Who to Target in Fundraising? A Field Experiment on Gift Exchange

Tobias Cagala, Ulrich Glogowsky, Johannes Rincke*

February 15, 2019

Abstract

This paper studies the optimal targeting of gifts in fundraising. We implement a randomized field experiment in a setting where a fundraiser sends out annual solicitation letters to a large population of individuals. Three treatment groups are compared: a control group, a gift treatment group, and a gift-plus-recognition treatment group. We identify a distinct heterogeneity in responses to the gift treatment: both the extensive margin and the intensive margin effects are hump-shaped with respect to donors’ baseline willingness to donate, translating into a hump-shaped average treatment effect. Due to the heterogeneity, up-front gifts are cost effective only for individuals with a moderate baseline probability to donate. Targeting the gift to those individuals leads to net revenue gains of 37.3 percent relative to a uniform provision of the gift. We also demonstrate that gifts are not profitable if they are framed in terms of recognition for past donations. The latter finding suggests that behavior in gift exchange is indeed driven by reciprocity concerns.

JEL codes: C93; D64; D82; H41; L31
Keywords: Fundraising; Charitable Giving; Gift Exchange; Targeting

*Cagala: University of Erlangen-Nuremberg (tobias.cagala@fau.de); Glogowsky: University of Erlangen-Nuremberg (ulrich.glogowsky@fau.de); Rincke: University of Erlangen-Nuremberg (johannes.rincke@fau.de). We thank seminar participants at various places for helpful comments. We are grateful for financial support from the Hans-Frisch-Stiftung and the Emerging Field Initiative at University of Erlangen-Nuremberg.
1 Introduction

Among U.S. charities, fundraising expenditures correspond to 12 percent of private donations (Andreoni and Payne, 2011). Given the relative importance of fundraising costs, inefficient fundraising may lead to a situation where charities waste a significant share of their resources. The ability to design optimal fundraising strategies is therefore of crucial importance for the overall efficiency of the philanthropic sector. Furthermore, individuals are sensitive to how charities spend their funds. Specifically, the evidence suggests a strong aversion of donors to charities’ overhead costs (Tinkelman and Mankaney, 2007; Gneezy et al., 2014). As a result, inefficient fundraising poses a threat to a charity’s reputation among donors, and thus the potential to raise donations in the future.

Given the pressure on charities to spend less on nonprogrammatic costs, it is surprising that there is little systematic evidence on how to target fundraising tools most effectively. This paper contributes to closing this gap by studying the fundraising strategy of gift exchange. Most importantly, we provide evidence on how charities can exploit information on individuals’ past donation behavior for the optimal targeting of gifts. We consider two forms of gifts: unconditional (up-front) gifts and gifts which are made in response to donations and thus signal recognition of donors’ pro-social behavior.

To assess the value of information on past donations for targeted gift exchange, we consider a setting where a fundraiser sends out annual solicitation letters to a large population of potential donors irrespective of individuals’ past donation behavior. Using panel data on donations for up to nine years, we predict each individual’s probability to give in the baseline. This probability can be interpreted as a measure for individuals’ baseline willingness to donate. We proceed by setting up a large-scale randomized field experiment in cooperation with the fundraiser. The experiment aims at understanding the heterogeneity in responses to gifts with respect to individuals’ baseline willingness to donate.

Before the experiment, the fundraiser used solicitation letters without gifts. In the experiment, we sent out letter treatments to a control group (letter without gift) and two treatment groups. Individuals in the gift treatment group received the letter and a small gift (three folded postcards with three envelopes). The letter was identical to the letter received by control group individuals, except for one sentence stating that the fundraiser ‘would like to provide the included postcards as a gift’. Individuals in the gift-plus-recognition treatment group received a letter that combined a gift with explicit recognition for pro-social behavior in the past. The treatment was identical to the gift treatment except for framing the gift as a response to past donations. This was achieved by including a sentence in the letter, stating that the fundraiser ‘would like to provide the included postcards as a gift to acknowledge your contributions in previous years’. Based on the idea of sequential reciprocity (Dufwenberg and Kirchsteiger, 2004), we
expect that recipients in the gift treatment perceive the fundraiser’s behavior as more kind compared to the gift-plus-recognition treatment. We thus predict a stronger positive treatment response in the gift treatment relative to the gift-plus-recognition treatment.

Our main contributions are as follows. First, we identify a distinct heterogeneity in responses to the gift treatment with respect to donors’ baseline willingness to donate. Using the approach of Staub (2014), we decompose the average treatment effect into an extensive margin and an intensive margin effect. We find both effects to be hump-shaped with respect to individuals’ baseline probability to give, translating into a hump-shaped average treatment effect. In our setting, the heterogeneity in the treatment response implies that up-front gifts are cost effective only for a subset of individuals with a moderate baseline probability to donate. In contrast, it is not profitable for the charity to provide the gift to individuals with a baseline probability to donate below or above certain threshold levels. We estimate that a mailing with an optimal targeting of the gift results in a net revenue gain of 37.3 percent relative to a uniform mailing comprising the gift.\footnote{The calculation assumes that in both scenarios, the charity would send a letter to all individuals.} This finding highlights the importance of targeting for efficient fundraising.

Second, we extend previous work on gift exchange (Falk, 2007; Alpizar et al., 2008a,b; Landry et al., 2010) by showing that combining a gift with recognition for past contributions wipes out most of the positive effect of providing a gift alone. The treatment effect of the gift-plus-recognition letter is small and insignificant, implying that gifts which are provided conditional on past donations are likely not profitable when applied to all warm-list individuals.\footnote{Warm-list individual are defined as individuals who donated at least once prior to the treatment year. Cold-list individuals are defined as individuals who never made a donation before the treatment year.} This finding is in line with a sequential-reciprocity interpretation along the lines of Dufwenberg and Kirchsteiger (2004). A final piece of evidence relates to the intertemporal substitution of donations in response to gifts. We demonstrate that an initial gift treatment in one year has no discernible effect on donations in the following year if the fundraiser uses a standard mailing without gift in the second year. This means that in our context, providing gifts does not result in an intertemporal substitution of donations.

The remainder of the paper is organized as follows. Section 2 explains our research design, Section 3 discusses our findings, and Section 4 concludes.

\section{Experimental Design and Conceptual Framework}

\subsection{Background and Data Sources}

We implemented our natural field experiment in cooperation with a fundraiser operating within the structure of the Catholic church in an urban area in Germany. The fundraiser
sends out annual solicitation letters to all households living in the urban area where at least one household member is member of the Catholic church. The solicitation letter highlights the fundraiser’s cause, which is to support the local church parishes in maintaining their clergy houses, parish centers, and churches, and asks recipients for a donation. Attached to the letter is a bank transfer form pre-filled with the fundraiser’s bank account information and the recipient’s name. Donations are made exclusively via bank transfer, and the fundraiser does not provide any information about individual donations to the church parishes. Donations are thus made in a private information setting.

Our analysis rests on data that combine a vector of recipients’ individual characteristics obtained from administrative records of the Catholic church (such as marital status, gender, household composition, and age) with comprehensive information on individual donations. We implemented the natural field experiment in two consecutive years, 2014 and 2015. In addition to individual donations in those two years, we exploit detailed information on the universe of individual contributions to the fundraiser in the pre-treatment years 2005–2013. Because the fundraiser sends out the solicitation letter to all households irrespective of previous donations (and has done so in all pre-treatment years that we cover), the data on individual donations cover the universe of church members in the area studied. This allows us to go beyond the typical distinction between cold-list and warm-list individuals. Instead, we use the rich panel data on individual pre-treatment donations to proxy for the continuous variation in individuals’ baseline willingness to donate in a large population of individuals. Using this measure, we study the heterogeneity in treatment effects along individuals’ baseline donation behavior.

2.2 Baseline Willingness to Donate

We define the baseline willingness to donate for individual $i$, $\hat{R}_{i,t+1}$, as $i$’s predicted donation probability in the treatment period $t + 1$ absent any treatment. Exploiting our comprehensive pre-treatment data, we obtain this proxy from a Probit regression that exploits past donation behavior. To derive the prediction, we regress a dummy variable $R_{i,t}$, taking the value one if individual $i$ donates in period $t$ and zero otherwise, on $L$ own lags and lags of an indicator $S_{i,t-l}$, specifying whether $i$ was living in the urban area

---

3Using pre-treatment payment behavior as a baseline characteristic relates our work to Dwenger et al. (2015) and Boyer et al. (2015).
in the period $t - l$:

$$R_{i,t} = \alpha + \sum_{l=1}^{L} \beta_l \cdot R_{i,t-l} + \sum_{l=1}^{L} \gamma_l \cdot S_{i,t-l} + u_{i,t}. \quad (1)$$

Our measure for individual $i$’s baseline willingness to donate in the treatment period $t + 1$ is recovered as the first-step out-of-sample predictions $\hat{R}_{i,t+1}$ from this regression. Counting predictions with a probability greater than 0.5 as predicting a strictly positive donation, $\hat{R}_{i,t+1}$ correctly predicts 86% of decisions in the control group.

### 2.3 The Natural Field Experiment

In the natural field experiment, we manipulated the content of the solicitation letter. In 2014, the first year with a randomized intervention, we assigned a stratified sample of 6,923 individuals to a control group and a gift treatment group. The strata were defined over quintiles of individuals’ baseline willingness to give, household type indicators, and age. This stratification strategy ensures that there is common support along individuals’ baseline willingness to donate in all treatments, which facilitates analysis by subgroup.

The control group letter informed recipients about the fundraiser’s cause (maintaining clergy houses, parish centers, and churches) and asked them to donate. Letters to individuals in the gift treatment group included a gift. The gift consisted of three different postcards showing flower motives, together with three envelopes (see Figure 1). The letter in the gift treatment group was identical to the control group letter apart from the fact that recipients were informed in a short remark following the two main paragraphs that the included postcards were a gift. The remark stated that the fundraiser “would like to provide the included postcards as a gift”. The sampling frame for the year 2014 comprised individuals from the cold list (people who had never donated to the fundraiser pre-treatment) and from the warm list (individuals who had donated at least once to the fundraiser pre-treatment).

In 2015, the second year with a randomized intervention, we assigned 5,417 individuals to a control group and two treatment groups. The first treatment replicated the gift treatment from the previous year, with exactly the same gift included in the letter and the exact same remark on the purpose of the gift. The second treatment was a gift-plus-

---

4We choose a specification with 5 lags, which maximizes the Bayesian information criterion and achieves the highest accuracy in out-of-sample predictions. For measuring out-of-sample predictive power, we exclude the last available year in the estimation and evaluate the share of correctly predicted decisions for this year after estimating the model on the remaining data.

5The four household types are defined as single Catholic female, single Catholic male, Catholic female married to non-Catholic spouse, married Catholic male.

6The specification of the gift relates our design to Falk (2007), who also used postcards and envelopes. The additional cost for the fundraiser to send out a gift letter (relative to sending out a standard letter without gift) was €1.16. This comprises €0.47 for the postcards and envelopes, and €0.69 additional costs for preparing the mailings and higher postage.
recognition treatment. It was identical to the gift treatment in all respects, including the gift provided, apart from the fact that the letter stated that the fundraiser “would like to provide the included postcards as a gift to acknowledge your contributions in previous years”. Hence, the difference between the gift treatment and the gift-plus-recognition treatment is that the letter in the gift-plus-recognition framed the gift as a response to past donations. In contrast to 2014, the sampling frame for the year 2015 comprised only individuals from the warm list.

Figure 2 summarizes the sampling frame. It is important to note that individuals who were treated in 2014 were not treated in 2015. This means that all individuals who found themselves in either the gift or the gift-plus-recognition treatment groups in either year were treated for the first time. This allows us to pool observations from 2014 and 2015 when studying the effects of the gift treatment. Given that individuals were treated only once, we can also test for a potential intertemporal substitution (or crowding-out) effect of providing gifts.

Table 4 in the appendix gives an overview on individual characteristics by treatment. The treatment groups are balanced in all observable dimensions.

3 Responses to Gifts

3.1 Aggregate Treatment Effects

Table 1 summarizes our main results and shows how charitable giving is causally affected by the gift (columns 1 to 8) and the gift-plus-recognition treatment (columns 9 and 10). It contains the relative frequency of donation, the amount donated, and the corresponding unconditional OLS estimations of treatment effects. Considering the gift treatment, the table not only contains results for the overall sample pooled across the years 2014 and 2015, but also results for the pooled cold list and the pooled warm list sample. This sample split allows us to study the heterogeneity regarding individuals' baseline willingness to donate in an aggregate way. To ensure comparability between the gift and the gift-plus-recognition treatment, the table also shows the causal effect of the gift treatment for the warm list sample in 2015.

Three insights emerge from Table 1. First, including a gift not only raises the relative frequency of donations but also significantly affects the level of donations. In the overall sample (columns 1 and 2), the relative frequency of people who donate in the control condition is 8.6%. This fraction increases to 10% under the gift condition. Thus, the gifts significantly increase the frequency by about 16%, suggesting that by providing potential donors with gifts, charities can trigger a reciprocal response. This is in line with earlier

7Because we oversampled warm list individuals, we present weighted averages across cold list and warm list individuals for the overall sample.
literature on the importance of reciprocity as a driver for donations (Hoffman, 1996; Fehr and Gächter, 2000; Falk, 2007). However, the overall treatment effect on donations is small: our estimate suggests that a full roll-out of the gift treatment would lead to a significant increase of average donations by €0.47. This estimate implies that, after taking the cost of the gift of €1.16 per letter into account, gifts are not profitable in the overall sample.

Second, we identify a distinct heterogeneity in the response to the gift treatment with respect to individuals’ baseline willingness to donate. Columns (3) to (6) show that on average, the gift triggers an increase in donations among warm-list (cold-list) individuals of €1.99 (€0.19), or 12.3% (126%) relative to average donations of €16.20 (€0.15) in the control group. Hence, both, cold-list and warm-list individuals significantly respond to the gift treatment. However, in contrast to the findings reported in Landry et al. (2010), the response of cold-list individuals is much smaller than that of warm-list individuals in absolute terms. Regarding the profitability of gifts, we conclude that providing the gift covers the costs only for warm-list individuals, while the gift is clearly not profitable in the cold list. The finding that the effectiveness of upfront gifts in fundraising strongly depends on individuals’ baseline characteristics highlights the importance of appropriate gift targeting. In subsection 3.3, we will further explore the question of optimal targeting of gifts in a more comprehensive way.

The third finding emerging from Table 1 is that combining a gift with recognition for past contributions wipes out the positive effect of providing a gift alone. This finding speaks to an emerging literature on the importance of framing effects in gift exchange (Newman and Shen, 2012; Landry et al., 2012; Chao, 2015). The average effect of the treatment on the amount donated is also insignificant, implying that the gift-plus-recognition treatment is likely not profitable when applied to all warm-list individuals. This finding suggests that positive responses to gifts in fundraising are indeed triggered by sequential reciprocity along the lines of Dufwenberg and Kirchsteiger (2004).

Our sampling frame allows us to also shed light on the persistence of the response to gifts. Table 5 in the appendix depicts the treatment effects of the gift treatment in the year 2014, evaluated in 2015. The table clearly shows that one year after the treatment, there is no difference between treatment and control groups. This means there is no evidence of intertemporal substitution in response to gift exchange. This finding is in line with Falk (2007).

In order to assess the take-up of the treatments, we conducted a small-scale telephone survey \( N = 101 \) among recipients of the solicitation letter shortly after the mail-out in the year 2014. We find that almost 90 percent of those receiving a letter confirmed they had read it. While our analysis focuses throughout on intent-to-treat effects, the high take-up rate among survey participants suggests that the corresponding average treatment effects should only be slightly scaled up. For convenience, we label the effects
discussed throughout the paper as (average) treatment effects.

3.2 Heterogeneity in Responses to Gifts

Our analysis of the heterogeneity in the response to gifts starts with the application of non-parametric methods that avoid distributional assumptions or functional form restrictions. We proceed in two steps. First, we calculate the difference in average donations in bins of potential donors’ baseline willingness to donate between the treatment and the control group. Second, we fit local linear regressions to the differences to illustrate the heterogeneity in the treatment effect over the willingness to donate. Figure 3 illustrates the outcome by depicting differences in average donations between gift and control. The bubbles included in the figure result from averaging over all differences in average donation using bins (in terms of the baseline willingness to donate) of size 0.1. The size of the bubbles reflects the number of observations in each bin. The figure also illustrates the estimated treatment effects resulting from the local linear regressions (red line) and a 95% confidence band based on a bootstrap.

The figure demonstrates that the treatment effect is significantly different from zero up to a baseline willingness to donate equivalent to a score value of 0.78. Starting from a small effect in absolute terms for individuals with a low baseline willingness to donate, the treatment effect increases and peaks for score values around 0.5. The treatment effect then declines, resulting in a hump-shaped form.

If we take the cost per gift (gray dashed line) into account, we find that the treatment effect on donations exceeds the cost of the gift for the middle part of the distribution with respect to the baseline willingness to donate. The implication in terms of an optimal targeting of gifts is straightforward: when rolling out a gift-exchange scheme, fundraisers can increase revenue by targeting gifts to warm-list individuals with a moderate baseline willingness to donate. In contrast, providing gifts to individuals below or above the threshold levels implied by Figure 3 might result in the charity losing money. In our context, we estimate that a mailing with an optimal targeting of the gift towards individuals with score values between 0.31 and 0.64 results in a net revenue gain of 37.3 percent relative to a uniform mailing comprising the gift.

Having shown that the treatment effect is hump-shaped with respect to the baseline willingness to donate, the implications for efficient fundraising and targeting are clear. This finding highlights the importance of targeting for efficient fundraising and extends previous work on gift exchange focusing on average treatment effects (Falk, 2007; Alpizar et al., 2008a,b; Landry et al., 2010).

---

8Note that the score measuring the baseline willingness to donate is distributed across a finite number of values between zero and one.
9The calculation assumes that the charity sends a letter to all individuals anyway. With optimal targeting, individuals with a score below 0.31 or above 0.64 would receive a letter without gift, while individuals with score values between 0.31 and 0.64 would receive the letter comprising the gift.
willingness to donate, we now turn to explaining the observed heterogeneity. This exercise will help us to build a better intuition for the effects at work and to transfer the insights from our experiment to other settings.

3.3 Extensive vs. Intensive Margin Responses

3.3.1 Decomposition

In order to better understand the different channels leading to the hump-shaped overall effect, we decompose the treatment effect into two parts: the extensive margin effect (EME) and the intensive margin effect (IME). For simplicity we omit the time index in this section and note that all variables refer to the respective treatment year \( t + 1 \).

Table 2 illustrates the origins of the IME and the EME by distinguishing between four types of individuals. The table adopts the potential outcomes notation. We denote individual \( i \)'s donation in the control group with \( Y_{0i} \) and her donation in the treatment group with \( Y_{1i} \). This leads to three distinct groups of individuals. The first group comprises the nonparticipants (NP). Nonparticipants do not make a donation irrespective of receiving the treatment. The second group are switchers. We can distinguish between two types of switchers, \( S_1 \) and \( S_2 \). The first type of switcher \( S_1 \) donates only if receiving the treatment. The second type \( S_2 \) donates if she does not receive the treatment, but does not donate if she receives the treatment. The last group comprises participants (P) who make a donation irrespective of receiving the treatment. The IME is defined as the part of the treatment effect which is driven by the response of participants. In contrast, the EME is the net effect of switchers on donations.

One way to estimate the IME is by comparing average donations in treatment and control group conditional on strictly positive donations. However, as Angrist and Pischke (2009) note, the estimated effect is biased if the group of donors in the treatment also comprises switchers (\( S_1 \)) whose average donations differ from participants. We follow Staub (2014) to correct for a possible bias and construct estimates of IME and EME allowing for a causal interpretation. To this end, we assume

\[
Y_{0i} \leq Y_{1i} \quad \forall \ i. \tag{2}
\]

The assumption requires that all individuals either do not respond to the treatment or increase their donation in response to the treatment. Under this assumption we can use a Tobit model to calculate the causal IME and EME.\textsuperscript{10}

Equation (3) shows the decomposition of the treatment effect into IME and EME.\textsuperscript{10}

\textsuperscript{10}Note that our Tobit model also estimates causal effects under the weaker assumption of monotonicity. Under monotonicity, the effect of the treatment can go either way as long as the direction is the same for all individuals. We adopt the stronger assumption (2) only for the sake of exposition.
under assumption (2).

\[ E(Y_{1i}) - E(Y_{0i}) = \]
\[ E(Y_{1i} - Y_{0i} | Y_{0i} > 0, Y_{1i} > 0) Pr(Y_{0i} > 0, Y_{1i} > 0) + E(Y_{1i} | Y_{0i} = 0, Y_{1i} > 0) Pr(Y_{0i} = 0, Y_{1i} > 0) \]  
(3)

The first right-hand side term shows the IME as the expectation of the treatment effect on participants’ donations, weighted by the population fraction of participants. The second right-hand side term shows the EME, i.e. the population fractions of switchers \((S_1)\) weighted by switchers’ expected donation. Because (2) rules out the existence of the second type of switchers \((S_2)\) and because nonparticipants’ expected donation is zero, those groups do not contribute to (3).

Following Staub (2014), we estimate the Tobit model

\[ Y_i = \max[0, \kappa_0 + \kappa_1 G_i + \kappa_2 \hat{R}_i + \kappa_3 G_i \times \hat{R}_i + U_i], \quad U_i \sim N(0, \sigma^2). \]  
(4)

Individual \(i\) chooses either a strictly positive donation or the corner solution of no donation. Her decision depends on a constant term \(\kappa_0\), her willingness to donate \(\hat{R}_i\), the gift dummy \(G_i\), and a random element \(U_i\). The model allows for heterogeneity in the response to the gift by including a linear interaction between \(G_i\) and \(\hat{R}_i\). We further assume that \(U_i\) is normally distributed with mean zero and variance \(\sigma^2\).

We recover the expectations and population fractions at each value of the score \(\hat{R}_i = \hat{r}\) in (3) from the Tobit estimates. Denote the evaluation of the standard normal probability density function (cumulative density function) at \((\kappa_0 + \kappa_2 \hat{r})/\hat{\sigma}\) with \(\phi_0\) (\(\Phi_0\)) and the evaluation at \((\kappa_0 + \kappa_1 + \kappa_2 \hat{r} + \kappa_3 \hat{r})/\hat{\sigma}\) with \(\phi_1\) (\(\Phi_1\)). This yields

\[ Pr(Y_{0i} > 0, Y_{1i} > 0, \hat{R}_i = \hat{r}) = \Phi_0, \]  
(5)

\[ Pr(Y_{0i} = 0, Y_{1i} > 0, \hat{R}_i = \hat{r}) = \Phi_1 - \Phi_0, \]  
(6)

\[ E(Y_{1i} - Y_{0i} | Y_{0i} > 0, Y_{1i} > 0, \hat{R}_i = \hat{r}) = \kappa_1 + \kappa_3 \hat{r}; \]  
(7)

\[ E(Y_{1i} | Y_{0i} = 0, Y_{1i} > 0, \hat{R}_i = \hat{r}) = \kappa_0 + \kappa_1 + \kappa_2 \hat{r} + \kappa_3 \hat{r} + \hat{\sigma} \frac{\phi_1 - \phi_0}{\Phi_1 - \Phi_0}. \]  
(8)

Intuitively, \(\Phi_0\) (\(\Phi_1\)) is the predicted probability that a randomly drawn individual makes a donation in the control (gift) treatment. From those probabilities we recover the population fractions of participants and switchers in (5) and (6). Equations (7) and (8) yield the average treatment effects on participants and on switchers. Under our assumptions, all of these quantities have a causal interpretation and together result in the treatment effect in (3).
3.3.2 Results: Extensive vs Intensive Margin Responses to Gifts

Table 3 shows the estimates of the Tobit model for gift vs control. Column (1) presents unconditional estimates, while column (2) controls for the strata variables, age, and the household type. Because the marginal treatment effects depend on covariates in a non-linear way, we report effects on the latent variable and only interpret the signs and significance of coefficients.\(^\text{11}\) We then graphically evaluate average treatment effects for all values of the baseline willingness to donate.

The Tobit estimates indicate a positive and significant effect of gifts on donations. Moreover, the treatment effect declines, ceteris paribus, with an increasing baseline willingness to donate. A straightforward intuition for the negative interaction can be derived if we interpret the baseline willingness to donate as a proxy for individuals’ intrinsic motivation. For a potential donor with low intrinsic motivation, receiving a gift is a stronger signal of the charities’ kindness towards her than for a highly intrinsically motivated donor. In the Dufwenberg and Kirchsteiger (2004) model, such a difference in signal-strength could be driven by different believes about the maximal and minimal possible kindness of the fundraiser, against which the motive of the gift is evaluated. In the following, we always use the unconditional estimates in column (1). All results are robust to using the conditional estimates in column (2).

Figure 4 shows the decomposition of the treatment effect of gifts (red line) into IME (gray area) and EME (blue area). The figure also includes 95% confidence bands based on a bootstrap. Two key results emerge from Figure 4: first, we note that the treatment effect follows a hump shape with respect to the baseline willingness to donate that is similar to the nonparametric estimates in Figure 3. Second, the decomposition reveals that the treatment effect is mainly driven by an intensive margin response to the treatment. On average, the share of the IME over the domain of the score is 88%.

In the next step, we provide further details on IME and EME that are useful in explaining the origins of the hump-shape of the treatment effect. We build this analysis on a graphical representation of our estimates of population fractions and expectations in equation (3). Figure 5 shows the population fractions (blue dashed lines) and expected changes in donations (red solid lines) for the intensive margin effect (Panel A) and the extensive margin effect (Panel B). For both margins, the effect of the gift on the expected donation decreases with donors’ baseline willingness to donate. This relates to the negative sign of the interaction term in the Tobit model. For the intensive margin, the share of participants increases with the baseline willingness to donate. The increase in the share of participants immediately follows from the fact that we use the baseline probability of contributing to proxy for the willingness to donate. The hump-shape in Figure 4 results from the larger share of participants and the decreasing expected change in donations.

\(^{11}\)With this approach, we follow the suggestions for inference in nonlinear models in Greene (2010).
working in opposite directions.

Turning to the extensive margin, the hump shape in the treatment effect on the population fraction of switchers follows from the shape of the standard normal cdf. Individuals with a baseline willingness to donate of 0.5 (i.e., a baseline probability of 50%) are located in the middle of the cdf. Intuitively, these potential donors are indifferent between giving and not giving. Because they are located at the steepest part of the cdf, treating indifferent individuals is most effective in nudging potential donors into becoming actual donors.

As we have seen from Table 1, the gift-plus-recognition treatment results in insignificant average treatment effects. For completeness, Figure 7 in the appendix reports the non-parametric evidence on the heterogeneity of the treatment effect. We again see a hump-shaped response, but the treatment effect is significantly different from zero only for a small range in the middle of the distribution, and the treatment is not profitable over the full range of the score. Also for completeness, Figure 8 in the appendix provides the non-parametric evidence on the persistence of the gift treatment in 2014, evaluated in the year 2015.

### 3.4 Applying Our Insights to Other Contexts

The back-of-the-envelope calculation based on the findings for the gift treatment suggests that in our context, the optimal targeting of gifts can lead to revenue gains of about 37 percent relative to a scheme with a uniform provision of gifts. In the following, we want to discuss briefly how the methods and results discussed before could be applied to other contexts.

A fundraiser who seeks to optimize the targeting of a specific fundraising vehicle can proceed in two simple steps. First, one needs to trace out the heterogeneous treatment effect of the fundraising vehicle over potential donors’ baseline propensity to donate in a random sample of the population of interest.\(^{12}\) Second, the fundraiser should roll out the campaign by targeting the fundraising vehicle to all individuals in the sub-population known to exhibit positive treatment effects.

Not all charities will be able to conduct analyzes as the one described before in order to target their fundraising activities. This raises the issue of whether the optimal starting point for a fundraising campaign in terms of donors’ baseline willingness to donate can be identified based on simpler techniques. Figure 6 shows how this might work. The figure depicts the baseline propensity to donate at the maximum treatment effect (ver-

\(^{12}\)The derivation of potential donors’ baseline willingness to donate can be simplified along two lines. First, instead of producing point predictions of the baseline propensity to donate on the basis of past behavior, the fundraiser can rank individuals with respect to their expected baseline willingness to donate using the data that is available (e.g., age, gender, etc.). Second, instead of estimating heterogeneous treatment effects, the fundraiser can compare mean donations between treatment and control group in bins of potential donors’ baseline willingness to donate, i.e. in predefined sections of the ranking.
tical axis) for different parameter values (horizontal axis). The setup of the figure thus allows to look for the optimal target individual. To construct the figure, we first vary the treatment effect of gifts on the latent variable and the interaction term in our empirical model (4). For each parameter combination, we locate the baseline propensity to donate at the maximum treatment effect by numerical optimization. To allow for an intuitive illustration, the figure shows the relative treatment effect defined as the effect for an individual with a baseline propensity to donate of zero ($score = 0$) divided by the treatment effect for an individual with a baseline propensity to donate of one ($score = 1$). For illustrative purposes, the gray dashed lines illustrate the relative treatment effect and the maximum treatment effect for our estimates from the gift treatment (compare Figure 4 for the treatment effects at the extreme values of the score and the position of the peak). Importantly, many charities will be able to come up with a reasonable estimate of the relative treatment effect, and will thus be able to identify a range of optimal target individuals (provided that the charity has some idea how to estimate individuals’ baseline probability to donate).

The mechanics of identifying optimal targets are straightforward. Suppose the hump in Figure 4 moves to the right. Then the treatment effect at $score = 0$ declines relative to the treatment effect at $score = 1$. This means that in Figure 6, we move to the left along the horizontal axis, arriving at a maximum treatment effect that is closer to a baseline propensity to donate of one. Consequently, the optimal strategy is, ceteris paribus, to target individuals with a higher baseline propensity to donate. In contrast, if the hump in Figure 4 moves to the left, Figure 6 tells us that the optimal strategy, ceteris paribus, targets individuals with a lower baseline propensity to donate.

4 Conclusion

Charities use a number of costly fundraising tools. Despite the fact that overall fundraising expenditures correspond to a large fraction of private donations, there is little systematic evidence on the optimal targeting of fundraising strategies. We contribute to closing this gap by exploiting data from a fundraiser who sends out annual solicitation letters to a large population of individuals irrespective of individuals past donation behavior. Using rich panel data on individual donations in the baseline, we derive a proxy for individuals’ baseline willingness to donate absent any treatment. Using this proxy, we study the heterogeneity in treatment responses in a large-scale randomized field experiment on gift exchange.

Our key finding is a distinct heterogeneity in responses to gifts with respect to donors’ baseline willingness to donate. We decompose the average treatment effect into an extensive margin and an intensive margin effect and find both effects to be hump-shaped with respect to individuals’ baseline probability to give. This translates into a hump-
shaped average treatment effect. The strong heterogeneity in the treatment response implies that up-front gifts are cost effective only for individuals with a moderate baseline probability to donate. In contrast, it is not profitable for the charity to provide the gift to individuals with a baseline probability to donate below or above certain threshold levels. The possible revenue gains associated with switching from a uniform provision of gifts to a scheme with optimal targeting are substantial: for our context, we estimate a net revenue gain from optimal targeting of 37.3 percent relative to a uniform scheme. This finding highlights the importance of targeting for efficient fundraising. We also show that combining a gift with recognition for past contributions is counterproductive. Our data suggest that the positive effect of providing a gift alone is wiped out if combined with recognition. This suggests that positive responses by donors to gifts received from fundraisers are driven by reciprocity concerns Dufwenberg and Kirchsteiger (2004).

Understanding heterogenous responses of donors to different fundraising vehicles can make charities more effective. The methods discussed in this paper can be applied to a wide range of fundraising strategies and should thus be fruitful in a variety of contexts. We hope that the ideas put forward in this paper can serve as a first step towards a more comprehensive analysis of optimal fundraising techniques.

References


Table 1: Summary of Outcomes

<table>
<thead>
<tr>
<th></th>
<th>Overall Sample</th>
<th>Cold List</th>
<th>Warm List</th>
<th>Warm List</th>
<th>Gift + Recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Years 2014 and 2015 Pooled</td>
<td>Years 2014 and 2015 Pooled</td>
<td>Years 2014 and 2015 Pooled</td>
<td>Year 2015</td>
<td>Year 2015</td>
</tr>
<tr>
<td>Relative Frequency of Donations</td>
<td>(1) 0.102 (0.003)</td>
<td>(2) 0.087 (0.003)</td>
<td>(3) 0.016 (0.002)</td>
<td>(4) 0.008 (0.002)</td>
<td>(5) 0.574 (0.009)</td>
</tr>
<tr>
<td>Treatment Effect on Frequency</td>
<td>(0.004) 0.014***</td>
<td>(0.003) 0.008***</td>
<td>(0.013) 0.049***</td>
<td>(0.016) 0.069***</td>
<td>(0.017) 0.020</td>
</tr>
<tr>
<td>Mean Donation</td>
<td>(1) 3.07€ (0.138)</td>
<td>(2) 2.60€ (0.108)</td>
<td>(3) 0.344€ (0.099)</td>
<td>(4) 0.152€ (0.044)</td>
<td>(5) 18.19€ (0.587)</td>
</tr>
<tr>
<td>Treatment Effect on Donation</td>
<td>(0.176) 0.467***</td>
<td>(0.108) 0.191*</td>
<td>(0.804) 1.99**</td>
<td>(1.05) 2.49**</td>
<td>(1.03) 0.89</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>5266</td>
<td>5269</td>
<td>2283</td>
<td>2286</td>
<td>2986</td>
</tr>
</tbody>
</table>

Notes: This table reports means for the relative frequency to donate and for donations. The table also shows unconditional OLS estimations of treatment effects. Standard errors are in parentheses. All effects are robust against controlling for observable strata and other confounders. Specifically, we control for household type, age, a dummy for the year 2014 (if estimation uses data from 2014 and 2015), and a full set of parish dummies. Treatment effects on the frequency of donations are also robust to Probit estimation. *** denotes significance at the 1%, ** at the 5%, and * at the 10% level.
<table>
<thead>
<tr>
<th>Group</th>
<th>Name</th>
<th>Potential Outcomes</th>
<th>Margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP</td>
<td>Nonparticipants</td>
<td>$Y_{0f} = 0, Y_{1f} = 0$</td>
<td></td>
</tr>
<tr>
<td>$S_1$</td>
<td>Switchers</td>
<td>$Y_{0f} = 0, Y_{1f} &gt; 0$</td>
<td>extensive</td>
</tr>
<tr>
<td>$S_2$</td>
<td>Switchers</td>
<td>$Y_{0f} &gt; 0, Y_{1f} = 0$</td>
<td></td>
</tr>
<tr>
<td>$P$</td>
<td>Participants</td>
<td>$Y_{0f} &gt; 0, Y_{1f} &gt; 0$</td>
<td>intensive</td>
</tr>
</tbody>
</table>

Notes: This table provides the definitions of the four types of individuals underlying the decomposition exercise discussed in Section 3.3.1.
### Table 3: Gift vs Control: Tobit Estimation

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gift</td>
<td>13.4***</td>
<td>13.4***</td>
</tr>
<tr>
<td></td>
<td>(2.89)</td>
<td>(2.89)</td>
</tr>
<tr>
<td>Gift x Baseline Probability of Donation</td>
<td>-12.1***</td>
<td>-12.0***</td>
</tr>
<tr>
<td></td>
<td>(4.09)</td>
<td>(4.08)</td>
</tr>
<tr>
<td>Baseline Probability of Donation</td>
<td>137.4***</td>
<td>136.7***</td>
</tr>
<tr>
<td></td>
<td>(3.37)</td>
<td>(3.44)</td>
</tr>
<tr>
<td>Constant</td>
<td>-95.4***</td>
<td>-97.1***</td>
</tr>
<tr>
<td></td>
<td>(2.54)</td>
<td>(3.44)</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>10535</td>
<td>10535</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-18559.3</td>
<td>-18549.5</td>
</tr>
</tbody>
</table>

Notes: This table reports the coefficients from tobit estimations of the effect of the gift treatment on donations. Column (1) shows the results of the estimation without any additional controls. Column (2) displays the corresponding results for an estimation where we additionally control for the strata variables age and household type.
Figure 1: The Gift: Folded Cards Plus Envelopes

Notes: The gift consists of three different folded cards showing flower motives from paintings of Albrecht Dürer plus three envelopes.
Figure 2: Sampling Frame

Notes: The figure shows the sampling scheme that covered two consecutive years. In 2014, the letter with the gift treatment was sent to a sample comprising cold-list and warm-list individuals. The control group therefore comprises cold-list and warm-list individuals as well. In 2015, only warm-list individuals were treated. Individuals who were treated in 2014 were not treated in 2015. So all individuals who found themselves in either the gift or the gift-plus-recognition treatment groups in either year were treated for the first time. This allows us to pool observations from 2014 and 2015 when studying the effects of the gift treatment.
Notes: The figure shows differences in average donations (bubbles) and the estimated treatment effect resulting from local linear regressions (red line) with a 95% non-parametric confidence band. We fit the local linear regressions (bandwidth 0.2; bandwidth choice based on ROT) to differences in average donations between treatment and control group, calculating average donations in bins (in terms of the baseline willingness to donate) of size 0.05. The bubbles included in the figure are for differences in average donations using bins of size 0.1. The size of the bubbles reflects the number of observations (gray numbers) in each bin. To construct the nonparametric 95% confidence band, we run a bootstrap with 1000 repetitions. The bootstrap draws individuals with replacement from our sample taking the stratified randomization into account.
Notes: The figure shows the average treatment effect (red line), intensive margin effect (gray area), and extensive margin effect (blue area). To construct the nonparametric 95% confidence bands (dashed lines), we run a bootstrap with 1000 repetitions. The bootstrap draws individuals with replacement from our sample taking the stratified randomization into account and reports confidence bands based on the normal distribution.
Figure 5: Elements of Decomposition

Panel A: Intensive Margin Effect

Panel B: Extensive Margin Effect

Notes: The figure shows estimates of the population fractions (blue dashed lines) and expectations (red solid lines) in equation (3). Panel A illustrates results for the intensive margin and Panel B shows results for the extensive margin.
Notes: Note: The figure shows the baseline propensity to donate at the maximum treatment effect over different parameter combinations. We implement all combinations of parameter values of $\kappa_1$ and $\kappa_3$ in the Tobit model (4) that lie in symmetric bands around the parameter estimates within $\hat{\kappa}_j \pm \hat{\kappa}_j^2$ for $j \in 1, 3$. For each parameter combination, we locate the baseline propensity to donate at the maximum treatment effect by numerical optimization. Instead of depicting the parameter values directly, the horizontal axis shows the relative treatment effect defined as the effect for an individual with a baseline propensity to donate of zero ($\text{score} = 0$) divided by the treatment effect for an individual with a baseline propensity to donate of one ($\text{score} = 1$). We checked if changes in the constant term ($\kappa_0$) and corresponding changes in the population fraction of participants alters our results. Increasing the fraction of participants shifts the scatter plot upwards and a lower fraction shifts the plot downwards but does not change the relationship between the relative treatment effect and the baseline propensity to donate at the maximum treatment effect.
### Table A1: Individual Characteristics by Treatment Assignment

<table>
<thead>
<tr>
<th></th>
<th>Number of Observations</th>
<th>Age</th>
<th>Female [yes=1]</th>
<th>Married [yes=1]</th>
<th>Predicted Baseline Probability to Contribute</th>
<th>Dummy Year 2014 [yes=1]</th>
<th>$F$-test on Joint Significance Relative to Control, p-value, Cold List Only</th>
<th>$F$-test on Joint Significance Relative to Control, p-value, Warm List Only</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Pooled Sample (Year 2014 and Year 2015)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control Group</td>
<td>5266</td>
<td>59.0</td>
<td>0.517</td>
<td>0.408</td>
<td>0.353</td>
<td>0.657</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[58.4;59.5]</td>
<td>[0.503;0.530]</td>
<td>[0.394;0.421]</td>
<td>[0.343;0.363] &amp; [0.644;0.670]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gift Treatment</td>
<td>5269</td>
<td>59.0</td>
<td>0.515</td>
<td>0.410</td>
<td>0.354</td>
<td>0.657</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[58.4;59.5]</td>
<td>[0.502;0.529]</td>
<td>[0.396;0.423]</td>
<td>[0.344;0.364] &amp; [0.644;0.670]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel B: Sample Year 2015</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control Group</td>
<td>1806</td>
<td>68.9</td>
<td>0.524</td>
<td>0.464</td>
<td>0.654</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[68.1;69.6]</td>
<td>[0.501;0.547]</td>
<td>[0.441;0.487]</td>
<td>[0.643;0.666]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gift Treatment</td>
<td>1806</td>
<td>68.8</td>
<td>0.524</td>
<td>0.472</td>
<td>0.656</td>
<td>-</td>
<td>0.990</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[68.0;69.6]</td>
<td>[0.501;0.547]</td>
<td>[0.449;0.495]</td>
<td>[0.644;0.667]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gift-plus-Recognition Treatment</td>
<td>1805</td>
<td>69.0</td>
<td>0.526</td>
<td>0.468</td>
<td>0.655</td>
<td>-</td>
<td>0.992</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[68.3;69.8]</td>
<td>[0.503;0.549]</td>
<td>[0.445;0.491]</td>
<td>[0.644;0.667]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** This table presents randomization checks. Panel A demonstrates that the pooled sample (year 2014 and year 2015) is balanced with respect to the variables used to define strata (age, gender, marital status, and predicted baseline probability to contribute) and with respect to the share of observations from the year 2014. Columns (2) to (6) present the baseline averages for the different observable characteristics and 95% confidence intervals in squared brackets. Column (7) shows the $p$-value of an $F$-Test, testing whether the observable characteristics are jointly significant in predicting assignment to the gift treatment relative to the control group. Column (8), Panel A reports the corresponding $p$-value for the subsample of cold-list individuals only, while Column (9) shows the $p$-value for the warm list. Panel B reports the same checks for the sample of observations from the year 2015, where only warm-list individuals were included in the experiment. Panel B includes the Gift-plus-Recognition Treatment, which was only implemented in 2015.
### Table A2: Gift Treatment, Persistency of Treatment Effects

<table>
<thead>
<tr>
<th></th>
<th>Gift Treatment (2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall Sample</td>
</tr>
<tr>
<td></td>
<td>Treatment</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Relative Frequency of Donations (2015)</td>
<td>0.080</td>
</tr>
<tr>
<td></td>
<td>0.076</td>
</tr>
<tr>
<td>Treatment Effect on Frequency (2015)</td>
<td>0.004</td>
</tr>
<tr>
<td>Mean Donation (2015)</td>
<td>2.63€</td>
</tr>
<tr>
<td></td>
<td>2.34€</td>
</tr>
<tr>
<td>Treatment Effect on Donation (2015)</td>
<td>0.295</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>3107</td>
</tr>
</tbody>
</table>

Notes: This table shows the persistency of the gift treatment effect. It reports means for the relative frequency to donate and for donations in the year 2015 for the gift treatment and the control group in 2014. Both groups received the control group letter in 2015. The table also shows unconditional OLS estimations of treatment effects. Standard errors are in parentheses. All effects are robust against controlling for observable strata and other confounders. Specifically, we control for household type, age, and a full set of parish dummies. Treatment effects on the frequency of donation are also robust to Probit estimation.
Notes: The figure shows differences in average donations (bubbles) and the estimated treatment effect resulting from local linear regressions (red line) with a 95% non-parametric confidence band. We fit the local linear regressions (bandwidth 0.2; bandwidth choice based on ROT) to differences in average donations between treatment and control group, calculating average donations in bins (in terms of the baseline willingness to donate) of size 0.05. The bubbles included in the figure are for differences in average donations using bins of size 0.1. The size of the bubbles reflects the number of observations (gray numbers) in each bin. To construct the nonparametric 95% confidence band, we run a bootstrap with 1000 repetitions. The bootstrap draws individuals with replacement from our sample taking the stratified randomization into account.
Figure 8: Heterogeneity: Persistence of Treatment Effect of Gifts on Donations

Notes: The figure shows differences in average donations (bubbles) and the estimated treatment effect resulting from local linear regressions (red line) with a 95% non-parametric confidence band. We fit the local linear regressions (bandwidth 0.2; bandwidth choice based on ROT) to differences in average donations between treatment and control group, calculating average donations in bins (in terms of the baseline willingness to donate) of size 0.05. The bubbles included in the figure are for differences in average donations using bins of size 0.1. The size of the bubbles reflects the number of observations (gray numbers) in each bin. To construct the nonparametric 95% confidence band, we run a bootstrap with 1000 repetitions. The bootstrap draws individuals with replacement from our sample taking the stratified randomization into account.