THE EFFECT OF EFFECTIVENESS:
DONOR RESPONSE TO AID EFFECTIVENESS IN A DIRECT MAIL FUNDRAISING EXPERIMENT

Dean Karlan
Daniel H. Wood

The Effect of Effectiveness: Donor Response to Aid Effectiveness in a Direct Mail Fundraising
Experiment
Dean Karlan and Daniel H. Wood
NBER Working Paper No. 20047
April 2014
JEL No. D64,H41,L31,O12

ABSTRACT

We test how donors respond to new information about a charity's effectiveness. Freedom from Hunger
implemented a test of its direct marketing solicitations, varying letters by whether they include a discussion
of their program's impact as measured by scientific research. The base script, used for both treatment
and control, included a standard qualitative story about an individual beneficiary. Adding scientific
impact information has no effect on whether someone donates, or how much, in the full sample. However,
we find that amongst recent prior donors (those we posit more likely to open the mail and thus notice
the treatment), large prior donors increase the likelihood of giving in response to information on aid
effectiveness, whereas small prior donors decrease their giving. We motivate the analysis and experiment
with a theoretical model that highlights two predictions. First, larger gift amounts, holding education
and income constant, is a proxy for altruism giving (as it is associated with giving more to fewer charities)
versus warm glow giving (giving less to more charities). Second, those motivated by altruism will
respond positively to appeals based on evidence, whereas those motivated by warm glow may respond
negatively to appeals based on evidence as it turns off the emotional trigger for giving, or highlights
uncertainty in aid effectiveness.

Dean Karlan
Department of Economics
Yale University
P.O. Box 208269
New Haven, CT 06520-8629
and NBER
dean.karlan@yale.edu

Daniel H. Wood
222 Sirrine Hall
John E. Walker Department of Economics
Clemson University
Clemson, SC 29634
dhwood@gmail.com

1. Introduction

Understanding why people choose to donate to charity is difficult; people give for different and multiple reasons. However, donors are typically consistent on one claim when prompted: they like the idea of giving to charities that are effective. Are such statements cheap talk, or are donors' behaviors consistent with this stated preference? If donors do not respond, or respond negatively, this would predict an underinvestment from the public sector in impact analysis (Pritchett 2002). Pritchett (2014) also poignantly argues that the randomized trial movement within development, ironically, is faith-based in the sense that the advocates do not have rigorous evidence of their impact on the development process. Here we put forward evidence to the contrary, albeit on a small scale: do donors respond favorably to evidence from randomized trials? The short answer: some do, some do not.

Following a paradigm put forward by Kahneman (2003) we explore a model of giving that incorporates two motivations for giving, altruism (akin to Kahneman's System II decisions, which are deliberate, effortful, reasoned and focused on impact) and warm glow (akin to Kahneman's System I decisions, which are intuitive, effortless and reactive). The model makes an important prediction: that individuals driven by altruism, holding all else equal (such as wealth and education), will donate larger amounts to fewer charities, and individuals driven by warm glow will do the reverse, i.e. donate smaller amounts to more charities. Likewise, individuals driven by altruism will respond favorably to information about the effectiveness of a particular charity, whereas those driven by warm glow, i.e. more emotionally driven triggers, may actually reduce giving.

It has long been recognized that altruism cannot be the entire explanation for charitable giving as it would lead to complete crowd-out of donations in response to other funding sources, which is not borne out by most estimates of crowd-out (Andreoni 2006). Recent experiments provide more direct evidence on the warm-glow motives that are part of our model. Null (2011) looks at how members of service clubs divided $100 among three charities. Most participants in Null (2011) revealed warm glow motives by giving to multiple charities, which is incompatible with risk-neutral altruism, (risk neutral altruists would give the entire $100 to the charity with the highest expected impact). In another direct test of warm-glow preferences, Crumpler and Grossman observed that most

subjects gave to a charity even though their donation crowded-out one-for-one a donation by the experimenters (Crumpler and Grossman 2008). Thus people seem to get utility from the act of giving in addition to utility from the effects of their giving.

We present the results of an experiment that provided donors with information about a charity's effectiveness. Our research both tackles a practical problem (how to communicate aid effectiveness to donors) and provides insight into identifying motivations for giving. Our work builds on prior work such as Yoruk (2013), which studies, with a regression discontinuity approach, how donors respond to Charity Navigator's star rating system. Charity Navigator uses mostly financial and governance data (i.e., not data of impact of the work of the charity) to inform the general public, using a 5 star categorical system. Yoruk finds that for otherwise similar charities, a one-star rating increase leads to 19.5 percent higher donations if the charities are relatively unknown, but that the rating increases have no effect on better-known charities.[2]

We collaborated with Freedom from Hunger to conduct two rounds of direct-mail marketing to its prior donors. Freedom from Hunger ("FFH") is a US-based nonprofit organization that provides technical advisory services to microfinance institutions (MFIs) in developing countries. In the first experimental wave, the control group received an emotional appeal focused on a specific beneficiary, along with a narrative explaining how FFH ultimately helped the individual. The treatment group received a similar emotional appeal (trimmed by one paragraph), with an added paragraph about scientific research on FFH's impact. The second wave was identical in design, except that the treatment group narrative included more specifics on the research, and briefly discussed randomized trials and their value as impact assessment tools.

We find that *average* donation behavior does not change when previous donors are presented with evidence of the charity's effectiveness in achieving its goals. However, we find that the aggregate effect masks different responses by small and large prior donors: large prior donors (those who had given $100 or more previously to FFH) donate more and small prior donors donate less in response to being told about the

---

[2] Many leaders in the philanthropic space (including the leaders of Charity Navigator, see http://www.overheadmyth.com) have criticized the use of overhead and management ratios, but little is available to donors beyond such data on a comprehensive level (e.g., Givewell.org, an alternative charity evaluator, focuses strictly on evidence of impact per dollar donated and room for growth, but in each year has named typically between 3 and 10 charities, in a limited number of causes).

quantitative effectiveness of the charity. Both the positive impact on large donors and negative effect on small donors are more pronounced among recent prior donors, who we interpret to be those most likely to actually open the envelopes. Although we argue that the size of the gift is a proxy for altruism versus warm glow giving, we discuss at the end alternative explanations, explanations which we are not able to dismiss with our data.

2. Motivation and Model

Gift size, holding all else constant, may tell us something about the underlying motivation for giving. Small gifts[3] may be more likely given casually, simply to participate or to appease social pressure (DellaVigna, List, and Malmendier 2012), without much concern for effectiveness. On the other hand, large donations may be a by-product of a more thoughtful and analytically-driven evaluation, driven by an altruistic desire to maximize impact of their donation. Our data are limited in their ability to discriminate perfectly the underlying mechanism. However, we use a simple model to make the case that donation size is a proxy (no doubt imperfect, empirically) for motivation, and that providing information on aid effectiveness has heterogeneous impact on giving, depending on motivation. We will discuss alternative interpretations in the discussion and conclusion.

Base Model

We present a model of multiple donations inspired by research that deliberation can interfere with emotional impulses for giving (Small, Loewenstein, and Slovic 2007). The model predicts that the dominant motive for small donations is different than for large donations, and explains why small and large donors respond differently to our treatment. Donors get utility from donations in one or both of two ways: (1) altruism, in which the donation affects utility through the increase in social welfare it gives rise to, and (2) warm-glow, in which the act of donating increases utility directly. The important distinction between altruism and warm-glow, as Null (2011) points out, is that altruists view charities as perfect substitutes and so respond to differences in charity efficiency,

---

[3] Note that "larger" in our context is $100 or more, but still strictly within the domain of FFH's direct-mail marketing, i.e., not so large as to be the by-product of major gifts personalized outreach and meetings.

while warm-glow donors only value donating and are not responsive to efficiency. We are thus shutting down other motivations for giving, for example social status (although to the extent that social status is binary or with diminishing returns to scale, and not linear or with increasing returns to scale, it would yield similar predictions as warm-glow).

Model: The Donation Decision

Our subjects divide their income *y* into a donation to a charity, *g*, and consumption net of giving, *c*. Utility from charitable giving stems from two sources, altruism and warm-glow. Utility is quasi-linear with constant marginal utility from consumption:

$$U_i = c + \beta_i * A(\gamma g) + (1 - \beta_i) * e_i * W(g).$$

β parameterizes the importance of altruism relative to warm-glow for a given subject *i*, γ parameterizes the charity's effectiveness (or more precisely, the perception of the charity's effectiveness, which we assume to be the same for all *i*) [4], and e_i parameterizes how emotionally connected the subject is to the charity. Altruism and warm-glow are linearly independent for convenience; our results would not change if they were complements as long as the complementarity was not too strong.

The important characteristics of warm-glow utility are that marginal utility diminishes quickly and that it is independent of the effectiveness of the charity. For example, giving beyond a certain threshold might provide little utility benefit. This threshold could be the minimum amount requested by the charity or $1 or some other threshold internal to the subject. We assume that warm-glow utility is a step function W(g) taking the value 1 when $g \geq \underline{g}$ and the value 0 otherwise, weighted by emotional attachment, *e*.

The altruism component of utility *A* is a strictly increasing, strictly concave function, whose argument is the increase in social welfare produced by the subject's donation. For small gifts, the marginal social benefit of each unit donated will be approximately constant and γ measures that slope. A more effective charity has a higher

---

[4] A more realistic assumption about γ would be to model individuals as believing that γ comes from a distribution. All of our results would remain in such a framework.

γ. Note that by making γ fixed and known exactly per charity, rather than a source of uncertainty, we are shutting down risk aversion. If γ were stochastic, then altruistic donors may give to multiple charities due to risk aversion.

Let γ_0 and e_0 be the representative individual's beliefs about the effectiveness of FFH and emotional connection to FFH before the experiment and in the control condition. We make some further assumptions about the characteristics of the utility function.

Assumption 1: $\gamma_0 A'(\gamma_0 \underline{g}) > 1$.

Assumption 2: $-x A''(x)/A'(x) < 1$ *for all x.*

Assumption 3: $e_0 > \underline{g}$.

Assumption 1 implies that the optimal donation of purely altruistic donors is higher than the optimal donation of purely warm-glow donors. Assumption 2 states that the curvature of $A(x)$ is not "too high", which implies that an increase in γ for purely altruistic donors leads them to increase their donation. Altruistic donors will always respond to a rise in γ with a donation that increases welfare more than their donation for lower γ, but depending on $A(x)$, this might involve a smaller monetary donation since each dollar has become more effective. Assumption 3 implies that perfectly warm-glow ($\beta = 0$) donors make a positive donation.

We first establish our claim that small and large donations tend to be driven by different motives. Under Assumption 1, more altruistic donors make larger donations, holding everything else equal. Given this, altruism will be more important for large donors, i.e., they will tend to have higher β than smaller donors.

Proposition 1: The optimal donation g^* is weakly increasing in the importance of altruism β and for high enough importance of altruism β is strictly increasing in β.

Proof: The optimal donation g^* is a critical point of $U(g)$, so $g^* \epsilon \{0, \underline{g}, g_{FOC}\}$, where g_{FOC} is the solution to $\gamma A'(\gamma g) = 1$.

First note that g^* is never at $g=0$. From Assumption 3, perfectly warm-glow donors with $\beta = 0$ prefer $g = \underline{g}$ to $g = 0$. From Assumption 1, perfect altruistic donors with $\beta = 1$ maximize utility at $g^* \geq \underline{g}$, and since utility for these donors is strictly increasing for $g^* \epsilon [0, g^*]$, $U(\underline{g})\big|_{\beta=0} > U(0)$. For interior β, $U(\underline{g})$ is a convex combination of $U(\underline{g})\big|_{\beta=0}$ and $U(\underline{g})\big|_{\beta=1}$, all donors prefer $g = \underline{g}$ to $g = 0$.

Because A is strictly concave, if $\lim_{g\to\underline{g}} U^{'}(g) > 0$, then $U(g_{FOC}) > U\left(\underline{g}\right)$. $U^{'}(g) = \beta\gamma A^{'}(\gamma g)$ so Assumption 1 implies that above some minimum β, $g^* = g_{FOC}$. For β past this threshold, from the Implicit Function theorem, at $g \neq \underline{g}$,

$$\frac{\partial g^*}{\partial \beta} = -\frac{\gamma A^{'}(\gamma g^*)}{\beta\gamma^2 A^{''}(\gamma g^*)}.$$

The denominator is negative and the numerator is positive *so g^* is strictly increasing in β.* **QED.**

We now explore what conditions are necessary for variation in β to drive our heterogonous treatment response. The experimental treatment's additional information about research is likely to increase how effective the subject perceives each unit donated to be, to $\gamma_1 > \gamma_0$, but also may reduce the emotional connection the subject feels to the charity, to $e_{i1} < e_{i0}$. The latter change means that low β subjects for whom emotional connection is especially important will possibly decrease their donations, and definitely will not increase their donations. However, the response of high β subjects is more ambiguous. Intuitively, an increase in γ is effectively a decrease in the price of social welfare, and it will have both income and substitution effects. If the substitution effect dominates, high β subjects will increase their donations. Which effect dominates depends on the curvature of A.

Proposition 2. For high enough importance of altruism β, donation amount g^* is strictly increasing in evidence of effectiveness γ.

Proof: There is some β cutoff such that for $\beta \geq \underline{\beta}$, $g = g_{FOC}$. Then, from the implicit function theorem,

$$\frac{\partial g^*}{\partial \gamma}\bigg|_{\beta \geq \underline{\beta}} = -\frac{A'(\gamma g^*) + g^*\gamma A''(\gamma g^*)}{\gamma^2 A''(\gamma g^*)}.$$

The numerator is strictly positive given Assumption 2 and the comparative static is continuous in β, so for high β, $\partial g^* / \partial \gamma$ is still positive. **QED.**

Proposition 3. A treatment that provides better evidence about effectiveness ($\gamma_1 > \gamma_0$) but reduces emotional connection ($e_{i1} < e_{i0}$) causes sufficiently altruistic individuals to strictly increase their donation and other individuals to weakly reduce their donation.

Proof: At $\beta = 1$, the overall effect is an increase in g^*, so for β above some $\underline{\beta} < 1$, the effect of the γ increase will dominate that of the e decrease. At $\beta = 0$, optimal giving g^* is $\underline{g}$ if $e > \underline{g}$, so if $e1 < \underline{g} < e0$, perfectly warm-glow motivated subjects completely reduce their giving, and if $\underline{g} < e_1$, they maintain the same level of giving. **QED**.

Given the quasi-linear utility form we assume, the only case in which the experimental treatment would induce higher giving is for individuals with high β and with low curvature of A. Assumptions 1 and 2 are the only assumptions consistent with observing large prior donors increasing their donations, because we know some people have β and A combinations that cause $\partial g^*/\partial \beta$ to be positive. This suggests different motives for small and large donors cause them to respond different to evidence of effectiveness.[5]

Multiple Charities

---

[5] Alternatively differences in the curvature of altruism functions alone might be responsible for the heterogeneous response and not variation in β. If for some individuals, A is more curved, that could produce both smaller donations and decreased donations in response to positive evidence about the charity's effectiveness. This could be thought of as an alternative way of modeling warm-glow giving, but since prior economics and social psychology research both support the view of warm-glow as a separate motive from altruism, we prefer our model to this alternative model.

Our base model only considers trading off private consumption with giving to a particular charity. In fact for altruistic subjects this tradeoff misses a dimension that may help to interpret subjects' responses: which charities to support. We extend the model to show how this second dimension could reinforce our hypothesis that variation in β drives our results. In this extension, warm-glow motives lead subjects to spread their donations across multiple charities, in contrast to altruistic motives, which lead to consolidating on the "best" charity.

Now assume that subjects divide their income y into giving to two charities -- g_1 and g_2 respectively -- and consumption net of giving, c. Utility remains quasi-linear:

$$U = c + \beta * A(\gamma_1 \, g_1 + \gamma_2 \, g_2) + (1 - \beta) * [e_1 W(g_1) + e_2 W(g_2)]$$

If $\gamma_1 > \gamma_2$, then charity 1 is "more effective" than charity 2. The subject gets a separate warm glow from each charity he or she supports, but the source of the net increase in social welfare does not matter.

The subject's donations will never exceed the warm-glow threshold $\underline{g}$ for more than one charity, because the most efficient way to maximize the altruistic component of utility is by donating more to the highest γ charity. Then $g_1^*(\gamma_{1,\gamma_2})$ is discontinuous at $\gamma_1 = \gamma_2$ and at that point an increase to $\gamma_1 > \gamma_2$, which makes charity 1 the most efficient charity, causes a large shift in donation towards that charity. This large increase in giving to the more effective charity would be somewhat attenuated due to risk aversion if beliefs about the charities' effectivenesses were not point estimates but instead were distributions (see Null 2011).

3. Experimental Design

The experiment was conducted with FFH, a non-profit focused on providing technical advisory services to microfinance institutions (MFIs) in developing countries. FFH's work was previously evaluated using a randomized controlled trial in Peru (Karlan and Valdivia 2011) and determined to be effective at improving business practices and smoothing microenterprise revenue. This evaluation provided the source of the information on charitable effectiveness presented in the direct mail campaign.

The experiment was conducted in two rounds, in June 2007 and October 2008. A middle round was attempted in March 2008 but an error in the randomization led us to drop these results from the analysis.[6] All rounds were conducted as part of regularly scheduled direct-mail fundraising campaigns, and all subjects were recent donors, defined as those who had given at least once to FFH in either the year of the experiment or the previous three calendar years. The designs of the mailers sent in each round were similar but not identical. In accordance with FFH's policy, letters were mailed first class in closed envelopes for donors that had donated $100+ or mailed non-profit in window envelopes for donors that had previously donated $1-$99.99.

In the first wave, June 2007, mailers were sent to 16,889 individuals who donated to FFH at least once between 2004 and 2007. Subjects were randomly assigned to receive one of three different types of solicitation, with the randomization stratified based on most recent donation year and group of previous donation amount ($1-$24.99, $25-$49.99, $50-$99.99, $100-$249.99, $250-$499, and $500+). All individuals were mailed a renewal letter requesting donations and providing an update on one of FFH's regional programs. Of the 16,889 in the sample, 5,628 individuals received only the renewal letter, with no added insert. An additional 5,630 were assigned to also receive a one-page insert with an emotional appeal and a personal story of one of the program's beneficiaries along with a final paragraph suggesting that FFH had helped him or her. The last 5,631 received an identical insert with the exception of the final paragraph, which instead mentioned studies that used "rigorous scientific methodologies" (the exact script is in the appendix) demonstrating the positive impact of the particular FFH program.

In the second wave, October 2008, mailers were sent to 17,784 individuals who donated to FFH at least once between 2005 and 2008. Randomization was stratified on size of most recent donation (above/below $100), experimental status in the June 2007 round, and whether or not the donor had donated in 2008. Again, all individuals were mailed a renewal letter requesting donations and providing an update on one of FFH's regional programs. Of the sample, 5,960 were in the control group and received a renewal

---

[6] We discovered upon receipt of the data that the March 2008 round was not randomized, but rather the timing of prior giving determined which letter was received. We considered using a regression discontinuity approach, but were deterred by imprecision of the discontinuity that we were not able to unravel, as well as the irony of using non-experimental analysis (for something that could be easily randomized) to identify the impact of using experimental impact evaluations to guide donor decisions.

letter with an emotional appeal to an identifiable victim. A further 5,903 received an identical renewal letter, except that there were additional paragraphs detailing the effectiveness of FFH programs in helping people like the previously introduced victim. The final 5,921 received the same letter as the other 5,903 in the treatment group, except that their letters explicitly cited Yale-affiliated researchers as the source of the statistics on the program's effectiveness. The assignment to treatment was conditional on whether they were a large prior donor or not, based on results from wave one, as per the procedures detailed in Hahn, Hirano and Karlan (2011).

The pooled sample consists of 34,673 requests for a charitable donation. These requests were sent to a total of 21,643 donors of which 13,030 were included in both June 2007 and October 2008, 4,754 were only in the October 2008 round, and 3,859 were only in the June 2007 round. There is not sufficient power to examine interaction effects across the two waves. For those in both rounds, given the time gap in between the mailers, each individual-round is treated as one observation.

4. Experimental Results

Pooled Data

Table 1 provides an overview of OLS regressions with the pooled results for two outcomes: making any donation (within about five months of the mailer), and donation amount (including non-response as zero, and again limiting to the five months following the mailer). All regressions controlled for the wave of data, and whether the individual was a large prior donor (i.e., the stratification variable).

Column 3 shows that large prior donors respond to evidence of recipient effectiveness by being more likely to donate (1 percentage point, se=1.2pp) while column 4 shows them giving larger amounts ($4.45 more on average, se=$7.31). Conversely, column 3 shows that small donors respond negatively to the treatment, becoming less likely to donate (-0.6 percentage points, se=0.4pp) and giving only slight more ($1.57, se=$2.73). However, none of the coefficients are statistically significant at the 10% level.

Importantly, we want to examine the heterogeneous treatment effect after including controls for heterogeneous treatment effects with respect to income and education. Although we do not have individual-level data on income and education, we

do have census data matched on zip-code. Table 1 Columns 5 and 6 include controls for treatment interacted with income and education, to examine whether the heterogeneous treatment effect with respect to being a large prior donor remains important. However, the standard errors increase, the individual point estimates are still not significant, and the test for heterogeneity, specifically the test that the treatment effect for large donors is the same as the treatment effect for small donors, is not rejected (p-value = 0.178 for donation and 0.694 for donation amount).

In Table 2, we present a similar analysis of the pooled results, but break down the sample based on donors' prior donation frequency. This, we conjecture, is essentially reducing noise, as it removes people who are less likely to open the letter at all.

Columns 3, 4, 9, and 10 show a more pronounced effect of the treatment among individuals who frequently donate to FFH, while the pattern is muted among infrequent donors. Frequent donors of large gifts had donation rates which were 2.2 percentage points (se=1.9pp) higher than the control group and donated $12.98 (se=$6.06) more on average. Frequent donors of small gifts were less likely (1.4 percentage points, se=0.8pp) to donate and donated less on average ($0.81, se=$2.53) than the control group. Infrequent donors across the board, however, did not see any significant changes in giving rates or donation amounts. These differences perhaps reflect that frequent donors were more likely to read the solicitation than infrequent donors. The differential result between large and small donors, for the recent donors, is statistically significant (p-value of 0.079 for comparison of the treatment effects on likelihood of giving, and p-value of 0.036 for comparison of the treatment effects on amount given).

Columns 5, 6 and 11, 12 then conduct the similar analysis on heterogeneity with respect to large prior donors, except now with controls for heterogeneous treatment effects with respect to zip-code level average education and income. As with the full sample analysis in Table 1, the results weaken statistically due to increased standard errors, but here we are able to reject the hypothesis that the treatment effects for large prior donors are the same as those for small prior donors (p-value = 0.065 for likelihood of donation, and p-value of 0.034 for donation amount).

Robustness Checks

We also present in Appendix Tables 1 and 2 the same regressions from Tables 1 and 2 using a probit for the binary "any donation" outcome variable and a tobit specification for the "donation amount" outcome variable because it is censored at zeroThe results are similar, but the standard errors are larger, in the tobit and probit specifications. Furthermore, we check in Appendix Tables 3 and 4 the effect of having the name "Yale" versus simply "university", and find that the sub-treatment test comparing the two subgroups does not yield consistently differential effects, although it does appear potentially to somewhat exacerbate the treatment effects and heterogeneity described above.

5. Discussion and Conclusion

Naturally other explanations exist for the results from our experiments. The weakest aspect of our interpretation is the usage of small versus large prior donors as a proxy for altruistic versus warm glow giving. Alternatively, for instance, large prior donors may pay more attention to the marketing material. This would predict that large donors give more, but however would not predict that small donors respond negatively (instead, it would predict that no effect on small donors). Second, large prior donors may be wealthier, and perhaps wealthier donors are more educated, more able to understand the importance of rigorous evaluations. Although the demographic data at the zip-5 code level enable tests of whether heterogeneity in giving is driven by education or income, rather than prior donation amount, this is clearly a noisy proxy for education at the individual level. We do find the varying response by prior donation amount remains robust even after controlling for heterogeneity by education and income. Furthermore, this confound could explain why larger givers respond positively to the research treatment, but does not explain why smaller donors respond negatively relative to the control.

Our finding that smaller prior donors respond to information on charitable effectiveness by donating less frequently and in smaller amounts is consistent with other research showing that emotional impulses for giving shut down in the presence of analytical information. Indeed, controlled laboratory experiments have produced insights

that suggest that emotionally triggered generosity may be dampened by appeals that include statistical or deliberative information. For example, people donate less to feed a malnourished child when statistics that put this child in the larger context of famine in Africa are mentioned (Small, Loewenstein, and Slovic 2007). Similarly, people expressed diminishing willingness to fund clean water that would suffice to save the lives of 4500 people in a refugee camp threatened by cholera as the population of the camp increased (Fethersonhaugh et al. 1997). Small, Loewenstein, and Slovic suggests that deliberate thinking decreases the emotional appeal of identifiable victims without a commensurate increase in motivation to give to statistical victims, which causes a drop in donations. However, neither of these experiments focused on effectiveness of the charity, but rather focused on the depth of the need, and the number of people in crisis. Furthermore, we believe it is helpful to test these questions in non-laboratory conditions, and hence our evaluation builds off of these prior studies and tests these findings through the normal operations of a nonprofit organization, without risk of individuals behaving differently because they are aware that their responses will influence a research study (see Levitt and List 2007 for a discussion of these methodological issues).

Naturally, we are not the first researchers to examine, and find, differences in charitable giving between small and large donors. Using a panel data set on charitable donations, Reinstein (2011) finds that larger donors have more "expenditure substitution" in charitable giving. Reinstein identifies his effects through correlations in the residuals of fixed-effect regressions on donations to particular categories of charity. He finds that a temporary shock such as a personal appeal that increases donations to one charity decreases donations to other charities for large donors but has little effect on other donation decisions by small donors. This behavior is consistent with the differences that we observe between small and large donors, and with these differences stemming from different motivations for charitable giving. Reinstein suggests that small donors are responding primary to temporary shocks or personal appeals, while large donors have other motives.

Experimental evidence that small donors have different motivations than large donors also comes from a recent field experiment by DellaVigna, List, and Malmedier (2012). Individual donations were observed during a door-to-door fund-raising campaign.

One-third of addresses were simply visited by fund-raisers, while another third were informed the day before the visit that the visit would occur, and a final third were informed and given a check-box form that they could use to opt out of the visit. The experimenters find that allowing subjects to avoid the fund-raisers reduced the share of subjects answering the door and also reduced giving by small donors, but not larger donations. Their interpretation is that small donors are primarily motivated by social pressure or avoidance of an annoyance (someone selling something, in this case a charity, at the door). While we find different motivations by small donors as well, our point mostly is to categorize "altruism" separately from non-altruistic motivations such as both social pressure, annoyance avoidance and warm-glow.

Lastly, we also observe many tiny donations in our data that are hard to reconcile with purely altruistic motives. Small donations can cost more to process than the gift itself, and so rational altruists would never make particularly small donations. Nonetheless, 64 donations in our dataset are for exactly one dollar, and 678 are for ten dollars or less. This is normal in retail fundraising. While this could be a by-product of underestimating processing costs and the outcome of a maximization process trading off personal consumption utility and altruistic utility for someone of low income, we conjecture this is more likely evidence of symbolic, warm-glow giving.

We find that presenting positive information about charitable effectiveness increases the likelihood of giving to a major U.S. charity for large prior donors, but turned off small prior donors. This heterogeneity is important, we believe, and is consistent with a model in which large donors (holding all else equal, including income and wealth) are more driven by altruism and small donors more driven by warm glow motives. Altruistic donors, we posit, are more driven by the actual impact of their donation, and thus information to reinforce or enhance perceived impacts will drive higher donations. On the other hand, for warm glow donors, information on impacts may actually deter giving by distracting the letter recipient from the emotionally powerful messages that typically trigger warm glow and instead put forward a more deliberative, analytical appeal which simply does not work for such individuals. This distinction is much along the lines of Kahneman (2003), in which System I decisions (peripheral decisions which use intuition and mere reaction, but no deliberation) are "warm glow"

decisions, and System II decisions (deliberative decisions requiring conscious reasoning and thought) are "altruism" decisions.

Works Cited

Andreoni, James. 2006. "Philanthropy." In *Handbook of Giving, Reciprocity and Altruism, Handbooks in Economics*, edited by Louis-Andre Gerard-Varet, Serge-Christophe Kolm, and Jean Mercier Ythier.

Crumpler, Heidi, and Philip J Grossman. 2008. "An Experimental Test of Warm Glow Giving." *Journal of Public Economics* 92 (5): 1011–21.

DellaVigna, Stefano, John A List, and Ulrike Malmendier. 2012. "Testing for Altruism and Social Pressure in Charitable Giving." *The Quarterly Journal of Economics* 127 (1): 1–56.

Fethersonhaugh, David, Paul Slovic, Stephen M. Johnson, and James Friedrich. 1997. "Insensitivity to the Value of Human Life: A Study of Psychophysical Numbing." *Journal of Risk and Uncertainty* 14 (3): 283–300.

Hahn, Jinyong, Keisuke Hirano, and Dean Karlan. 2011. "Adaptive Experimental Design Using the Propensity Score." *Journal of Business & Economic Statistics* 29 (1): 96–108.

Kahneman, D. 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *American Economic Review* 93 (5): 1449–75.

Karlan, Dean, and Martin Valdivia. 2011. "Teaching Entrepreneurship: Impact of Business Training on Microfinance Clients and Institutions." *Review of Economics and Statistics* 93 (2): 510–27.

Levitt, Steven, and John List. 2007. "What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?" *Journal of Economic Perspectives* 21 (2): 153–74.

Null, Clair. 2011. "Warm Glow, Information, and Inefficient Charitable Giving." *Journal of Public Economics* 95 (5): 455–65.

Pritchett, Lant. 2002. "It Pays to Be Ignorant: A Simple Political Economy of Rigorous Program Evaluation." *The Journal of Policy Reform* 5 (4): 251–69. doi:10.1080/1384128032000096832.

———. 2014. "An Homage to the Randomistas on the Occasion of the J-PAL 10th Anniversary: Development as a Faith-Based Activity." *Center for Global Development*. Accessed March 12. http://www.cgdev.org/blog/homage-randomistas-occasion-j-pal-10th-anniversary-development-faith-based-activity.

Reinstein, David A. 2011. "Does One Charitable Contribution Come at the Expense of Another?" *The BE Journal of Economic Analysis & Policy* 11 (1).

Small, Deborah A, George Loewenstein, and Paul Slovic. 2007. "Sympathy and Callousness: The Impact of Deliberative Thought on Donations to Identifiable and Statistical Victims." *Organizational Behavior and Human Decision Processes* 102 (2): 143–53.

Yoruk, Baris. 2013. "Charity Ratings". University at Albany, SUNY, Department of Economics. http://ideas.repec.org/p/nya/albaec/13-05.html.

Appendix I: Test from Direct Marketing Solicitations

Wave 1: June 2007

Research Mailer
The following text and a picture of an old woman identified as Sebastiana, surrounded by a pink border:

> She's known nothing but abject poverty her entire life. Why on earth should Sebastiana have hope now? After forty-two years of toil in the unforgiving land of the high Andes, Sebastiana looks much older than her years. She has borne nine children and is alone to care for them after losing her husband six years ago. But a few months ago, Sebastiana joined a women's group sponsored by Freedom from Hunger. There she received a loan of $64 and training on how to grow her small, home-based business.
>
> But does she really have a right to hope for something different? According to studies on our programs in Peru that used rigorous scientific methodologies, women who have received both loans and business education saw their profits grow, even when compared to women who just received loans for their businesses. But the real difference comes when times are slow. The study showed that women in Freedom from Hunger's Credit with Education program kept their profits strong – ensuring that their families would not suffer, but thrive.

Control Mailer
The following text and a picture of an old woman identified as Sebastiana, surrounded by a pink border:

> She's known nothing but abject poverty her entire life. Why on earth should Sebastiana have hope now? After forty-two years of toil in the unforgiving land of the high Andes, Sebastiana looks much older than her years. She has borne nine children and is alone to care for them after losing her husband six years ago. But a few months ago, Sebastiana joined a women's group sponsored by Freedom from Hunger. There she received a loan of $64 and training on how to grow her small, home-based business.
>
> But does she really have a right to hope for something different? Like Sophia and Carmen before her, the good news is, yes! Because of caring people like you, Freedom from Hunger was able to offer Sebastiana a self-help path toward achieving her dream of getting "a little land to farm" and pass down to her children. As Sebastiana's young son, Aurelio, runs up to hug her, she says, "I do whatever I can for my children."

Wave 2 October 2008
Treatment Mailer

In order to know that our programs work for people like Rita, we look for more than anecdotal evidence. *That is why we have coordinated with independent researchers [at Yale University] to conduct scientifically rigorous impact studies of our programs.* In Peru they found that women who were offered our Credit with Education program had 16% higher profits in their businesses than those who were not, and they increased profits in bad months by 27%! This is particularly important because it means our program helped women generate more stable incomes throughout the year.

These independent researchers used a randomized evaluation, the methodology routinely used in medicine, to measure the impact of our programs on things like business growth, children's health, investment in education, and women's empowerment.

Control mailer

Many people would have met Rita and decided she was too poor to repay a loan. Five hungry children and a small plot of mango trees don't count as collateral. But Freedom from Hunger knows that women like Rita are ready to end hunger in their own families and in their communities.

Treatment postscript at bottom of letter

Rita is one of more than a million women Freedom from Hunger serves. We work hard to deliver services that make a difference – and we employ rigorous research[, like the evaluation by Yale University,] to keep us on track, to maximize our impact on women and their children all over the world. Your prior gifts have made these measurable impacts possible. Please continue to help us and make a gift today!

Control postscript at bottom of letter

Rita is one of more than a million women Freedom from Hunger serves – women who continue to prove the power of credit and education in the hands of a determined mother. Your prior support has been an essential ingredient in this worldwide recipe for financial security. Please continue to help us and make a gift today!

Outcome:	Any donation (1)	Amount given (2)	Any donation (3)	Amount given (4)	Any donation (5)	Amount given (6)
Received insert emphasizing research (A)	-0.004	1.923				
	(0.004)	(2.557)				
Received research insert X not large prior donor (B)			-0.006	1.572	0.002	8.31
			(0.004)	(2.730)	(0.020)	(11.979)
Received research insert X large prior donor (C)			0.010	4.445	0.019	11.388
			(0.012)	(7.313)	(0.023)	(14.086)
Received insert emphasizing research X education					0.000	-0.505
					(0.002)	(1.065)
Received insert emphasizing research X income					0.000	0.000
					(0.000)	(0.000)
Constant	0.169***	16.704	0.170***	16.921	0.135***	-4.925
	(0.036)	(22.296)	(0.036)	(22.304)	(0.045)	(27.467)
Controls for stratification variables	Yes	Yes	Yes	Yes	Yes	Yes
Control for large prior donor	No	No	Yes	Yes	Yes	Yes
Controls for education and income	No	No	No	No	Yes	Yes
T-test p-value: (B) = (C)			0.200	0.713	0.178	0.694
Observations	34673	34673	34673	34673	34673	34673
Mean of dependent variable for small donors	0.166	6.777	0.166	6.777	0.166	6.777
Mean of dependent variable for large donors	0.239	67.24	0.239	67.24	0.239	67.24

Notes: Standard errors in parentheses. All regressions include controls for the groups into which respondents were stratified for the randomization. For round 1, the randomization was stratified based on prior giving amount ($1-$24.99, $25-$49.99, $50-$99.99, $100-$249.99, $250-$500, and $500+) and recency of donation (2004/2005 or 2006/2007). For round 2, the randomization was stratified based on prior giving amount (above/below 100), whether they donated in 2008, and treatment status in prior rounds. * p<0.10, ** p<0.05, *** p<.01

Table 2: Analysis Separated by Recent and Past donors - OLS

Outcome:	Recent donors (>0 gifts in the last year)						Past donors (0 gifts in the last year)					
	Any donation	Amount given	Any donation	Amount given	Any donation	Amount given	Any donation	Amount given	Any donation	Amount given	Any donation	Amount given
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Received insert emphasizing research (A)	-0.009	1.236					0.001	2.646				
	(0.007)	(2.332)					(0.004)	(4.406)				
Received research insert X not large prior donor (B)			-0.014*	-0.805	-0.015	12.302			0.001	3.662	0.021	4.800
			(0.008)	(2.526)	(0.035)	(11.309)			(0.004)	(4.640)	(0.019)	(20.049)
Received research insert X large prior donor (C)			0.022	12.978**	0.023	26.306**			0.002	-6.677	0.022	-5.349
			(0.019)	(6.058)	(0.040)	(12.859)			(0.013)	(14.052)	(0.023)	(24.577)
Received insert emphasizing research X education					0.002	-1.104					-0.002	0.140
					(0.003)	(0.992)					(0.002)	(1.806)
Received insert emphasizing research X income					0.000	0.000					0.000	0.000
					(0.000)	(0.000)					(0.000)	(0.000)
Constant	0.485***	68.217**	0.489***	69.725**	0.560***	56.543*	0.075**	1.434	0.075**	0.842	-0.003	-23.674
	(0.091)	(29.596)	(0.092)	(29.602)	(0.102)	(33.076)	(0.030)	(31.519)	(0.030)	(31.531)	(0.039)	(41.890)
Controls for stratification variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control for large prior donor	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Controls for education and income	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
T-test p-value: (B) = (C)			0.079	0.036	0.065	0.034			0.956	0.485	0.950	0.494
Observations	16611	16611	16611	16611	16611	16611	18062	18062	18062	18062	18062	18062
Mean of dependent variable for small donors	0.275	9.171	0.275	9.171	0.275	9.171	0.071	4.696	0.071	4.696	0.071	4.696
Mean of dependent variable for large donors	0.328	92.165	0.328	92.165	0.328	92.165	0.115	32.801	0.115	32.801	0.115	32.801
p-value from test of equality of coefficients of (A) for the two above models, with and without prior gifts in the past year (e.g., (1)<>(7))	0.222	0.728										
p-value from test of equality of coefficients of (B) for the two above models, with and without prior gifts in the past year (e.g., (3)<>(9))			0.083	0.137	0.362	0.656						
p-value from test of equality of coefficients of (C) for the two above models, with and without prior gifts in the past year (e.g., (3)<>(9))			0.432	0.418	0.993	0.242						

Notes: Standard errors in parentheses. All regressions include controls for the groups into which respondents were stratified for the randomization. For round 1, the randomization was stratified based on prior giving amount ($1-$24.99, $25-$49.99, $50-$99.99, $100-$249.99, $250-$500, and $500+) and recency of donation (2004/2005 or 2006/2007). For round 2, the randomization was stratified based on prior giving amount (above/below 100), whether they donated in 2008, and treatment status in prior rounds. * p<0.10, ** p<0.05, *** p<.01

Appendix 1: Treatment Effects on Combined Rounds - Probit and Tobit

Outcome:	Any amount (probit)	Amount given (tobit)	Any amount (probit)	Amount given (tobit)	Any amount (probit)	Amount given (tobit)
	(1)	(2)	(4)	(5)	(7)	(8)
Received insert emphasizing research	-0.004	-1.211				
	(0.004)	(9.948)				
Received research insert X not large prior donor			-0.006	-5.637	0.001	20.360
			(0.004)	(10.833)	(0.020)	(46.762)
Received research insert X large prior donor			0.009	22.595	0.017	50.235
			(0.011)	(25.127)	(0.024)	(53.305)
Received insert emphasizing research X education					0.000	-0.945
					(0.002)	(4.130)
Received insert emphasizing research X income					0.000	0.000
					(0.000)	(0.001)
Constant		-514.614***		-513.156***		-623.515***
		(13.354)		(13.422)		(62.521)
Controls for stratification variables	Yes	Yes	Yes	Yes	Yes	Yes
Control for large prior donor	No	No	Yes	Yes	Yes	Yes
Controls for education and income	No	No	No	No	Yes	Yes
Observations	34668	34673	34668	34673	34668	34673
Mean of dependent variable for small donors	0.166	6.777	0.166	6.777	0.166	6.777
Mean of dependent variable for large donors	0.239	67.24	0.239	67.24	0.239	67.24

Notes: Standard errors in parentheses. All regressions include controls for the groups into which respondents were stratified for the randomization. For round 1, the randomization was stratified based on prior giving amount ($1-$24.99, $25-$49.99, $50-$99.99, $100-$249.99, $250-$500, and $500+) and recency of donation (2004/2005 or 2006/2007). For round 2, the randomization was stratified based on prior giving amount (above/below 100), whether

Appendix 2: Analysis Separated by Recent and Past donors - Probit and Tobit

Outcome:	Recent donors (>0 gifts in the last year)						Past donors (0 gifts in the last year)					
	Any donation (probit)	Amount given (tobit)	Any donation (probit)	Amount given (tobit)	Any donation (probit)	Amount given (tobit)	Any donation (probit)	Amount given (tobit)	Any donation (probit)	Amount given (tobit)	Any donation (probit)	Amount given (tobit)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Received insert emphasizing research (A)	-0.022	-1.521					0.051	10.631				
	(0.031)	(4.816)					(0.048)	(8.155)				
Received research insert X not large prior donor (A)			-0.044	-6.263	-0.040	-8.624			0.045	6.667	0.286	41.936
			(0.033)	(5.249)	(0.137)	(21.494)			(0.051)	(8.690)	(0.217)	(37.216)
Received research insert X large prior donor (B)			0.116	23.734*	0.124	22.793			0.100	38.352*	0.358	75.968*
			(0.084)	(12.140)	(0.161)	(24.754)			(0.149)	(23.291)	(0.266)	(44.470)
Received insert emphasizing research X education					0.006	1.776					-0.015	-2.190
					(0.012)	(1.939)					(0.020)	(3.321)
Received insert emphasizing research X income					0.000	-0.000*					0.000	0.000
					(0.000)	(0.000)					(0.000)	(0.000)
Constant		158.333***		158.207***		157.978		169.424***		169.015***		168.816***
		(3.444)		(3.441)		(3.435)		(7.852)		(7.832)		(7.821)
Controls for stratification variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control for large prior donor	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Controls for education and income	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
Observations	16558	16610	16558	16610	16558	16610	17812	18062	17812	18062	17812	18062
Mean of dependent variable for small donors	0.082	1.794	0.082	1.794	0.082	1.794	0.018	0.487	0.018	0.487	0.018	0.487
Mean of dependent variable for large donors	0.069	13.915	0.069	13.915	0.069	13.915	0.024	4.328	0.024	4.328	0.024	4.328

Notes: Standard errors in parentheses. All regressions include controls for the groups into which respondents were stratified for the randomization. For round 1, the randomization was stratified based on prior giving amount ($1-$24.99, $25-$49.99, $50-$99.99, $100-$249.99, $250-$500, and $500+) and recency of donation (2004/2005 or 2006/2007). For round 2, the randomization was stratified based on prior giving amount (above/below 100), whether they donated in 2008, and treatment status in prior rounds. * p<0.10, ** p<0.05, *** p<.01

Appendix 3: Effect of Yale Insert on Combined Rounds 1 and 2 - OLS

Outcome:	Any amount (1)	Amount given (2)	Any amount (3)	Amount given (4)	Any amount (5)	Amount given (6)
Received insert emphasizing research and Yale insert	0.000	-1.538				
	(0.006)	(3.825)				
Received insert emphasizing research and no Yale insert	-0.006	3.121				
	(0.004)	(2.740)				
Received insert emphasizing research X large prior donor X Yale insert (C)			-0.002	2.633	0.008	9.559
			(0.017)	(10.620)	(0.026)	(16.015)
Received insert emphasizing research X large prior donor X no Yale insert (D)			0.015	5.138	0.024	12.207
			(0.013)	(7.885)	(0.024)	(14.409)
Received insert emphasizing research X not large prior donor X Yale insert (E)			0.000	-2.155	0.009	4.642
			(0.007)	(4.101)	(0.020)	(12.351)
Received insert emphasizing research X not large prior donor X no Yale insert(F)			-0.008*	2.843	0.000	9.677
			(0.005)	(2.923)	(0.020)	(12.030)
Received insert emphasizing research X education					0.000	-0.515
					(0.002)	(1.065)
Received insert emphasizing research X income					0.000	0.000
					(0.000)	(0.000)
Constant	0.006	-3.121	0.155	-269.373	0.112	-290.865
	(0.367)	(224.624)	(10633.624)	(6.52e+06)	(10630.713)	(6.51e+06)
Controls for stratification variables	Yes	Yes	Yes	Yes	Yes	Yes
Control for large prior donor	No	No	Yes	Yes	Yes	Yes
Controls for education and income	No	No	No	No	Yes	Yes
T-test p-value: (C) = (E)			0.917	0.677	0.953	0.671
T-test p-value: (D) = (F)			0.093	0.791	0.083	0.772
T-test p-value: (C) = (D)			0.354	0.817	0.351	0.806
T-test p-value: (E) = (F)			0.186	0.225	0.19	0.221
Observations	34673	34673	34673	34673	34673	34673
Mean of dependent variable for small donors	0.166	6.777	0.166	6.777	0.166	6.777
Mean of dependent variable for large donors	0.239	67.24	0.239	67.24	0.239	67.24

Notes: Standard errors in parentheses. All regressions include controls for the groups into which respondents were stratified for the randomization. For round 1, the randomization was stratified based on prior giving amount ($1-$24.99, $25-$49.99, $50-$99.99, $100-$249.99, $250-$500, and $500+) and recency of donation (2004/2005 or 2006/2007). For round 2, the randomization was stratified based on prior giving amount (above/below 100), whether they donated in 2008, and treatment status in prior rounds. * p<0.10, ** p<0.05, *** p<.01

Appendix 4: Effect of Yale Insert on Combined Rounds 1 and 2 - Probit and Tobit

Outcome:	Any amount (probit)	Amount given (tobit)	Any amount (probit)	Amount given (tobit)	Any amount (probit)	Amount given (tobit)
	(1)	(2)	(3)	(4)	(5)	(6)
Received insert emphasizing research and Yale insert	0.018	3.176				
	(0.037)	(6.057)				
Received insert emphasizing research and no Yale insert	-0.012	0.628				
	(0.027)	(4.485)				
Received insert emphasizing research X large prior donor X Yale insert (C)			0.056	13.230	0.095	14.990
			(0.102)	(15.255)	(0.152)	(23.844)
Received insert emphasizing research X large prior donor X no Yale insert (D)			0.123	31.454***	0.163	33.204
			(0.076)	(11.424)	(0.137)	(21.765)
Received insert emphasizing research X not large prior donor X Yale insert (E)			0.013	1.610	0.047	2.179
			(0.040)	(6.590)	(0.116)	(18.959)
Received insert emphasizing research X not large prior donor X no Yale insert(F)			-0.032	-4.951	0.003	-4.263
			(0.029)	(4.880)	(0.113)	(18.452)
Received insert emphasizing research X education					0.001	1.042
					(0.010)	(1.657)
Received insert emphasizing research X income					0.000	0.000
					(0.000)	(0.000)
Constant		164.164***		163.933***		163.755***
		(3.248)		(3.242)		(3.239)
Controls for stratification variables	Yes	Yes	Yes	Yes	Yes	Yes
Control for large prior donor	No	No	Yes	Yes	Yes	Yes
Controls for education and income	No	No	No	No	Yes	Yes
Observations	34586	34673	34586	34673	34586	34673
Mean of dependent variable for small donors	0.048	0.048	0.048	0.048	0.048	0.048
Mean of dependent variable for large donors	0.049	0.048	0.049	0.048	0.049	0.048

Notes: Standard errors in parentheses. All regressions include controls for the groups into which respondents were stratified for the randomization. For round 1, the randomization was stratified based on prior giving amount ($1-$24.99, $25-$49.99, $50-$99.99, $100-$249.99, $250-$500, and $500+) and recency of donation (2004/2005 or 2006/2007). For round 2, the randomization was stratified based on prior giving amount (above/below 100), whether they donated in 2008, and treatment status in prior rounds. * p<0.10, ** p<0.05, *** p<.01