NBER WORKING PAPER SERIES

TESTING MOTIVES FOR CHARITABLE GIVING:
A REVEALED-PREFERENCE METHODOLOGY WITH EXPERIMENTAL EVIDENCE

Rahul Deb
Robert S. Gazzale
Matthew J. Kotchen

Working Paper 18029
http://www.nber.org/papers/w18029

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2012

Testing Motives for Charitable Giving: A Revealed-Preference Methodology with Experimental Evidence
Rahul Deb, Robert S. Gazzale, and Matthew J. Kotchen
NBER Working Paper No. 18029
May 2012
JEL No. C91,D01,D64,H41

ABSTRACT

A large economics literature seeks to understand the reasons why individuals make charitable contributions. Fundamental features of most models of charitable giving are the inclusion of externalities induced by other agents and the Lancasterian characteristics approach to specifying utility functions. This paper develops a general, revealed-preference methodology for testing a variety of preference structures that allow for both externalities and characteristics. The tests are simple linear programs that are transparent, computationally efficient, and straightforward to implement. We show how the technique applies to standard models of privately provided public goods and novel models that account for social comparisons based on relative consumption and donations among individuals. We also conduct an original experiment that enables nonparametric tests of many models on a single data set. The results provide the first revealed-preference evidence on the importance of social comparisons when individuals make charitable contributions. Models that include preferences for either relative consumption or donations yield greater explanatory power than the standard model of impure altruism.

Rahul Deb
Department of Economics
University of Toronto
150 St George St, Toronto, ON, Canada
rahul.deb@utoronto.ca

Robert S. Gazzale
Department of Economics
Williams College
24 Hopkins Hall Dr
Williamstown, MA 01267
rgazzale@williams.edu

Matthew J. Kotchen
School of Forestry & Environmental Studies,
School of Management,
and Department of Economics
Yale University
195 Prospect Street
New Haven, CT 06511
and NBER
matthew.kotchen@yale.edu

1 Introduction

There exists a large economics literature that seeks to understand why individuals make charitable contributions. One explanation, based on the standard model of private provision of a pure public good, assumes that individuals benefit from the aggregate level of a public good and may thus have an incentive to make voluntary contributions. The pure public good model has well-known predictions about the crowding out of donations and pervasive free-riding.[1] More recent research attempts to reconcile the standard model's stark predictions with the observation that people are typically more generous than the model predicts. Many studies assume that individuals obtain a private benefit from some aspect of their own giving, and this encourages donations beyond that which would occur based on public benefits alone. Different interpretations of the private benefit range from a feeling of "warm-glow" satisfaction, social approval, prestige and signalling about income.[2]

In this paper, we develop a revealed-preference methodology for testing general models of charitable giving and provide experimental evidence on different motives driving actual donations. Laboratory experiments are an increasingly common way to test for specific motives that underlie charitable behavior. Vesterlund (2006) describes how the control that experimental methods afford the researcher has broadened the scope of empirical analysis beyond studies of crowding out to consider social norms, rules, and different ways of accounting for others' behavior.[3] Moreover, Vesterlund (2012) argues that "the objective is no longer to determine whether individuals are selfish or cooperative, but instead whether giving can be viewed as rational, and if so what set of preferences are consistent with the observed pattern of giving" (p. 2).

One advantage of laboratory experiments in particular is the relative ease of employing revealed-preference methods. Andreoni and Miller (2002) were the first to use revealed preferences to test for a particular form of altruism. They consider individual preferences of the form $U_i(x_i, y_i)$, where x_i and y_i are payoffs in a dictator game for oneself and another anonymous subject. After asking subjects to make repeated choices with varying endowments and relative prices, they find that the specified utility function, which is considered altruistic because it accounts for another's payoff, rationalizes the vast majority of subject behavior.

[1] A useful general reference for the standard model is Bergstrom, Blume and Varian (1986), while other papers provide more detailed studies of the model's important properties with respect to crowding out (Warr 1982; Roberts 1984) and free-riding (Andreoni 1988).

[2] Many of these models can be characterized broadly as refinements or extensions on the impure public good model (Cornes and Sandler 1984), which is based on joint production of private and public benefits. References on the specific examples listed are Andreoni (1989, 1990), Hollander (1990), Harbaugh (1998), and Glazer and Konrad (1986).

[3] Experimental studies are not limited to the laboratory, however, and recent field experiments have illuminated several important aspects of the market for charitable giving (see List 2011).

While revealed-preference methods have the distinct advantage of allowing nonparametric tests of different preference structures, researchers seeking to use them on a broader set of motives for charitable behavior must confront an additional set of challenges. These arise because classical revealed-preference techniques do not readily accommodate two features that are central to most theoretical models of charitable giving: externalities among agents and Lancasterian (1971) characteristics within utility functions. Consider the model of impure altruism (Andreoni 1989, 1990) with preferences $U_i(x_i, y_i + Y_{-i}, y_i)$, where y_i is one's charitable contribution and the additional term Y_{-i} is the sum of contributions by all others. This is essentially a characteristics model because y_i enters the utility function in two places: the second argument as a contribution to the public good and the third argument as the private, warm-glow benefit. The existence of a public-good externality from giving, through Y_{-i}, is also fundamental to the model and enters in only one of the characteristics.

To address these challenges, the first contribution of this paper is a general, revealed-preference methodology for testing a variety of preference structures that may underlie charitable giving. The approach builds on recent innovations in revealed-preference theory that allow for the incorporation of externalities (Carvajal 2010; Deb 2009) and characteristics (Blow, Browning, and Crawford 2008) into a standard model of consumption. Empirically, the tests are simple linear programs that are transparent, computationally efficient, and straightforward to implement. While fully nonparametric and thus eliminating the need for many ad-hoc assumptions about functional form, the approach allows for complete heterogeneity across agents because the analysis is conducted separately upon the repeated choices of each individual.

The second contribution of the paper is consideration of several existing and novel models of charitable giving within our revealed-preference framework. We show how the tests apply to standard models of warm-glow giving, pure altruism, and impure altruism. Through these cases, it is clear how the framework can readily accommodate models with externalities and possibly joint production of private and public benefits from individual contributions. A recent paper by Korenok, Millner and Rozzolini (2011) focuses on impure altruism and employs a related approach based on the observation that artificially constructed prices can be used to test the model. While their analysis has some features in common with ours, as we will explain, the revealed-preference methodology developed here has meaningful computational advantages and, more importantly, a level of generality that admits a range of preferences beyond impure altruism. The methodology thus provides a tool that should be useful for researchers employing revealed-preference analysis even beyond that considered in the present paper.

The other preference structures that we do consider, however, account for social comparisons when individuals make decisions about charitable giving. In one case, we specify a model in which individuals are motivated to donate because of concerns about how their own donation compares to that of others. Shang and Croson (2006, 2009) report the results of several field experiments that include social comparisons, and they find that donors to a public radio station tend to adjust their contribution levels toward that of the social comparison. Our analysis is complementary in that we provide a close link to theory and show how tests for the importance of social comparisons can exploit multiple choices of the same individual rather than a cross section among individuals. We also consider an alternative model in which the relevant social comparison for charitable giving is concern about relative private consumption. The model of relative consumption is distinct in our case because it implicitly takes account of relative wealth when individuals interpret comparative donation levels. In effect, this model specifies preferences that depend on the "philanthropy" of others, as opposed to solely the absolute amount donated. There is, of course, a large literature on the importance of relative consumption for explaining consumer behavior, but we are not aware of any other study that considers a model of relative consumption to explain charitable giving.[4]

The third contribution of this paper is experimental results that demonstrate the applicability of our revealed-preference framework and highlight the importance of social comparisons for understanding charitable contributions. A distinct feature of our experimental design is that several models are testable on a single data set. Subjects in a laboratory setting face allocation choices based on the division of tokens between themselves and a charity. Through a series of choices for each subject, we vary the endowment of tokens and the value per token for private consumption and charitable giving. The experimental design differs from that of Andreoni and Miller (2002) and Korenok *et al.* (2011) because a small, local, non-profit organization receives the value of donated tokens, as opposed to another anonymous subject in the lab.[5] More fundamental to our experimental design, however, is that the subjects of primary interest were informed of the choices made by others in an earlier round when faced with the same token endowment and relative prices. This simple design allows both crowding out and social comparisons to affect subject choices, thereby enabling revealed-preference tests of the models described previously and combinations thereof.

The experimental results provide new evidence on the importance of social comparisons

[4] References on the importance of relative consumption to consumer behavior include the classic work of Veblen (1899), the more formal theory of Duesenberry (1949), and the popular writings of Frank (1985, 1999).

[5] Another study based on revealed-preference analysis in a dictator game is Fisman, Kariv and Markovits (2007). They use a graphical interface to replicate the results of Andreoni and Miller (2002) and conduct a further experiment to distinguish preferences for giving from preferences that account for trade-offs between the payoffs of others.

to explain charitable giving. As a benchmark, and after making adjustments to control for differences in the power of revealed-preference tests, we find that the standard model of impure altruism rationalizes the choices of 72 percent of our subjects. The model performs substantially better than the special cases of warm-glow giving (50 percent) and altruism consistent with provision of a pure public good (58 percent). But impure altruism performs less well than models based on concerns about relative donations (80 percent) and relative private consumption (81 percent). Empirically, these results, along with robustness checks that we discuss, provide the first revealed-preference evidence on the importance of social comparisons to the understanding of charitable giving.

2 Theoretical Framework

This section develops our theoretical framework in two steps. We begin with specification of a general utility function that nests different models for private provision of a public good, including the standard models and novel ones that account for relative preferences. We then illustrate how Lancasterian characteristics and externalities, both of which are fundamental to the models we consider, complicate revealed-preference analysis. Finally, we establish a theorem that enables revealed-preference tests of any model based on preferences that satisfy properties of the general utility function.

2.1 The Utility Function

There are $i = 1, ..., N$ agents in the economy. Each agent is endowed with wealth w_i that can be divided between consumption of a private good x_i and donations to a public good y_i. Prices are denoted p_x and p_y. We define the vectors $\mathbf{x} = (x_1, \ldots, x_N)$ and $\mathbf{y} = (y_1, \ldots, y_N)$, along with $\mathbf{x}_{-i}$ and $\mathbf{y}_{-i}$ equal to the respective vector excluding the element for agent i. Capital letters denote sums such that $X = \sum_i x_i$ and $X_{-i} = \sum_{j \neq i} x_j$, with Y and Y_{-i} defined analogously.

We consider preferences of the general form

$$U_i \left[x_i,\ c_i(x_i, \mathbf{x}_{-i}),\ y_i + Y_{-i},\ d_i(y_i, \mathbf{y}_{-i}) \right], \tag{1}$$

where the utility function is concave and weakly increasing in all four arguments, but strictly increasing in x_i and y_i. The functions $c_i(\cdot, \cdot)$ and $d_i(\cdot, \cdot)$ are assumed to be concave in their first argument (x_i and y_i, respectively) and continuous in all arguments. An important feature of the utility function is that both consumption and donation have multiple nonlinear characteristics. The quantities x_i and y_i provide the agent utility through the amount of private consumption and the total level of public good provision, respectively. These same quantities also provide utility to agent i through the functions $c_i(\cdot,\cdot)$ and $d_i(\cdot,\cdot)$, which also depend on the corresponding quantities $\mathbf{x}_{-i}$ and $\mathbf{y}_{-i}$ for all other agents in the economy.

Notice that the utility function in (1) is a special case of the more general specification $U_i[x_i, \mathbf{x}_{-i}, y_i, \mathbf{y}_{-i}]$. While the more general specification could, in principle, be the starting point for our analysis, it is known that such general preferences impose only trivial restrictions for revealed-preference tests in models with externalities (Carvajal 2010; Deb 2009). Intuitively, this occurs because changes in the level of externalities create a large degree of flexibility in utility functions to rationalize revealed preferences. By adopting the utility function in (1), however, we impose structure on the agents' preferences that is both reasonable and intuitive in the context of private provision of a public good. This, as we will show, allows us to consider a class of preferences that are general enough to subsume several existing and novel models while simultaneously imposing nontrivial restrictions to test on observed data. Moreover, the specific cases that we consider demonstrate how the framework remains general enough to account for multiple externalities and multiple characteristics arising from x_i and y_i.

We begin with the standard models of privately provided public goods. Keeping only the first and third arguments of (1) yields preferences of the form $U_i(x_i, Y)$, which are consistent with the classic public goods model (e.g., Bergstrom *et al.* 1986). Agents obtain utility from their own private consumption and the aggregate level of the public good. Following convention, we refer to this case as *pure altruism*. If, instead, we keep only the first and fourth arguments under the additional assumption that $d_i(y_i, \mathbf{y}_{-i}) = y_i$, the result is a model of *warm glow*, with preferences $U_i(x_i, y_i)$ that are consistent with those specified by Andreoni and Miller (2002).[6] Moreover, considering both of these cases simultaneously generates preferences of the form $U_i(x_i, Y, y_i)$, and these match those for the model of *impure altruism* (Andreoni 1988, 1990).

The preferences specified in (1) can also accommodate social comparisons whereby agents are concerned with how their own donation, private consumption, or both compare with those of others. The argument $d_i(y_i, \mathbf{y}_{-i})$ is intended to capture preferences for relative donation, as the functional form can account for one's own donation and any possible subset of others' donations.[7] The way we model such concerns is to specify $d_i(y_i, \mathbf{y}_{-i}) = y_i -$

[6] Recall that Andreoni and Miller (2002) refer to these preferences as altruistic because y_i reflects the payoff to the responder in a dictator game. In the context of private provision of a public good, however, the preferences are more commonly referred to as consistent with warm glow, reflecting that fact that donations may arise even without any concern for the overall level of the public good.

[7] Romano and Yildirim (2001) consider special cases of $d_i(y_i, \mathbf{y}_{-i})$ to study phenomena such as the "snob" and "bandwagon" effects in a sequential move game of charitable giving.

$\frac{Y_{-i}}{N-1}$, which implies that individuals care about how their donation compares with the mean donation of other agents in the economy. When this term is active in the utility function, we refer to it as *relative donation*. Note that an advantage of our setup, which we exploit later, is that relative donation can be combined with pure altruism and warm glow. It turns out, however, that no additional revealed-preference restrictions are imposed when combining relative donation with impure altruism, which we show formally later in the paper.

We also consider social comparisons based on relative private consumption. To gain intuition for why relative consumption might be important and different from relative donation, consider an individual trying to decide how much she will donate to a local public good. Suppose she is considering a \$10 donation, but finds out that a friend with twice her income has also donated \$10. From a social comparison perspective, it is easy to envision how an agent might be reluctant to donate as much when others in the economy are proportionally less generous. It is, however, important to recognize that the idea here is distinct from concern about relative donation. To see how, change the scenario so that the friend still donates \$10 but now has half the relative income. This situation, in contrast, could easily encourage the agent to donate more. But, importantly, both of these scenarios are treated identically in models where agents care only about the donations of others, while ignoring others' wealth and therefore levels of private consumption. Introducing relative consumption to models of charitable giving thus endows agents with preferences that may depend on the relative philanthropy of others, rather than the absolute amount donated.

The argument $c_i(x_i, \mathbf{x}_{-i})$ is intended to capture preferences for relative consumption. While there exists a substantial literature on the importance of relative consumption in consumer behavior (e.g., Duesenberry 1949; Boskin and Sheshinki 1978; Layard 1980; Ljungqvist and Uhlig 2000; Luttmer 2005), we are aware of only one model that considers it in the context of public goods, but the focus is on implications for taxation (Aronsson and Johansson-Stenman 2008). Following convention in the literature and in parallel with our treatment of relative donation, we specify $c_i(x_i, \mathbf{x}_{-i}) = x_i - \frac{X_{-i}}{N-1}$, which implies that agents care about how their private consumption compares with average consumption in the economy when deciding how much to donate.[8] When this term is active in the utility function, we refer to it as *relative consumption*. In what follows, we consider cases in which relative consumption is combined with pure altruism, warm glow, impure altruism, and relative donation.

[8] An alternative way to model both relative consumption and donation is to assume that agents care about the ratio of their own consumption or donation to the corresponding mean of others in the economy. This would imply $c_i(x_i, \mathbf{x}_{-i}) = \frac{x_i}{X/N}$ and $d_i(y_i, \mathbf{y}_{-i}) = \frac{y_i}{Y/N}$. While our theoretical framework readily accommodates these alternatives, and we also conducted the revealed-preference analysis for these specifications, we do not report the results here because they are very similar, and we think the specifications themselves have less intuitive appeal. Later in the paper, we also consider the possibility of social comparisons based on difference aversion.

Before turning to our methodology for conducting revealed-preference tests, we illustrate why the models just discussed, with the exception of warm glow, add complications to the standard revealed-preference framework. Using the example of relative consumption of the form $U_i(x_i, x_i - \frac{X_{-i}}{N-1}, Y)$, Figure 1 illustrates the budget frontier of an agent that seeks to maximize utility subject to $p_x x_i + p_y y_i \leq w_i$. The frontier is simply the line segment AB in the three-dimensional characteristics space. The point A denotes the allocation when all w_i is spent on x_i, and B denotes the allocation when all w_i is spend on y_i. Now consider an observed choice on the interior, say at point C. While the indifference curve must be tangent to AB, the relative prices of all three characteristics, which are necessary to test revealed preferences, are not defined by the budget frontier alone. Utility maximization implies that there exists a plane DBE that includes AB and is also tangent to the agent's indifference curve. The gradients of this plane, which depend on the budget frontier and the agent's utility function, define shadow prices of the characteristics that depend not only on the observed prices p_x and p_y, but also w_i and the two externalities of Y_{-i} and $\frac{X_{-i}}{N-1}$.[9] Hence, conducting revealed-preference tests in this environment, where any of the exogenous variables $(p_x, p_y, w_i, Y_{-i}, \frac{X_{-i}}{N-1})$ can be changing, hinges on whether there exists a well-behaved utility function and shadow prices for different allocation choices that are consistent with rational choice. This is, of course, different than standard tests for which budget sets are clearly defined with exogenous prices and income, and application of the Generalized Axiom of Revealed Preference (GARP) is relatively straightforward. We now turn to our methodological approach for carrying out such tests, which also accommodates the possibility of corner solutions.

2.2 Revealed-Preference Tests

We describe the conceptual framework for revealed-preference tests of models consistent with the preferences specified in (1). The approach accounts for characteristics and externalities of the type discussed previously. By definition, agent i's allocation choice (x_i, y_i) is a best response to the choices of the other agents $(\mathbf{x}_{-i}, \mathbf{y}_{-i})$ if

$$(x_i, y_i) \in \underset{(x,y)}{\arg\max} \left\{ U_i \left[x,\ c_i(x, \mathbf{x}_{-i}),\ y + Y_{-i},\ d_i(y, \mathbf{y}_{-i}) \right] : p_x x + p_y y \leq w_i \right\}.$$

[9] In the public goods literature, these unobserved shadow prices are sometimes referred to as virtual prices (see Cornes and Sandler 1996).

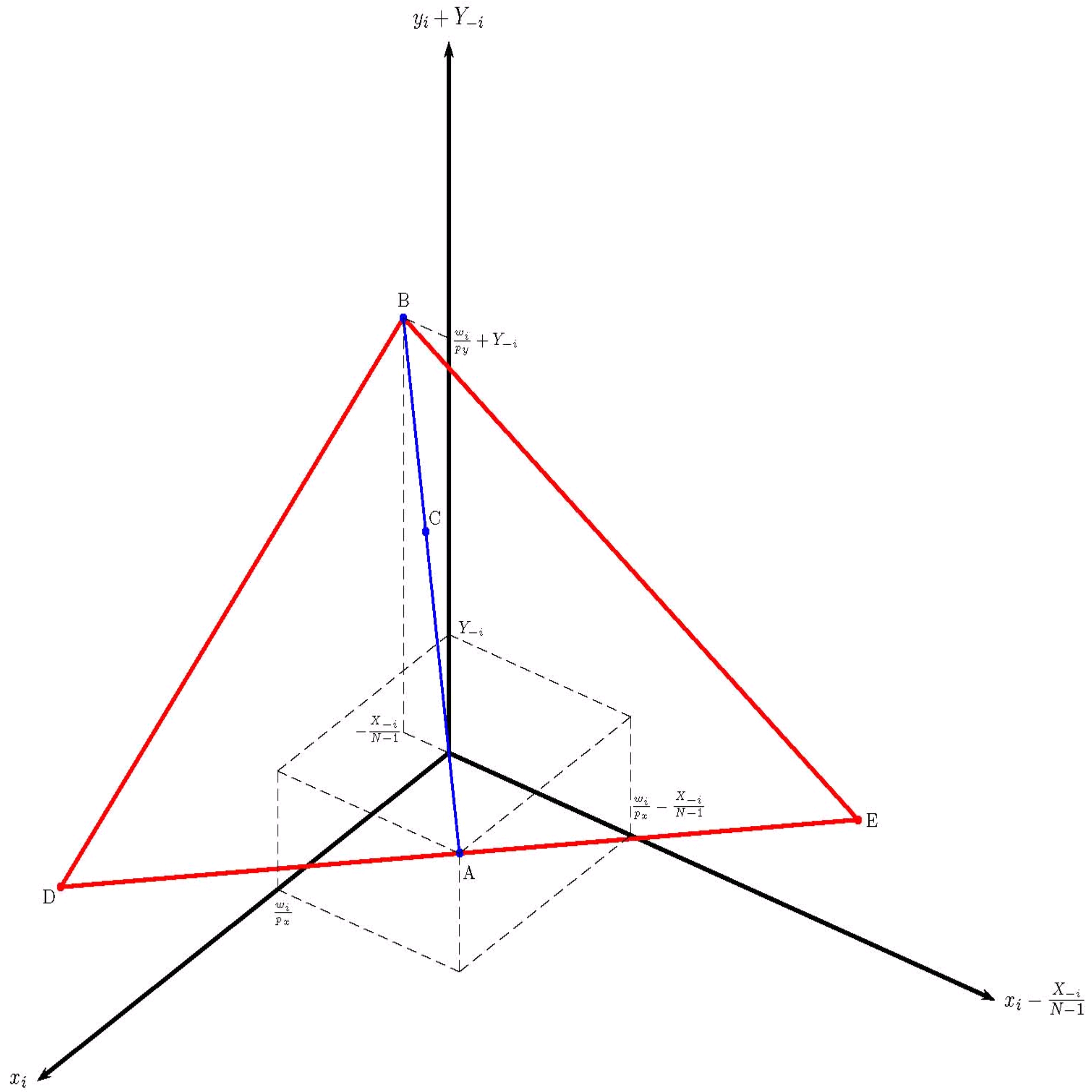

Figure 1: Budget set in characteristics space for the model of relative consumption

We denote agent i's set of best responses as $B_i(\mathbf{x}_{-i}, \mathbf{y}_{-i}, p_x, p_y, w_i)$. A vector of choices $(\mathbf{x}^*, \mathbf{y}^*)$ is thus an equilibrium if for all $i = 1, ..., N$, it holds that

$$(x_i^*, y_i^*) \in B_i(\mathbf{x}_{-i}^*, \mathbf{y}_{-i}^*, p_x, p_y, w_i).$$

Because the model's setup constitutes a concave game, equilibrium existence is guaranteed (Rosen 1965).[10] While there may be more than one equilibrium, establishing uniqueness is not necessary for our purposes.

In general, the revealed-preference approach involves the examination of a panel of choices made by agents across different budget sets. Consider a series of choices from $t = 1, ..., T$ in which each agent i faces changing prices, endowments, and choices made by other agents. The set of T choices for all N agents produces a data set of the form $\{(p_x^t, p_y^t, \mathbf{x}^t, \mathbf{y}^t)\}_{t=1}^T$, which characterizes all choices and exogenous variables. Note that endowments are defined implicitly as $w_i^t = p_x^t x_i^t + p_y^t y_i^t$. For any such data set, it is straightforward within the context of our model to define the notion of rationalization that provides the basis for revealed-preference tests.

Definition 1 *Given a data set $D = \{(p_x^t, p_y^t, \mathbf{x}^t, \mathbf{y}^t)\}_{t=1}^T$ and functions $\{c_i, d_i\}_{i=1}^N$, agent i's choices in D are rationalized if there exists a time-invariant utility function U_i such that for all t, the observed data satisfies*

$$(x_i^t, y_i^t) \in B_i(\mathbf{x}_{-i}^t, \mathbf{y}_{-i}^t, p_x^t, p_y^t, p_x^t x_i^t + p_y^t y_i^t).$$

Moreover, the entire data set D is rationalized if the choices of all agents $i = 1, ..., N$ are rationalized.

It follows that rationalization of an agent's choices involves finding a utility function such that the best response of an agent to the choices of others with respect to that function yields all of the agent's observed choices. This, in turn, implies that if the choices of all agents can be rationalized, there exist preferences such that the observed data corresponds to an equilibrium. Note that rationalization allows for complete heterogeneity across agents. The analysis is done separately for each individual and therefore imposes no requirement that U_i be the same for different individuals. Though we consider homogenous specifications for c_i and d_i in our experiment (discussed in the next section), the framework is general enough to admit heterogeneity of these functions as well. Allowing for such heterogeneity is, of course, one of the primary advantages of studying behavior using revealed preferences.

[10] The assumptions on U_i, c_i, and d_i imply that each agent i's utility function is continuous in $(\mathbf{x}, \mathbf{y})$ and concave in (x_i, y_i) for all fixed $(\mathbf{x}_{-i}, \mathbf{y}_{-i})$. Moreover, each agent's strategy set is compact and convex. Together, these conditions imply a concave game.

The following theorem, which we prove in Appendix A, formally states the conditions that underlie the revealed-preference tests of our model.

Theorem 1 *Given a data set* $D = \{(p_x^t, p_y^t, \mathbf{x}^t, \mathbf{y}^t)\}_{t=1}^T$ *and functions* $\{c_i, d_i\}_{i=1}^N$*, where* $c_i^t = c_i\left(x_i^t, \mathbf{x}_{-i}^t\right)$ *and* $d_i^t = d_i\left(y_i^t, \mathbf{y}_{-i}^t\right)$*, the following statements are equivalent:*

1. *There exists a utility function* U_i *of the form (1) that rationalizes the choices of agent* i *in* D*.*

2. *The following inequalities have non-negative solutions for* κ_i^t*,* π_i^t*,* γ_i^t*,* η_i^t *and positive solutions for* λ_i^t*,* U_i^t *for all* t *and* t'*:*

$$U_i^{t'} \leq U_i^t + \kappa_i^t[x_i^{t'} - x_i^t] + \pi_i^t[c_i^{t'} - c_i^t] + \gamma_i^t[Y^{t'} - Y^t] + \eta_i^t[d_i^{t'} - d_i^t],$$
$$\kappa_i^t + \pi_i^t \frac{\partial c_i(x_i^t, \mathbf{x}_{-i}^t)}{\partial x_i} \leq \lambda_i^t p_x^t \quad \text{holding with equality if } x_i^t > 0,$$
$$\gamma_i^t + \eta_i^t \frac{\partial d_i(y_i^t, \mathbf{y}_{-i}^t)}{\partial y_i} \leq \lambda_i^t p_y^t \quad \text{holding with equality if } y_i^t > 0.$$

The theorem states that an agent's choices can be rationalized if and only if the a set of linear inequalities has a solution. The inequalities are based on the first-order conditions, the concavity restrictions on the utility function, and the relative consumption and donation functions. Commonly referred to as Afriat (1967) inequalities, the conditions enable explicit construction of utility levels and the marginal utility of income associated with each agent's observation t; that is, they define a utility level $U_i^t = U_i[x_i^t, c_i\left(x_i^t, \mathbf{x}_{-i}^t\right), y_i^t + Y_{-i}^t, d_i\left(y_i^t, \mathbf{y}_{-i}^t\right)]$ and a marginal utility of income λ_i^t associated with the endowment $p_x^t x_i^t + p_y^t y_i^t$ for each observed $(\mathbf{x}^t, \mathbf{y}^t)$. The unobserved shadow prices of characteristics are reflected, in part, through the values of κ_i^t, π_i^t, γ_i^t, and η_i^t, which themselves represent marginal utilities for the corresponding characteristrics. Simply dividing them by the Lagrange multiplier on the budget constraint, λ_i^t, therefore, reveals the shadow prices. A useful feature of the theorem's proof, as shown in Appendix A, is that it is constructive, meaning that when an agent's choices can be rationalized, the proof provides a candidate utility function. Because the inequalities are linear in the unknowns, it is also simple and computationally efficient to verify whether they have a solution.

We use the example of impure altruism to illustrate the key inequalities for revealed-preference tests of that particular model.

Example 1 *Recall that impure altruism implies* $c_i(x_i, \mathbf{x}_{-i}) = 0$ *and* $d_i(y_i, \mathbf{y}_{-i}) = y_i$ *for the utility function specified in (1). The inequalities corresponding with the conditions in*

Theorem 1 are

$$
\begin{aligned}
U_i^{t'} &\leq U_i^t + \kappa_i^t \left[x_i^{t'} - x_i^t\right] + \gamma_i^t \left[Y^{t'} - Y^t\right] + \eta_i^t \left[y_i^{t'} - y_i^t\right], \\
\kappa_i^t &\leq \lambda_i^t p_x^t \quad \textit{holding with equality if } x_i^t > 0, \\
\gamma_i^t + \eta_i^t &\leq \lambda_i^t p_y^t \quad \textit{holding with equality if } y_i^t > 0.
\end{aligned}
$$

This example shows that for a data set $D = \{(p_x^t, p_y^t, \mathbf{x}^t, \mathbf{y}^t)\}_{t=1}^T$ agent i's choices can be rationalized by the impure altruism model if and only if the derived system of linear inequalities has a solution for non-negative κ_i^t, γ_i^t, η_i^t and positive λ_i^t and U_i^t. If a solution exists, we cannot reject optimizing behavior; whereas if a solution does not exist, the agent's choices are inconsistent with optimizing behavior. It is therefore possible to derive pass rates for different models among agents to compare how models are more or less successful at explaining the repeated choices of subjects. Because it is straightforward, we do not derive the explicit inequalities for testing the other models discussed previously, but we use them when carrying out the tests reported later in the paper.

It is worth mentioning that the impure altruism application of our theorem is related to the revealed-preference tests in Korenok *et al.* (2011). They provide separate necessary and sufficient conditions (their theorem and result, respectively) for impure altruism to rationalize a given data set. Their sufficient condition states that in order for the data to satisfy GARP, there must be shadow prices for the characteristics such that the data can be rationalized in characteristic space. The intuition follows from our previous discussion of Figure 1. Our approach differs, however, in that a solution to the inequalities in Example 1 provides exactly such shadow prices : $\frac{\gamma_i^t}{\lambda_i^t}$ for altruism and $\frac{\eta_i^t}{\lambda_i^t}$ for warm glow, in addition to $\frac{\kappa_i^t}{\lambda_i^t} = p_x^t$ for private consumption. It is thus possible to show that their sufficient condition is closely related to the inequalities in our example. Beyond the fact that our framework is more general than impure altruism, a further difference between approaches is the ease of application. Following the steps of Korenok *et al.* (2011), one must search over the space of shadow prices to find a price vector that satisfies rational choice, but there is no general algorithm to find these shadow prices in a finite number of steps.[11] Moreover, if such prices are not found, it remains unclear whether the reason is because the search algorithm failed or because they do not exist. In contrast, our inequalities are in the form of a linear program, and there are well-known algorithms to check feasibility and solve systems of linear inequalities.

[11] The computational problem is similar to the one encountered by Varian (1983), who derives revealed-preference tests of weak functional separability.

3 Experiment Design

We design an experiment that allows us to test and differentiate among models of charitable giving using our revealed-preference framework. Each subject is tasked with making a series of allocation choices between oneself and donating to a charitable cause. While our experiment has several features in common with Andreoni and Miller (2002) and Korenok *et al.* (2011), one important difference is that we study giving to a local non-profit organization rather than another anonymous subject in the lab. Using notation from the previous section, subjects are asked to make allocation choices $\{x_i^t, y_i^t\}_{t=1}^T$ in scenarios with changing values of the subject's endowment (w_i^t), prices of private consumption (p_x^t) and charitable donation (p_y^t), others' private consumption (x_{-i}^t), and others' charitable giving (y_{-i}^t).[12]

The fact that we study how subjects respond to the choices of other subjects necessitates an experimental design with two distinct cohorts, denoted A and B. The primary purpose of Cohort-A, as we will explain, is to generate allocation choices for subsequent use with Cohort-B. All subjects in our actual experiment were volunteers among the undergraduate student population at Williams College in Williamstown, Massachusetts, and all experimental sessions took place during May and June 2011, with recruitment using an online system (Greiner 2004). Because our experimental procedures did not vary greatly between cohorts, we first describe those for Cohort-A and then identify important differences for Cohort-B.

Cohort-A subjects were asked to make choices about splitting an endowment between oneself and the Hoosic River Watershed Association (HooRWA), which is a non-profit organization dedicated to the restoration, conservation, and enjoyment of the Hoosic River and its watershed. The Hoosic River flows by the Williams campus, and students are generally familiar with HooRWA because of its presence in a small town and offices in a non-college building near the center of campus. Because HooRWA is a relatively small non-profit organization, it is reasonable to assume that subjects might consider their donations to be meaningful. At the beginning of each session, subjects received a copy of the instructions (included in Appendix B) that were read aloud. After hearing the instructions and before making their choices, subjects were required to successfully calculate the earnings associated with two hypothetical scenarios, and these review questions were conducted using the software program z-Tree (Fischbacher 2007).

The actual choices—20 in total for each subject—were made using pencil and paper. For each choice, the subject was given a token endowment and informed of the value per token

[12] Because our experimental scenarios involve only two agents, we can simplify notation at this point such that $X_{-i}^t = \mathbf{x}_{-i}^t = x_{-i}^t$ and $Y_{-i}^t = \mathbf{y}_{-i}^t = y_{-i}^t$.

for private consumption and charitable donation. Tokens translated into a different number of points for consumption and donation, and each point was worth 10 cents. Each choice was also associated with separate and unchangeable endowments of points for HooRWA and another randomly paired subject in the current session, referred to as a match.[13] Each subject was then asked to divide tokens between those to hold for private consumption and those to pass for donation to HooRWA. The following summarizes the details and presentation of an example scenario:

Initial points		Tokens	
Match	HooRWA	to divide	Your choice
40	20	50	*Hold* ___@1 point each, and *Pass* ___@2 points each

Each subject filled out a choice sheet that included the 20 scenarios in randomized order. Table 1 lists the 20 Cohort-A scenarios, along with the mean number and percentage of tokens passed for each scenario. Cohort-A consisted of 36 subjects from three equally sized sessions of 12 participants. The sessions lasted approximately one hour, and payment to subjects was based on one randomly selected scenario, which determined payments to the subject, HooRWA, and the subject's match. A Cohort-A subject's payment thus consisted of two parts: points per tokens kept in the randomly selected scenario plus the points from serving as another subject's match. Payments were made at the end of the session using a double-blind procedure (Hoffman, McCabe and Smith 1996) to limit giving induced by strategic altruism. The total payment per subject, which included the two parts, was \$17.61 on average, and the average payment to HooRWA was \$8.96 per subject, which included the initial points plus the donated points.

As mentioned previously, Cohort-B is the main focus of our experiment, and the primary purpose of Cohort-A was to generate scenarios for Cohort-B. Readers will have noticed that our Cohort-A initial points for both the randomly matched subject and HooRWA were synthetic constructs for x^t_{-i} and y^t_{-i}. Because they were not based on the choices of actual subjects, Cohort-A is of limited (though useful, as we will show) value for testing the importance of social comparisons. Cohort-B differs because we use the previous choices of subjects to produce real values for x^t_{-i} and y^t_{-i} that conform to the No Deception Rule in economic experiments.[14] In other words, the task of Cohort-B subjects closely mirrored that of Cohort-A

[13] The reason for including a match in Cohort-A will become clear when we describe Cohort-B, as it is important for subjects in both cohorts to face similarly structured scenarios.

[14] This norm is largely based on the belief that deception will adversely impact the ability of subsequent experimenters to maintain experimental control (e.g., Friedman and Sunder 1994). In fact, Karlan, Jamison and Schecter (2008) find that deception affects both a subject's likelihood of returning for subsequent experiments and the choices that are made conditional on returning.

Table 1: Cohort-A scenarios and summary statistics of tokens passed

	Initial points		Tokens	Points per		Average	
						tokens	percent
Scenario	Match	HooRWA	to divide	held	passed	passed	passed
1	10	58	60	5	1	4.83	8.06
2	40	60	70	4	1	9.67	13.81
3	16	46	50	4	1	4.97	9.94
4	33	49	60	3	1	6.00	10.00
5	30	70	80	3	1	8.81	11.01
6	20	56	32	5	2	2.75	8.59
7	52	64	90	2	1	9.39	10.46
8	60	90	120	2	1	13.47	11.23
9	39	84	55	3	2	4.83	8.79
10	36	94	65	2	2	4.58	7.05
11	28	132	80	2	2	5.22	6.53
12	44	69	45	2	3	5.75	12.78
13	54	72	90	1	2	11.53	12.81
14	46	128	110	1	2	11.47	10.43
15	22	135	38	2	5	5.08	13.38
16	58	126	100	1	3	12.61	12.61
17	50	120	90	1	3	11.39	12.65
18	70	80	90	1	4	11.50	12.78
19	67	52	80	1	4	13.39	16.74
20	62	40	70	1	5	12.19	17.42

Notes: Cohort-A includes 36 subjects making choices on all 20 scenarios.

with the important exception that the initial points given to Cohort-B are actual allocation decisions made by Cohort-A subjects. Specifically, Cohort-A and Cohort-B subjects faced 20 scenarios with the same endowments and prices, but each scenario for Cohort-B was associated with a previous Cohort-A subject's chosen points for private consumption and HooRWA when the subject faced the same endowment and prices. The following summarizes the details and presentation of an example Cohort-B scenario:

Previous Participant Choice		Initial	
Held	HooRWA	HooRWA	Tokens
Points/Tokens	Points/Tokens	Points	To Divide
40 points / 40 tokens	20 points / 10 tokens	20	50

Your choice
Hold ___@1 point each, and *Pass* ___@2 points each.

It is also important to mention that we explicitly informed Cohort-B subjects that their choices would not be presented to subsequent subjects.

The first six columns of Table 2 report the 20 scenarios that all Cohort-B subjects received in randomized order. There were 120 subjects in Cohort-B from 9 sessions with a mean size just over 13 and ranging between 12 and 17. Table 2 also reports the average number and percentage of tokens passed for each scenario. While the Previous Participant Choice for private consumption and HooRWA are based on actual decisions, the Cohort-A choices we that presented to Cohort-B were not selected at random nor presented as such. Rather, our objective was to select "realistic" choices that would give subjects ample opportunity to make allocation decisions that are inconsistent with our theoretical models of interest. To accomplish this, we experimented with Bronars (1987) *ex ante* test for the likelihood that random and uniformly distributed choices in each of the scenarios would result in a panel of choices inconsistent with each specified utility model. Our approach was somewhat *ad hoc* given that we are testing several different models, but as we discuss in the next section, the scenarios that we put forth in Cohort-B produce Bronars results with sufficient power to ensure plenty of scope for rejecting models.

Appendix C includes the Cohort-B instruction sheet, and the procedures closely followed that for Cohort-A. One difference was that to compensate for the fact that Cohort-A subjects received a match payment ($4.20 on average), each Cohort-B subject received a $5 participation payment in addition to the point earnings on the one randomly selected scenario. Another point of clarification about payoffs is that for Cohort-B's randomly selected

Table 2: Cohort-B scenarios and summary statistics of tokens passed

Scenario	Previous subject's chosen points: Consumption	Previous subject's chosen points: HooRWA	Tokens to divide	Points per held	Points per passed	Average tokens passed	Average percent passed
1	175	25	60	5	1	10.17	19.94
2	100	45	70	4	1	12.33	17.61
3	200	0	50	4	1	8.83	17.67
4	120	20	60	3	1	11.91	19.85
5	150	30	80	3	1	15.88	19.84
6	115	18	32	5	2	5.28	16.51
7	104	38	90	2	1	17.38	19.31
8	160	40	120	2	1	21.15	17.63
9	117	32	55	3	2	10.27	18.67
10	66	64	65	2	2	13.79	21.22
11	146	14	80	2	2	16.15	20.19
12	80	15	45	2	3	10.20	22.67
13	62	56	90	1	2	19.27	21.41
14	105	10	110	1	2	24.49	22.27
15	76	0	38	2	5	10.08	26.54
16	75	75	100	1	3	20.90	20.90
17	70	60	90	1	3	18.99	21.10
18	88	8	90	1	4	22.08	24.53
19	74	24	80	1	4	19.98	24.97
20	60	50	70	1	5	18.18	25.98

Notes: Cohort-B includes 120 subjects making choices on all 20 scenarios. Tokens to divide and Points per columns are the same as those in Table 1 for Cohort-A. Previous subject's chosen points are based on selected choices from Cohort-A.

scenarios, additional payments were made to HooRWA (20 points in the example) but not to the previous participant (40 points in the example). The payment per Cohort-B subject was \$17.68 on average, and the average payment to HooRWA was \$6.30 per subject. Both figures are quite similar to those from Cohort-A.

4 Experiment Results

We focus analysis of the results on how different models of charitable giving rationalize the choices of our experimental subjects. As discussed previously, we consider standard models of privately provided public goods, along with novel models that account for social comparisons. We focus throughout on Cohort-B, but consider some comparisons with Cohort-A as part of robustness checks at the end of the section.

4.1 Preliminaries

With experimental studies, it is often useful to begin with comparisons of subject behavior to other experiments as a check of representativeness, but direct comparisons are not possible in our case because the experimental design has several unique features. We nevertheless compare selected results with those in Andreoni and Miller (2002) and Korenok *et al.* (2011). Andreoni and Miller's (2002) experiment is based on a panel of choices in a dictator game with changing prices and endowments, and Korenok *et al.* (2011) have a similar design that also includes changing initial endowments of the recipient. Recall that our design differs because (*i*) subjects are making donations to a local non-profit organization rather than another anonymous subject in the lab, and (*ii*) subjects are informed of another's choices for private consumption and donation when faced with the same choice scenario.

Despite differences in the experimental design, several of the standard comparisons are surprisingly similar. The "perfectly selfish" strategy of keeping all tokens for oneself in all 20 choice scenarios was played by 22.7 percent of our subjects, compared to the identical percentage and 25.2 percent for Andreoni and Miller (2002) and Korenok *et al.* (2011), respectively.[15] Another commonly referenced strategy is to maximize the aggregate payoff, in which case points held and passed are perfect substitutes. We find that 2.5 percent of our subjects played this strategy compared to 6.2 and 1.6 percent for the other studies, respectively. When faced with a price ratio of $p_x/p_y = 1$ (i.e., scenarios 10 and 11 in Table

[15] Unless otherwise indicated, percentages for our experiment are based on 119 Cohort-B subjects. We dropped one Cohort-B subject because of subject confusion. The subject wrote down the difference between the additional number of tokens desired to pass and the number the previous participant passed, meaning that the subject was erroneously constrained to passing at least as many as the previous participant.

2), the token pass rate of our subjects was 20.6 percent compared to 23.0 and 22.9 percent for the other studies, respectively, when faced with the same prices for holding and passing tokens.

Though not a central part of our analysis, we also estimate an average demand function for donations in order to verify that our subjects make choices consistent with standard economic theory on charitable giving. In particular, we estimate the following fixed-effects model:

$$y_i^t = \alpha p_{yi}^t + \beta p_{xi}^t + \phi w_i^t + \theta y_{-i}^t + \nu_i + \varepsilon_{it},$$

where y_{it} is the contribution of subject i in scenario t measured in points; p_{yi}^t is the price of contribution points in terms of tokens (i.e., the inverse of points per passed in Table 2); p_{xi}^t is the price of private consumption points in terms of tokens (i.e., the inverse of points per held in Table 2); w_i^t is the token endowment; y_{-i}^t is the previous participant choice of HooRWA points; ν_i is a subject-specific intercept; and ε_{it} is the error term. A useful feature of this specification is that coefficients are identified off of variation within subjects. Table 3 reports the coefficient estimates and corresponding elasticities. We find that the results are consistent with theory. While the income effect is statistically insignificant, donations are decreasing in the price of making a donation, increasing in the price of private consumption (consistent with donations being a gross substitute for private consumption), and decreasing in the previous participant's contribution level. In terms of magnitudes, the estimates imply that the crowding out at 14 cents per dollar is substantially less than one-for-one, and the price elasticity of giving is approximately -0.4.

We now turn attention to the use of revealed preferences for testing models. When conducting such tests, it is important to ensure that they have "power"; that is, useful tests are those that provide subjects ample opportunity to make choices that are inconsistent with the model being tested. The Bronars' (1987) Power Index is the most commonly used criteria for evaluating the power of revealed-preference tests. It produces the probability that a random and uniformly distributed set of choices on the budget sets for a series of choice scenarios will violate revealed-preference tests.

Table 4 lists the different models that we consider, by name and utility function, along with Bronars results for each, given the scenarios presented to Cohort-B of our experiment. To facilitate interpretation, we report a modified version of the standard Bronar's Index: the percentage of random draws that are consistent with the corresponding model, rather than the proportion that are inconsistent. We find a high degree of power across all models. For example, based on simulations of random and uniformly distributed choices of 50,000 subjects, only 0.04 percent are consistent with warm-glow preferences, meaning that 99.96 percent are inconsistent with the model. With more general utility functions, the power

Table 3: Fixed-effects model of demand for donations

	Coefficients	Elasticities
Price of donation (p_{yi}^t)	-25.762***	-0.419***
	(3.363)	(0.055)
Price of private consumption (p_{xi}^t)	61.222***	0.995***
	(9.758)	(0.159)
Token endowment (w_i^t)	-0.053	-0.107
	(0.055)	(0.111)
Previous participant donation (y_{-i}^t)	-0.144***	-0.123***
	(0.035)	(0.030)
Subject fixed effects	Yes	
Number of subjects	119	
Observations	2,380	
R-squared (within)	0.264	

Notes: The dependent variable is the subject's contribution to HooRWA in points (y_i^t). Standard errors clustered on each subject are reported in parentheses. One, two, and three asterisk(s) indicate statistical significance at the 90-, 95- and 99 percent levels, respectively.

declines, but even for relative consumption + relative donation, only 20.37 percent of the simulated subjects are consistent with the model. While different utility functions yield different degrees of power, the general strength of the tests reported in Table 4 is due to our selection of budget sets with a large number of intersections and sufficient variation in the Cohort-A choices that we report in the scenarios for Cohort-B.

4.2 Revealed-Preference Tests and Model Comparisons

The last column of Table 4 reports the percent of Cohort-B subjects that actually made choices consistent with each model.[16] These percentages are derived by implementation of the tests based on Theorem 1. We carried out the analysis in MATLAB using a linear programme solver. The solver tests whether the feasible region corresponding to the linear inequalities of Theorem 1 is non-empty, in which case there is a solution to the inequalities, and the choices can be rationalized. In all cases, the percentage of subjects whose choices are rationalized by the model is substantially higher than the *ex ante* Bronars results, indicating

[16] As mentioned in footnote 15, the percentages are based on 119 subjects. Eight choices among the total of $119 \times 20 = 2,380$ required a bit of data cleaning due to simple mathematical error on the part of subjects allocating too few or many tokens. For these observations we conducted all analyses under two alternative assumptions: that tokens passed was correct and adjusted tokens held accordingly, and that tokens held was correct and adjusted tokens passed accordingly. We find that the results do not change under the two different assumptions.

Table 4: Models and Cohort-B Bronars and revealed-preference results

Model	Utility function	Bronars results	% of Subjects rationalized
Warm glow	$U_i(x_i, y_i)$	0.04	49.58
Pure altruism	$U_i(x_i, Y)$	0.29	57.98
Impure altruism	$U_i(x_i, Y, y_i)$	1.70	73.95
Relative donation	$U_i(x_i, Y, y_i - \frac{Y_{-i}}{N-1})$	3.89	84.03
Relative consumption	$U_i(x_i, x_i - \frac{X_{-i}}{N-1}, Y)$	4.59	85.71
Relative consumption + Impure altruism	$U_i(x_i, x_i - \frac{X_{-i}}{N-1}, Y, y_i)$	11.18	91.60
Relative consumption + Relative donation	$U_i(x_i, x_i - \frac{X_{-i}}{N-1}, Y, y_i - \frac{Y_{-i}}{N-1})$	20.37	94.12

Notes: We report a modified version of the standard Bronars index: the percentage of random draws that are consistent with the corresponding model, rather than the proportion that are inconsistent. These results are based on 50,000 replications. The percentage of subjects rationalized by each model is based on 119 Cohort-B subjects.

that the models meaningfully explain subject behavior. For example, nearly 50 percent of the subjects made choices consistent with warm-glow preferences, and relative consumption + relative donation rationalizes the choices of more than 94 percent of the subjects. These numbers compare to 0.04 and 20.37 percent for the Bonars results, respectively. Taken as a whole, this pattern of results suggests that, despite differences among models, optimization in one form or another helps to explain a substantial amount of the subjects' charitable giving. The set of results in Table 4 also demonstrate how our framework can be used to carry out revealed-preference tests based on an array of models with both externalities and characteristics.

We now focus on comparisons among the models of charitable giving in order to make judgments about which perform better. In doing so, it is important to recognize that we are testing models with varying degrees of power and generality on the same data set. This means that valid comparisons among models must account not only for the percentage of data rationalized, but also the differing power of tests. A recent paper by Beatty and Crawford (2011) provides a methodology for making such comparisons, and we follow their recommendation here. Let a denote the Bronars measure that we report in Table 4. The percentage a can be interpreted as a measure of the set of choices defined by the revealed-preference restrictions relative to the set of all possible choices. If a is 100, the revealed-preference test almost surely imposes no restrictions; whereas if a is 0, choices can almost surely never pass the revealed-preference test. The explanatory power of a given model must therefore depend on the percentage of data rationalized, denoted r, and the "target" area in consumption space a. There are many possible functions defined on (r, a) that could be chosen, with some intuitive candidates being $r - a$ and r/a. Beatty and Crawford

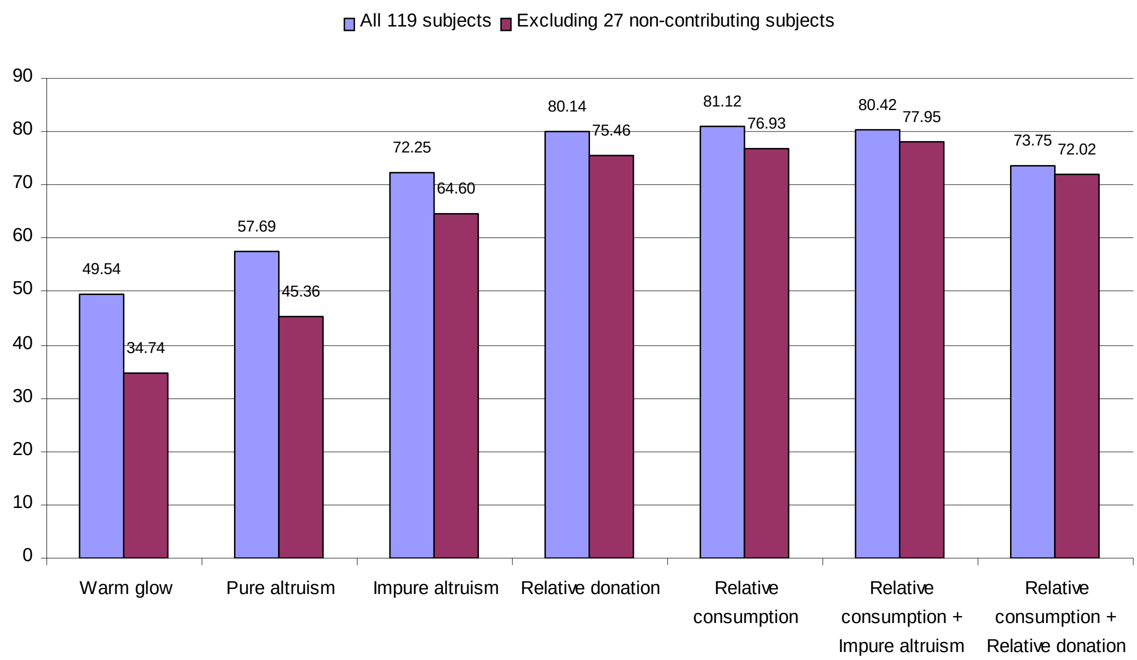

Figure 2: Power adjusted $(r - a)$ percent of Cohort-B subjects that are consistent with each model, by all subjects and excluding non-contributing subjects.

(2011) argue, however, that a desirable measure should satisfy the three basic axioms of monotonicity, equivalence, and aggregability. They also prove that only the function $r - a$ (and affine transformations of it) satisfies the axioms and recommend using the measure for evaluating the performance of models. We therefore base our comparisons on this measure and conclude that models with a higher $r - a$ perform better.[17]

Figure 2 illustrates these power adjusted results for all models. The differences, $r - a$ for each model, are shown for all Cohort-B subjects as the first set of histogram bars. Also shown in Figure 2 are bars that exclude the 27 "purely selfish" subjects that never donated in any of the 20 choice scenarios. We report the second set of results as a simple point of comparison because all of the models rationalize the choices of non-donating subjects, making differences among models appear less stark.

We begin with the standard models. Warm glow and pure altruism are distinct models

[17]It is possible to conduct approximate rationality tests in our setting as well. One approach would be to calculate the Critical Cost Efficiency Index (CCEI) of Afriat (1972). Intuitively, this approach involves finding the degree to which the budget sets of subjects need to be relaxed in order to rationalize their choices. We do not, however, conduct this analysis for two reasons. First, our focus in on demonstrating applicability of our theoretical framework and comparing various models rather than testing and justifying a single model; and for comparing models, an approximate measure is not necessary. Second, Beatty and Crawford (2011) make the important observation that relaxing budget sets also changes the power of tests, and taking this into account reveals that the CCEI approach is more *ad hoc* than employing standard thresholds (e.g., the conventional 95-percent rule) often implies.

in the sense that neither subsumes the other. This follows because with warm glow, two scenarios in which x_i and y_i are the same but Y_{-i} differs will yield the same level of utility, whereas the same scenarios will yield two different levels of utility for pure altruism. Similarly, two scenarios in which x_i and $y_i + Y_{-i}$ are the same but y_i differs will yield the same utility for pure altruism and different levels of utility for warm glow. A well-known difference is that pure altruism allows for crowding out while warm glow does not. The power of both tests on our data is very strong, and both models fit the data reasonably well, rationalizing 50 and 58 percent of the data for warm glow and pure altruism, respectively (Table 4). After making the power adjustments (Figure 2), we find that pure altruism fits the data better by 8.2 and 10.6 percentage points with and without the "selfish" subjects, respectively. These results suggest that, in the context of our experiment, donations appear to operate like a public good because crowding out plays a role in explaining donation levels.

The comparison of these two models with impure altruism is somewhat different because impure altruism nests the other two. This implies that impure altruism will rationalize all of the data that are rationalized by either warm glow or pure altruism. The question is thus whether the generalization meaningfully improves the goodness of fit after making the power adjustment. We find that it does, increasing the fit by 14.6 and 19.2 percentage points beyond pure altruism with and without the "selfish" subjects respectively. The contrast with warm glow is even more substantial, at 22.8 and 29.8 percentage points, respectively. Note that the latter conclusion—that impure altruism has greater explanatory power than warm glow—accords with the results of Korenok *et al.*'s (2011) experiment. It is worth keeping in mind, however, that our results are based on giving to a charity, while their results are based on giving to another subject in the lab. Overall, our tests of these standard models indicate that crowding out plays an important role in charitable giving, and allowing the crowding out to be less than one-for-one, as with impure altruism, strengthens the conclusion even more.

We now turn to the more novel models that account for social comparisons. First consider the model of relative donation, which rationalizes 84 percent of the data before making the power adjustment. After the adjustment, the numbers are 80.1 and 75.5 percent with and without the "selfish" subjects, respectively. This is 7.9 and 10.9 percentage points more than the adjusted results for impure altruism; however, it is important to consider whether the models are independent or one nests the other. In this case, it is a bit more subtle than we have encountered previously, but it can be shown that relative donation is a generalization of impure altruism. To prove this, consider a utility function of the form

$$\tilde{U}_i[x_i, y_i + Y_{-i}, \alpha(y_i + Y_{-i}) + \beta(y_i - \tfrac{Y_{-i}}{N-1})],$$

which is non-decreasing and concave in all three arguments, and $\alpha, \beta \geq 0$ are constants. While it is straightforward to see that this utility function is a special case of relative donation, we can also show that it is a generalization of impure altruism. Imposing the restrictions that $\alpha = 1/N$ and $\beta = (N-1)/N$ yields the impure altruism utility function of $\tilde{U}_i = \tilde{U}_i[x_i, y_i + Y_{-i}, y_i]$. The important empirical question, therefore, is not whether relative donation rationalizes more of the data, but again, whether the difference is empirically meaningful. Based on the numbers referenced above, we conclude that the additional explanatory power of relative donation is meaningful, increasing the goodness of fit from impure altruism by more than half the amount that impure altruism does compared to pure altruism.

A notable feature of the model of relative consumption is that it does not nest impure altruism. While it is also independent of relative donation and warm glow, relative consumption does subsume pure altruism. This follows because none of the other models depend on X_{-i}, and relative consumption has two arguments that are identical to those for pure altruism. We find that relative consumption rationalizes 85.7 percent of the data. After making the power adjustment, the numbers are 81.1 and 76.9 percent with and without the "selfish" subjects, respectively. How does this compare to impure altruism, which might be considered a competing model? Relative consumption performs better with adjusted differences of 8.9 and 12.3 percentage points with and without the "selfish" subjects. We interpret these results as strong evidence that how one's level of philanthropy compares to others, scaled by income, helps to explain decisions about charitable giving; for both relative giving and endowments enter the model implicitly because of the exogenous levels of y^t_{-i} and x^t_{-i} that each agent i observes when making allocation choices.[18]

The other results shown in Table 4 and Figure 2 are further generalizations of the utility function. We consider relative consumption + impure altruism and relative consumption + relative donation. One reason for including these cases is to demonstrate the flexibility of the revealed-preference framework that we develop for testing a variety of different preference structures. The experimental results show that both generalizations rationalize more than 90 percent of the data (Table 4). But after making the power adjustments, these models have less explanatory power than relative consumption on its own (Figure 2). We conclude,

[18] We also use our framework to test an alternative model of relative consumption that considers difference aversion. This admits the possibility that individuals care not about whether their private consumption is more or less than that of others, but only about the absolute difference is private consumption. Specifically, we consider a utility function of the form $U_i[x_i, -(x_i - \frac{X_{-i}}{N-1})^2, Y]$ and find that it performs similar to impure altruism, but less well than relative consumption. The Bronars result is 0.72, and considering all subjects, the percent of subjects rationalized is 76.47, with a power adjusted percentage of 75.75. Excluding the non-contributing subjects, the power adjusted percentage is 68.84. It is also worth mentioning that a similar analysis can be carried out for difference aversion related to donations, and this underscores the generality of our framework for testing many types of models.

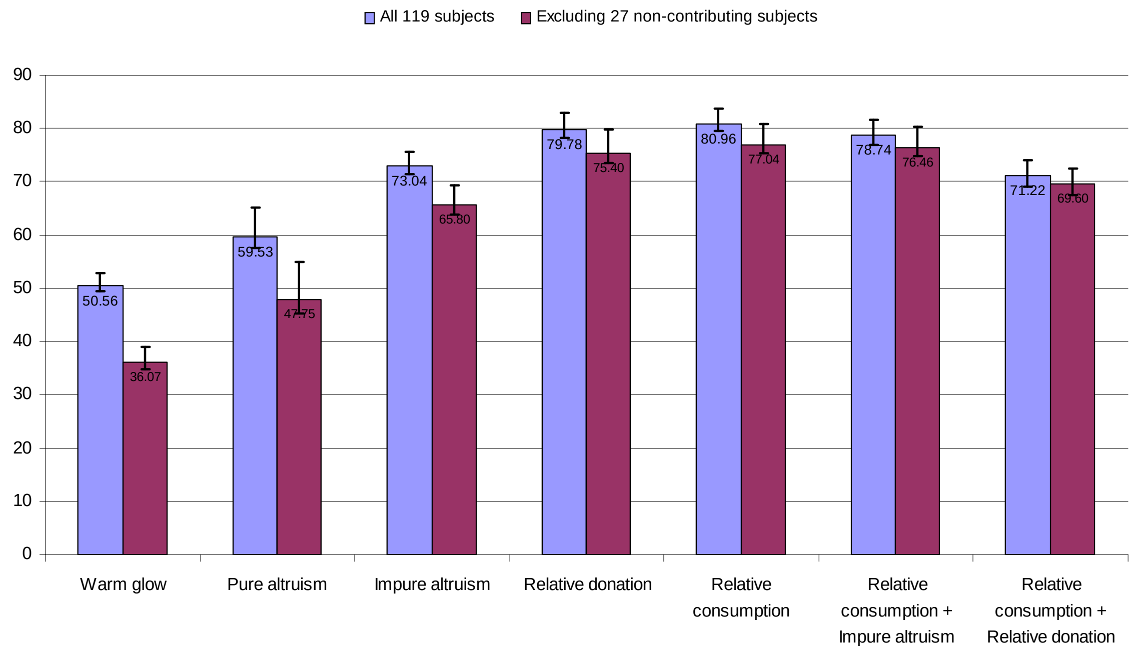

Figure 3: Means, minimums, and maximums for the power adjusted $(r - a)$ percent of Cohort-B subjects that are consistent with each model when dropping each of the 20 choice scenarios in different replications, by all subjects and excluding non-contributing subjects.

therefore, that these generalizations do not have "significant" effects.

4.3 Robustness Checks

A potential concern with our analysis is that one of the 20 choice scenarios is having undue influence over the results. This could affect the tests of a particular model, comparisons among them, or both. To evaluate sensitivity of our results to any particular scenario, we replicate the analysis 20 times, dropping one scenario each time. That is, we exclude each scenario once and conduct the analysis on the remaining 19 scenarios. Note that each replication requires new Bronars results and revealed-preference tests for each replication, as the power of each test differs with changes in the included set of scenarios.

Figure 3 summarizes the replication results. The histogram bars represent the means of the power adjusted results for the corresponding model, with and without the non-contributing subjects. Hence, the bars illustrate the mean results in parallel with the results in Figure 2. The whiskers on each bar indicate the minimum and maximum of the power adjusted results across the 20 replications. We find little variation in the explanatory power within models and no change in the comparisons across models based on the mean results.

Table 5: Cohort-A Bronars and revealed-preference results

Model	Bronars results	% of Subjects rationalized
Warm glow	0.04	47.22
Pure altruism	0.64	66.67
Impure altruism	3.94	91.67
Relative donation	8.02	94.44
Relative consumption	8.87	88.89
Relative consumption + Impure altruism	15.07	97.22
Relative consumption + Relative donation	22.21	97.22

Notes: We report a modified version of the standard Bronars index: the percentage of random draws that are consistent with the corresponding model, rather than the proportion that are inconsistent. These results are based on 50,000 replications. Percent of subjects rationalized by each model is based on 36 Cohort-A subjects.

We therefore conclude that any one scenario is not critical to the overall pattern of results.

For the final part of our analysis, we return to Cohort-A, which provides a useful comparison with Cohort-B given that our main experimental results are on the importance of relative consumption and donation. Recall that the difference between cohorts is the way that subjects were informed about x^t_{-i} and y^t_{-i}. These values were simply asserted for Cohort-A as part of the experimental design, while they were reported (without deception) as the result of a previous subject's choices to Cohort-B.[19] Given this difference, it is reasonable to expect that while concerns about relative consumption and donation help explain Cohort-B behavior, that same pattern should not be apparent in Cohort-A, as the comparisons for these subjects are not with the choices of another subject. While we recognize that the Cohort-A sample size is relatively small, we nevertheless make the comparison because it produces a useful counterfactual where values for x^t_{-i} and y^t_{-i} do not arise from another subject's choices.

Table 5 reports the Cohort-A Bronars results and the percentage of subjects rationalized by each model in parallel with the results for Cohort-B in Table 4. The more useful comparisons, however, are the power adjusted results illustrated in Figure 4. With respect to the standard models, we find a similar pattern to that shown previously for Cohort-B: the explanatory power increases as we move from warm glow to pure altruism, and even more so for impure altruism. But the pattern differs in relation to the additional explanatory power

[19] It is an aside but worth mentioning that none of our Cohort-B subjects mirror the distribution of allocation choices of the previous Cohort-A subject. Thus, while we have shown that Cohort-B subjects respond to the social comparisons, they do not abdicate their allocation responsibilities and simply mirror the behavior of others.

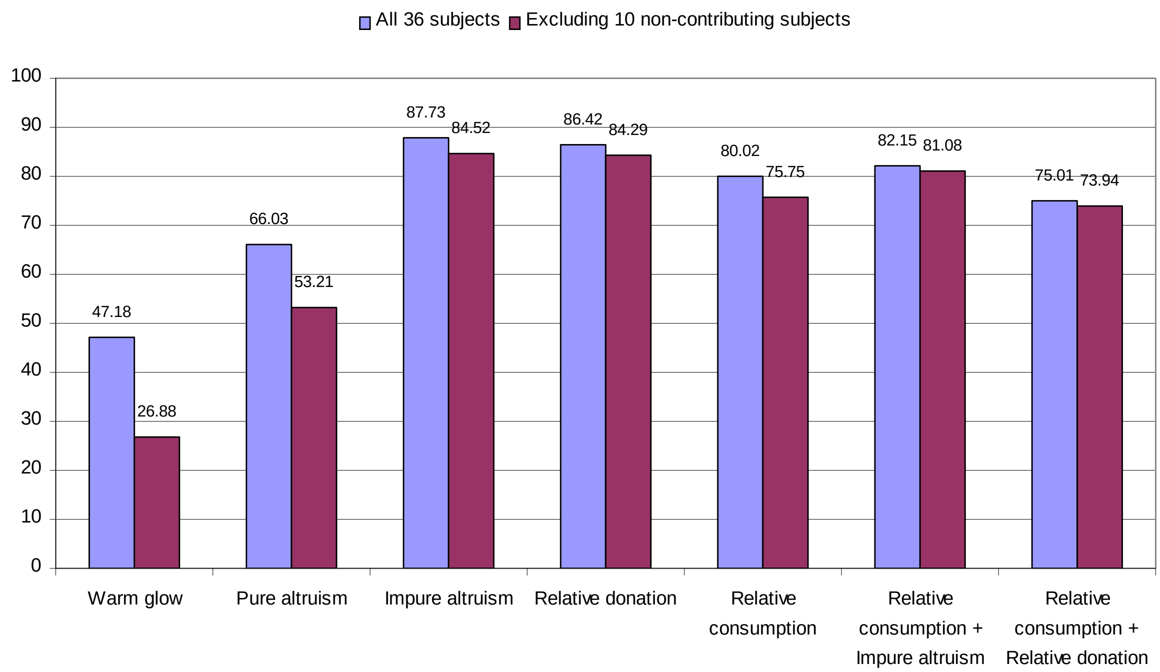

Figure 4: Power adjusted $(r - a)$ percent of Cohort-A subjects that are consistent with each model, by all subjects and excluding non-contributing subjects.

of models that account for relative donation, relative consumption, or both. We find that after making the Bonars adjustments these modes have less explanatory power than impure altruism. This result, of course, differs from that for Cohort-B, where relative consumption and donation added significant explanatory power. We interpret the contrasting results between cohorts as further evidence in support of the finding that social comparisons—based on choices made by others in a similar environment—are important explanatory factors of charitable giving.

5 Conclusion

The methodological contribution in this paper is a general, revealed-preference approach for testing models of charitable giving. The approach differs from standard tests of GARP because it accommodates the characteristics approach to specifying utility functions and externalities imposed by other agents—two features that are common to most models of charitable giving. At the most general level, the approach requires only that utility functions be concave, weakly increasing, and continuous in the externalities. But, as we have shown, imposing further structure is both reasonable and intuitive in the context of charitable giving, and it leads to nontrivial testable restrictions that increase the value of revealed-preference tests.

While we have considered a number of standard and novel preference structures throughout the paper, many more are possible and readily accommodated within our framework. We therefore hope the techniques demonstrated here open the door to greater use of revealed-preference analysis in future research on charitable giving. Towards this end, it is worth mentioning that the revealed-preference techniques described herein can also be applied to cross-sectional data sets, whereby one can pool observations that are similar in observables to form the analog of a panel used here. Moreover, our main theorem for conducting revealed-preference tests is even more general than our experimental application: while the of our experiment is based on the second player in a two-player sequential game, the theory applies equally to multi-agent simultaneous games.

Our experimental design shows how revealed preferences can be used to test several models on the same data set. The results provide evidence on the importance of social comparisons for understanding decisions about the level of one's charitable contributions. Some of our results are consistent with those of Korenok *et al.* (2011) on the importance of impure altruism, but our application considers actual charitable giving rather than payoffs to another laboratory subject. In terms of explaining subject choices, we find that impure altruism performs markedly better than the special cases of warm-glow giving and pure altruism. The more novel findings, however, are that models based on relative preferences for either private consumption or levels of donation yield meaningful differences in the explanatory power over and above the standard model of impure altruism. Specific features of our experimental design and revealed-preference tests also ensure that subject behavior is not being driven by other social motives such as signaling about charity quality, prestige, and signaling about income. Notably, because subjects receive 20 choice scenarios in randomized order on the same sheet, key features such the charitable cause and choice environment are held constant, leaving social comparisons as the only variable other than standard parameters that affect budget constraints.

Finally, we conclude with reasons why one might expect social comparisons to play even a more important role on charitable behavior outside our laboratory setting. On the surface, it may seem that our experiment is biased against a finding that social comparisons are important. Subjects in Cohort-B, who provide the main results, are informed about the choices of a "previous participant," and it is quite reasonable to expect that subjects would want to resist manipulation of their choices based on the assumption that the experimenter selected particularly altruistic choices of the previous participant. In addition, the previous participant is anonymous, and studies in other settings have shown the importance of social comparisons when there is a more targeted group identity (e.g., Frey and Meier 2004; Shang and Croson 2006, 2009). It is thus compelling to expect that the importance of social comparisons is even more pronounced in real-world settings. We expect this is particularly true for preferences that include relative consumption, as they allow individuals to respond not only to the donations of others, but to others' level of philanthropy. To the best of knowledge, this effect has been largely ignored both in experimental and field work on charitable giving, and we think it provides an important subject for future research.

Appendix A: Proof of Theorem 1

This appendix provides a formal proof of Theorem 1 in the main text. We prove the result for an arbitrary agent i, recognizing that the same argument can be extended to all other agents.

Proof that (1)$\Longrightarrow$(2): We show that existence of a utility function U_i that rationalizes the data implies a solution to the inequalities. Utility maximization requires that

$$(x_i^t, y_i^t) \in \arg\max_{x_i, y_i} \left\{ U_i \left[x_i, c_i \left(x_i, \mathbf{x}_{-i}^t \right), y_i + Y_{-i}^t, d \left(y_i, \mathbf{y}_{-i}^t \right) \right] : p_x^t x_i + p_y^t y_i \leq p_x^t x_i^t + p_y^t y_i^t \right\}.$$

The observed choices (x_i^t, y_i^t) must satisfy the first-order conditions

$$x_i^t : U_{i1}^t + U_{i2}^t c_{i1}^t \leq \mu^t p_x^t$$

and

$$y_i^t : U_{i3}^t + U_{i4}^t d_{i1}^t \leq \mu^t p_y^t,$$

where numerical subscripts indicate partial derivatives with respect to the corresponding argument, and μ^t is the Lagrangrian multiplier for the budget constraint. If the utility function is not differentiable, derivatives can be replaced with subderivatives, which will exist because U_i is concave. Each of the first-order conditions will hold with equality if x_i^t and y_i^t are positive, respectively. We define the following: $\kappa_i^t = U_{i1}^t$, $\pi_i^t = U_{i2}^t$, $\gamma_i^t = U_{i3}^t$, $\eta_i^t = U_{i4}^t$, and $\lambda_i^t = \mu^t$. Note that $\kappa_i^t, \pi_i^t, \gamma_i^t, \eta_i^t \geq 0$ and $\lambda_i^t > 0$. Substituting these parameter values into the first-order conditions proves two of the three inequalities. For the final inequality, concavity of U_i implies that for any t and t', it must hold that

$$U_i^{t'} \leq U_i^t + \kappa_i^t [x_i^{t'} - x_i^t] + \pi_i^t [c_i^{t'} - c_i^t] + \gamma_i^t [Y^{t'} - Y^t] + \eta_i^t [d_i^{t'} - d_i^t],$$

which completes this direction of the proof.

Proof that (2)$\Longrightarrow$(1): We define agent i's utility function as follows:

$$\begin{aligned} U_i &= U_i \left[x_i, c_i \left(x_i, \mathbf{x}_{-i} \right), y_i + Y_{-i}, d_i \left(y_i, \mathbf{y}_{-i} \right) \right] \\ &= \min_{1 \leq t \leq T} \{ U_i^t + \kappa_i^t (x_i - x_i^t) + \pi_i^t \left[c_i \left(x_i, \mathbf{x}_{-i} \right) - c_i^t \right] \\ &\quad + \gamma_i^t [(y_i + Y_{-i}) - (y_i^t + Y_{-i}^t)] + \eta_i^t [d_i \left(y_i, \mathbf{y}_{-i} \right) - d_i^t] \}. \end{aligned}$$

This function is concave in all arguments because it is the lower envelope of linear functions. It is standard to show that $U_i \left(x_i^t, c_i \left(x_i^t, \mathbf{x}_{-i}^t \right), y_i^t + Y_{-i}^t, d_i \left(y_i^t, \mathbf{y}_{-i}^t \right) \right) = U_i^t$ as follows. By definition, there is some $1 \leq t' \leq T$ such that

$$\begin{aligned}
U_i\left[x_i^t, c\left(x_i^t, \mathbf{x}_{-i}^t\right), y_i^t + Y_{-i}^t, d\left(y_i^t, \mathbf{y}_{-i}^t\right)\right] &= U_i^{t'} + \kappa_i^{t'}[x_i^t - x_i^{t'}] + \pi_i^{t'}[c_i^t - c_i^{t'}] \\
&\quad + \gamma_i^{t'}[(y_i^t + Y_{-i}^t) - (y_i^{t'} + Y_{-i}^{t'})] + \eta_i^{t'}[d_i^t - d_i^{t'}] \\
&\leq U_i^t + \kappa_i^t[x_i^t - x_i^t] + \pi_i^t[c_i^t - c_i^t] \\
&\quad + \gamma_i^t[(y_i^t + Y_{-i}^t) - (y_i^t + Y_{-i}^t)] + \eta_i^t[d_i^t - d_i^t] \\
&= U_i^t,
\end{aligned}$$

where the inequality cannot be strict because it would violate the first inequality of condition (2). We observe that since c_i and d_i are concave in the first argument, for any $(\mathbf{x}_{-i}, \mathbf{y}_{-i})$, the following inequalities must hold for all $(x_i', y_i'), (x_i'', y_i'')$:

$$c_i(x_i'', \mathbf{x}_{-i}) - c_i(x_i', \mathbf{x}_{-i}) \leq c_{i1}(x_i', \mathbf{x}_{-i})[x_i'' - x_i'],$$

$$d_i(y_i'', \mathbf{y}_{-i}) - d_i\left(y_i', \mathbf{y}_{-i}\right) \leq d_{i1}(y_i', \mathbf{y}_{-i})[y_i'' - y_i'].$$

To complete the proof, we must now show that the observed choices of agent i for $t = 1, ..., T$ maximize the constructed utility function U_i. Consider any bundle (x_i, y_i) such that $p_x^t x_i + p_y^t y_i \leq p_x^t x_i^t + p_y^t y_i^t$. It follows by definition of U_i and concavity of c_i and d_i that

$$\begin{aligned}
U_i[x_i, c_i(x_i, \mathbf{x}_{-i}^t), y_i + Y_{-i}^t, d_i(y_i, \mathbf{y}_{-i}^t)] &\leq U_i^t + \kappa_i^t[x_i - x_i^t] + \pi_i^t[c_i(x_i, \mathbf{x}_{-i}^t) - c_i^t] \\
&\quad + \gamma_i^t[(y_i + Y_{-i}^t) - (y_i^t + Y_{-i}^t)] + \eta_i^t[d_i(y_i, \mathbf{y}_{-i}^t) - d_i^t] \\
&\leq U_i^t + [\kappa_i^t + \pi_i^t c_{i1}^t][x_i - x_i^t] + [\gamma_i^t + \eta_i^t d_{i1}^t][y_i - y_i^t].
\end{aligned}$$

The second two inequalities of condition (2) imply that either $\kappa_i^t + \pi_i^t c_{i1}^t = \lambda_i^t p_x^t$ or $x_i^t = 0$, and that either $\gamma_i^t + \eta_i^t d_{i1}^t = \lambda_i^t p_y^t$ or $y_i^t = 0$. Substituting these expressions into the previous inequality yields

$$\begin{aligned}
U_i\left[x_i, c_i\left(x_i, \mathbf{x}_{-i}^t\right), y_i + Y_{-i}^t, d_i\left(y_i, \mathbf{y}_{-i}^t\right)\right] &\leq U_i^t + \lambda_i^t[(p_x^t x_i + p_y^t y_i) - (p_x^t x_i^t + p_y^t y_i^t)] \\
&\leq U_i^t.
\end{aligned}$$

But because a utility of U_i^t can be achieved by choosing (x_i^t, y_i^t), the inequality shows that (x_i^t, y_i^t) is a best response, and this completes the proof.

Appendix B: Instruction Sheet for Cohort-A

Welcome

This is an experiment about decision making. You will be paid for participating, and the amount of money you will earn depends on the decisions that you make. Your decisions will also affect the amount of money the experimenters will donate to the Hoosic River Watershed Association (HooRWA), a local non-profit dedicated to the restoration, conservation and enjoyment of the Hoosic River and its watershed, through education, research, and advocacy. The entire experiment should be completed within an hour. At the end of the experiment you will be paid privately and in cash for your decisions. A research foundation has provided the funds for this experiment.

Claim Check

In a few moments, you will choose an envelope. In it, you will find a **Choice Sheet** with a unique number in the upper-right corner, and your **Claim Check**, a purple slip of paper with this same number. Each participant has a different number. At the end of the experiment, your money will be in an envelope whose number matches your Claim Check.

Your Identity

We asked for your name when you arrived so we could keep track of who has participated. We will not ask you again to reveal your identity. Neither the experimenters nor the other participants will be able to link you to any of your decisions. In order to keep your decisions private, *please do not reveal your choices or claim check number to any other participant.*

Your Choices

In this session, each subject has a match. Let us call you Participant B. A subject will be randomly selected to be your match, Participant C. You will be the match of 1 different randomly selected participant, Participant A. You will not know the identity of either of these participants.

You will make a choice in each of 20 scenarios. At the end of the experiment, one of the 20 scenarios for each participant will be randomly selected for payment. The choice that you made in the selected scenario determines the number of points you and HooRWA earn. Based on which scenario is chosen, *but not the choice you make in that scenario*, Subject C (your match) also earns points. Your total points are the sum of the points you receive in your selected scenario as well as the points you receive in Subject A's selected scenario. Every point is worth 10 cents. For example if you earn 58 points, you earn \$5.80. If HooRWA earns 58 points, the experimenter donates \$5.80 to HooRWA.

In each scenario, HooRWA and your match each start with a number of points. These points vary from scenario to scenario. You are given a set of tokens, and you are asked to divide the tokens between yourself and HooRWA. As you divide the tokens, you and HooRWA each earn additional points. The number of tokens you have to divide, as well as the number of points you earn for each token you hold and the number of points HooRWA earns for each token you pass to HooRWA, varies from scenario to scenario.

To summarize, you choose how to divide tokens in each of 20 scenarios, where each scenario specifies:

- The number of points your match earns if this scenario is chosen for payment.
- The number of points HooRWA starts with.
- The number of tokens you have to divide between yourself and HooRWA.
- The number of points you get for each token you hold.
- The number of additional points HooRWA gets for each token you pass to HooRWA.

Example

Initial points		Tokens	
Match	HooRWA	to divide	Your choice
40	20	50	*Hold* ___@1 point each, and *Pass* ___@2 points each

In this example, HooRWA starts with 20 points, and your match has 40 points. You must divide 50 tokens, and your match receives 40 points regardless of how you divide these tokens. You can keep all the tokens, keep some and pass some to HooRWA, or pass all the tokens. In this example, you will receive 1 point for every token you hold, and HooRWA will receive an additional 2 points for every token you pass. Remember, each point is worth $0.10. For example:

- If you hold 50 tokens and pass 0:
  - Your match winds up with 40 points, and earns $4.00 if this scenario is selected for payment.
  - You receive 50 points from 50 tokens at 1 point each, and thus earn $5.00 if this scenario is selected for payment.
  - HooRWA receives no additional points. As HooRWA started with 20 points, it earns $2.00 if this scenario is selected for payment.
- Alternatively, if you hold 0 tokens and pass 50:
  - Your match winds up with 40 points, and earns $4.00 if this scenario is selected for payment.
  - You receive no points, and thus earn $0.00 if this scenario is selected for payment.
  - HooRWA receives 100 additional points from 50 tokens at 2 points each. As HooRWA started with 20 points, it now has 120 points and earns $12.00 if this scenario is selected for payment.
- You could, however, choose any number between 0 and 50 to hold. For example, if you hold 29 tokens and pass 21:
  - Your match winds up with 40 points, and earns $4.00 if this scenario is selected for payment.
  - You receive 29 points from 29 tokens at 1 point each, and thus earn $2.90 if this scenario is selected for payment.

– HooRWA receives 42 additional points from 21 tokens at 2 points each. As HooRWA started with 20 points, it now has 62 points and earns \$6.20 if this scenario is selected for payment.

Further Examples

In a few moments, your computer screen will display a couple of further examples. In each case, we specify a hypothetical distribution of tokens, and ask you to calculate the final number of points and earnings for you, HooRWA, and your match. If you have any questions about the examples or how to calculate the final number of points or earnings, please raise your hand and an experimenter will assist you. Once everyone has successfully answered all of the questions, we shall finish the instructions.

For both these examples and the choices you will make later in the session, please feel free to use your calculator, or the one provided by the experimenter, to calculate points and to assure that all of the tokens have been allocated.

How You Will Be Paid

The payment procedure satisfies three goals: 1) Ensure that you get the appropriate earnings; 2) Ensure that you are confident that we follow the promised procedures; and 3) Ensure that it will never be possible to link your identity to your choices. To satisfy all of these goals, the payment procedure has a number of steps. I encourage you to ask any questions you have after I describe the payment procedure.

First, I point out that all participants make decisions for the same 20 scenarios, although each participant's Choice Sheet lists them in a different random order.

After you have made all 20 choices, I am going to ask you to place your Choice Sheet back in the envelope, making sure you keep your Claim Check. After all participants have done so, I will pass around a box and ask you to place your envelope in the box. I will then shuffle all of the envelopes, and ask for one volunteer to monitor the next steps.

The monitor and I will take envelopes to the next room. Based on the order of the shuffled forms, we shall use a table of random numbers to select one of each participant's choices to carry out. For the scenario chosen for your form, you get your points for the tokens you hold, HooRWA gets its points (initial points plus points for tokens you pass), and the subject whose Choice Sheet follows yours in the shuffled stack gets the points indicated for Match. This means that your total points for the session will be the sum of your points on the randomly selected scenario chosen for your Choice Sheet and Match's points on the randomly selected scenario chosen for the preceding Choice Sheet. You get 10 cents per point, and we will put your money in an envelope, write your Claim Check number on it, and seal it. We will also add up all of HooRWA's points. I will write a check to HooRWA, place it in a stamped envelope, and give it to the monitor to mail.

The monitor and I will return to this room. I will pass around the stack of envelopes, asking each participant to take the envelope corresponding to his or her claim check, but not yet open it. Once every participant has the envelope corresponding to his or her claim check, the session is over, and you are free to leave and of course, open your envelope.

Next Steps

Are there any questions? I will now let you choose an envelope in which you will find your Claim Check and Choice Sheet. Once you have put your Claim Check in a safe place, you may start making your choices. I ask that for each scenario, you make sure the number of tokens you hold plus the number you pass adds up to the number you have to divide. After you have made all of your choices, please put your Choice Sheet back in the envelope. I ask that when you are done you click the OK button on your computer monitor. This will allow me to know when all participants have finished.

Appendix C: Instruction Sheet for Cohort-B

The instruction sheet for Cohort-B was similar to that for Cohort-A . The organization of sections was identical. Here we include only those sections that differed for Cohort-B.

Your Choices

In this session, you will make a choice in each of 20 scenarios. At the end of the experiment, one of the 20 scenarios for each participant will be randomly selected for payment. The choice that you made in the selected scenario determines the number of points you and HooRWA earn. Every point is worth 10 cents. For example if you earn 58 points, you earn $5.80. If HooRWA earns 58 points, the experimenter donates $5.80 to HooRWA.

In each scenario, HooRWA starts with a number of points, which varies from scenario to scenario. You are given a set of tokens, and you are asked to divide the tokens between yourself and HooRWA. As you divide the tokens, you and HooRWA each earn points. The number of tokens you have to divide, as well as the number of points you earn for each token you hold and the number of additional points HooRWA earns for each token you pass to HooRWA, varies from scenario to scenario.

HooRWA's initial points come from actual choices participants made in a previous session. These participants had the same number of tokens to divide, faced the same number of points for each token held and passed, and had 1 of their 20 choices selected for payment for themselves and HooRWA. For each scenario, we selected one choice from a set of previous choices to match with each of yours. Your Choice Sheet will explicitly indicate for each scenario how a previous participant decided to divide his or her tokens in that scenario.

To summarize, you choose how to divide tokens in each of 20 scenarios, where each scenario specifies:

- The number of tokens you have to divide between yourself and HooRWA.
- The number of points you get for each token you hold.
- The number of additional points HooRWA gets for each token you pass to HooRWA.
- The number of points HooRWA starts with, which is equal to the number of tokens a randomly selected previous participant chose to pass times the number of points HooRWA gets for each token.
- The number of tokens a randomly selected previous participant chose to hold, and the resulting number of points that subject earned if the scenario was selected for payment.

Please note that while we are using the actual choices of previous participants to determine the number of points HooRWA starts with in each scenario, we will not use your choices in any subsequent session.

Example

Previous Participant Choice		Initial HooRWA Points	Tokens To Divide
Held	HooRWA		
Points/Tokens	Points/Tokens		
40 points / 40 tokens	20 points / 10 tokens	20	50

Your choice
Hold ___@1 point each, and *Pass* ___@2 points each

In this example, HooRWA starts with 20 points, and you must divide 50 tokens which are worth 1 point each to you and 2 points each to HooRWA. You can keep all the tokens, keep some and pass some to HooRWA, or pass all the tokens. Remember, each point is worth \$0.10. When faced with this scenario, a subject in a previous session held 40 tokens and passed 10 tokens, resulting in 40 points (\$4.00) for the participant and 20 points (\$2.00) for HooRWA. Therefore, HooRWA starts with 20 points in your scenario. The amount you and HooRWA earn depend on your choice. For example:

- If you hold 50 tokens and pass 0:
  - You receive 50 points from 50 tokens at 1 point each, and thus earn \$5.00 if this scenario is selected for payment.
  - HooRWA receives no additional points. As HooRWA started with 20 points, it earns \$2.00 if this scenario is selected for payment.
- Alternatively, if you hold 0 tokens and pass 50:
  - You receive no points, and thus earn \$0.00 if this scenario is selected for payment.
  - HooRWA receives 100 additional points from 50 tokens at 2 points each. As HooRWA started with 20 points, it now has 120 points and earns \$12.00 if this scenario is selected for payment.
- You could, however, choose any number between 0 and 50 to hold. For example, if you hold 29 tokens and pass 21:
  - You receive 29 points from 29 tokens at 1 point each, and thus earn \$2.90 if this scenario is selected for payment.
  - HooRWA receives 42 additional points from 21 tokens at 2 points each. As HooRWA started with 20 points, it now has 62 points and earns \$6.20 if this scenario is selected for payment.

How You Will Be Paid

[Only the third paragraph differed in this section.] The monitor and I will take envelopes to the next room. Based on the order of the shuffled forms, we shall use a table of random numbers to select one of each participant's choices to carry out. For the scenario chosen for your form, you get your points for tokens you hold and HooRWA gets its points (initial points plus points for tokens you pass). You get 10 cents per point plus your $5 participation fee, and we will put your money in an envelope, write your Claim Check number on it, and seal it. We will also add up all of HooRWA's points. I will write a check to HooRWA, place it in a stamped envelope, and give it to the monitor to mail.

References

Afriat, S. (1967) "The Construction of a Utility Function from Expenditure Data." *International Economic Review*, 8:67-77.

Afriat, S. (1972) "Efficiency Estimates of Production Functions." *International Economic Review*, 13:568-598.

Andreoni, J. (1988) "Privately Provided Public Goods in a Large Economy: The Limits of Altruism." *Journal of Public Economics*, 35:57-73.

Andreoni, J. (1989) "Giving With Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy*, 97:1447-1458.

Andreoni, J. (1990) "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *Economic Journal*, 100:464–77.

Andreoni, J. and J. Miller. (2002) "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica*, 70:737-53.

Aronsson,T. and O. Johansson-Stenman. (2008) "When the Joneses' Consumption Hurts: Optimal Public Good Provision and Nonlinear Income Taxation" *Journal of Public Economics*, 92:986-97.

Beatty, T. K. M. and I. A. Crawford. (2011) "How Demanding Is the Revealed Preference Approach to Demand?" *American Economic Review*, 101:2782-95.

Bergstrom, T. C., L. E. Blume, and H. R. Varian. (1986) "On the Private Provision of Public Goods." *Journal of Public Economics*, 29:25-49.

Blow, L., M. Browning, and I. Crawford. (2008) "Revealed Preference Analysis of Characteristics Models." *Review of Economic Studies*, 75:371-89.

Boskin. M. J. and E. Sheshinski. (1978) "Optimal Redistributive Taxation When Individual Welfare Depends Upon Relative Income." *Quarterly Journal of Economics*, 92:589-601.

Bronars, S. G. (1987) "The Power of Nonparametric Tests of Preference Maximization [The Nonparametric Approach to Demand Analysis]." *Econometrica*, 55:693–98.

Carvajal, A. (2010) "The Testable Implications of Competitive Equilibrium in Economies with Externalities." *Economic Theory*, 45:349-78.

Cornes, R. and T. Sandler. (1984) "Easy Riders, Joint Production, and Public Goods." *Economic Journal,* 94:580-98.

Cornes, R. and T. Sandler. (1996) *The Theory of Externalities, Public Goods and Club Goods*, 2nd edition. New York: Cambridge University Press.

Deb, R. (2009) "A Testable Model of Consumption with Externalities." *Journal of Economic Theory*, 144:1804-16.

Duesenberry, J. (1949) *Income, Savings and the Theory of Consumer Behavior*. Cambridge: Harvard University Press.

Fischbacher, U. (2007) "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics*, 10:171-8.

Frank, R. (1985) *Choosing the Right Pond: Human Behavior and the Quest for Status*. New York: Oxford University Press.

Frank, R. (1999) *Luxury Fever: Money and Happiness in a Era of Excess*. New York: The Free Press.

Frey, B. S. and S. Meier. (2004) "Social Comparisons and Pro-social Behavior." *American Economic Review*, 94: 1717:22.

Friedman, D. and S. Sunder. (1994) *Experimental Methods: A Primer for Economists*. Cambridge: Cambridge University Press.

Fisman, R., S. Kariv, and D. Markovits. (2007) "Individuals Preferences for Giving." *American Economic Review*, 97:1858-76.

Glazer, A. and K. A. Konrad. (1986) "A Signalling Explanation for Charity." *American Economic Review,* 86:1019-28.

Greiner, B. (2004) "An Online Recruitment System for Economic Experiments," in Kurt Kremer and Volker Macho, eds.,*Forschung und wissenschaftliches Rechnen 2003*, Vol. 63 of GWDG-Bericht, Göttingen, Germany: Gesellschaft für wissenschaftliche Datenverarbeitung mbh, pp. 79–93.

Harbaugh, W. T. (1998) "What Do Donations Buy? A Model of Philanthropy Based on Prestige and Warm Glow." *Journal of Public Economics,* 67:269-84.

Hoffman, E., K. A. McCabe, and V. L. Smith. (1996) "Social Distance and Other-Regarding Behavior in Dictator Games." *American Economic Review*, 86:653-60.

Hollander, H. (1990) "A Social Exchange Approach to Voluntary Cooperation." *American Economic Review,* 80:1157-67.

Jamison, J., D. Karlan, and L. Schechter. (2008) "To Deceive or Not to Deceive: The Effect of Deception on Behavior in Future Laboratory Experiments." *Journal of Economic Behavior & Organization*, 68:477-88

Korenok, O., E. L. Millner, and L. Razzolini. (2011) "Impure Altruism in Dictators' Giving." Working paper. Department of Economics, Virginia Commonwealth University.

Lancaster, K. (1971) *Consumer Demand: A New Approach.* New York: Columbia University Press.

Layard, R. (1980) "Human Satisfactions and Public Policy." *Economic Journal*, 90:737-50.

List, J. A. (2011) "The Market for Charitable Giving." *Journal of Economic Perspectives*, 25:157-80.

Ljungqvist, L. and H. Uhlig. (2000) "Tax Policy and Aggregate Demand Management under Catching up with the Joneses." *American Economic Review*, 90:356-66.

Luttmer, E. (2005) "Neighbors As Negatives: Relative Earnings and Well-Being." *Quarterly Journal of Economics*, 120:963–1002.

Roberts, R. D. (1984) "A Positive Model of Private Charity and Public Transfers." *Journal of Political Economy*, 92:136-48.

Romano, R. and H. Yildirim. (2001) "Why Charities Announce Donations: A Positive Perspective." *Journal of Public Economics*, 81:423-47.

Rosen, J. B. (1965) "Existence and Uniqueness of Equilibrium Points For Concave N-Person Games." *Econometrica*, 33:520-34.

Shang, J. and R. Croson. (2006) "The Impact of Social Comparisons on Nonprofit Fund Raising," in R. Mark Isaac, Douglas D. Davis (ed.) *Experiments Investigating Fundraising and Charitable Contributors* (Research in Experimental Economics, Volume 11), Bingley: Emerald Group Publishing Limited, 143-156.

Shang, J. and R. Croson. (2009) "A Field Experiment In Charitable Contribution: The Impact of Social Information On The Voluntary Provision of Public Goods." *Economic Journal*, 119:1422-39.

Varian, H. R. (1983) "Non-Parametric Tests of Consumer Behavior." *Review of Economic Studies*, 50:99-110.

Vesterlund, L. (2006) "Why do people give?" in (W. Powell and R.S. Steinberg, eds.), *The Nonprofit Sector: A Research Handbook*, 2nd Edition, 568–87, New Haven: Yale University Press.

Vesterlund, L. (2012) "Voluntary Giving to Public Goods: Moving Beyond the Linear VCM." Working paper. Department of Economics, University of Pittsburgh.

Warr, P. G. (1982) "Pareto Optimal Redistribution and Private Charity." *Journal of Public Economics*, 19:131-8.