The Demand for Products Linked to Public Goods: Evidence from an Online Field Experiment [*]

Brian McManus
University of North Carolina

Richard Bennet
CCA Global Partners

May 2009

We conduct a field experiment at a nonprofit organization's online store to study how demand changes when consumers' purchases generate revenue for a charitable cause. Consumers respond strongly when their purchases generate small donations by an anonymous outside group, but responses are substantially weaker when the outside donations are relatively large. Responses are also strong when the outside donation requires a personal donation which consumers generally decline. Overall, increasing the salience of financial incentives appears to dampen consumers' responses to charitable messages. We also present evidence that the donation pledges reduce price sensitivity and have positive long-term effects on demand.

JEL Codes: C93, D64, L30, M14

Keywords: Field Experiments, Charity-Linked Products, Corporate Social Responsibility, E-Commerce

[*] We thank Amar Cheema, Dan Elfenbein, Amanda Friedenberg, Mark Isaac, and Ambar Rao for many helpful comments. We also thank the anonymous nonprofit organization which generously provided the experimental setting and data. This research was supported by grants from the Kauffman Foundation and the NET Institute.

Authors' email addresses: mcmanusb@email.unc.edu, rbennet@ccaglobal.com. Corresponding author: McManus. Mailing address: Department of Economics, University of North Carolina, Chapel Hill NC 27599-3305.

1. Introduction

An increasing number of firms and organizations are selling products linked to social causes.[1] For-profit firms do this through donation pledges (e.g. Target and Whole Foods) or commitments to green or fair-trade production (e.g. Starbucks). Non-profit organizations sell products to supplement revenue from direct donations (e.g. the Nature Conservancy offers apparel and a magazine). There may be many motivations for this activity. On the supply side, a for-profit firm may offer these products for strategic reasons, perhaps because consumers or employees value the charity attribute, while non-profits may view the products as a way to advertize their causes and expand their base of supporters. On the demand side, positive responses from consumers may come from a desire to contribute to a public good, a personal utility benefit from acting charitably, or an inference that socially-linked products are of relatively high quality. With a variety of potential motivations for their actions, consumers' observable responses to these products may be complex as well.

To improve our understanding of consumers' responses to socially-linked products, we analyze data from a field experiment conducted at the online store of a large anonymous nonprofit organization (NPO). We observe how consumers' purchasing decisions are affected by variation in the revenue their choices generate for the NPO's charitable mission. To interpret consumers' choices, which are counterintuitive in some cases, we offer a theoretical framework for describing how consumers' preferences interact with the incentives they face within the experiment. We also study whether consumers' price sensitivity is affected by variation in the revenue their purchases generate for charitable mission, and whether the experiment's immediate effects on demand lead consumers to substitute away from future purchases or direct charitable donations.

Our examination of actions both within and beyond the experimental period is new to field studies in this area. Numerous surveys and classroom experiments have established that consumers will express a willingness to favor products sold by socially responsible firms. The controlled and low-cost nature of these studies permits inference on some fine distinctions about consumers' responses, for example what types of products are most amenable to social links (Strahilevitz and Myers, 1998), whether positive or negative news has a greater impact on demand (Sen and Bhattacharya, 2001), and how consumers interpret profit donation pledges

[1] For a description of recent trends, see Bonini, Mendonca, and Oppenheim (2006).

(Olsen, Pracejus, and Brown, 2003). Field studies, by contrast, are more costly to conduct as experiments and more likely to contain identification challenges when they employ observational data. As a result, the existing literature utilizing field data has tended to focus on testing simpler hypotheses of whether social links, in general, affect demand. Elfenbein and McManus (2008) and Popkowski Leszczyc and Rothkopf (2006) study charity auction markets to establish that consumers will submit higher bids in a charity auction compared to a non-charity auction for the same product. Hiscox and Smyth (2005) use a field experiment to demonstrate that consumers will favor products with a "fair trade" label. In these field studies, it is difficult to infer what aspects of a charity announcement are most important: Is it the presence of any offer at all? The precise financial terms of the charity pledge? An additional challenge lies in uncovering the broader effects of a consumer's purchase of a charity-linked product, whether through the consumer's later demand for merchandise or his donations to the same cause.

Our study also complements the substantial literature on consumers' choices in making direct charitable donations. Recent research on donation choices, largely with field experiments, demonstrates that consumers behave in interesting and sometimes surprising ways, especially with regard to varying information about others' actions. Eckel and Grossman (2003) find that financially equivalent inducements for donations can bring different results depending on how the inducements are presented, with matching pledges from other donors out-performing rebates. Karlan and List (2007) also report that matching pledges can stimulate donations significantly, but they find that substantially different matching incentives yield similar responses from donors. Shang and Croson (2006) offer further evidence that one consumer's actions can stimulate another's giving; they find that public radio donors increase their donations when told that another individual recently made a large gift. While the immediate effects of information and matching pledges can be substantial, Meier (2007a) finds that consumers who received a short-term matching stimulus reduced their later donations so that their overall contributions were no greater than those of a control group that received no stimulus. In all, these studies suggest that financial incentives matter, but the information conveyed through experimental treatments may have a greater effect. The data from our experiment suggests the same.

For our study, we observed the actions of over 100,000 consumers at the NPO's online store during March and April 2007. When each new consumer arrived at the store, a "cookie" was placed on his or her internet browser to record and preserve the random assignment of the

consumer into an experimental treatment. The experiment varied the appearance of the store's front page. Consumers in a control group received the standard storefront. Others were told that $1 or $5 would be donated by an anonymous outside group if the consumer purchased at least $10 in merchandise. Yet other consumers could trigger $1, $5, or $10 donations from the outside group by purchasing at the store and donating to the NPO. All consumers received either standard store prices or discounted prices. For 12 months following the experiment, the NPO tracked consumers who purchased from the store, and we observe their subsequent store purchases and donations to the NPO.

Consumers who received donation pledges generated 20% more revenue, on average, than consumers who received no pledge.[2] Although a positive demand response to the donations pledges was expected, there were several surprising aspects of the results. First, despite the structure of the incentives, the additional revenue largely came from increased order sizes. While consumers who purchased under the control message spent an average of $43, those who received a donation pledge spent $51. The structure of our pledge implies that this difference did nothing to increase the outside donation to the NPO. This happened even if the pledged donation was small ($1), and more strikingly if the pledged donation was not triggered at all because the consumer did not complete the required personal donation portion of an offer. We interpret the donation pledges as stimulating a change in consumers' perceived value of supporting the NPO through their own purchases. Not all donation pledges had this effect. Consumers who observed the richest offer – a $5 outside donation with a purchase but no consumer donation required – had the weakest response among experimental treatments. Consumers did not respond to these $5 pledges unless their orders risked being too small to trigger the pledged donation. We describe these consumers as focusing on the financial terms of the pledge rather than using the pledge to update their preferences for supporting the NPO in general. This suggests that consumers' responses to socially-linked products may be very sensitive to how a firm or non-profit describes the public goods benefit of purchasing the products. This also provides a potential explanation for the literature's broad range of estimates for charitable giving's price elasticity, which Karlan and List (2007) describe. Small differences across studies in how price variation is generated may have a large impact on how consumers interpret the variation.

[2] With this statistic and others referenced in the Introduction, we describe activity in the lower 95% of all store purchases. The data are skewed by a small number of large orders that are from businesses rather than households. See Section 3 for discussion of this approach to the data.

During the year following the experimental period, consumers who initially purchased under a donation pledge were more likely to return to the store and order again. This suggests that the brief stimulation of the experiment did not merely shift consumer spending away from future time periods. We also found very little overlap between the store customer population and the NPO's direct donors, suggesting minimal substitution between intensified store activity and donations. Our evidence on consumers' responses to changing prices is mixed. The median consumer appears to be less price-sensitive when exposed to messages about pledged donations, but the average effects show no significant difference between consumers who received the control and a donation pledge.

The rest of the paper proceeds as follows. We introduce the experimental setting and procedures in Section 2, and in Section 3 we discuss the incentives generated by the experiment. In Section 4 we describe the data, and Section 5 contains our econometric analysis. In Section 6 we offer a concluding discussion and provide suggestions for future work.

2. Experimental setting and structure

2.A. The online store

At its online store the NPO offers a variety of books, functional equipment consistent with the NPO mission, and apparel. The store also includes a donation opportunity with which consumers can "buy" donations in $10 increments. In addition to the central online store at which our experiment ran, some of the NPO's products can be purchased through local branches of the organization. In total, sales at the main store account for about 15% of NPO merchandise sold through all channels. Merchandise sales are a small fraction of overall NPO annual revenue, which exceeds $500 million.

The NPO advertises the store on its home page and through emails to its past donors, customers, and volunteers. In addition, the NPO places store ads on internet search engines and in other websites' advertising banners. Around the time of the experiment the store received approximately 60,000 visits per month yielding 1,200 orders. A typical store customer purchases two or three items during a visit, presumably for household use, but some customers represent firms or organizations that require a large quantity of items. During spring 2007 excluding the weeks of the experiment, the median order size was $28 and approximately 97% of orders

included at least $10 of merchandise. During the same period, a limited number of consumers (about 1%) included donations with their orders.

Consumers who enter the store are shown a large graphic which displays a promotional message. The store's front page always includes a reminder that consumers' orders support the NPO's charitable mission. To the side of the front page's central graphic there is a list of product category links that take consumers to summary pages which display the pictures and prices of items within the product categories. From these summary pages, a consumer can click on an individual item to see a detailed picture, product description, and price, and the consumer can elect to place one or more units of the item in his shopping cart.

2.B. Experimental treatments

We conducted the experiment from March 6, 2007 until April 15, 2007.[3] During this period, each store visitor had a cookie placed in his or her internet browser to assign an experimental treatment and identification code. We use this code to track store visitors who return multiple times within the experimental period.[4] As long as the visitor's cookie was not deleted from his browser, each time he returned to the store he observed the same experimental treatment. Consumers who deleted cookies from their browsers or visited the store from multiple computers (e.g., home and work) were assigned a new cookie and likely placed into a different experimental treatment. Unfortunately, we cannot measure the number of consumers who observed multiple experimental treatments due to this characteristic of cookies. In the discussion below we describe aspects of the study and data that indicate this is unlikely to be a major problem.

There are two components to each experimental treatment. First, each consumer received one of six messages regarding the charitable value of purchasing at the NPO's online store. Second, each consumer was assigned either the regular online store prices or a set of discounted prices. In total, there were 12 experimental treatments in the experiment (6 messages × 2 price levels). The 12 treatments were assigned with equal probability. During the experimental period the NPO suspended all other price promotions for its products.

[3] The experiment's starting date and duration were decided jointly with the NPO, and was limited in part by the NPO's desire to vary the store's promotional offers, which generally were suspended during the experiment.

[4] See Chan, Li, and MacKie-Mason (2006) on the variety of methods available for online field experiments.

2.B.i. Donation messages

Consumers' messages regarding charitable value were delivered through the central graphic of the store's front page. In the control case ("Null"), the central graphic displayed text indicating that store activity benefits the NPO's charitable mission. In Figure 1 we display the relevant portion of this central graphic, with identifying information about the NPO obscured. The extended section of obscured text is a description of the NPO's mission, and the partial image on the right of the Figure is a young girl wearing NPO apparel. The first two non-Null messages promised consumers that an additional donation would be made if they purchased at least $10 in merchandise at the online store.[5] The donation values were $1 and $5, which we identify here as "P:1" and "P:5," respectively, and as "P:δ" when referenced as a pair. The donations were pledged by the semi-anonymous "WU Support Fund" (WUSF) and were paid by the researchers. See Figure 2 for the relevant portion of the P:1 graphic. The P:5 graphic is identical except for the donation amount. In the final set of messages, consumers were promised that if they purchased $10 or more in merchandise plus donated $10 or more at the store, the WUSF would donate $1, $5, or $10. We identify these treatments as "P+D:1," "P+D:5," and "P+D:10," and as "P+D:δ" when referenced as a group. In Figure 3 we display the central graphic for the P+D:1 message. The P+D:5 and P+D:10 graphics are different only in their donation amounts. Each non-Null treatment included a block of fine print that described the WUSF ("a group of private individuals committed to assisting the [NPO] in achieving its mission"), stated that the donation offer was limited to one per customer for the duration of the promotion, and specified that there was no maximum total donation from the WUSF.

In addition to the WUSF donation pledges on the front page's central graphic, we placed condensed versions of each message near the top of the store's product-level pages. In Figures 4 and 5 we display these condensed messages for the P:1 and P+D:1 treatments. Consumers who were assigned the Null treatment received no message in this space. We included these reminders in case a consumer clicked past the store's front page without reading the text in the central graphic.

[5] Together with the NPO, we decided that our donation pledges might not appear credible if they were larger than the least expensive items at the online store. Unfortunately, the $10 threshold is not an optimal size for testing consumers' interpretation of the WUSF messages.

In order to handle potential consumer confusion about the donation messages, the NPO provided a phone number for customer service. The NPO's customer service employees were prepared to give callers additional information about the donation pledges, but the NPO reported that no consumers called to request this information.[6] The absence of phone calls provides some evidence that consumer exposure to multiple experimental treatments is not a serious concern.

To understand the benefits of using the lump-sum WUSF donations rather than another mechanism that generates charity value, consider a few alternatives. First, it is possible to simply tell consumers that a $30 item includes an implicit donation of either $3 or $6 to the NPO's mission. One potential problem is that only one (at most) of the two donation amounts can be true without outside funding. An additional problem is that without describing the presence of outside funding, consumers may make different inferences about the underlying quality of the product being sold – in this example that the item's cost is either $27 or $24. Second, we might have used varying percentage donations from the WUSF, for example that the WUSF would donate 1%, 5%, or 10% of a customer's order value. While this scheme has the benefit of resembling some firms' actual pledges of donations in proportion to sales, paying pledged WUSF donations in these cases could be prohibitively costly due to a small number of store customers who place very large orders (several thousand dollars). In Section 3, once we have introduced a model of consumers' choices, we describe additional benefits of the lump-sum WUSF donations in allowing us to distinguish among possible explanations for consumers' responses to the experiment.

2.B.ii. Price variation

To assess how price sensitivity changes with the strength of charity association, we randomly assigned consumers to either the regular (non-sale) store prices or discounted prices. Of the 45 items offered at the store, we received permission to adjust prices on 20 items, which represents about half of the store's sales. When the discount was applied, the prices of these 20 items were reduced by 8-20% relative to the control.[7] Focusing on items that were potentially discounted during the experimental period, the (weighted) average regular price was $32.20 and the average

[6] Customer service representatives were also instructed offer a consumer the discounted set of prices if the consumer saw both price levels and was confused by the difference. No consumers called with questions about prices.

[7] Variation in discounting percentages is generally due the goal of reducing prices in whole-dollar increments. The NPO's marketing group provided extensive input in selecting the discounted prices.

discounted price was $28.54. The 25 items that were not subject to discounting had an average price of $15. Other than the price levels, the discounted prices had the same appearance as the regular prices. No additional messages about sale prices were displayed.

3. Incentives generated by the experiment

We now provide a brief discussion of how consumers might respond to the experimental treatments. We first consider what would occur if consumers' preferences and information about the NPO are stable, and consumers respond to the financial terms of the treatments in a conventional way. We do not consider all permutations of the treatments here, but instead provide an overview of some important patterns that would appear in the data. We then comment on how consumers' actions would be different if the experiment's main impact is to generate information for consumers about the value of supporting the NPO. In the contrast to the more conventional incentive model, we describe the information's impact as changing the consumers' preferences over revenue transferred to the NPO.

Assume that a consumer has the opportunity to purchase $x \in \{0, 1, 2\}$ units from the NPO's online store at a price of p per unit. In addition to utility from the product itself, $v(x)$, the consumer receives a warm-glow benefit from three potential revenue sources to the NPO. First, there is revenue to the NPO's charitable mission from the "profit" on x units purchased, $(p–c)x$, where c is the NPO's marginal cost. Second, there is a possible donation from the WUSF. Let δ represent the WUSF's offered donation, and let δ_A be the actual transfer from the WUSF to the NPO, where $\delta = \delta_A$ if the consumer meets the terms of the WUSF offer and $\delta_A = 0$ otherwise. Third, there is the consumer's own direct donations, d, to the NPO. These donation-related benefits generate utility through the function $w((p–c)x, \delta_A, d)$. Let w_j represent the derivative of w with respect to the j^{th} argument, and let w_{jk} be the cross-partial derivative between arguments j and k. If a consumer views NPO revenue from different sources as substitutes, then $w_{jk} < 0$ across pairs of distinct arguments. Assume that v and w are increasing in each argument, and $v(0) = w(0, 0, 0) = 0$. By "conventional" responses to financial incentives, we mean that a consumer's value of w evaluated at any (x, δ_A, d) depends only on the values of (x, δ_A, d), and not on a WUSF offer with unmet terms so that $\delta \neq \delta_A$. We combine v and w linearly along with spending on x and d to obtain a consumer's utility:

$$U(x, \delta_A, d) = v(x) + w((p–c)x, \delta_A, d) – px – d.$$

A consumer who does not purchase from the NPO store may receive positive utility by donating $d > 0$. To account for variation across consumers in their observed choices of (x, d), we assume that consumers are heterogeneous in their functions v and w.

We begin by focusing on the relationship between x and δ_A, holding fixed a consumer's preferences in v and w.[8] Assume that a quantity of $x = 1$ is sufficient to trigger the donation in a P:δ message. For a given p, an increase in δ_A raises the benefit of purchasing $x > 0$ relative to $x = 0$. Thus, we expect the share of store visitors who purchase to be greater with P:1 than under the Null, and larger still with P:5. If, however, $w_{12} < 0$ because an increase in δ_A reduces the marginal benefit of a consumer's own implicit donation through x, then consumers will purchase less (conditional on $x > 0$) as δ_A increases from \$0 to \$1 to \$5. If $w_{12} = 0$ because consumers do not view the WUSF's donations and their own as substitutes, then consumers who would have purchased under the Null should not change their choice of x. Changes in d under the P:δ messages could come from non-zero cross-effects between x or δ_A and d.

The P+D:δ treatments are designed to stimulate consumers' donation decisions while they are in the NPO store. The pledged WUSF donation effectively lowers a consumer's "price" (\$10) of a \$11, \$15, or \$20 donation. A consumer who would choose $x = 2$ and $d = 0$ in the absence of a P+D:δ offer may do so, in part, because of the positive utility received from the implicit donation in the NPO's products. Under a P+D:δ message, the appeal of contributing $d > 0$ rises, but a consumer who now chooses $d > 0$ in order to trigger the donation δ_A may have a smaller charity benefit from the second unit of x (i.e., if $w_{13} < 0$). The impact of δ_A can compound the likelihood of a reduction in x if $w_{12} < 0$ as well. In addition, we note that a consumer with stable preferences who is exposed to a P+D:δ message yet does not donate should not change his value of x relative to the Null.[9]

The impact of price variation is clearest in its effect on utility conditional on x. When price falls, there is a direct benefit to U through a consumer's expenditures (px) and an offsetting

[8] In addition to the effects we describe here, a WUSF donation is likely to draw-in a different distribution of tastes among consumers who choose to purchase. These consumers' optimal values of x may be different than the xs of consumers who would purchase under the Null. We abstract away from this concern due to the apparent minimal effect of WUSF messages on the extensive margin, which we describe in Section 4.

[9] If NPO revenue is perceived as a public good, the consumer may reduce his choice of x if the marginal value of his implicit contribution is reduced by the belief that other consumers will change their donation and purchasing behavior (in addition to revenue from the WUSF).

change to the consumer's implicit donation, captured by w. In all, the impact of a reduction in p on U is $-(w_1 - 1)x$, where the term w_1 dampens utility's sensitivity to price variation. This change determines how the WUSF treatments affect price sensitivity at the extensive margin, where the difference in utility from $x = 0$ or $x = 1$ is captured by $-(w_1 - 1)$. For a consumer with stable preferences, donations of δ_A from the WUSF or d from the consumer himself may reduce w_1 if $w_{12} < 0$ or $w_{13} < 0$, which increases price sensitivity. If there is no crowding-out across donation sources or types, then elasticity at this margin is unchanged relative to the Null.[10] Price sensitivity on the intensive margin is more difficult to characterize because consumers considering a switch from $x = 1$ to $x = 2$ are subject to both a marginal and inframarginal effect of changing prices on w.[11] Without further assumptions on w, it is uncertain whether a price reduction will increase or decrease the difference in w a consumer receives from increasing x. Thus, we cannot predict whether price elasticity on this margin should vary across experimental treatments, and determining this is an empirical question.

We may consider the intertemporal effects of the experiment by assuming that consumers make choices over x and d repeatedly over time. There are several ways that the model could be extended for the dynamic choice problem. One possibility is that U is defined for a period longer than the experiment, and large values of x and d selected during the experiment reduce the marginal benefit from implicit or direct donations later. Alternatively, we could interpret U as a flow utility for each round of x and d choices, and consumers with a fixed budget for charitable spending who increase x and d during the experiment would adjust their later activity. When preferences are stable and U is concave in its arguments, either interpretation of the dynamic choice problem predicts that an increase in x, d, or δ_A during the experiment will result in a decrease in a consumer's future donations or purchases.

In contrast to the responses as described above, where consumers with stable preferences respond conventionally to the financial terms of the experimental treatments, it is possible that the NPO store visitors process the WUSF messages differently. As in Shang and Croson (2006),

[10] When a WUSF donation is made, a different collection of consumers (with different preferences) may be on the extensive margin relative to the Null. Differences across treatments in the density of consumers on the extensive margin may affect the observed changes to purchasing behavior across price treatments.

[11] Consumers considering the difference in U from $x = 1$ and $x = 2$ must calculate $w(2(p-c), \bullet) - w((p-c), \bullet)$, but the impact of changing prices on this w difference depends on the curvature of w. For example, if $w = \log[(p-c)x]$ then the difference in w is unaffected by changing prices, while other w specifications could cause the difference to increase or decrease with prices.

consumers who receive information about other agents' actions may update their own preferences regarding a charitable organization or its products.[12] A WUSF message may change the properties of w, especially through the benefit consumers receive from their own implicit (x) or direct (d) donations, regardless of whether they generate the donation δ_A.

We use the experimental data to distinguish which sort of interpretation of WUSF messages dominates in the consumer population. Two aspects of the experiment allow us to make this evaluation. First, if consumers with stable w functions met the terms of a P+D:δ message, we predict that their choices of x will decrease or remain unchanged. If the presence of the message increases the consumer's benefit from the implicit donation in x, then the P+D:δ message has the opposite impact on x. Second, the lump-sum benefit δ_A in the P:δ messages is predicted to reduce or not change values of x among consumers with $x > 0$. However, an increase in w_1 in the presence of a P:δ message can lead to an increase in a consumer's x choice. By contrast, if we had created the P:δ messages with a percentage donation (e.g. 5% of a consumer's total order), an increase in x could be ascribed to either a shift in the marginal financial incentives or a change to w_1.[13]

Similar distinctions may be noted in consumers' responses to price variation and their activity beyond the experiment. If the WUSF message shifts a consumer's preferences in w, this increases the dampening effect of price changes on utility at the extensive margin, reducing demand elasticity rather than increasing it or keeping it fixed. Likewise, the stable preferences interpretation of the model predicts that post-experiment expenditures should fall, while an increase in w_1 and/or w_3 leads to increased later activity.

4. Data

4.A. Reports from the online store

After the experiment ended, the NPO provided data on activity at the online store. These data capture activity at two levels. First, we observe basic information about every visit ("session") to the online store. For each session, we observe the time and date it was initiated, the

[12] "Updated preferences" may be interpreted as the reduced form of a model in which consumers use information, including WUSF pledges, to form expectations about the quality of the NPO mission.

[13] This is a challenge in interpreting the results in Karlan and List (2007), who employ a proportional donation-matching scheme.

assignment into an experimental treatment, the visitor's Internet Protocol (IP) address, the visitor's approximate location based on their IP address, and the individual products viewed by the visitor.[14] Through the cookie stored on a consumer's browser, we are able to track consumers' repeat visits to the store. The second level of data is for consumers' purchases. For each consumer who placed an order at the store, we observe all session-level details listed above, the items selected and prices paid, any donation activity within the store, and the consumer's billing zip code. We use consumers' email addresses (which are entered during check-out) as an additional way to link individual consumers' activity across multiple transactions at the store.

The personal information, browser cookies, and identification codes of online store visitors were not used by the NPO or the researchers for any purpose other than the academic study we describe in this paper.

4.B. Consumer activity at the store

During the experimental period of March 6 to April 15, 2007 there were about 6,200 sessions initiated for each of the twelve treatments. These initiated sessions represent approximately 5,800 consumers per treatment that we examine in this paper. Some of the difference between the total sessions and studied consumers is due to consumers making multiple visits to the store before purchasing.[15] On average, consumers who did not order visited the site 1.07 times during the experimental period, while consumers who ordered visited 1.59 times. Other store visits are eliminated because of short-term disruptions in our treatment-assignment mechanism. Finally, for each consumer who purchased multiple times within the experimental period (and under a single treatment), we aggregate all orders into a single observation.

[14] Unfortunately, we were unable to link consumers' IP addresses to their geographic locations for our empirical analysis. The IP address provides the location of the consumer's Internet Service Provider (ISP), which can be in a different city or state that the consumer himself. In unreported analysis, we compared zip code-level demographics for consumers who purchased under our various treatments, and we found that consumers' locations were not significantly related to their choices under the WUSF messages.

[15] We investigated sessions (and subsequent orders) that were initiated in rapid succession from the same IP address, with the store visitor apparently deleting their browser cookies between each session. This may be due to testing of the website by the NPO or ourselves, or it may be consumers who realize that the store appearance changes each time they delete cookies and begin a new session at the store. About 4% of store orders follow from a session associated with this sort of activity, and these store orders are distributed fairly evenly across experimental treatments (including price discounting). This suggests that few store customers were behaving opportunistically with regard to inserting themselves into experimental treatments with more favorable terms.

About 2% of store visitors placed orders, with an average order size of $87.70. This implies that each unique store visitor generated $1.76 in revenue for the NPO. These values and similar statistics in this paper include only merchandise and exclude shipping charges. Due to a small number of very large orders, the median order is $37, considerably below the mean.[16] In general, the largest orders are placed on behalf of firms rather than households, but no perfect method exists to identify which orders are from firms. To focus on households' choices, which are more relevant for our objectives and the literature to which we are contributing, we drop orders above the 95th percentile of order values. After eliminating the top 5% of non-zero order values, the average order falls to $50 for a per-visitor average of $0.95. For our comparisons across experimental treatments, the threshold is always defined separately at the 95th percentile within each group of observations.[17]

In the top panel of Table 1 we display the number of visitors and their order values for the six experimental messages, including the Null. Regular and discounted price treatments are combined within messages in this portion of the Table. Consumers' purchase probabilities are marginally greater in most of the non-Null treatments, but none of the differences are large. The more striking difference is between the average order values in Null and non-Null messages. Revenue per store visitor is $0.65 greater when consumers receive a non-Null message. This is due to a difference of $31.50 per customer conditional on an order. The difference in order sizes between the Null and non-Null messages is $8.59 when we drop the largest 5% of orders from each subsample. With and without dropping large orders, the median order value is $40 for orders placed under each WUSF message except P:5, while both the P:5 and Null messages have median values just under $30.

The average order values across messages generally follow the pattern in the median order values. Excluding the largest 5% of orders, the order value per visitor for the P:1 treatment exceeds the Null by 27%, while the same statistic for P:5 is 5% greater than the Null. Each P+D:δ message has an average value per visit that is substantially larger than the Null, varying from 13% to 32% greater. In Figure 6 we display the cumulative distribution functions for three groups of orders. The order distribution under the Null (blue solid line) is strictly to the left of

[16] About 1% of orders include over $1,000 of merchandise.

[17] By defining the threshold separately for each group of data, we minimize bias due to differences across experimental treatments in what share of observations within the treatment would be above a single threshold.

the orders placed under a P:1 or P+D:δ message (green dashed line). This pattern is repeated if P:1 or any P+D:δ message is compared individually to the Null. The distribution of P:5 orders (red dotted line) follows the other non-Null treatments until about $25, beyond which it follows the distribution of Null message orders for greater values. We add the dashed line at $10 to ease examination of the distributions relative to the threshold for triggering a WUSF donation.

Despite increases in their order values, consumers in the P+D:δ treatments predominantly chose to omit donations from their store visits. As shown in Table 1, about 4% of consumers who ordered under a P+D:δ treatment made a donation, with the greatest number donating within the P+D:10 message. Consumers with P+D:δ messages who donated had orders that averaged $4.51 less than those of consumers who did not donate. It is important to emphasize that when consumers in P+D:δ treatments did not make their own donation, then no WUSF donation was triggered.[18] This regularity, along with the stronger consumer reaction to P:1 than P:5, presents a significant puzzle within the data. As we describe in Section 3, this response is difficult to explain with a single model of stable consumer preferences but more reasonable if some WUSF messages effectively shifted w_1.

In Panel 2 of Table 1 we display average order values within the price treatments. As expected, lower prices lead to both an increased purchase probability and an increased order size per store visitor. The displayed difference in order sizes understates the magnitude of consumer response to the price reductions. The revenue figures in the Table are reduced by price discounts; we account for this in the analysis below. An additional factor that complicates interpretation of Panel 2 is that the statistics include products that were never discounted. When we focus on the products that are included in the discounting portion of the experiment, we find that the chance of a store visitor purchasing at least one of these products increases from 1.06% without a discount to 1.26% with the discount in place. This is a greater proportional difference in purchase probability than when all merchandise orders are included.

Our experimental structure included equal division of consumers into the Null and each non-Null message. For some of the analysis below it is useful to compare a larger group of Null-like consumers to those who received a non-Null message. To do this, we augment the

[18] In all, WUSF offers generated a $1,636 donation to the NPO.

experimental data with store data from March and April outside of the experimental period.[19] Consumers who purchased from the store during these portions of March and April saw store graphics and prices similar to those in the Null portion of the experiment. Store activity during the supplemental period was similar to activity under the Null. See Panel 3 of Table 1 for summary statistics on the supplemental period. Of 31,388 visitors during this period, 2.3% of consumers placed an order (2% for the Null), and the average order value in the lower 95% of orders was $42.39 ($42.91 for the Null). Differences in the percentage of visitors who purchase may be due to promotions running during the supplemental period or general consumer shopping patterns during these portions of March and April.

4.C. Actions after the experiment

We supplement the data on consumers' choices in March and April 2007 with information on the same consumers' interaction with the NPO between May 2007 and March 2008. Due to the anonymity of consumers who visit the store and leave without purchasing, we are limited to consumers who purchased during the experiment and supplemental periods.[20] We study two aspects of these consumers' choices. First, we collect data on the consumers' transactions at the online store. Second, we received information on donation dates and amounts from the NPO database of online donors, which does not include donations made at the store. In all, we identified 86 consumers who were active with the NPO outside of the experimental period. Most of these consumers (69) were repeat customers of the store but not donors, and only 1 consumer appeared in both the follow-up store data and the donation data.

In examining a consumer's actions after the experimental and supplemental periods, we focus on transactions that occurred between 60 and 300 days after the consumer's last action during March and April 2007. We do this so that there is no overlap between subjects in the experimental period and supplemental period.[21]

---

[19] We exclude consumers who visit the store on April 16-17 to minimize the impact of consumers who first visited the store during the experimental period and were exposed to one of our WUSF messages. Due to data limitations, we also exclude consumers who entered the store on March 1 and after April 27.

[20] After the experimental period no new cookies were placed on consumers' web browsers, and the cookies that were placed during the experiment expired soon after its conclusion.

[21] For example, we do not want a consumer who purchases on March 10 and returns to the store on April 20 to be part of both the experimental and supplemental groups. But if the March 10 consumer's actions are omitted from the analysis until 60 days after purchase, then a consumer who first arrives at the store on April 23 should be treated in the same way.

On Table 2 we provide summary statistics on these consumers' purchases outside of the experimental and supplemental periods. Consumers who ordered multiple times after the experiment have their orders summed. In the Table and all related analysis we omit observations on consumers who were in the top 5% of orders during the experiment and supplemental periods. Consumers who saw the Null message or were in the supplemental group purchased at a rate of 2.6%, while 4.1% of consumers who received a WUSF message returned to purchase. Conditional on returning to the store, consumers who received WUSF messages placed substantially larger orders.

Considerably fewer store consumers appear in the NPO's roll of online donors. This donor database contains over 2 million names, and it accounts for a large share of the organization's donation revenue. Despite this, only 24 of the 2,116 customers from March and April 2007 were located in the donor database, with 18 consumers making donations after April 2007. There were 12 consumers (1% of 1169) who saw a WUSF message and donated later, and 6 consumers (.7% of 947) who donated after seeing the Null message or were a part of the supplemental group. While we forgo further empirical analysis of donation patterns because of the small number of observed consumers, the scarcity of these consumers may be considered a result in itself. This suggests that the NPO's collections of store customers and donors are largely distinct, and charity-oriented promotions at the online store are likely to have little impact on donation activity. The WUSF messages' sizable impact on order sizes has little opportunity to affect donation revenue.

5. Detailed analysis and testing

We now examine more closely the impact of the experiment on behavior. In evaluating the overall impact of the experiment, our central analysis is at the session level, which captures changes at both the extensive and intensive margins. We supplement this with detailed analysis of choices conditional on an order, as this provides an opportunity to further analyze consumers' responses. We also examine consumers' price sensitivity and whether exposure to WUSF messages affected behavior after the experiment.

5.A Impact of the experiment on order revenue

In this analysis our primary variable of interest is *r*, the total revenue from a visitor to the online store. Unless noted, we use Ordinary Least Squares (OLS) to estimate the effect of subsets of our experimental treatments on *r*. In most of the analysis below we focus on the transactions of consumers in the lower 95% of all store customers with $r > 0$. While the NPO's charitable mission benefits from store "profit" rather total revenue, consumers' demand responses to WUSF messages are demonstrated effectively with *r*. Order revenue allows us to account for the wide variety of objects for sale at the store, so that Section 3's variable *x* is more similar to *r* than *x* is to the quantity of items purchased. We do not observe the store's cost of filling orders.

We begin by presenting a set of session-level results in Table 3. For Specification 1 we regress *r* on a simple indicator of whether a store visitor saw a WUSF donation pledge. The model's constant recovers the average transaction value under the Null, 0.81, as we reported in the summary statistics of Table 1. Consumers who received any non-Null treatment spent an average of \$0.16 more than consumers with the Null ($p = .08$). In Specification 2 we divide the WUSF treatments by whether a consumer's own donation was required (P+D:δ) or not (P:δ). We find that impact of the P+D:δ treatments is significantly different from zero and larger in magnitude than the P:δ treatments, which have a positive but insignificant impact on revenue per store visitor. In Specification 3 we further separate the treatment messages by their individual content, and we find that the two messages with \$5 WUSF donations had the smallest impacts on store revenue; their coefficient estimates are positive but statistically insignificant. The estimates for the remaining messages range from \$.22 to \$.26 and each is significantly different from zero. In the remaining columns of the Table we display the impact of discounted prices on store revenue (Specification 4) and the results from a full division of treatments by WUSF message and price discount (Specification 5).

The results on Table 3 are driven by increases in spending conditional on purchasing from the store. Additional analysis (not reported here) shows no significant impact of WUSF messages on the probability of a consumer placing an order, regardless of the level of aggregation across WUSF messages. This may have occurred because the positive impact of WUSF pledges on some visitors was balanced by other visitors responding negatively to the text-heavy aesthetic of the storefront when a WUSF message was displayed. Among consumers who ordered, the central remaining question is why they had a stronger positive response to the

(untriggered) P+D:δ messages and P:1 message than the relatively rich P:5 message.[22] We suspect that the P:1 message and all three P+D:δ messages acted only to shift consumers' perceptions of the store as an effective way to support the NPO, as if w_1 increased in Section 3's model. The specific details of these WUSF offers were either unnoticed or ignored. The P:5 message, by contrast, was simple and large enough to attract the consumers' attention and encourage them to think about the specific terms of the WUSF offer, which is closer to the model in Section 3 with unchanging preferences for NPO revenue. In the analysis below we provide evidence to support this view.

To explore these ideas further, we now turn to examining order revenue conditional on an order. This approach allows a greater focus on order sizes and requires minimal sacrifice in ignoring the extensive margin. In addition, the stable versus changing preference interpretations offered in Section 3 have their sharpest differences in order sizes conditional on a positive order. Specifications 1-3 of Table 4 replicate the structure of Specifications 1-3 of Table 3, and the qualitative attributes of the results are very similar. Within specific WUSF messages, the P:1, P+D:1, and P+D:10 treatments have significant impacts on order size, and the magnitude of these effects are similar to each other (between $10.50 and $12.50). The impact of the P+D:5 message is positive ($7.05) but not significantly different from zero. The P:5 message stands out with an estimated coefficient that is nearly zero ($0.40). When we divide consumers with P+D:δ messages by whether they made a donation, we find no significant difference in order size between those who did and did not donate.[23]

We verify the robustness of these results in the remainder of Table 4, where we take alternative approaches to the data. In Specifications 4-6 we truncate the data at the 90th percentile rather than the 95th. The estimated coefficients all fall in magnitude, but the standard errors do too as we are eliminating a set of observations that have a substantial impact on the data's variance. The key result is the parameter on P+D:5 displayed in Specification 6. This coefficient takes a value ($5.54) that is statistically significant and now more similar to the P+D:1 and P+D:10 coefficients, which are $7.59 and $7.25, respectively. This pattern is

[22] We show below that the P+D:5 message is more naturally grouped with the other P+D:δ messages than with P:5.

[23] That donating consumers placed slightly smaller orders, on average, reinforces the notion that some consumers experienced a shift in their utility for purchasing from the NPO store. Their increased orders cannot be explained as attempts to meet the terms of the P+D:δ message.

repeated in Specifications 7-9, where we use the log of order value as the dependent variable. Each WUSF message other than P:5 has a positive and significant impact on the log of order size, with the magnitudes ranging from 21.6% to 23.5%. The impact of the P:5 message, by contrast, is estimated at 6.6% and insignificantly different from zero. The differences between Specifications 7-9 and the others on this Table – in the narrower range of coefficient values and the sizes of standard errors – illustrates the impact of outlier transactions on order values.

With Table 4 we established that consumers' responses to the P:5 message are different from their responses to other WUSF messages, which in turn are similar to each other. We now consider more closely what is happening within the P:5 pledge. In the analysis so far, our results suggest that consumers' responses to the P:5 message may be no different than their responses to the Null. This would be surprising in light of consumers' responses to the P:1 message and either of the conjectures in Section 3 (stable or intensified preferences) on how the P:5 message might affect choices.

If consumers are simply ignoring the P:5 message, then we would expect them to have the same probability as Null recipients of falling below the $10 purchasing threshold necessary to trigger the WUSF donation. Overall, a small fraction of store orders (about 3% in total and 6% under the Null) have values below $10. In Specifications 1-3 of Table 5 we present results on the probability that a consumer's order is below $10. The main result is in Specification 3, which shows that consumers assigned to P:5 were significantly more likely to exceed the $10 threshold than those who received the Null. This is consistent with consumers responding to the marginal (financial) incentives of the P:5 message, which have their strongest impact when purchase size would be small under the Null.[24] Consumers who received other WUSF messages respond similarly, which may be due to an overall shift in the marginal utility from NPO revenue (w_1). Given the structure of WUSF offers, these results are closely related to Section 3's predictions for choices at the extensive margin.[25] Similar to the responses of consumers with a P:5 message, there were no consumers who received a P+D:δ message, donated, and failed to reach the $10 merchandise threshold.

[24] Recall from Table 1 that store visitors who received the P:5 message purchased with greater probability than recipients of the Null and most other WUSF messages, which is also predicted by the model with stable preferences.

[25] When we estimate the probability that a store visitor placed an order greater than $10, the effect of each WUSF message is positive but insignificantly different from zero.

In the remaining specifications on Table 5 we continue to investigate whether consumers who received the P:5 message behaved differently than those who received the Null. In additional analysis not reported here, we have found that consumers with WUSF pledges generally increased their orders by increasing the quantity of items purchased. Substitution to higher-priced items was less pronounced.[26] With this in mind, we divide consumers' orders by whether they included any merchandise item with a price below $10. Consumers who are not interested in items priced below $10 have no chance to fall below the threshold for a WUSF donation. Consumers with one or more sub-$10 items may have purchased only a single sub-$10 item in the absence of WUSF pledge.

In this analysis we use the log of order revenue because of the effectiveness of this transformation in reducing outliers' impact, as in Table 4. In Specification 4 we find that a WUSF message added 54%, on average, to the order size of a person who selected a sub-$10 item, and Specifications 5 and 6 show that this impact was fairly similar across WUSF messages including P:5. Turning to the consumers who purchased no sub-$10 items, we find that the P:5 message lags the other WUSF messages. The average impact of all WUSF messages, shown in Specification 7, is smaller (12.2%) than in Specification 4. In Specification 9 we find that the messages other than P:5 have effects that range from 14.4% to 18.5% and are statistically significant, while the effect of the P:5 message is estimated to be -3.9%. The negative point estimate, while not significantly different from zero, is consistent with a crowding-out of x on the intensive margin, which is predicted by the model with stable preferences and $w_{12} < 0$. Taken together, the results of Table 5 add evidence to the view that consumers who received the P:5 message noticed its content and responded in cases when doing so mattered for the WUSF donation, but consumers did not change their actions when there was no (outside) benefit to the NPO from doing so.

[26] Orders under a WUSF pledge contain 0.48 (s.e. 0.18) more items than orders under the Null, but items in WUSF orders have an average price only $0.63 (s.e. 1.16) greater than the Null average.

5.B Price sensitivity

A goal of the experiment was to determine whether the presence of social messages influences price sensitivity.[27] We analyze this issue by testing whether WUSF pledges affected consumers' responses to discounted prices. In order to have our analysis reflect changes in quantities purchased while also accounting for the variety of items offered at the online store, we make an adjustment to the measurement of consumers' choices. Rather than continuing to look at the total (dollar) spending of each consumer, we examine the amount that a consumer *would have spent* if she faced the undiscounted prices for the choices she actually made. This allows us to evaluate whether a consumer buys "more" in dollar-weighted quantity units when she sees lower prices, and thus permits calculation of demand elasticity.

On Table 6 we analyze the impact of the price variation. We present four sets of results. On the left side of the Table we include transactions on all items in the store. On the right side of the Table we include choices on only the items that were included in the experiment's pricing component. This includes about half of all product selections and transactions made during the experiment. Consumers who only purchased items that were excluded from the pricing experiment are treated in the same way as consumers who visited the store and purchased nothing. Similarly, for consumers who selected some never-discounted and some sometimes-discounted items, only the latter items are included. In some models we divide non-Null treatments by whether a P:5 message was displayed, as this is the main factor that distinguishes consumer responses in order revenue, described in Section 5.1.

In the top half of Table 6 we examine purchase probabilities at the session level, which we estimate with a probit model. The analysis of all transactions, on the left, reveals few useful results. The standard errors are large relative to the parameter estimates, and none of the coefficients are significantly different from zero. Specification 1 shows a positive but insignificant increase in store visitors' probability of placing an order. Specifications 2 and 3 also yield similar results – various combinations of discounts and messages had no significant effect on purchase probability. The analysis of sometimes-discounted items, on the right, is more informative. In Specification 4 we report a positive and significant increase in the probability of an order when the consumer observes lower prices. The remaining Specifications

[27] Hiscox and Smyth (2005) study consumers' responses to price changes for products with fair trade labels, and they find that higher prices lead to greater quantity demanded. They attribute this result to consumers giving more credence to fair trade labels when prices are high.

in this panel show that the presence of WUSF messages did not significantly affect order probabilities. The coefficients on the interaction terms of Specifications 4 and 5 are small in absolute value, suggesting little impact of WUSF messages on price sensitivity.

In the lower half of Table 6 we examine the impact of discounts on total spending by session, using undiscounted prices to calculate price-weighted units selected by consumers. We find that when all sales are grouped together regardless of WUSF message (Specification 7), discounts have a significant and positive effect, increasing order size by about 22% per store visitor. The interaction coefficients in Specifications 8 and 9 indicate that the WUSF messages generated a positive but statistically insignificant increase in the price sensitivity of online store customers. Our estimates in the lower right portion of the Table are more precise, as expected, but fail to uncover significant variation in price sensitivity with exposure to WUSF messages. The overall measure of price sensitivity in Specification 10 is positive and significant at $p < 0.01$. This increase of 36.8% in quantity, when paired with an average discount size of 11.4%, implies a demand elasticity of -3.1 for items at the NPO store.[28]

Although the average effects of price discounts, as measured in Table 6, do not uncover significant differences in demand elasticity with and without WUSF messages, these results may be driven by the price sensitivity of a few large customers.[29] We now turn to an informal analysis of how price variation affected the median consumer within an experimental treatment. We focus on choices over products that were included in the discounting portion of the experiment, and we consider pre-discounted order values rather than actual revenue.

There were 57 consumers who purchased potentially-discounted items after receiving the Null and observing undiscounted prices. The median order value was $29.85 for these consumers. Consumers who purchased with discounted prices and the Null message were more numerous (64) and had a greater median order value ($38.93). If each consumer who purchased under the Null was to place an order equal to the median value from his or her respective group,

[28] During discussions with the online store managers we discussed their costs of acquiring and processing items available at the store, and we learned that the estimated demand elasticity implies that current pricing practices approximately satisfy the inverse elasticity pricing rule for profit maximization. Prior to our conversations, the managers were unaware of the rule and had little familiarity with demand elasticity as a method to calculate price sensitivity. The "profit" maximizing pricing strategy is optimal if the store's objective it to maximize revenue that can be transferred from the store to the NPO's charitable efforts.

[29] In contrast to the donation pledges of the WUSF messages, the discounting portion of this experiment shifted financial incentives for both large and small purchases.

then the total order value would increase by 46.4% with discounting. Likewise, the 69 consumers who purchased under P:5 without discounted prices had a median order of $22, while the 64 consumers with a P:5 message and discounts placed a median order of $34.95, which implies that order value would increase by 47.4% if all orders were at the median. By contrast, the sensitivity at the median was lower for consumers who received a WUSF message other than P:5. Regardless of whether they received discounted prices, these consumers placed orders with a median value of $39.95, and the increase in the number of ordering consumers with discounts (from 231 to 288) would yield an increase of 24.7% in order value if these consumers all purchased at the median. In contrast to the results on Table 6, this suggests that charitable messages may reduce price sensitivity among households making relatively small purchases, provided they cause consumers to update their preferences on a charity's quality.

5.C Long-term effects of the experiment

While the WUSF messages led to a substantial increase in store revenue during the experiment, a potential concern is that this increased activity comes at the expense of other support that consumers may offer the NPO. Consumers could substitute intertemporally and reduce future spending at the store after they purchase under a WUSF message, as occurred in Meier's (2007a) study of stimulating direct donations. Consumers may also reduce direct donations to the NPO because they see those donations as a substitute for NPO merchandise.[30] In terms of the model in Section 3, substitution across periods appears more likely if the WUSF message stimulates a temporary demand response under stable preferences. We address this issue by examining consumers' behavior in the year that followed our experimental messages. While it would be ideal to analyze all consumers who viewed an experimental message while visiting the online store, we are limited to consumers who purchased from the store. Our examination of consumers who purchased is aided by the minimal impact of the WUSF messages on order probabilities, so each observed group contains roughly the same percentage of treated consumers (i.e. those who initiated sessions at the store).

Of the 1,328 consumers who purchased under the experiment and were in the lower 95% of orders (divided by type of message or Null), 50 returned to the store in the following year and

[30] Duncan (1999) explores the closely related topic of consumer substitution between donations of time and money, and he uncovers some intrapersonal crowding-out between the two activities.

placed another order. Of these 50, only 4 received the Null message during the experiment. These 4 returning consumers come from a group of 216 who purchased under the Null and were in the lower 95% of orders. Once we add the lower 95% of the supplemental group described in Section 4, we observe a total of 69 consumers who made follow-up purchases from the store, and a total of 2,011 consumers from March and April to consider as potentially-returning customers.

We test for differences between consumers who received WUSF messages and others in Table 7. We begin by examining the probability of returning to purchase. In Specification 1 we separate consumers who are in the Null and supplemental groups, and we group together all consumers who placed an order after receiving a WUSF message. We find that the supplemental group returns at a slightly higher rate than the Null group, but this difference is not significantly different from zero. By contrast, the consumers who purchased with a WUSF message are nearly twice as likely to return to the store and purchase at least once. In Specification 2 we combine the Null and supplemental groups and again find that the consumers who saw a WUSF message are more likely to return to the store and purchase again. Following the division of consumers we employed in Section 5.2, in Specifications 3 and 4 we divide consumers with WUSF messages by whether they received the P:5 version. We find that the consumers who we describe above as having shifted preferences return to the store at a significantly higher rate, while those with P:5 messages do not differ significantly from the Null group.

In the remainder of Table 7 we examine the revenue from consumers who returned to the store. The dependent variable is the sum of an individual's spending during all return visits to the store, and we include the $0 outcomes of consumers who chose not to return. In Specifications 5 and 6 we report that consumers who received a WUSF message spent significantly more than those who purchased under the Null or in the supplemental period. In Specifications 7 and 8 we again divide consumers with WUSF messages by whether they received P:5, and the results reinforce those on the probability of returning to the store. Consumers who received a non-Null message other than P:5 spent significantly more at the store following the experiment, as we might expect if their utility benefit from store purchases, w_1, increased due to exposure to the WUSF message. While consumers who received the P:5 message spent more, on average, than consumers in the Null or supplemental group, the differences are not statistically significant.

6. Conclusions

In our field experiment, we find that consumers respond strongly and positively to products that generate revenue for social causes or public goods. Consumers' responses to these products, however, depend on the details of how this revenue is generated. We find that consumers' demand responses are strongest when the explicit financial benefit to the NPO is small or requires extra actions that the consumers do not take. When the NPO's additional financial benefit was relatively large and simple to trigger, the demand response was nearly zero. The difference in responses appears to be due to consumers becoming sensitive to financial incentives only when doing so is justified by the financial stakes and the consumer's likely actions. When the financial incentives are not salient, consumers instead change their behavior in a way consistent with updated preferences or information about the NPO cause.

The unexpected ordering of consumers' responses echoes the central results from the literature on intrinsic and extrinsic motivation.[31] Intrinsic motivation may be more powerful than extrinsic incentives, but the former may be crowded-out when the latter is introduced. For example, Gneezy and Rustichini (2000) find that workers' motivation can be greater when an activity is unpaid rather than meagerly compensated, and Frey and Oberholzer-Gee (1997) report that Swiss citizens are less willing to live near a nuclear waste facility when they are offered compensation for doing so. In our paper, the preference-stimulating nature of some WUSF messages is the relatively strong effect, and the impact of the WUSF messages is greatly reduced when consumers' attention is drawn to the financial terms of a message.

Our results are also related to recent findings on direct donation decisions, in which the choices of one individual may be sensitive to the nature and extent of other consumers' actions in supporting the same public good. The results in Eckel and Grossman (2003), Karlan and List (2007), and Shang and Croson (2006), described in the Introduction, are consistent with consumers being aware that they hold imperfect information about a charitable organization's or cause's quality, and positive information on others' support leads consumers to update their preferences for supporting the cause. The WUSF pledges in our study may have the same effect. The presence of a P+D:1 pledge, for example, may be relevant to consumers only in that it indicates that some other individuals thought it was worthwhile to create such an offer, which increases the utility value of supporting the NPO.

[31] See Meier (2007b) for a review of this literature.

For for-profit firms considering associations with social causes, our results are clear but perhaps dispiriting for the beneficiary charities. A small donation pledge can stimulate demand more than a large one, and a vague promise of good deeds can be more effective than a specific pledge. These prescriptions, however, depend on how consumers interpret messages about charitable actions. This is an important area for future research. In addition, studies of longer duration and of greater scope would provide useful data on whether the demand stimulation reported here can be replicated as part of a long-term strategy for a firm or nonprofit. While we found only positive or neutral long-run effects from our experimental treatments, our results came from the activity of a relatively small number of consumers who were exposed to a brief demand stimulus.

References

Bagnoli, Mark and Susan G. Watts (2003): "Selling to Socially Responsible Consumers: Competition and the Private Provision of Public Goods." *Journal of Economics and Management Strategy* 12, pp. 419-445.

Bonini, Mendonca, and Oppenheim (2006): "When Social Issues Become Strategic," *The McKinsey Quarterly* (2), pp. 20-32.

Chen, Yan; Xin Li; and Jeffrey K. MacKie-Mason (2006): "Online Fund-Raising Mechanisms: A Field Experiment." *Contributions to Economic Analysis & Policy* Vol. 5 (2), Article 4.

Duncan, Brian (1999): "Modeling Charitable Contributions of Time and Money," *Journal of Public Economics* 72(2), pp. 213-42.

Eckel, Catherine C. and Philip J. Grossman (2003): "Rebates Versus Matching: Does How We Subsidize Charitable Contributions Matter?" *Journal of Public Economics* 87, pp. 681-701.

Elfenbein, Daniel and Brian McManus (2008): "A Greater Price for a Greater Good? Evidence that Consumers Pay More for Charity-Linked Products," Washington University working paper.

Frey, Bruno S. and Felix Oberholzer-Gee (1997): "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out." *American Economic Review* 87(4), pp. 746-55.

Gneezy, Uri and Aldo Rustichini (2000): "Pay Enough or Don't Pay At All," *Quarterly Journal of Economics* 115(3), pp. 791-810.

Hiscox, Michael J. and Nicholas F.B. Smyth (2005): "Is There Consumer Demand for Improved Labor Standards? Evidence from Field Experiments in Social Labeling." Harvard University working paper.

Karlan, Dean and John A. List (2007): "Does Price Matter in Charitable Giving? Evidence from a Large-Scale Natural Field Experiment," *American Economic Review* 97(5), pp. 1774-1793.

Meier, Stephan (2007a): "Do Subsidies Increase Charitable Giving in the Long Run? Matching Donations in a Field Experiment." *Journal of the European Economic Association*, 5(6), pp. 1203-1222.

Meier, Stephan (2007b): "A Survey of Economic Theories and Field Evidence on Pro-Social Behavior" in Bruno S. Frey and Alois Stutzer, eds. *Economics and Psychology: A Promising New Field*. Cambridge: MIT Press. pp. 51-88.

Olsen, G. Douglas; John W. Pracejus; and Normal R. Brown (2003): "When Profit Equals Price: Consumer Confusion about Donation Amounts in Cause-Related Marketing." *Journal of Public Policy and Marketing* 22 (2), pp.170-80.

Popkowski Leszczyc, Peter T. L. and Rothkopf, Michael H. (2006), "Charitable Intent and Bidding in Charity Auctions," University of Alberta working paper.

Sen, Sankar and CB Bhattacharya (2001): "Does Doing Good Always Lead to Doing Better? Consumer Responses to Corporate Social Responsibility," *Journal of Marketing Research* 38, pp. 225-43.

Shang, Jen and Rachel Croson (2006): "Field Experiments in Charitable Contribution: The Impact of Social Influence on the Voluntary Provision of Public Goods." University of Texas at Dallas working paper.

Strahilevitz, Michal and John G. Myers (1998): "Donations to Charity as Purchase Incentives: How Well They Work May Depend on What You Are Trying to Sell." *Journal of Consumer Research* 24, pp. 434-46.

Figure 1

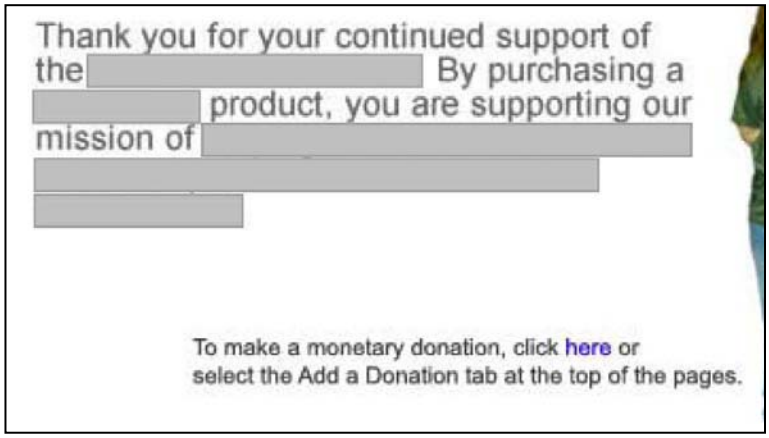

Figure 2

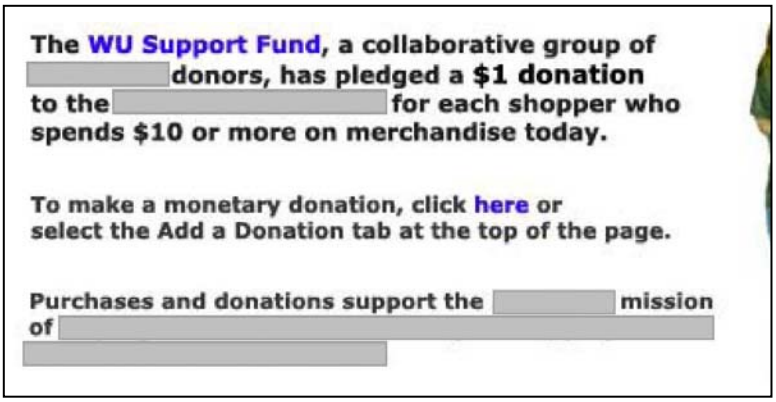

Figure 3

Figure 4

Figure 5

Figure 6
Cumulative distributions of order revenue

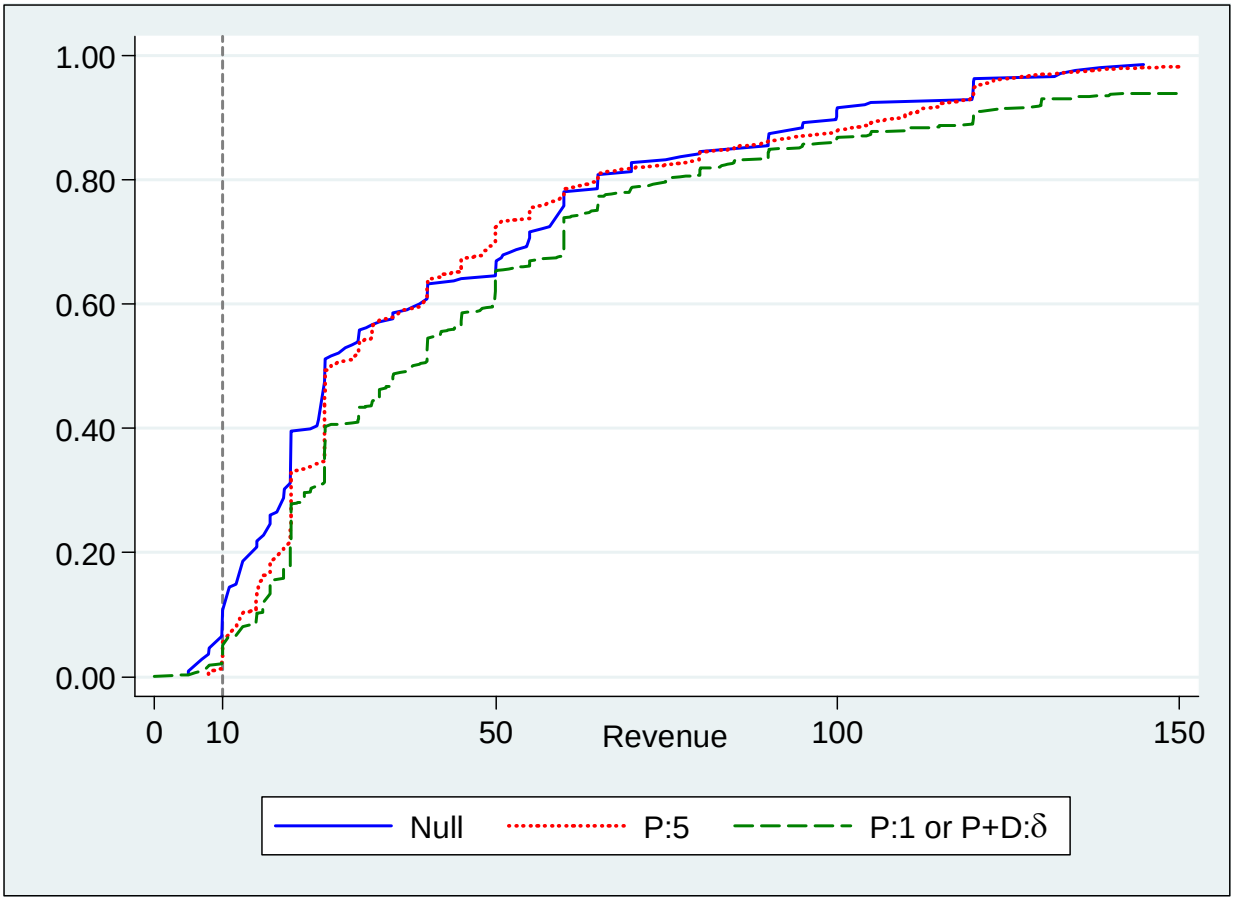

Table 1
Summary statistics

	Number of Visitors	Number of Orders	Purchase Probability	Average Order per Visit	Avg. Order per Visit, Large Orders Excluded	Average Order Revenue	Median Order Revenue	Avg. Order, Large Orders Excluded	Number of Donations
				Panel 1: Activity by message					
Message Type									
Null	11,397	226	1.98%	$1.216	$0.810	$61.31	$28.00	$42.91	2
Any non-Null	58,073	1,169	2.01%	$1.867	$0.968	$92.81	$39.90	$51.50	35
P:1	11,887	232	1.95%	$1.701	$1.030	$87.14	$39.95	$55.36	2
P:5	11,752	243	2.07%	$1.647	$0.852	$79.63	$29.90	$43.32	3
P+D:1	11,457	233	2.03%	$2.529	$1.074	$124.90	$40.00	$55.64	7
P+D:5	11,489	221	1.92%	$1.435	$0.914	$74.62	$40.00	$49.97	7
P+D:10	11,488	240	2.09%	$2.034	$1.062	$97.35	$39.95	$53.46	16
				Panel 2: Activity by pricing treatment					
Discount?									
No	35,111	688	1.96%	$1.671	$0.886	$85.25	$37.90	$47.54	17
Yes	34,359	706	2.05%	$1.851	$1.005	$90.09	$37.00	$51.44	20
				Panel 3: Supplemental data					
Not during experiment	31,388	720	2.30%	$1.467	$0.924	$63.97	$28.00	$42.39	9

Notes: "Large Orders Excluded" statistics are calculated after eliminating orders above the 95th percentile of non-zero orders within the subsample utilized in each row. The supplemental data are from visits to the store in March and April that occurred outside of the experimental period.

Table 2
Post-experiment purchases

	Consumers with orders	Share of subjects	Average total orders	Median total orders
Messages				
Null	4	.0185	$37.08	$39.25
P:1/5	14	.0310	$126.90	$95.26
P+D:1/5/10	32	.0484	$100.66	$43.48
Supplemental data	19	.0277	$52.35	$32.95
Null + Supplement	23	.0256	$49.69	$32.95

Notes: Consumers included in this Table are in the lower 95% of orders during the experiment and supplemental periods. All transactions occurred between 60 and 300 days after the consumer's last action at the store during March and April 2007.

Table 3
Impact of treatments on order revenue ($) by session

Specification	1	2	3	4	5
Constant	.810*** (.085)	.810*** (.085)	.810*** (.085)	.886*** (.048)	.826*** (.122)
Any Message	.163* (.093)				
Any P:δ Message		.123 (.103)			
Any P+D:δ Message		.198** (.098)			
P:1			.220* (.122)		.083 (.170)
P:5			.041 (.122)		-.019 (.172)
P+D:1			.264** (.123)		.092 (.173)
P+D:5			.104 (.123)		.243 (.173)
P+D:10			.252** (.123)		.212 (.173)
Discount				.119* (.068)	-.032 (.174)
Discount × P:1					.251 (.170)
Discount × P:5					.092 (.171)
Discount × P+D:1					.312 * (.173)
Discount × P+D:5					-.314* (.173)
Discount × P+D:10					.050 (.173)

Notes: Omitted category is Null message and No Discount, as appropriate. Total number of observations is 69,401. *** Indicates 1% significance. ** Indicates 5% significance. * Indicates 10% significance.

Table 4
Impact of treatments on order revenue ($) conditional on order

Dependent Variable	Order Revenue ($)			Order Revenue ($)			Log(Order Revenue)		
Conditions	Lower 95% of Orders			Lower 90% of Orders			Lower 95% of Orders		
Specification	1	2	3	4	5	6	7	8	9
Constant	42.92***	42.92***	42.92***	38.10***	38.10***	38.10***	3.427***	3.427***	3.427***
	(2.95)	(2.96)	(3.06)	(2.16)	(2.20)	(2.22)	(.054)	(.054)	(.054)
Any Message	7.94**			5.31**			.190***		
	(3.22)			(2.36)			(.059)		
Any P:δ Message		5.86			4.20			.146**	
		(3.60)			(2.67)			(.065)	
Any P+D:δ Message		9.64***			6.65***			.224***	
		(3.41)			(2.53)			(.062)	
P:1			12.44***			7.83**			.234***
			(4.30)			(3.12)			(.076)
P:5			0.40			0.17			.066
			(4.25)			(3.09)			(.075)
P+D:1			12.47***			7.59**			.216***
			(4.30)			(3.12)			(.076)
P+D:5			7.05			5.54*			.225***
			(4.36)			(3.16)			(.077)
P+D:10			10.54**			7.25**			.235***
			(4.27)			(3.10)			(.075)
N	1,327	1,327	1,327	1,258	1,258	1,258	1,325	1,325	1,325

Notes: Omitted category is Null message. *** Indicates 1% significance. ** Indicates 5% significance. * Indicates 10% significance

Table 5
Further details on order size ($) conditional on order

Dependent Variable	Order is < $10			Log(Order Revenue)			Log(Order Revenue)		
Conditions/Method	Probit Marginal probabilities			Order in lower 95%, includes items below $10			Order in lower 95%, no items below $10		
Specification	1	2	3	4	5	6	7	8	9
Constant				3.000*** (.154)	3.000*** (.154)	3.000*** (.154)	3.510*** (.057)	3.510*** (.057)	3.510*** (.057)
Any Message	-.036*** (.017)			.540*** (.168)			.122** (.062)		
Any P:δ Message		-.026** (.010)			.571*** (.179)			.064 (.069)	
Any P+D:δ Message		-.031*** (.012)			.571*** (.185)			.165** (.065)	
P:1			-.015 (.009)			.498** (.216)			.185** (.080)
P:5			-.021** (.008)			.601*** (.213)			-.039 (.079)
P+D:1			-.015 (.009)			.453** (.217)			.171** (.080)
P+D:5			-.033*** (.007)			.642*** (.225)			.144* (.080)
P+D:10			-.021* (.008)			.526** (.225)			.174** (.079)
N	1,395	1,395	1,395	215	215	215	1,112	1,112	1,112

Notes: Omitted category is Null message. *** Indicates 1% significance. ** Indicates 5% significance. * Indicates 10% significance

Table 6
Price sensitivity and donation pledges

	Probability of purchase: Any item			Probability of purchase: Discounted item		
Specification	1	2	3	4	5	6
Predicted Probability	.020	.020	.020	.012	.012	.012
Any Message		-.0012			.0004	
		(.0021)			(.0016)	
P:5 Message			-.0004			.0020
			(.0026)			(.0022)
Any Message Except P:5			-.0009			-.00002
			(.0021)			(.0016)
Discount	.00095	-.0016	-.0016	.0020**	.0016	.0016
	(.00106)	(.0026)	(.0026)	(.0008)	(.0020)	(.0020)
Discount × Any Msg		.0030			.0005	
		(.0030)			(.0022)	
Discount × P:5 Message			.0026			-.0020
			(.0040)			(.0024)
Discount × Any Msg Except P:5			.0032			.0012
			(.0031)			(.0023)

	Undiscounted Order Value, Lower 95%			Undiscounted Order Value of Discounted Items, Lower 95%		
Specification	7	8	9	10	11	12
Constant	.886***	.826***	.826***	.513***	.432***	.432***
	(.050)	(.127)	(.127)	(.042)	(.115)	(.115)
Any Message		.127			.149	
		(.139)			(.126)	
P:5 Message			-.019			.065
			(.179)			(.162)
Any Message Except P:5			.165			.170
			(.142)			(.128)
Discount	.197***	.028	.028	.189***	.097	.097
	(.072)	(.181)	(.181)	(.060)	(.164)	(.164)
Discount × Any Msg.		.114			.083	
		(.197)			(.179)	
Discount × P:5 Message			.125			.030
			(.253)			(.230)
Discount × Any Msg Except P:5			.112			.097
			(.202)			(.182)

Notes: Omitted category is Null message and No Discount, as appropriate. Total number of observations is 69,470 for probabilities and 69,401 for order values. Probit estimates are marginal probabilities.
*** Indicates 1% significance. ** Indicates 5% significance. * Indicates 10% significance.

Table 7
Post-experiment activity

	Probability of Later Order				Average Revenue from Later Orders			
Specification	1	2	3	4	5	6	7	8
Constant					.687 (1.885)	1.270 (.923)	.686 (1.885)	1.270 (.923)
Predicted Probability	.033	.035	.033	.033				
Supplemental Period	.0134 (.0185)		.0132 (.0182)		.768 (2.162)		.0768 (2.162)	
Any Message	.0255* (.0151)	.0158* (.0079)			3.812* (2.060)	3.228*** (1.242)		
P:5 Message			.0046 (.0217)	-.0052 (.0135)			2.620 (2.616)	2.037 (2.036)
Any Message Except P:5			.0317* (.0177)	.0208** (.0090)			4.122** (2.104)	3.539*** (1.313)
Null combined with supplemental period?	N	Y	N	Y	N	Y	N	Y

Notes: Omitted category is Null message or Null + Supplement, as appropriate. Total number of observations is 2,011. *** Indicates 1% significance. ** Indicates 5% significance. * Indicates 10% significance.