Warm glow, information, and inefficient charitable giving ☆

C. Null *

Department of Global Health, 1518 Clifton Road, Emory University, Atlanta GA 30322, USA

ABSTRACT

More than 200 donors participated in a framed field experiment which consisted of a series of decisions about how to divide a gift between a set of similar charities. Most subjects simultaneously gave to multiple charities with similar missions even when the social benefit of gifts to different charities were not equal, as proxied by the matching rates applied to subjects' gifts. Taking each subject's preferences over the set of charities as given, these choices resulted in substantial inefficiencies: subjects forfeited social surplus (matching funds) equal to 25% of the value of their gifts. Suggestive evidence indicates that warm glow utility derived from the act of making a gift, which can lead to a love of variety even among similar charities, and risk aversion over the social value of charitable gifts are both important factors motivating donors who make socially inefficient gifts. Additionally, few subjects were willing to pay for information that could have enabled them to increase the social benefit of their gifts, although many of these subjects also forfeited potential personal gains in an investment decision, casting some doubt on this interpretation. The possibility that the personal value of information might not be equal to the social value might help explain why there are so few rigorous evaluations of aid programs: such evaluations are costly to charities and might not be valued by donors.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Charitable giving in the U.S. is big business. In 2008, Americans are estimated to have made donations totaling $307 billion, or 2% of GDP (Giving USA Foundation, 2009). However, while the total sum of money being donated is quite large, it is divided across myriad charities in a very disaggregated way by many individual donors. Over 65% of households are estimated to have given to charity in 2008 (Giving USA Foundation, 2009) and a recent random sample of Americans shows that most gifts are relatively small, with two-thirds of the reported gifts less than $100 and a median gift of $50 (The Center on Philanthropy at Indiana University, 2007). There are over 500,000 public charities in the U.S. registered with the IRS, and over 60% of these are small operations with less than $100,000 in annual revenue (National Center for Charitable Statisics at The Urban Institute, 2009). With so many donors and charities and no social planner or market mechanism to solve coordination problems, the potential for inefficient allocations is a serious concern.

Nonetheless, we know relatively little about *how* donors choose which charities to support. We can infer that any donor who simultaneously gives to more than one charity does so because her expected marginal utility from each charity is equal. In the next section I argue that there are two aspects of the donation allocation decision that could lead to equal marginal utilities, even between charities that produce the same public good and thus might otherwise be substitutes. First, a donor might be motivated by something other than the output her gift to a charity produces, deriving private "warm glow" utility directly from the act of making the gift (Andreoni, 1990). Second, charitable contributions can be thought of as credence goods since the donor never knows the true value of her gift in terms of what the charity produced (Darby and Karni, 1973). As in the standard investment decision framework, donors who are risk averse over the social value of contributions might choose a portfolio of charitable giving that has a lower expected productivity in exchange for a reduction in the variance of charitable output.

In observational data it would be impossible to determine if a donor gives to multiple charities because doing so maximizes her warm glow utility or because she is diversifying her charitable portfolio as a result of risk aversion, or because of some combination of the two factors. Similarly, it would be difficult to precisely quantify charities' productivities. To overcome both of these identification challenges, I designed and conducted a framed field experiment which allows me to take each donor's preferences over a set of charities as given, based on her initial allocations.[1] Then, by changing

[1] According to the taxonomy of experiments proposed by Harrison and List (2004) a framed field experiment imposes a set of rules but makes use of field context and a nonstandard subject pool.

the matching rates received by the charities, a proxy for the social benefit of donations, and inducing risk over whether or not the matching funds would be received by the charity, I can identify whether donors are solely motivated by the social benefit of their gifts or whether other considerations such as warm glow and risk aversion lead to inefficient allocations. I focus on forgone matching funds as a quantifiable measure of inefficiency in the experiment: if a donor initially gives to two charities when they have the same matching rate, and then continues to give to both of them in a later scenario even if one of them now receives a much higher matching rate than the other, the magnitude of the inefficiency is defined as the difference in the matching rates times the size of the gift to the charity with the lower rate (the forgone matching funds).

In this modified dictator game, subjects were first asked to divide a gift between a charity they currently supported and a set of three charities which have similar mission statements (CARE, Mercy Corp, and/or Oxfam America). In the second stage of the budgeting process, when they were asked to allocate their gift between the three charities which are arguably substitutes, 70% of the subjects gave money to more than one charity. When one of the charities that a subject supported became exogenously more productive (in the sense that it had a higher matching rate), very few subjects perfectly substituted into that charity, giving it their entire gift. Rather, the majority of subjects weakly substituted, moving a larger share of their total giving into a higher-valued charity but continuing to allocate at least some of their total giving to the now lesser-valued charities. In the process, subjects who did not perfectly substitute contributed to major inefficiencies in the allocation of gifts across charities (taking initial preferences as given), forfeiting matching funds equal to 25% of total un-matched giving. Weak substitution could be consistent with either risk aversion or warm glow, and although the experimental design did not conclusively separate these two motivations, decisions that involve risk over matching rates suggest that both likely play a role in donors' decisions.

An alternate test of how much subjects cared about the social benefit of their gifts indicates that relatively few subjects were willing to pay for information about matching rates that could have enabled them to increase the value of their gifts. When they were told the distribution of matching rates but not how the rates would be assigned to charities, only 40% of subjects were willing to give up a small portion of their endowments in order to find out which charity would receive the highest rate; the rest preferred to allocate their gifts without knowing what they would be worth to the charities. While many of these subjects also chose not to purchase personally profitable information in an investment scenario, casting some doubt on the results about the value of information to donors, this should nonetheless be concerning for those who argue that rigorous evaluations of aid projects are too rare (Duflo, 2004; Savedoff and Levine, 2006). The possibility that donors might place so little value on information about the relative social benefits of different programs could help explain why so few charities are willing to undertake costly evaluations of their projects.

This paper bridges the broad research fields of charitable giving and generosity in experimental dictator games, extending the existing literature by exploring the importance of risk aversion and a more general form of warm glow as factors that influence donors' decisions of which charities to support.[2] In that regard, this research complements

two recent publications that also used dictator games in framed field experiments with unconventional recipients (welfare beneficiaries and diabetic smokers) in order to study dictators' sensitivity to how their gifts would be spent, though neither of these papers tested as rigorously for motivations as I do (Fong, 2007; Jacobsson et al., 2007). Several authors have investigated substitution in charitable giving, but this work has been more frequently concerned with crowding out of private donations by public contributions (Andreoni, 1989; Ribar and Wilhelm, 2002; Andreoni and Payne, 2008) rather than with a single donor's choice between gifts to multiple charities. Reinstein (2006) and (2007) are the only other studies, aside from this one, to consider one individual's substitution patterns between charities, but neither of these papers is able to assess the magnitude of inefficiencies or the likely cause of imperfect substitution as my experiment allows me to do. The experimental nature of my data allows me to probe reasons *why* donors do or do not substitute between charities, and the fact that all of the experimental subjects are donors outside the laboratory strengthens the relevance of their experimental choices.

In the next section, I explain a simple theoretical framework for understanding how donors with different types of preferences will allocate their gifts in the experiment, which is described in Section 3. In Section 4, I present the empirical results. Section 5 discusses limitations of the experimental design and alternative interpretations of the results and Section 6 concludes.

2. Theoretical framework

The basic goal of this study is to characterize donors' preferences over charities which could be considered substitutes in that they produce the same thing. These charities use the donations they receive in order to produce the same public good Y which is an argument of the donor's utility function.[3] Ultimately, a donor chooses the bundle of consumption goods and charitable donations that maximizes her utility, but here I abstract away from the question of private consumption versus charity and focus on only the allocation of donations across charities, as in the second step of a two-stage budgeting process, in keeping with the experimental setup.

Specifically, let $Y = f(g_i + G_{-i})$ where g_i denotes donor i's gift, and gifts from all others are summed as the quantity, G_{-i}. These gifts are then converted into the public good according to the production function $f(\cdot)$. Following Andreoni (1990) we can define a pure altruist as a donor i whose utility function over private consumption c_i and the public good Y is $u(c_i, Y)$, whereas a purely warm glow donor's utility function is $u(c_i, g_i)$ such that the donor cares only about her gift to the charity g_i and not at all about the public good.[4] Warm glow implies that someone else's donation is not a perfect substitute for one's own donation and helps explain the fact that private contributions to public goods are not perfectly crowded out by public contributions, as would be the case if all donors were purely altruistic.[5]

[2] Camerer (2003) provides an overview of dictator games such as those used in laboratory tests of altruism and social preferences by Andreoni and Miller (2002), Charness and Rabin (2002), and Fisman et al. (2007). Eckel and Grossman (2003) find that subjects in dictator games are more generous to charities when contributions are matched rather than subsidized, with important implications for income tax policy. Karlan and List (2007) and Meier (2007) provide evidence from field experiments that donors are responsive to the "price" of their gifts in terms of matching rates. Other field experiments have compared various fundraising mechanisms such as lotteries, challenge grants, and matching grants (Landry et al., 2006; List and Lucking-Reiley, 2002; Rondeau and List, 2008).

[3] For example, CARE, Mercy Corps, and Oxfam America (the charities used in the experiment) all have very similar mission statements and could be considered to "produce" poverty alleviation, which is a public good for everyone who cares about the welfare of the world's poor (i.e. whenever *anyone* makes a gift to one of these charities, *everyone* who cares about the welfare of the poor is better off).

[4] Duncan (2004) proposes an alternative model of "impact philanthropy" in which donors derive utility from personally increasing the level of public good production, drawing on both the altruistic and warm glow theories. As such, impact philanthropy is based on beliefs about others' gifts and the initial endowment of the public good, neither of which are meaningfully different across the various experimental scenarios used to classify preference types in this paper. For this reason, impact philanthropy does not offer useful predictions for choices in the experiment.

[5] In perhaps the most direct possible test, Crumpler and Grossman (2008) designed an experiment in which subjects' contributions to charities perfectly crowded out contributions made by the experimenters. Nonetheless, over half of the experimental subjects chose to contribute from their own experimental payouts, even though they could not increase the net amount the charities received from the experiment. Previously, in another laboratory study, Andreoni (1993) found that private contributions to a public good were not perfectly crowded out by taxes.

For simplicity, assume that each dollar given to charity j by donor i produces m_j units of the public good. While it is likely that the charities' production functions exhibit diminishing marginal returns to donations globally, it is not unreasonable to assume that they are approximately linear over the range of gifts that any one donor would contribute. Clearly, a purely altruistic donor will make gifts only to the charity(ies) with the highest value of m. In contrast, I assume that the marginal warm glow utility of donations to charity j is decreasing.[6] A purely warm glow donor considers only the marginal warm glow utility she would derive from giving to each of the n charities when making her allocation decision, and is likely to make smaller gifts to several charities because the marginal utility of a gift to any single charity is decreasing in the size of the gift. More realistically, a donor who is motivated by both altruism and warm glow will weakly substitute toward a more productive charity but persist in giving to multiple charities even if their marginal public good productivities are not equal, because for such a donor the marginal warm glow utility she derives from multiple gifts outweighs the inefficiencies in public good production that arise from supporting less productive charities.

However, donors do not know the true value of m for any given charity, either because they do not know the shape of the charity's production function or because they have imperfect knowledge about others' gifts and thus do not know where the charity is currently located along its production function. As an alternative to modeling charitable gifts as simply another consumption good, due to uncertainty over what gifts will be worth in terms of public good production the donor's decision is more accurately modeled as the choice of a portfolio of risky assets in which donors maximize their expected utility based on their priors about the distribution of m for each charity. By definition, a donor who is risk averse over the production of the public good will be willing to accept a lower expected level of public good production as long as the variance of public good production is also lower. This reduction in risk can be achieved by giving to charities whose marginal social productivities are not positively correlated.

In the experimental setting, it is possible to exogenously shift the matching rates (productivities) and observe how donors respond. In the notation of the framework above, donors who are risk neutral and purely altruistic will perfectly substitute. Determining whether donors who only weakly substitute do so because of warm glow or risk aversion is not as straightforward. As described in the next section, I use variation in the degree of risk over matching rates to identify risk averse donors, and categorize the warm-glow donors as those who do not perfectly substitute but also do not respond to additional risk.

3. The experiment

The experiment was designed to 1) identify the degree to which decisions about allocations between charities are motivated by a desire to maximize production of the public good versus other factors such as warm glow or risk aversion and 2) quantify the extent of inefficiencies that resulted from weak substitution in an empirical setting. The identification strategy relies on observing each subject's sequence of allocation decisions in which there was exogenous variation in the level and degree of risk over matching rates applied to gifts to the charities. To address concerns about external validity, I recruited members of service clubs (Kiwanis, Lions, and Rotary) to participate in the experiment, so that everyone in the subject pool was known to be a donor in real life.[7]

Prior to the experiment subjects were informed that after the experiment one decision for one subject at the session would be randomly selected to be paid (anonymously) in real money according to the chosen subject's decisions in the experiment. Provisions for protecting the subjects' confidentiality were explained, informed consent was obtained, and then the subjects progressed through the experiment (full text available in Online Appendix A). The author read the instructions aloud but they were also printed on the response forms so that subjects could proceed at their own pace. Subjects recorded their decisions, their response forms were collected, and then the author gave a guest lecture on the topic of the research at the club meeting. There was a manila envelope containing a thank you note for each participant identification number, which subjects could claim prior to departure; the subject whose decisions were randomly selected for payment found the payments in this envelope. More details on the logistics of the experiment are provided in Online Appendix B.

In the first stage of the budgeting process, subjects divided a $100 gift from the author's research grant between their service club and a group of three charities (CARE, Mercy Corps, and Oxfam). Subjects were told that gifts to the charities would be matched by the research grant at a rate of 50¢ per dollar given and that they would be able to specify which of the charities in particular they wanted to give the money to in following decisions.[8,9] Subjects were given the charities' mission statements as listed on the independent charity evaluator CharityNavigator.org and assured that all three of these secular charities had received the highest ranking for being "fiscally responsible and financially healthy".

After having set the total amount that they wanted to give to the charities, subjects were asked to specify how they wanted to divide the gift between the three charities in eight situations with varying matching rates and degrees of risk. As promised, the matching rate was always at least 50¢ and in some cases was more. The variation in matching rates proxied for changes in the charities' productivities, in the sense that a higher matching rate led to a larger gift for the charity and thus increased production of the public good.[10] Specifically, each subject allocated her gift under the following scenarios:

1. Equal $1 matching rates for all three charities.
2. Matching rates of 50¢, $1, and $1.50 (order randomly assigned to the three charities).

3. Matching rates of 50¢, $1.25, and $1 (paired with the random assignment in the prior decision).
4. Equal matching rates as in decision 1 but with 20% risk that each charity would not receive the matching funds (risks across charities were independent, determined by 3 separate rolls of a 10-sided die).
5. Unequal matching rates as in decision 2 but with 20% risk that each charity would not receive the matching funds.
6. Equal matching rates with risk as in decision 4 but with the option to purchase delivery guarantees (i.e. accept a guaranteed but lower matching rate).[11]
7. Unequal matching rates with risk as in decision 5 but with the option to purchase delivery guarantees (i.e. accept a guaranteed but lower matching rate).
8. Matching rates of $1.50 for two charities and $3 for one charity (unspecified assignment of rates to charities); information about which matching rate was assigned to each charity could be purchased for $5 deducted from the total gift – alternately, the subject could opt not to purchase the information and divide her gift without knowing the assignment of matching rates to charities.

Decision 1 defines each subject's set of "preferred charities" (those that the subject gives to when matching rates are the same for all three charities). Decisions 2 and 3 were designed to test whether subjects who initially gave to multiple charities would treat their preferred charities as perfect substitutes when the matching rates changed.[12] Among subjects who did not perfectly substitute, these decisions allow me to check that preferences are consistent with standard utility maximization theory as well as gauge the extent of the inefficiencies that arise when subjects do not perfectly substitute. Decisions 4–7 were designed to explore the degree to which allocations were motivated by risk aversion. In decision 4, subjects who are risk averse will spread their gifts equally between their preferred charities. In decision 5, subjects who are sufficiently risk averse will forfeit higher matching rates in order to allocate their gifts more equally. In decisions 6 and 7, the delivery guarantee option allows sufficiently risk averse subjects to revert to their preferred allocations as in decisions 1 and 2. Finally, decision 8 tests whether subjects care about the social benefit of their gifts.

A follow-up questionnaire elicited basic demographic data (age and sex), and asked subjects if they had ever "heard of/supported with time or money/or had a personal connection to" the three charities prior to the experiment.

4. Results

A total of 205 subjects participated in the study over the course of 16 sessions at Kiwanis, Lions, and Rotary club meetings in the East Bay region of California (details provided in Online Appendix B). Participating clubs ranged in size from fewer than ten to more than thirty members, but were generally demographically similar with the exception of one club which was a college campus affiliate of Rotary and had much younger members than the other clubs. Two-thirds of the 151 subjects who provided demographic information were

male. Ages ranged from 18 to 87 with a mean and median of 52 and 56 years, respectively.

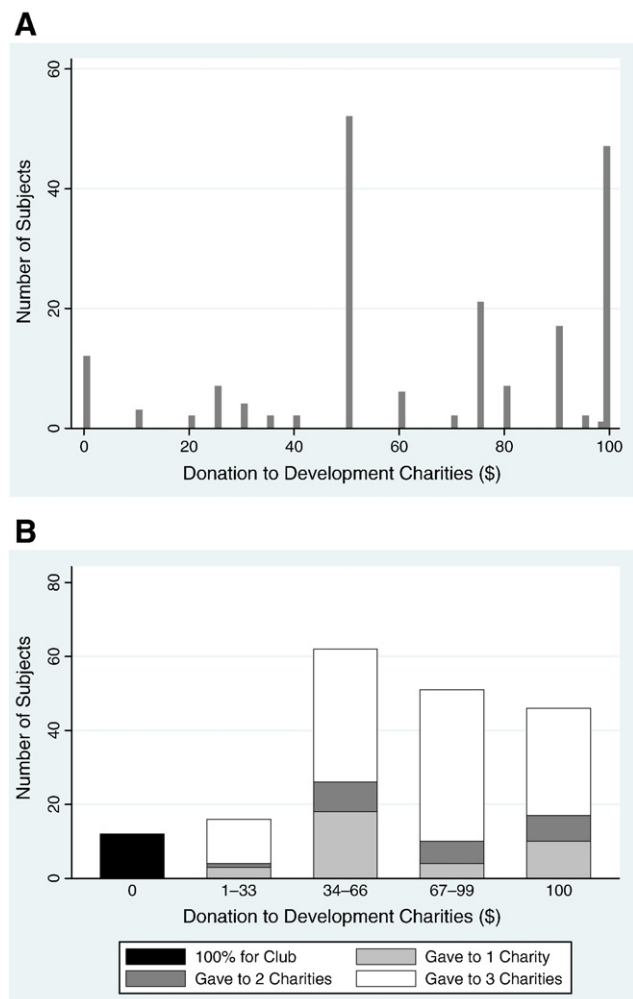### 4.1. Budget first-stage and initial allocations

Subjects in the experiment were quite generous to the group of charities. Panel A of Fig. 1 shows a histogram of gifts to the charities from the first stage of the budgeting process. Around half of the subjects gave either $50 or $100 to the charities, with almost all of the rest choosing gifts that were multiples of either $10 or $25.

Only 35 subjects (17% of the total sample) gave their entire gift to a single charity when the matching rates were equal. As shown in Panel B of Fig. 1, the majority of subjects (70% of the total sample) split their gifts across all three of the charities.

4.2. Responses to changes in matching rates

4.2.1. Substitution

Subjects' responses to changes in matching rates are categorized in Table 1. While perfect substitution was quite rare, weak substitution was quite common. More detail on the extent of substitution is shown

Fig. 1. First stage budget decisions and initial allocations across charities. Notes: Subjects' responses when asked to divide $100 between their club and a set of three charities (CARE, Mercy Corps, and Oxfam America), under the conditions that 1) in the subsequent decisions they would be able to choose which charity(ies) of the three in particular to give the money to and 2) gifts to the charities would be matched at a rate of 50¢ per dollar. The number of charities receiving gifts from a subject is calculated based on responses to decision 1, in which all three charities had the same matching rate. Eighteen of the 205 subjects skipped this section of the experiment.

---

[11] This is essentially an insurance policy, but I did not frame it that way because I was concerned that there might be a social norm of purchasing insurance without considering the financial value as carefully.

[12] Consistent with the theory described above, I assume that subjects experience warm glow from the amount of their donation that they allocate to a charity and not from the donation plus matching rates, in keeping with the assumption that matching rates proxy for productivity and the definition of altruism as utility over the public good. This distinction is perhaps less clear in the context of the experiment since the public good is also measured in dollars, but by definition utility over one's donation plus the attendant matching rates is altruism. Subjects were asked to write down the amount of their donation they wanted to allocate to each charity, not the amount of the donation plus matching funds.

Table 1
Substitution in response to changes in matching rates.

	Summarizing results from decisions 2–3 (unequal matching rates) relative to decision 1 (equal matching rates)
Perfectly substituted	14
Weakly substituted	62
Allocated less efficiently	23
Mixed evidence	6
Never changed allocation	44
Preference for 1 charity	23
Gave only to club	12
Missing data	21
Total number of subjects	205

Notes: In decision 2 matching rates were 50¢, $1, and $1.50 (order randomly assigned to the three charities); in decision 3 matching rates were 50¢, $1.25, and $1 (order paired with the random assignment in the prior decision). Subjects are categorized as having perfectly substituted if in both decisions they moved their entire gift into the one of their preferred charities that had the highest matching rate. Weak substitution occurred when subjects moved some but not all of their gift into a preferred charity with a higher matching rate. Subjects who gave more of their gift to a charity whose matching rate was lower than that of other preferred charities are referred to as having allocated less efficiently. Subjects who perfectly substituted once and allocated less efficiently once are counted as having "mixed evidence" on substitution.

in Fig. 2 which compares the share of each subject's total gift given to a particular charity before and after the matching rates change.[13] In the figure, perfect substitution appears as observations for which the highest-valued preferred charity lies along the $y = 1$ line, i.e. the highest-valued charity received 100% of the subject's gift when matching rates differed between the three charities, whereas observations between the 45° line and the $y = 1$ line depict weak substitution, i.e. the highest-valued preferred charity's gift increased when it was more highly valued relative to decision 1 but did not reach 100% of the subject's gift. In contrast, observations for which the highest-valued charity falls below the 45° line represent subjects who actually reallocated some of their gifts into charities that were *less* valuable when the matching rates were changed, treating the social value the charities produced as complements.[14] Finally, observations along the 45° line correspond to cases in which the subject left the charity's share of the total gift unchanged between decision 1 (equal matching rates) and decisions 2 and 3 (unequal matching rates). As is clear from Fig. 2, the majority of subjects weakly but not perfectly substituted in response to changes in matching rates, with those farther from the $y = 1$ line generating larger inefficiencies. In general, the subjects' responses were quite consistent between decisions 2 and 3.

4.2.2. Constant allocations

Forty-four subjects (21% of the total sample) maintained the exact same allocations even when matching rates changed. There are three possible explanations for this behavior: first, the subjects could have been exerting minimal effort in the experiment and not seriously considering the different scenarios; second, they could have been confused by the experimental protocol; or third, they could have been truly impervious to changes in the matching rates and degree of risk associated with gifts to the three charities (an extreme form of warm

---

[13] Data on charities with the highest matching rates is both necessary and sufficient to show perfect substitution. Weak substitution is sometimes only evident from data on charities with the lowest matching rates since subjects might have only reallocated between the two lesser-valued charities, leaving the share received by the charity with the highest matching rate unchanged.

[14] Alternatively, subjects who appear to be treating the charities as complements (by giving less to the one with the higher matching rate) could also be exhibiting risk aversion if they perceive that there is uncertainty over the process by which gifts and matching funds are converted into the public good. Choices in the decisions with risk over matching rates allow me to distinguish between subjects who view the charities as complements and those who give less to charities with higher matching rates because of risk aversion, as discussed in the next subsection.
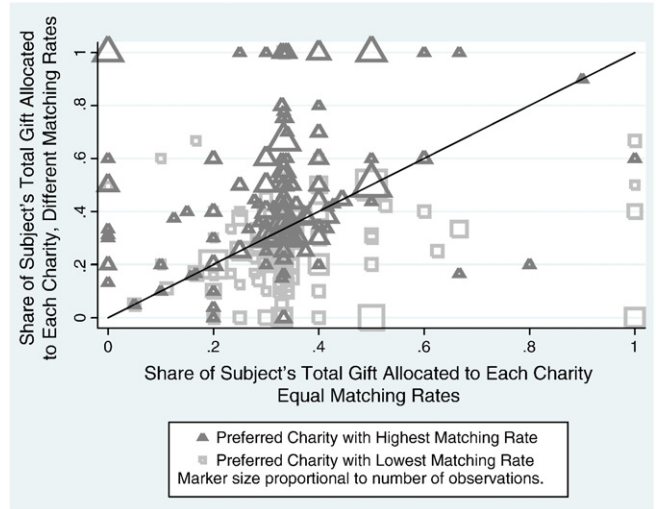
Fig. 2. Substitution across charities in response to changes in matching rates. Notes: Data on allocations from decisions 1 (*x* axis) and 2–3 (*y* axis). Graph excludes observations in which the subject never gave any of the gift to the charity in question or always gave the entire gift to the charity in question. See Section 4.2.1 for a description of the substitution patterns depicted in this figure.

glow). Unfortunately, it is impossible to distinguish between these three hypotheses with the available data, although a robustness check described in Section 4.6 suggests that confusion might be an important factor. Demographic characteristics of these subjects also provide few clues as to why they did not respond to changes in the matching rates (details available from the author upon request).

Regardless of the factors that led these subjects to be unresponsive to changes in matching rates, it is clear that such behavior is inefficient from the point of view of maximizing production of the public good (in this case, matched contributions to the charities). Moreover, it seems quite plausible that in real-world decisions about charitable giving, donors might also be effort-minimizers, confused, or motivated entirely by warm glow. To be conservative, I exclude these 44 subjects when trying to determine why subjects did not perfectly substitute so that my estimates are a lower bound on the degree of warm glow and risk aversion (and the resulting aggregate inefficiencies) in this sample of donors.

4.2.3. Rational choices

Comparing subjects' allocations between the three budget sets, only four subjects' choices violate the Weak Axiom of Revealed Preference (WARP).[15] All other subjects' choices were consistent with WARP so there is no reason to believe the data were generated by a process other than standard utility maximization. Thus, the weak substitution observed in the experiment is a rational choice and can be interpreted as contributing to true inefficiency.

4.2.4. Inefficiency

The relative inelasticity of gifts with respect to changes in matching rates led to large sums of money being left on the table in the form of unclaimed matching funds. Taking as given preferences from the first decision in which the matching rates were equal and excluding subjects who did not respond to changes in matching rates, imperfect substitution resulted in forfeited matching funds ranging from $2.50 to $112.50 per subject, or 25% or more of their un-matched gift to the charities for half of the subjects way changes in matching

---

[15] This is similar to the rate of violations in Andreoni and Miller (2002) but much lower than the rate of violations in Fisman et al. (2007), likely because the former paper observed choices under 8-11 budget sets whereas the latter observed choices under 50 different budget sets.

rates. Summing up the losses over both decisions, the total of over $3600 represents almost 25% of the total un-matched gifts to charities.

4.3. Risk aversion

What causes subjects to substitute weakly but not perfectly? One possibility is that even though observable matching rates changed, subjects were reluctant to invest too heavily in a single charity when they have incomplete knowledge of its effectiveness. The experiment did not attempt to measure subjects' priors over the distributions of the three charities' relative productivities, but the exogenous risk that matching funds would not be delivered to the charities makes it possible to observe how they respond to additional risk. From this we can infer whether or not their reluctance to substitute could be at least partially driven by their original perceptions of the real-world risk associated with the three charities. Specifically, in decision 4, when the matching rates are equal, risk averse subjects will divide their gift as evenly as possible between their set of preferred charities since the risks associated with matching funds for each charity are independent. Subjects who do not adjust their allocations in response to the additional risk are unlikely to have originally split their gifts because of risk aversion.

Unfortunately, decision 4 was not very informative since so many subjects had initially split their gifts equally as shown in Panel A of Table 2.[16] In decision 5, when matching rates differed between charities, there was more scope for subjects to change their allocations, since fewer subjects had split their gifts equally in the absence of risk, but we cannot conclusively infer a lack of risk aversion based on this decision, since even risk averse subjects might not spread their gifts more equally when doing so requires the subject to sacrifice a higher matching rate. Panel B of Table 2 summarizes the results from decisions 4 and 5. While some subjects who did not perfectly substitute between multiple charities in the first three decisions might have been motivated by risk aversion, there is also a group of subjects who show no evidence of risk aversion (or even exhibited risk loving preferences) and who therefore must have chosen to support multiple charities because of warm glow.[17]

For the remaining subjects who did not perfectly substitute but for whom decisions 4 and 5 were uninformative, the delivery guarantee purchase decisions provide an indication of risk aversion regardless of the initial allocation of gifts.[18] Of course, finding that a subject chooses not to purchase a delivery guarantee does not necessarily mean that she is not risk averse, only that she is not sufficiently risk averse for the purchase to be utility maximizing given the cost of the guarantee and the degree of risk over matching funds. With that caveat in mind, based on the decision to purchase delivery guarantees, I find evidence of risk aversion for an additional 11 subjects for whom decisions 4 and 5 were uninformative.

Table 2
Reallocations in response to changes in risk among subjects who gave to multiple charities.

Panel A		
Reallocations across preferred charities:	Decision 4 (risk) relative to decision 1 (no risk) (equal matching rates)	Decision 5 (risk) relative to decision 2 (no risk) (unequal matching rates)
Divided gift more equally	9	19
Divided gift less equally	9	12
No change in how gift was divided	14	36
Gift initially divided equally	54	19
Missing data necessary for comparison	4	4
Total number of subjects (did not perfectly substitute)	91	91

Panel B	
	Summarizing results from decisions 4–5 (risk) relative to decisions 1–2 (no risk)
Evidence of risk aversion	18
Evidence of risk loving	12
Mixed evidence on risk preferences	7
No change in response to risk	11
Uninformative test	38
Missing data necessary for comparison	5
Total number of subjects (did not perfectly substitute)	91

Notes: Includes subjects from Table 1 who weakly substituted, allocated less efficiently, or had mixed evidence. Subjects are counted as risk averse if they reallocated their gifts across preferred charities more equally in at least one of the decisions and did not consolidate gifts in either decision, risk loving is defined as the reverse, and mixed evidence on risk preferences refers to subjects who were risk averse in one decision and risk loving in the other. Subjects categorized as "No change in response to risk" were those who could have allocated more equally in decision 4 when matching rates were equal but did not and also made no changes in their allocations in decision 5 when matching rates were not equal; those who had no capacity to reallocate more equally in decision 4 are counted as having an "Uninformative test" provided that they did not change their allocations in decision 5.

4.4. Willingness to pay for information

The final decision in the experiment investigated whether or not subjects were willing to pay for information about the assignment of matching rates to charities. This choice mimics the real-world trade-off between spending money on programs and spending money on evaluations in order to improve program efficiency.

Altruistic subjects whose donation was at least $20 and gave to all three charities, or whose total gift is greater than $35 and gave to two charities, should find it profitable to purchase the information. As shown in Fig. 3, the majority of subjects met these criteria, but fewer than half actually purchased the information.[19]

Unfortunately, many subjects have missing data on their allocations in decision 8 so it is difficult to say conclusively whether those who purchased the information would have made use of it.[20] Among those who purchased information and did record their allocation decisions, many reallocated money into the charity with the highest matching rate but again fewer than half of them substituted perfectly,

---

[16] To be conservative, I assume that subjects perceived $10 as the smallest increment in which gifts could be divided. For example, I would count a $100 gift that was divided $30/$30/$40 between the three charities as being equally divided. While some subjects did divide gifts in $5 increments and a few even divided in $1 increments, $10 increments were much more common in subjects' allocations.

[17] Among the 14 subjects who perfectly substituted in response to changes in matching rates and therefore appeared to be pure altruists, the extra risk was enough to induce seven (50%) of these subjects to spread their gifts more equally and sacrifice higher matching rates.

[18] In practice, subjects "purchased" the delivery guarantee by agreeing to accept a lower matching rate that was a sure thing. The price of a guarantee against the 20% risk of matching funds not being received was 12.5¢ for each 50¢ of matching funds, so that only sufficiently risk averse subjects should want to purchase a guarantee. A table in the experimental protocol clearly listed the price of the guarantee and the value of a $1 gift with the guarantee and without the guarantee, contingent upon the roll of a die.

[19] Of course, other subjects might also have been willing to purchase the information despite having not previously given to certain charities since the matching rates in this decision were much higher than those in prior decisions. Indeed, six subjects who would not have been predicted to purchase the information, on the basis that they had only given to a single charity in all prior decisions, nonetheless did so.

[20] The prevalence of missing data in this decision is likely due to the logistics of how the information purchase was handled in the experiment. Information about which charity would receive the $3 matching rate was only provided after everyone had made their purchase decisions. After this delay, many of the subjects who had purchased information did not make the effort to record their allocations in response to knowledge about the matching rates.
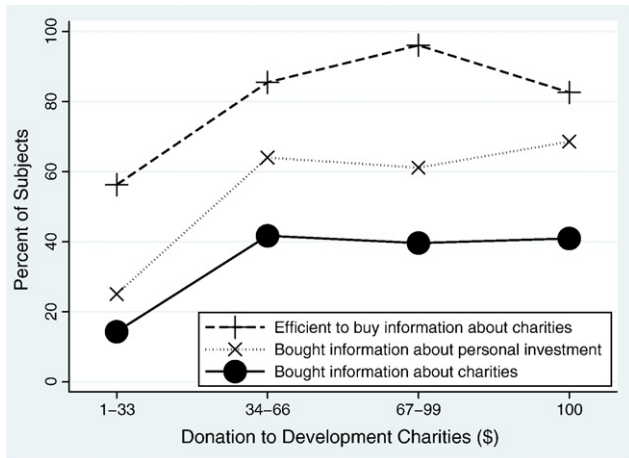
Fig. 3. Decisions to purchase information about pairings of matching rates to charities. Notes: Subjects' responses when asked if they wanted to spend $5 of their gift in order to find out which of the three charities would have a $3 matching rate, knowing that the other two would each receive matching rates of $1.50. "Efficient to buy information about charities" refers to subjects whose gift was at least $20 and gave to all three charities at some point during the experiment or whose gift was at least $35 and gave to two of the charities at some point during the experiment. See Section 4.4 for a description of the personal investment robustness check.

consistent with behavior in earlier decisions, even though the efficiency costs were much larger in this case. The bigger inefficiency in this case comes from subjects who chose not to purchase information. These subjects forfeited matching funds ranging from 30 to 150% of the value of their unmatched gifts, with the median donor sacrificing matching funds exactly equal to the value of her unmatched gift, a truly staggering sum.

Did subjects simply not care about the potential to substitute into the charity with the highest matching rate? Given the evidence from prior decisions, this seems quite plausible and the data are broadly consistent with this possibility. Referring back to substitution patterns in decisions 2 and 3, 70% of the subjects who had perfectly substituted purchased information whereas only 40% of those who weakly substituted did so.

Another possible explanation for why so few subjects purchased the information is that they simply did not understand the cost of not doing so. In order to gauge whether or not subjects could decide "correctly" in the absence of warm glow, a subset of subjects were asked to repeat decision 8 for personal profit, with the opportunity to keep for themselves the money earned from this one decision round.[21] Rather than having the option to give to three charities with different matching rates, subjects were asked to divide the amount of their donation between three generic investment options labeled simply A, B, and C, one of which had a $3 rate of return with the other two earning $1.50 per dollar invested. Again, information regarding the assignment of rates of return to investment options was available for purchase at a cost of $5 (deducted from the investment). This exercise indicates that indeed some subjects may not have understood the consequences of their decision not to purchase information since among those for whom it would have been personally profitable, roughly 40% opted not to purchase information as shown by the middle line in Fig. 3.[22] That said, a third of those who inefficiently chose not purchase information regarding charities' matching rates did purchase information in the investment decision, indicating that they had the ability to assess

whether or not the cost of the information was worthwhile in terms of the ability to earn higher rates of return. This discrepancy is illustrated in Fig. 3 as the vertical distance between the percent of subjects who purchased information about the personal investment and those who purchased information about the charities' matching rates.

Fig. 3 also suggests a concerning pattern in how subjects seem to have approached the information decisions. We would expect the percent of subjects who purchase information to be increasing in the size of their gifts, since the cost of the information was fixed but the financial benefit of the information is greatest for subjects whose gifts are larger; this is also true for the personal investment decision. The fact that these curves are so flat might be an indication that subjects thought only about the fixed cost of information and not how it related to the higher payouts they could earn, either for the charities or themselves.

I conclude that the results of the information purchase decision are not necessarily meaningful for all subjects, since some might not have been able to accurately judge the value of information in terms of the potential to capture higher matching rates. Nonetheless, roughly one-third of subjects who were capable of assessing the value of information inefficiently chose not to purchase, with major consequences in terms of forfeited matching funds. Since information actually reduces risk in this situation, these inefficient choices must result from warm glow. Puzzlingly, there seems to be much higher willingness to pay for a delivery guarantee that eliminates downside risk over matching rates than there is for information that enables subjects to lock in a higher matching rate when there is risk over which charity will be most valuable. Loss aversion and framing effects might help explain some of the difference, but this paradox warrants further investigation in future work.

4.5. Classifying preference types

Unfortunately, the experimental design does not allow me to conclusively determine the relative importance of warm glow versus risk aversion as motivating factors for inefficient giving. On one hand I cannot identify warm glow separately from risk aversion and on the other hand the experimental measures of risk aversion were uninformative for many subjects who had no scope for reallocating their gifts in response to exogenous risk. Nonetheless, by synthesizing the data on substitution patterns and responses to risk it is clear that both are likely relevant characteristics of the subjects' preferences, as shown in Table 3. Bootstrapped confidence intervals for the proportion of subjects of each type (excluding those who maintained constant allocations) indicate that warm glow and risk aversion are roughly equally important in subjects' choices not to perfectly

Table 3
Classifying preference types.

	No evidence of risk aversion	Risk averse
Pure altruist	Perfectly substitutes $n = 14$ [7–21% of sample]	Does not perfectly substitute, responds to risk $n = 42$ [31–51% of sample]
Both altruist and warm glow	Does not perfectly substitute, does not respond to risk $n = 46$ [35–55% of sample]	
Pure warm glow	Maintains constant allocations throughout $n = 44$	

Notes: "Perfectly substitutes" means that the subject gave all of her gift to the one of her preferred charities with the highest matching rate. "Responds to risk" refers to subjects who spread their gifts more equally among their preferred charities when exogenous risk over matching funds was introduced or purchased a delivery guarantee if the initial test was uninformative. Number of subjects in each cell synthesizes data in Tables 1 and 2. Twenty-four subjects were missing data and 35 gave their entire gift either to their club or a single one of the charities. Bootstrapped 95% confidence intervals for the proportion of subjects of a given type are shown in brackets, excluding subjects who gave constant allocations throughout.

---

[21] This question was added to the experiment after the first 5 sessions had been held.
[22] This is almost definitely an upper bound on the fraction of subjects who would make such a poor financial decision in real life. Some of these subjects probably did not read the question carefully, since the experiment was almost over and the wording closely resembled that of the prior question about charities. Another possibility is that this behavior is more evidence of the $\frac{1}{n}$ heuristic.
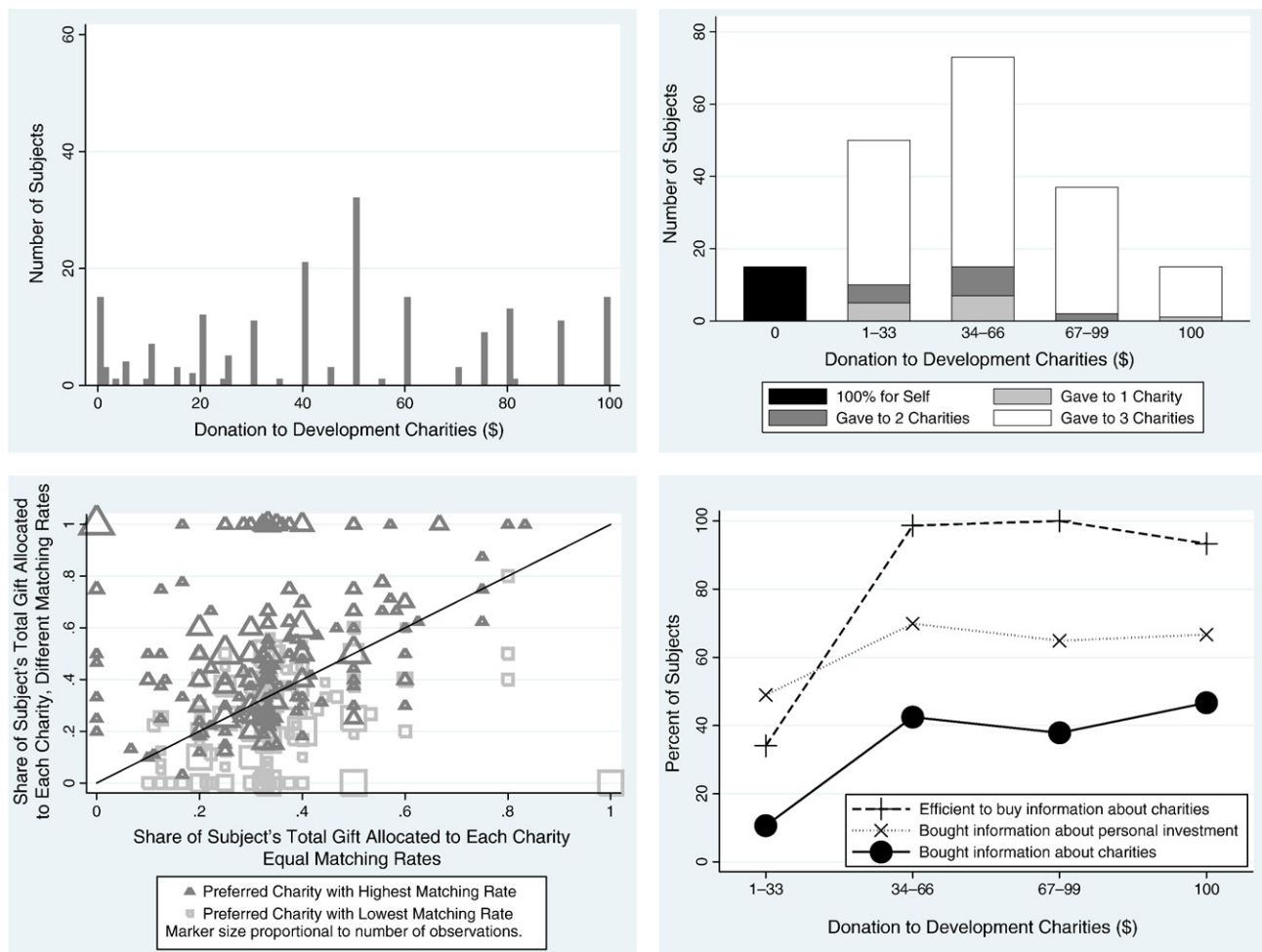
Fig. 4. Results from experienced subjects. Notes: The top two panels correspond to Fig. 1 and depict subjects' responses when asked to divide $100 between oneself and a set of three charities (CARE, Mercy Corps, and Oxfam America). See Figures 2 and 3 for details on the data and interpretations of the lower two panels.

substitute. In terms of both the dollar amounts and the percentages of subjects' gifts that were forfeited in terms of unclaimed matching funds, both motivations contributed equally to the aggregate inefficiencies.

4.6. Results from experienced subjects

The subjects participating in the experiment were all experienced donors, in contrast to most laboratory studies of altruism which rely on experienced subjects, the majority of whom are undergraduate economics and psychology students. One potential problem with the novel subject pool, however, is that they were generally unfamiliar with this sort of study. As a result they might not have fully understood that their decisions could have real monetary consequences or might have been confused about the details of the different decisions. As a robustness check, in this section I present evidence from an additional 191 experienced subjects who participated in the experiment and were unlikely to have been confused by the protocol.[23] In general, the experimental choices made by the two subject pools are remarkably similar, despite major demographic differences.[24]

The only modification to the experimental protocol described in Section 3 and Online Appendix A was that the outside option for the experienced subjects was to keep the gift for oneself, rather than keep it for one's club.[25] Unsurprisingly, experienced subjects showed less generosity toward the charities relative to the experienced donors' allocations, as evidenced by the differences between the upper left panel of Fig. 4 and Panel A of Fig. 1. One difference between the two subject pools' choices is that there was much more heterogeneity in the size of the gifts that the experienced subjects gave to the charities relative to the experienced donors, perhaps indicating that the experienced subjects took a more sophisticated approach to solving the allocation problem. Nonetheless, the experienced subjects were just as likely as the experienced donors to divide their gifts across all three charities in the baseline decision with equal matching rates as shown in the upper right panel of Fig. 4.

The experienced subjects' responses to changes in matching rates were also quite similar to those of the experienced donors, as shown in the lower left panel of Fig. 4.[26] Although some experienced subjects did perfectly substitute into whichever of their preferred charities had the highest matching rate, weak substitution was again much more common. At the same time, experienced subjects were twice as likely as among experienced donors to treat the charities as complements,

---

[23] Professional subjects were recruited through the Xlab at U.C. Berkeley.

[24] The average age among the experienced subjects was 21 and only 43% were male. I did not collect data on race from the experienced donors, but based on my observations at the club meetings, the vast majority were Caucasian, with only a few African, Hispanic, or Asian Americans. Among the experienced subjects 70% identified themselves as Asian American in their Xlab registration records.

[25] In accordance with Xlab policies, each subject was paid a $10 participation fee for the 30-minute experiment in addition to the randomly chosen decision for one subject that was paid in real money.

[26] Again, there were very few (6) subjects whose choices violated WARP.

moving a higher share of their gifts into charities that were less valuable. Like the experienced donors, the experienced subjects' substitution patterns led them to forfeit large sums of matching rates, equal to roughly 30% of the value of un-matched gifts.

A slightly higher share of the experienced subjects who did not perfectly substitute showed evidence of risk aversion based on decisions 4 and 5. As with the experienced donors, this test was uninformative for many experienced subjects, and several additional subjects were identified as risk averse based on their decision to purchase a delivery guarantee. Moreover, there is only weak evidence that delivery guarantees promoted more efficient allocations. Ultimately, combining results from all decision rounds involving risk over matching rates in Table 4, around a third of the experienced subjects who did not perfectly substitute are explained by warm glow, with the remainder showing evidence of risk aversion.

Finally, the lower right panel of Fig. 4 shows that the experienced subjects were no more willing to pay for information than their experienced donor counterparts. Only half of the experienced subjects who could have increased the social benefit of their gifts chose to purchase the information that would have allowed them to do so. Furthermore, although a surprisingly high fraction (32%) of the experienced subjects were not able to correctly assess the value of the information even when they stood to personally profit, half of the experienced subjects who were capable of assessing the value of the information nevertheless chose not to purchase information about charities' matching rates, consistent with the prevalence of warm glow.

In summary, experienced donors and experienced subjects made remarkably similar choices throughout the experiment, mediating concerns that the experienced donors might not have understood the experimental protocol. At the same time, part of the explanation for the consistency of the results between the two subject pools likely lies in the fact that 75% of the experienced subjects said they had given either time or money to at least one charity in the past year; in this regard the difference between the two subject pools is not so much that experienced subjects were less likely to also be experienced donors, but that the experienced subjects' gifts were more likely to be in terms of time rather than money.

5. Discussion

The results presented above are, of course, predicated on a number of assumptions regarding the theoretical model and how it relates to the experimental design. Likewise, certain aspects of subjects' choices in the experiments and other logistical considerations challenge the validity of the empirical results. In this section, I address the most significant of these concerns.

Table 4
Classifying preference types, experienced subjects.

	No evidence of risk aversion	Risk averse
Pure altruist	Perfectly substitutes $n = 22$ [9–21% of sample]	
Both altruist and warm glow	Does not perfectly substitute, does not respond to risk $n = 49$ [25–39% of sample]	Does not perfectly substitute, responds to risk $n = 84$ [45–62% of sample]
Pure warm glow	Maintains constant allocations throughout $n = 14$	

Notes: "Perfectly substitutes" means that the subject gave all of her gift to the one of her preferred charities with the highest matching rate. "Responds to risk" refers to subjects who spread their gifts more equally among their preferred charities when exogenous risk over matching funds was introduced or purchased a delivery guarantee if the initial test was uninformative. One subject was missing data and 21 either kept their entire gift or gave it to a single one of the charities. Bootstrapped 95% confidence intervals for the proportion of subjects of a given type are shown in brackets, excluding subjects who gave constant allocations throughout.

5.1. Assumptions

5.1.1. Is the marginal utility of warm glow decreasing in the size of the gift?

This is the fundamental assumption that drives the prediction that donors motivated by warm glow will want to make smaller gifts to more charities. When considering the validity of this assumption, it is important to remember that warm glow is the utility derived from the act of making the gift, *not* from what the gift produces. As a very crude example, think of warm glow as being equivalent to the number of thank-you letters a donor receives from the charities she supports – a larger gift to any given charity does nothing to increase her utility, whereas dividing her gift between two charities doubles it. *A priori*, I am not aware of any evidence to indicate whether or not we would expect larger gifts to be more satisfying independent of what they produce (including reputational effects as a "product" of a gift). However, given the preponderance of donors who make many yet small gifts both in survey data and the experiment, the assumption that this behavior is motivated at least in part from charity-specific warm glow seems plausible, since it would lead to precisely such behavior. Admittedly, the experimental results in this paper do not prove that warm glow is the source of inefficient allocations, but it does provide a useful framework for understanding why donors might make such choices. Regardless of the specific nature of donors' preferences, this experiment has at least documented the extent of the resulting inefficiencies and documented that access to information is not sufficient to improve efficiency.

5.1.2. Are matching rates in the experiment a valid proxy for productivity?

No explicit connection was made to suggest that subjects should consider the matching rates to be a proxy for charities' productivities, so it is possible that these results are a lower-bound, worst-case estimate of the extent of inefficient substitution. Perhaps if the decisions had been framed as a choice between charities with different productivities, rather than matching rates, subjects would not have made such inefficient choices. That said, the way in which matching rates would increase their contributions to the charities *was* clearly explained, and this is the relevant quantity in terms of donors' allocation decisions. Fundamentally, it does not matter how subjects interpreted the matching rates — the point is that they could have sent more money to the charities by exploiting the matching rates and yet they did not in many cases.

5.1.3. Is weak substitution inefficient?

The analysis is based on a very specific definition of inefficiency, namely that subjects did not maximize the total sums received by the charities, taking each individual's preferences as given. While I maintain that perfect substitution is the only socially optimal solution under these conditions, there could certainly be cases in the real world in which it would not be, for example if there is stochastic variation in any element of the charity's production function and the charity faces adjustment costs. Such situations could not plausibly justify imperfect substitution in the experimental set-up, however.

5.2. Empirical results

5.2.1. Weak incentives

Did subjects in the experiment take their decisions seriously, recognizing that they might have real monetary consequences? One cause for concern is that not all chosen experienced donors actually claimed the manila envelope with the donation checks that were issued based on their decisions. In addition, the inclination to divide gifts only in $10 increments could indicate lack of effort; this also compromises the power of the test as there is much less variation in the data. Finally, over 80% of the experienced donors who gave to

more than one charity divided their gift equally between whichever charities they supported, consistent with use of the $\frac{1}{n}$ heuristic that has been observed in other contexts, such as retirement fund portfolio choice (Benartzi and Thaler, 2001).

The fact that such a large share of the experienced donors maintained constant allocations throughout the experiment is likely an indication that the incentives were, indeed, too weak. From an experimental point of view, the most important difference between the two subject pools is the fact that this behavior was much less common among the experienced subjects (see the bottom row of Tables 3 and 4), perhaps because they were more convinced that their decisions could have real monetary consequences, even though the odds of that happening were lower since the experiments with experienced subjects were conducted in larger groups. This finding suggests that some of the experienced donors who adopted this behavior might have done so because of confusion or laziness, strengthening the argument for excluding such subjects from the analysis, although it introduces an important source of selection bias. The fact that when those who maintained constant allocations are excluded, choices were so consistent between the two subject pools bolsters the credibility of the data from the experienced donors, since the experienced subjects did appear to be better incentivized to take the decisions seriously. Nonetheless, whether or not those subjects who changed their allocations were truly optimizing is another question. This is another way in which the experimental results might overestimate the inefficiencies in real-world charitable giving.

5.2.2. Confidentiality

Despite emphasizing provisions to protect subjects' confidentiality at the beginning of the experiment and immediately prior to announcing which decisions would be paid, experienced donors whose decisions were chosen for payout frequently identified themselves at the conclusion of the experiment. This logistical aberration could not have affected the results, however, since at the time that subjects were making their choices, no one had yet identified themselves nor did they realize there would be the opportunity to do so, such that the experimental choices can be considered anonymous (the experiment was not a repeated game).

5.2.3. Challenges with identifying risk aversion

As has already been noted, there was a very narrow margin for detecting risk aversion since so many subjects had already divided their gifts as equally as possible. In addition, there were a number of cases in which choices were surprising (high numbers of subjects who appeared to be risk loving), internally inconsistent (subjects who exhibited evidence of both risk loving and risk aversion in decisions 4 and 5), or contradicted the theory (purchased delivery guarantees but then actually divided their gifts less efficiently than in the initial decisions without risk or subjects who did not purchase delivery guarantees but still changed their allocations in decisions 6 and 7 relative to 4 and 5, despite the fact that without the guarantee those two sets of decisions were identical). Some of these choices, which are problematic for the analysis, likely resulted from presenting subjects with too many different scenarios given the time constraints of the experimental set-up. Finally, there is an interesting discrepancy between the demand for delivery guarantees, which was quite high at a purchase rate of 65%, versus the low willingness to pay for information that could have allowed subjects to lock in a higher matching rate. Loss aversion and simple marketing (use of the word "guarantee") could help to explain this seeming inconsistency, but the issue warrants further research. Experiments that investigate risk preferences over charitable giving, similar to Chakravarty et al.'s (forthcoming) tests of risk preferences over gifts to other people, could be a fruitful area for research given the importance of risk aversion identified by the experiment described in this paper. In the absence of more powerful tests, and better evidence that subjects really understood the parameters of the decisions they were making, the results on risk aversion in this paper must be viewed in the context of the experiment's limitations.

6. Conclusions

In summary, warm glow and risk aversion both seem to be important determinants of the inefficiencies that arise from donors' choices about which charities to support. In the framed field experiment described in this paper, very few subjects perfectly substituted to whichever of their preferred charities had the highest social benefit, as proxied by matching rates applied to gifts. Specifically, when one of the charities they already supported became exogenously more productive in the experiment, subjects continued to give to lesser-valued charities, forfeiting social surplus (matching funds) equal to 25% of the value of their gifts. If these choices are indicative of how donors would respond to changes in the social benefit of their gifts in the real world, the results bode ill for efficiency of resource allocation across charities.

The experiment also sheds light on one potential reason why there are so few evaluations of aid projects in the real world. Just as half of the subjects in the experiment were unwilling to pay for information that could have increased the social benefit of their gifts, evaluations might be rare in part because they are costly to charities and largely ignored by donors. Moreover, the fact that those who made larger gifts were no more likely to purchase information suggests that subjects focused on the fixed cost of acquiring information rather than the benefits of knowing which programs are most effective.

Having established that warm glow and risk aversion influence donors' decisions about their charitable giving, we can begin to design and test policies that could help mitigate the inefficiencies that result from these aspects of donors' preferences. Unfortunately, it will be difficult to create policies that can mitigate warm glow inefficiencies across charities without also having unintended consequences for the level of giving, since warm glow can affect both the size of the pie as well as how it is divided. There is more hope for addressing inefficiencies from risk aversion, however, since many subjects in the experiment were willing to pay for an intervention that eliminated experimental risk over matching rates. Although such extreme guarantees are infeasible in the real world, information from charity evaluators such as CharityNavigator.org, the Better Business Bureau's Wise Giving Alliance, and the revised IRS form 990 (filed by tax-exempt organizations) could all be viewed as forms of "insurance", even if there is no way to entirely avoid the uncertainty surrounding the social benefit of charitable donations.

This paper has taken the first step of highlighting several ways in which the classical model of a purely altruistic donor who contributes to a single public good fails to predict donors' behavior in the real world, when faced with multiple charities to support and uncertainty over the social value of one's gift. Incorporating warm glow and risk aversion into our understanding of how donors choose which charities to support will make it possible to identify policies and institutions that can promote efficient resource allocation in this sector of the economy.

Acknowledgments

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jpubeco.2010.06.018.

References

Andreoni, J., 1989. Giving with impure altruism: applications to charity and Ricardian equivalence. Journal of Political Economy 97 (6), 1447–1458.

Andreoni, J., 1990. Impure altruism and donations to public goods: A theory of warm-glow giving. The Economic Journal 100 (401), 464–477.

Andreoni, J., 1993. An experimental test of the public-goods crowding-out hypothesis. The American Economic Review 83 (5), 1317–1327.

Andreoni, J., Miller, J., 2002. Giving according to garp: an experimental test of the consistency of preferences for altruism. Econometrica 70 (2), 737–753.

Andreoni, J., Payne, A., 2008. Crowding out both sides of the philanthropy market: Evidence from a panel of charities. Working Paper.

Benartzi, S., Thaler, R., 2001. Naive diversification strategies in defined contribution saving plans. The American Economic Review 91 (1), 79–98.

Camerer, C., 2003. Behavioral Game Theory: Experiments in Strategic Interaction. Princeton University Press.

Chakravarty, S., Harrison, G., Haruvy, E., Rutström, E., forthcoming. Are you risk averse over other people's money? Southern Economic Journal.

Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. The Quarterly Journal of Economics 117 (3), 817–869.

Crumpler, H., Grossman, P., 2008. An experimental test of warm glow giving. Journal of Public Economics 92, 1011–1021.

Darby, M., Karni, E., 1973. Free competition and the optimal amount of fraud. Journal of Law and Economics 16 (1), 67–88.

Duflo, E., 2004. Scaling up and evaluation. Annual World Bank Conference on Development Economics 2004: Accelerating Development 1(30228).

Duncan, B., 2004. A theory of impact philanthropy. Journal of Public Economics 88, 2159–2180.

Eckel, C., Grossman, P., 2003. Rebate versus matching: does how we subsidize charitable contributions matter? Journal of Public Economics 87, 681–701.

Fisman, R., Kariv, S., Markovits, D., 2007. Individual preferences for giving. The American Economic Review 97 (5), 1858–1876.

Fong, C., 2007. Evidence from an experiment on charity to welfare recipients: Reciprocity, altruism and the empathic responsiveness hypothesis. The Economic Journal 117, 1008–1024 July.

Giving USA Foundation, 2009. Giving USA 2009: The annual report on philanthropy for the year 2008. researched and written by the Center on Philanthropy at Indiana University. Indianapolis, IN.

Harrison, G., List, J., 2004. Field experiments. Journal of Economic Literature 42 (4), 1009–1055.

Jacobsson, F., Johannesson, M., Borgquist, L., 2007. Is altruism paternalistic? The Economic Journal 117, 761–781 April.

Karlan, D., List, J., 2007. Does price matter in charitable giving? Evidence from a large-scale natural field experiment. The American Economic Review 97 (5), 1774–1793.

Landry, C., Lange, A., List, J., Price, M., Rupp, N., 2006. Toward an understanding of the economics of charity: evidence from a field experiment. Quarterly Journal of Economics 121 (2), 747–782.

List, J., Lucking-Reiley, D., 2002. The effects of seed money and refunds on charitable giving: experimental evidence from a university capital campaign. Journal of Political Economy 110, 215–233.

Meier, S., 2007. Do subsidies increase charitable giving in the long-run? matching donations in a field experiment. Boston Federal Reserve Board Working Paper.

National Center for Charitable Statisics at The Urban Institute, 2009. Registered nonprofit organizations by level of total revenue. URL http://nccsdataweb.urban.org.

Reinstein, D., 2006. Does one contribution come at the expense of another? empirical evidence on substitution between charitable donations. University of Essex Economics Discussion Papers 618.

Reinstein, D., 2007. Substitution between (and motivations for) charitable contributions: An experimental study. University of Essex Economics Discussion Papers 648.

Ribar, D., Wilhelm, M., 2002. Altruistic and joy-of-giving motivations in charitable behavior. Journal of Political Economy 110 (2), 425–457.

Rondeau, D., List, J., 2008. Matching and challenge gifts to charity: evidence from laboratory and natural field experiments. Experimental Economics 11, 253–267.

Savedoff, W., Levine, R., 2006. Learning from development: the case for an international council to catalyze independent impact evaluations of social sector interventions. Center for Global Development Brief May.

The Center on Philanthropy at Indiana University, November 2007. American express charitable gift survey, available at http://home3.americanexpress.com/corp/pc/2007/pdf/aecgs.pdf.