regarding emotional benefit of donating money to a non-profit organization (warm glow behaviour). In addition, the experiment included a test of the proposition that norm-following behaviour can be attenuated by cues about the locality of a descriptive norm (as proposed by Goldstein et al., 2008).

The results can be summarized as follows: first, from our pre-test data and from our experiments we observe that an overwhelming portion of consumers prefers to keep their refunds to themselves, as the experimental interventions failed to generate any substantial increase in the overall amounts of money donated through the refund machines or an increase of individual donors. Although the experimental periods yielded 5 months with the highest overall donation levels ever recorded, neither the ratio of donors to customers nor the average donation levels were raised in any substantial fashion. Second, the cues of describing the behaviour of other consumers in a particular market – the local norm treatment – lead to small but consistent increases in donation levels. Third, the cue that addressed warm glow behaviour as an incentive for donating one's refund could be identified as one experimental condition that lead to an increase in donation levels across different designs and with help of different identification strategies. Both a difference-in-difference estimation and multi-level models provided some evidence for the impact of the (self-regarding) emotional benefit of donating money to a non-profit organization due to warm glow of giving.

Together, both findings contribute to a long debate within sociology and economics about the appropriate theoretical assumptions of a model of man (Opp, 1999; Thaler and Sunstein, 2008; Vanberg, 2008). In addition, we gather evidence about individual decision making in real-life settings that can be compared to a situation individuals face when they play dictator games in the lab. Here, we find clear evidence that individuals can be

considered self-regarding actors with rather stable preferences. There is only a small portion of consumers who actually donated their refund without a direct benefit. In parts, they did so within a decision context that highlighted the equally self-regarding benefit of experiencing an emotional ‘payoff’ from altruistic giving. By this, the investigation provided one of the rare cases of evidence for the existence of warm glow behaviour outside of the lab (see for an overview Andreoni, 2006).

Another contribution to the literature lies in the observation that the importance of situational circumstances for norm adherence may also translate to settings of consumer choices (as found in Goldstein et al., 2008). Cues about the empirical expectation with regard to the behaviour of other consumers within a particular supermarket (local norm) had a higher impact on donation levels in comparison to the framing of general empirical expectation (descriptive norm). One possible explanation is that the more general information about the behaviours of others caused a ‘crowding out’ of potential donors across these treatments (see Huck et al., 2015). More research is necessary whether the impact of so-called provincial norms holds across different contexts and in contrast to the mere description of the norm following of other individuals.

The limitations of our investigations are numerous. The availability of precise pre-test data was missing and detailed field data after the interventions started suffered from data processing problems. In retrospect, the evidence regarding the impact of the treatments could have been estimated more precisely if the start of the interventions would have postponed until several months to gather more precise pre-test data. Still, it probably would not have changed the main take away that customers are very reluctant in donating small amounts of money to charity. As usual, field experiments face challenges because exogenous variables are hard to control. Usage of the donation button may have been difficult and may have led to an erroneous refund decisions in the case of mishandling the procedure. Finally, customers may have a negative view of the non-profit organization that would have profited from the donations, perhaps because the German division of the catholic Order of Malta is perceived negatively in an area mostly populated by citizens of non-denomination. Studies have found that religious persons or individuals who show a stronger certainty with regard to their religious faith are also more likely to contribute to religious purposes (e.g. Corcoran, 2013). Unfortunately, we cannot control for the religiosity of customers, but we also do not expect any antagonistic perception of the well-known and widely respected work of the Order of Malta in Germany and Saxony in particular.

What can explain the rather low levels of donations? One explanation could be that the effort of storing empty bottles at home, bringing them

back to the supermarket and carrying them to the machine induces an additional effort that customers have to bear. Donating the refund simply sacrifices these efforts and incurs additional costs beyond the amount donated. One possible theoretical explanation is provided by the theory of goal-framing (Lindenberg, 2009; Lindenberg and Steg, 2007).¹⁰ There, individuals deal with social situations by focusing on certain foreground and background goals. The three main goals identified by the theory are a hedonic goal (short term), a gain goal (long term) and a normative goal (to 'act-appropriate', see Lindenberg, 2009: 57). Any of the three can be a foreground goal that is perceived as focal through a cognitive process of selective activation, hence 'people are made to be momentarily rather one-sided by the goal that is focal at the moment' (Lindenberg, 2009: 56). Therefore, only one of three goals can be the focal goal in a particular situation, with the other two serving the role as background goals. Which goal will represent the focal one depends on whether a cue (internal or external) activates a certain goal frame, from which all subsequent processes are governed. Therein, focal goals can be weakened or strengthened by the background goals.

At the refund machine, we can assume that customers arrive at supermarkets with a clear focus – satisfying the hedonic short-term goal of cashing out the refund. In accordance with the theory, one can also assume that hedonic short-term goals usually represent the strongest, normative goals the weakest incentives for certain actions, especially within a consumption setting (Lindenberg, 2009). Unless the normative goal receives 'more personal, social and institutional support than do the other overarching goals in order to be [...] dominant', the hedonic goal will constitute the dominant frame in the particular situation. If customers would experience external rewards from third parties (bystanders, the market staff) for donating money or would they experience sanctions in the form of vilification or dismissal as a result of cashing out the refund, the normative goal may overrule the a priori foreground goal. But given that the donations are performed anonymously, alone, without any form of third-party gratification or the opportunity to signal one's own trustworthiness through charitable giving (see Fehrler and Przepiorka, 2013), the low donation levels may not come as such a surprise.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

Supplemental material

Supplemental material for this article is available online.

Notes

1. Source: <https://www.nptrust.org/philanthropic-resources/charitable-giving-statistics/> (accessed 29 November 2017).
2. Numerous studies exist that investigate the impact of warm glow giving and social norms on individual decision making in laboratory experiments. See the overview in Andreoni (2006) and Kimbrough and Vostroknutov (2016). Some studies vary the pool of participants, for example, Liebe and Tuitc (2010) who investigate the effect of status on altruistic giving by playing dictator games with pupils from different schools. Lately, Chang et al. (2018) have argued that framing may occur because they evoke different norms, but they test their prediction in a laboratory experiment and not in the field.
3. We could not investigate psychological traits that are usually measured through surveys. We assume that cost–benefit aspect with regard to the tax deductibility of donations does not play a role in our setting. Also, the aspects of gains in reputation are absent, due to the conditions of decision. We return to this point in the concluding section. For general overviews on the determinants of philanthropy, see Bekkers and Wiepking (2011) and Sargeant and Woodliffe (2005).
4. For instance, the just recently published *Handbook of Field Experiments* (Banerjee and Duflo, 2017) does not cover a single field experiment that tests warm glow giving, although it covers several empirical investigations that are concerned with how individuals have preferences about the outcomes of others, thus social preferences. We do not use this term, as we follow the arguments made elsewhere (Kimbrough and Vostroknutov, 2016; Vanberg, 2008).
5. The Order of Malta is international catholic non-profit organization that provides care services for the elderly and sick. Another area where the organization is involved is social work and ambulant medical assistance and it also has an international branch that focuses on emergency aid and disaster management. In large parts, this work is performed voluntarily by 50,000 volunteers and 30,000 professional workers in Germany alone.
6. This formulation resembles survey items that attempt to measure an obligation to pay (see Liebe et al., 2011). Note these are our own translations from the German language.
7. We started our intervention in November 2015 after having contacted the manufacturer of the refund machines. We were assured that the precise data would be available, but it turned out that all machines had to undergo a manual software update to enable the provision of detailed monthly data. This costly procedure was performed only after the beginning of the first intervention, in the first week of January 2016. That is why the first experimental phase was prolonged by 2 months.
8. For example, if 2 months (e.g. June and July) were reported with missing data, total refund numbers from August in comparison to May were used to compute the average levels.
9. One meta-study on dictator games has reported that on average only 36% (in contrast to our 98.9%) do not send any money to the unknown recipient (Engel, 2011: 584).

10. The author would like to thank Rafael Wittek for this remark during a presentation of this study in the summer of 2016.

References

- Abadie A (2005) Semi-parametric difference-in-differences estimators. *The Review of Economic Studies* 72(1): 1–19.
- Adena M and Huck S (2017) Matching donations without crowding out? Some theoretical considerations a field and a lab experiment. *Journal of Public Economics* 148: 32–42.
- Altmann S, Falk A, Heidhues P, et al. (2014) *Defaults and donations: evidence from a field experiment*. Working Paper no. 5118, December. Munich: CESifo Group.
- Andreoni J (1989) Giving with impure altruism: applications to charity and Ricardian equivalence. *Journal of Political Economy* 97(6): 1447–1458.
- Andreoni J (1990) Impure altruism and donations to public goods: a theory of warm-glow giving. *Economic Journal* 100(401): 464–477.
- Andreoni J (1995) Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments. *The Quarterly Journal of Economics* 110(1): 1–21.
- Andreoni J (2006) Philanthropy. In: Kolm SC and Ythier JM (eds) *Handbook of the Economics of Giving, Altruism and Reciprocity*. Amsterdam: Elsevier, pp. 1201–1269.
- Andreoni J and Miller J (2002) Giving according to GARP: an experimental test of the consistency of preferences for altruism. *Econometrica* 70(2): 737–753.
- Andreoni J, Rao JM and Trachtman H (2017) Avoiding the ask: a field experiment on altruism, empathy, and charitable giving. *Journal of Political Economy* 125(3): 625–653.
- Angrist JD and Pischke JS (2008) *Mostly Harmless Econometrics: An Empiricists Companion*. Princeton, NJ: Princeton University Press.
- Balafoutas L and Nikiforakis N (2012) Norm enforcement in the city: a natural field experiment. *European Economic Review* 56(8): 1773–1785.
- Banerjee AV and Duflo E (2017) *Handbook of Economic Field Experiments*, vols. 1 and 2. Amsterdam: Elsevier.
- Bekkers R and Wiepking P (2011) A literature review of empirical studies of philanthropy. *Nonprofit and Voluntary Sector Quarterly* 40(5): 924–973.
- Berger J and Hevenstone D (2016) Norm enforcement in the city revisited: an international field experiment of altruistic punishment norm maintenance and broken windows. *Rationality and Society* 28(3): 299–319.
- Bicchieri C (2006) *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press.
- Bicchieri C (2010) Norms preferences and conditional behaviour. *Politics Philosophy & Economics* 9: 297–313.
- Card D and Krueger AB (1994) Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania. *The American Economic Review* 84(4): 772–793.

- Chang D, Chen R and Krupka E (2018) Rhetoric matters: a social norms explanation for the anomaly of framing. *WZB Discussion Paper SP II 204*: 1–52.
- Cialdini RB (2007) Descriptive social norms as underappreciated sources of social control. *Psychometrika* 72(2): 263–268.
- Cialdini RB, Demaine LJ, Sagarin BJ, et al. (2006) Managing social norms for persuasive impact. *Social Influence* 1(1): 3–15.
- Cialdini RB, Reno R and Kallgren CA (1990) A focus theory of normative conduct: recycling the concept of norms to reduce littering in public spaces. *Journal of Personality and Social Psychology* 58(6): 1015–1026.
- Corcoran KE (2013) Divine exchanges: applying social exchange theory to religious behavior. *Rationality and Society* 25(3): 335–369.
- DellaVigna S (2009) Psychology and economics: Evidence from the field. *Journal of Economic Literature* 47(2): 315–72.
- Deutscher Spendenrat (2017) *Bilanz Des Helfens 2017*. Berlin: Deutscher Spendenrat. Available at: http://www.spendenrat.de/wp-content/uploads/2017/02/Bilanz_des_Helfens_2017.pdf
- Elster J (1989) Social norms and economic theory. *Journal of Economic Perspectives* 3: 99–117.
- Engel C (2011) Dictator games: a meta study. *Experimental Economics* 14(4): 583–610.
- Fehr E and Camerer CF (2007) Social neuroeconomics: the neural circuitry of social preferences. *Trends in Cognitive Sciences* 11(10): 419–427.
- Fehrer S and Przepiorka W (2013) Charitable giving as a signal of trustworthiness: disentangling the signaling benefits of altruistic acts. *Evolution and Human Behavior* 34(2): 139–145.
- Frey BS and Meier S (2004) Social comparisons and pro-social behavior: testing conditional cooperation in a field experiment. *American Economic Review* 94(5): 1717–1722.
- Goldstein NJ, Cialdini RB and Griskevicius V (2008) A room with a viewpoint: using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research* 35(3): 472–482.
- Grossman Z and Van Der Weele JJ (2017) Self-image and willful ignorance in social decisions. *Journal of the European Economic Association* 15(1): 173–217.
- Harbaugh WT, Mayr U and Burghart DR (2007) Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* 316(5831): 1622–1625.
- Hechter M and Opp K-D (eds) (2001) *Social Norms*. New York: Russel Sage Foundation.
- Huck S and Rasul I (2011) Matched fundraising: evidence from a natural field experiment. *Journal of Public Economics* 95(5): 351–362.
- Huck S, Rasul I and Shephard A (2015) Comparing charitable fundraising schemes: evidence from a natural field experiment and a structural model. *American Economic Journal: Economic Policy* 7(2): 326–369.
- Karlan D and List JA (2007) Does price matter in charitable giving? Evidence from a large-scale natural field experiment. *American Economic Review* 97(5): 1774–1793.

- Karlan D, List JA and Shafir E (2011) Small matches and charitable giving: evidence from a natural field experiment. *Journal of Public Economics* 95(5–6): 344–350.
- Keizer K, Lindenberg S and Steg L (2008) The spreading of disorder. *Science* 322(5908): 1681–1685.
- Keuschnigg M and Wolbring T (2015) Disorder social capital and norm violation: three field experiments on the broken windows thesis. *Rationality and Society* 27(1): 96–126.
- Kimbrough EO and Vostroknutov A (2016) Norms make preferences social. *Journal of the European Economic Association* 14(3): 608–638.
- Knutsson M, Martinsson P and Wollbrant C (2013) Do people avoid opportunities to donate: a natural field experiment on recycling and charitable giving. *Journal of Economic Behavior & Organization* 93: 71–77.
- Krupka EL and Weber RA (2013) Identifying social norms using coordination games: why does dictator game sharing vary? *Journal of the European Economic Association* 11(3): 495–524.
- Levitt SD and List JA (2007) What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives* 21(2): 153–174.
- Liberman V, Samuels SM and Ross L (2004) The name of the game: predictive power of reputations versus situational labels in determining prisoners dilemma game moves. *Personality and Social Psychology Bulletin* 30(9): 1175–1185.
- Liebe U and Tutic A (2010) Status groups and altruistic behaviour in dictator games. *Rationality and Society* 22(3): 353–380.
- Liebe U, Preisendörfer P and Meyerhoff J (2011) To pay or not to pay: competing theories to explain individuals willingness to pay for public environmental goods. *Environment and Behavior* 43(1): 106–130.
- Lindenberg S (2009) Why framing should be all about the impact of goals on cognitions and evaluations. In: Hill P, Kalter F, Kopp J, et al. (eds) *Hartmut Essers Erklärende Soziologie: Kontroversen und Perspektiven*. Frankfurt/Main: Campus Verlag, pp. 53–79.
- Lindenberg S and Steg L (2007) Normative gain and hedonic goal frames guiding environmental behaviour. *Journal of Social Issues* 63(1): 117–137.
- List JA (2009) Social preferences: some thoughts from the field. *Annual Review of Economics* 1(1): 563–579.
- 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.
- Moll J, Krueger F, Zahn R, et al. (2006) Human fronto–mesolimbic networks guide decisions about charitable donation. *Proceedings of the National Academy of Sciences of the United States of America* 103(42): 15623–15628.
- Nelson TE, Clawson RA and Oxley ZM (1997) Media framing of a civil liberties conflict and its effect on tolerance. *American Political Science Review* 91(3): 567–583.

- Neumann R and Mehlkop G (2018) Pro-environmental behavior and the relationship between attitudes, costs, and situational framing – testing competing theories with data from the GESIS panel. *Zeitschrift Für Soziologie* 47(2): 101–118.
- Opp KD (1999) Contending conceptions of the theory of rational action. *Journal of Theoretical Politics* 11(2): 171–202.
- Ostrom E (2000) Collective action and the evolution of social norms. *Journal of Economic Perspectives* 14(3): 137–158.
- Palfrey TR and Prisbrey JE (1997) Anomalous behavior in public goods experiments: how much and why? *The American Economic Review* 87: 829–846.
- Przepiorka W and Berger J (2016) The sanctioning dilemma: a quasi-experiment on social norm enforcement in the train. *European Sociological Review* 32(3): 439–451.
- Sargeant A and Woodliffe L (2005) The antecedents of donor commitment to voluntary organizations. *Nonprofit Management and Leadership* 16(1): 61–78.
- Schulz JF, Thiemann P and Thöni C (2016) *Nudging generosity: choice architecture and cognitive factors in charitable giving*. IZA Discussion Paper no. 11097, October. Bonn: IZA-Institute of Labor Economics.
- Shadish WR, Cook TD and Campbell DD (2002) *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin Company.
- Shang J and Croson R (2009) A field experiment in charitable contribution: the impact of social information on the voluntary provision of public goods. *The Economic Journal* 119: 1422–1439.
- Thaler RH and Sunstein CR (2008) *Nudge: Improving Decisions about Health Wealth and Happiness*. New Haven, CT; London: Yale University Press.
- Tversky A and Kahneman D (1981) The framing of decisions and the psychology of choice. *Science* 211(4481): 453–458.
- Vanberg VJ (2008) On the economics of moral preferences. *American Journal of Economics and Sociology* 67(4): 605–628.
- Wooldridge JM (2015) *Introductory Econometrics: A Modern Approach*. Toronto, ON, Canada: Nelson Education.
- Wunderink SR (2000) The economics of consumers gifts and legacies to charitable organisations. *International Journal of Nonprofit and Voluntary Sector Marketing* 5(3): 268–287.

Appendix I

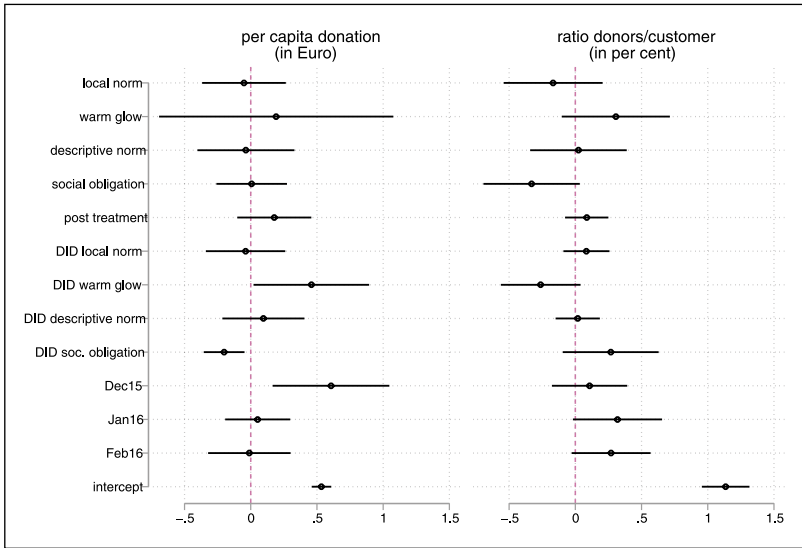


Figure 8. Difference-in-difference estimation results (with clustered standard errors) for the first experimental sequence of the matched control group design. The first results (left) cover the outcome variable “monthly per capita donations (in Euro)”; the second model (right) comprises the results on the outcome variable “ratio of donors/customers”.

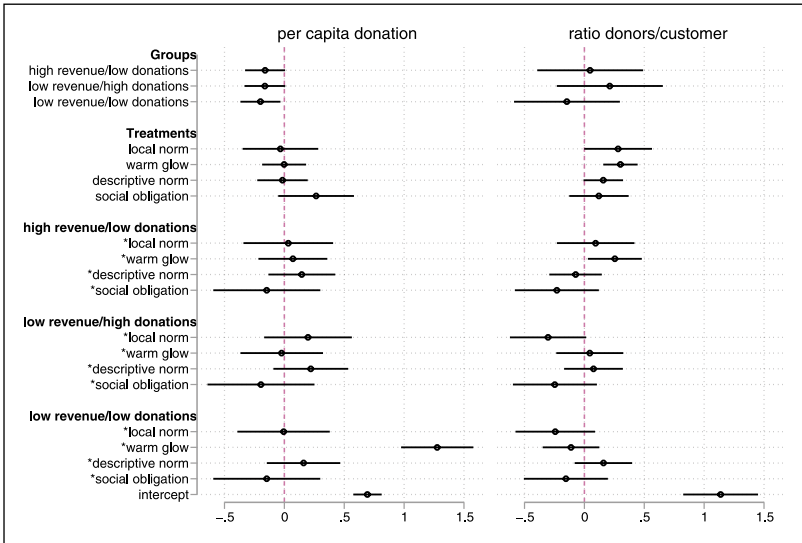


Figure 9. Estimation results from two varying-intercept models on per capita donation levels (in Euro) and on the ratio of donors/ customer (in per cent) over the entire experimental period.