Journal of Public Economics 120 (2014) 181–192

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Testing motives for charitable giving: A revealed-preference methodology with experimental evidence☆

Rahul Deb [a,*], Robert S. Gazzale [a], Matthew J. Kotchen [b]

[a] University of Toronto, Canada
[b] Yale University, United States

ARTICLE INFO

Article history:
Received 25 December 2013
Received in revised form 18 September 2014
Accepted 19 September 2014
Available online 28 September 2014

Keywords:
Revealed preference
Charitable giving
Other regarding preferences

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 other-regarding preferences based on relative consumption and donations among individuals. We also conduct an original experiment that enables testing and comparing many models on a single data set. Our experiment design allows us to focus on intrinsic motivations which are often hard to disentangle from other extrinsic or image effects in field data. 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 significantly greater explanatory power than the standard model of impure altruism.

1. Introduction

Why do individuals make charitable contributions? Despite a large economics literature on the subject, important questions remain. Standard explanations of private donations to a public good assume that individuals benefit from the aggregate level of the public good (Bergstrom et al., 1986) and may obtain an additional private benefit – commonly modeled as a "warm glow" – from their own giving (Cornes and Sandler, 1984; Andreoni, 1989, 1990). A notable feature of these standard models is that the contributions of others affect one's own charitable giving in only one way: crowding out through the total amount of the public good provided.

More recently, researchers have recognized that individuals may respond to the donations of others because of extrinsic, social concerns. For example, some studies seek to explain patterns of charitable giving based on reputation, signaling about income, and avoiding social pressure.[1] There is, however, a much smaller body of work on how the actions of others may affect *intrinsic* motives for charitable giving, and this is surprising given evidence on the importance of "other-regarding" preferences in dictator and bargaining games.[2] It is easy to imagine, for example, that an individual's warm glow depends on how her donation compares to the donations of others, while other individuals may be reluctant to contribute if their wealthier peers appear relatively uncharitable. Alternatively, an individual may feel guilty if she donates less than a social norm while simultaneously having no desire to be overly charitable.

In this paper, we develop a theoretical and experimental framework to test whether intrinsic motivations play an important role in charitable giving. In doing so, our aim is not only to show the existence of

☆ We would like to thank Laurens Cherchye, Ian Crawford, Bram de Rock, Thomas Demuynck, three anonymous referees and the editor Tim Cason for their comments which greatly improved the paper. We also gratefully acknowledge the helpful discussions with Jim Andreoni, Matt Polisson, Laura Razzolini and numerous seminar audiences.

* Corresponding author at: 150 St. George St., Toronto ON M5S 3G7, Canada.
E-mail addresses: rahul.deb@utoronto.ca (R. Deb), robert.gazzale@utoronto.ca (R.S. Gazzale), matthew.kotchen@yale.edu (M.J. Kotchen).

[1] See for example Hollander (1990), Glazer and Konrad (1986), Harbaugh (1998), Benabou and Tirole (2006), Ariely et al. (2009), and DellaVigna et al. (2012).

[2] Notable exceptions are Shang and Croson (2006, 2009) and Charness and Cheung (2013) that report the results of field experiments to study social comparisons. The former find that donors to a public radio station tend to adjust their contribution levels toward that of the social comparison. The latter show that varying the suggested contribution amount on a donation jar nonmonotonically affects contributions. Our analysis in the present paper, as we will show, is complementary in that we provide a close link to theory and show how tests for the importance of social comparisons (based on donations, private consumption, or both) can exploit multiple choices of the same individual rather than a cross section among individuals.

http://dx.doi.org/10.1016/j.jpubeco.2014.09.009
0047-2727/

such motives, but, additionally, to consider whether they are compatible with well-behaved preferences. We thus develop a revealed-preference methodology for testing different models of charitable giving. On one level, our theoretical framework nests the standard models of pure altruism, warm glow, and impure altruism (Bergstrom et al., 1986; Andreoni, 1989, 1990). More importantly, however, the framework readily accounts for models with other-regarding preferences based on concerns about relative donations, relative consumption, or both. Specifically, our specification of models, as we will show, allows for social comparisons in the well-established equity framework of Bolton and Ockenfels (2000) and the inequality framework of Fehr and Schmidt (1999).

We also provide evidence from a laboratory experiment showing that models incorporating other-regarding preferences are significantly better at explaining observed donations than the standard models. Our experimental approach exploits variation in a series of choices by each individual about charitable giving over different budget sets, relative prices, and information about the choices of others. With these data, we evaluate models based on whether there exists a concave and increasing utility function for each individual such that all of the observed choices are consistent with utility maximization. Hence, the revealed-preference tests on our experimental data provide "pure" tests of the various models, as the methodology is nonparametric, allows for complete heterogeneity across individuals, and is free of measurement error. While this procedure enables testing each model of charitable giving individually, we also show how specification tests are possible among models. Our statistical tests among competing models, as we will explain, account for differences in the power of revealed-preference tests according to adjustments proposed by Selten (1991) and Beatty and Crawford (2011).

Andreoni and Miller (2002) were the first to use revealed preferences to test for a particular form of altruism as an intrinsic motivation. 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, respectively. 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.

Notwithstanding these results, the use of revealed preferences to test a broad set of motives for charitable giving poses new challenges. These arise because classical revealed-preference techniques do not readily accommodate two features that are central not only to models of charitable giving, but also to models of other-regarding preferences. One is externalities among agents, and the other is the Lancasterian characteristics within utility functions (Lancaster, 1971). Consider a simple demonstrative model involving two individuals with preferences $U_i(x_i, y_i - y_j, y_i + y_j)$ where x_i is private consumption, and y_i and y_j are the respective individual's contributions to a public good. This is essentially a characteristics model because one's own contribution enters the utility function in two places: the third argument as a standard contribution to the public good, and the second argument as a component of concern about relative donations. Negative and positive externalities are also present in the second and third arguments, respectively. As we will show, the two features of characteristics and externalities are fundamental to both standard models of public good provision and alternative models that incorporate social comparisons based on relative donations, consumption, or both.

A significant contribution of this paper is that we address new methodological challenges with the development of a revealed-preference approach for testing a range of models on charitable giving that include externalities and Lancasterian characteristics. The approach builds on recent innovations in revealed-preference theory that allow for both externalities (Carvajal, 2010; Deb, 2009) and characteristics (Blow et al., 2008) into a standard model of consumption. Empirically, the tests are simple linear programs that are transparent, computationally efficient, and straightforward to implement.

Our experimental results demonstrate the applicability of our revealed-preference framework and highlight the importance of other-regarding preferences as intrinsic motivation for charitable contributions. A distinct feature of our experiment 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 local, charitable organization. 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. Fundamental to our experiment design is that the subjects of primary interest are 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 different choice models. Additionally, the laboratory setting allows us to focus on intrinsic motivations that are often difficult to disentangle from extrinsic effects in the field.[3] This is because our analysis is conducted separately on the choices made by each subject on multiple decision problems, which ensures that all other extrinsic factors such as beliefs about the quality of the charity, total donations by other subjects and donors *are held constant across the choice scenarios*.

We find new evidence on the importance of social comparisons as an intrinsic motivator for voluntary donations. Regarding the standard models, and after making power adjustments for revealed-preference tests, we find that impure altruism performs significantly better than the special cases of warm-glow giving and altruism consistent with provision of a pure public good. Importantly, however, impure altruism performs less well than alternative models based on intrinsic concerns about relative donations or relative private consumption. 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. While we consider a range of models in support of this conclusion, a strength of our methodology is its usefulness for revealed-preference analysis beyond the particular cases considered here. Indeed, we hope that our novel approach combined with the evidence herein will further research on the underlying motives for charitable contributions.

2. Theoretical framework

In this section, we develop our theoretical framework. While our experiment focuses on testing whether a single agent is best responding, we present the model in its full generality by allowing for multiple agents. We begin with the 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 other-regarding 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 $i = 1,\ldots,N$ 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.[4] Prices are denoted p_x and

[3] With respect to public goods, 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. Also consistent with our revealed-preference approach, 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).

[4] The model can be generalized to allow for multiple private and public goods for each agent i.

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

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

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 characteristics that are potentially nonlinear. 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.

While we have defined utility as a function of the total donation amount to the public good $y_i + Y_{-i}$ and not the total level of the public good provided, this does not imply that the agent's preferences are not determined by the latter. (This formulation was chosen to be consistent with our experiment in which we do not observe what subjects perceive to be the public good production technology.) For instance, if the agent thinks that the public good is produced by a linear, constant returns to scale technology, both preferences are clearly equivalent. More generally, note that our utility function allows for subjects to think that the public good is produced by a decreasing returns to scale production technology f_i.[5]

Notice that the utility function in Eq. (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 observed choices. By adopting the utility function in Eq. (1), 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 Eq. (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 Eq. (1) can also accommodate other-regarding preferences 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 $d_i(y_i, \mathbf{y}_{-i}) = y_i - \frac{Y_{-i}}{N-1}$, 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 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,[8] 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.[9] 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.

[5] This is because class of utility functions (1) and

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

are identical for concave f_i. To see one direction, note that taking $f_i(y_i + Y_{-i}) = y_i + Y_{-i}$ implies that the latter class subsumes the former. To see the other direction, note that when f_i is concave and nondecreasing, (2) is concave and weakly increasing in the four arguments $(x_i, c_i, y_i + Y_{-i}, d_i)$. Therefore, the former class subsumes the latter and hence they must be equivalent. In other words, when we are testing a subject's choices for consistency with preferences (1), we are equivalently testing whether there exists any concave f_i such that the subject's choices are consistent with (2).

[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.

[8] See for example Veblen (1899), Duesenberry (1949), Boskin and Sheshinski (1978), Layard (1980), Frank (1985, 1999), Ljungqvist and Uhlig (2000), and Luttmer (2005).

[9] An alternative way to model both relative consumption and donation, which is consistent with the equity framework of Bolton and Ockenfels (2000), 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. These result are similar, yet the models perform somewhat less well. Later in the paper, we also consider the possibility of social comparisons based on difference aversion.

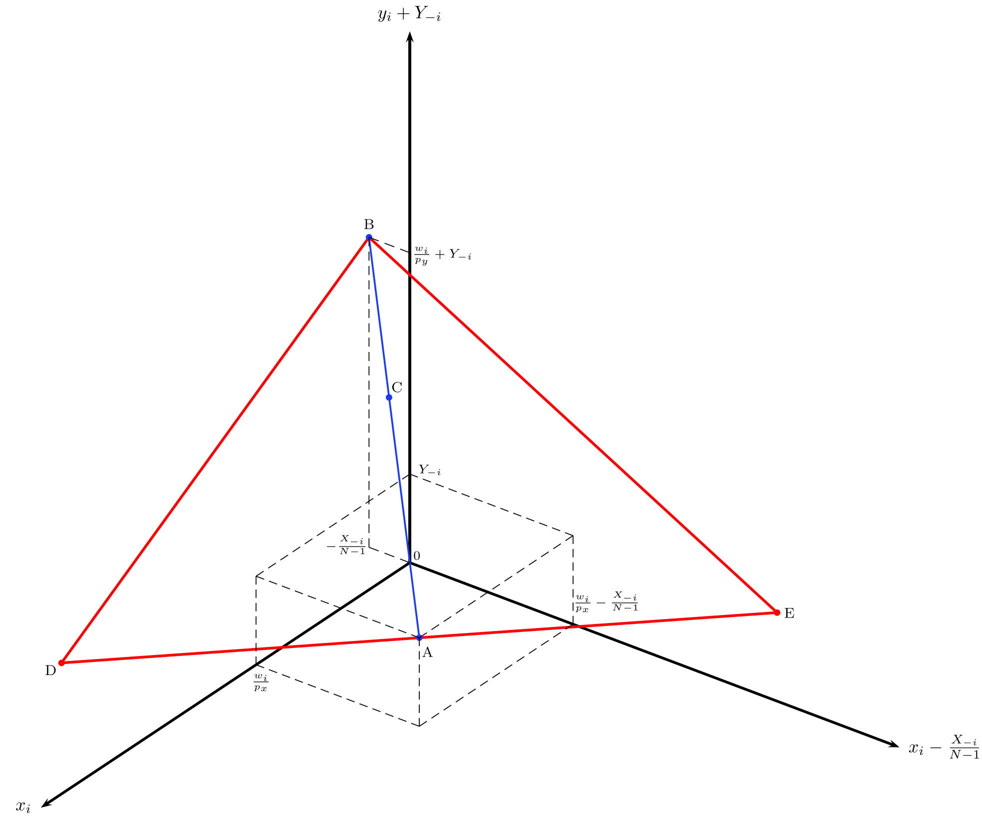

Fig. 1. Budget set in characteristics space for the model of relative consumption.

In our nonparametric framework, when we test for a particular specification of $c_i(\cdot)$ and $d_i(\cdot)$, it implies that we are testing for *all* parametric utility functions that are special cases. Therefore, if the model cannot rationalize the data, it rules out that the subject had *any* preference consistent with the model. For instance, if the data cannot be rationalized by a model with relative consumption $c_i(x_i, \mathbf{x}_{-i}) = x_i - \frac{X_{-i}}{N-1}$, it implies that, in particular, the subject does not have preferences of the Charness and Rabin (2002) form: $\rho \max\left\{x_i - \frac{X_{-i}}{N-1}, 0\right\} + \sigma \max\left\{\frac{X_{-i}}{N-1} - x_i, 0\right\}$ where $\sigma > \rho > 0$. Additionally, note that the requirement that U_i is monotone in the second argument does not imply that c_i has to be monotone in x_i. In other words, our model can accommodate preferences such as inequity aversion (Fehr and Schmidt, 1999). For this, we can consider $c_i(x_i, \mathbf{x}_{-i}) = -\left|x_i - \frac{X_{-i}}{N-1}\right|$ or $c_i(x_i, \mathbf{x}_{-i}) = -\left(x_i - \frac{X_{-i}}{N-1}\right)^2$ which are concave but not increasing in x_i. To summarize, it is possible to incorporate within our model many different other-regarding preferences which feature altruism, relative consumption, inequity aversion, spite, etc.

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 $U_i(x_i, x_i - \frac{X_{-i}}{N-1}, Y)$, Fig. 1 illustrates the budget frontier of an agent that seeks to maximize utility subject to $p_x x_i + p_y y_i \le w_i$. 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 on w_i and the two externalities of Y_{-i} and $\frac{X_{-i}}{N-1}$.[10] Hence, conducting revealed-preference tests in this environment, where any of the exogenous variables $\left(p_x, p_y, w_i, Y_{-i}, \frac{X_{-i}}{N-1}\right)$ 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 Eq. (1). 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

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

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 $i = 1, \ldots, N$, it holds that

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

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

Because the model's setup constitutes a concave game, equilibrium existence is guaranteed (Rosen, 1965).[11] 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 $t = 1, \ldots, T$ 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

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

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

The following theorem, the proof of which is in Appendix A, formally states the conditions that constitute the revealed-preference test of our model. Note that the test can be used to check whether any subset of agents are best responding (however, in our experiment, we test the choices of a single agent).[12]

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 $c_i^t := c_i(x_i^t, \mathbf{x}_{-i}^t)$ and $d_i^t := d_i(y_i^t, \mathbf{y}_{-i}^t)$, the following statements are equivalent:*

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

(2) *The following inequalities have non-negative solutions for $\kappa_i^t, \pi_i^t, \gamma_i^t$, and η_i^t and positive solutions for λ_i^t and U_i^t for all t and t′:*

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

$$\kappa_i^t + \pi_i^t \frac{\partial c_i\left(x_i^t, \mathbf{x}_{-i}^t\right)}{\partial x_i} \le \lambda_i^t p_x^t \text{ holding with equality if } x_i^t > 0,$$

$$\gamma_i^t + \eta_i^t \frac{\partial d_i\left(y_i^t, \mathbf{y}_{-i}^t\right)}{\partial y_i} \le \lambda_i^t p_y^t \text{ holding with equality if } y_i^t > 0.$$

The theorem states that an agent's choices can be rationalized if and only if the 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 $U_i^t = U_i[x_i^t, c_i(x_i^t, \mathbf{x}_{-i}^t), y_i^t + Y_{-i}^t, d_i(y_i^t, \mathbf{y}_{-i}^t)]$ 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 $\kappa_i^t, \pi_i^t, \gamma_i^t$, and η_i^t, which themselves represent marginal utilities for the corresponding characteristics. Simply dividing them by the Lagrange multiplier on the budget constraint, λ_i^t, 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, which is, of course, the revealed preference test.

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 Eq. (1). The inequalities corresponding with the conditions in Theorem 1 are

$$U_i^{t'} \le U_i^t + \kappa_i^t\left[x_i^{t'} - x_i^t\right] + \gamma_i^t\left[Y^{t'} - Y^t\right] + \eta_i^t\left[y_i^{t'} - y_i^t\right],$$

$$\kappa_i^t \le \lambda_i^t p_x^t \text{ holding with equality if } x_i^t > 0,$$

$$\gamma_i^t + \eta_i^t \le \lambda_i^t p_y^t \text{ holding with equality if } y_i^t > 0.$$

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, and 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 Fig. 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.[13] 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

[11] 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.

[12] For simplicity, the theorem assumes that c_i and d_i are differentiable in x_i but this assumption is not required. Since they are assumed to be concave in x_i, they are subdifferentiable which is sufficient for the test.

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

the form of a linear program, and there are well-known algorithms to check feasibility and solve systems of linear inequalities.

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. In this respect, our design is similar to that of Eckel and Grossman (1996) who showed that using a well-known charity rather that anonymous subjects increases the donation amounts in a dictator game. 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).[14]

The fact that we study how subjects respond to the choices of other subjects necessitates an experiment 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 were volunteers among the undergraduate student population at Williams College in Williamstown, Massachusetts, and all sessions took place during May and June 2011, with recruitment using an online system (Greiner, 2004).

Cohort-A subjects were asked to make choices about splitting an endowment between oneself and the Hoosic River Watershed Association (HooRWA), 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) which 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 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.[15] 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 to divide	Your choice
Match	HooRWA		
40	20	50	*Hold* ___@1 point each, and *Pass* ___@2 points each

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

Scenario	Initial points		Tokens to divide	Points per		Average	
	Match	HooRWA		Held	Passed	Tokens passed	Percent 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.

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 1 h, 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 et al., 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_{-i}^t and y_{-i}^t. Because they were not based on the choices of actual subjects, Cohort-A is of limited (though useful, as we will discuss) 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_{-i}^t and y_{-i}^t that conform to the No Deception Rule in economic experiments.[16] In other words, the task of Cohort-B subjects closely mirrored that of Cohort-A 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

[14] 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 $Y_{-i}^t = \mathbf{y}_{-i}^t = y_{-i}^t$.

[15] 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.

[16] 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, Jamison et al. (2008) find that deception affects both a subject's likelihood of returning for subsequent experiments and the choices that are made conditional on returning.

prices. The following summarizes the details and presentation of an example Cohort-B scenario.

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 our notation, this decision problem translates to $w_i = 50$, $p_x = 1$, $p_y = \frac{1}{2}$, $x_{-i} = 40$ and $y_{-i} = 20$. 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 the number of participants ranging from 12 to 17 subjects. 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 that we 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 the 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 the Bronars results with sufficient power to ensure plenty of scope for rejecting models.

Appendix C includes the Cohort-B instruction sheet. 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 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.

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

Scenario	Previous subject's chosen points		Tokens to divide	Points per		Average	
	Consumption	HooRWA		Held	Passed	Tokens passed	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.

Note that, despite both Cohort-A and Cohort-B subjects facing the same budget sets in each choice scenario, relative consumption and relative donation generate distinct preferences. To see this, consider the following example. Suppose a Cohort-B subject was informed that the Cohort-A subject donated exactly the same dollar amount $y^t_{-i} = k$ in every choice scenario. In this case, the lack of variation of y^t_{-i} over t implies that testing for relative donation is equivalent to testing for the simpler model of warm glow or pure altruism. This is because the preferences given by $U_i(x_i, y_i)$, $U_i(x_i, y_i + k)$ and $U_i(x_i, y_i + k, y_i - k)$ are identical (the utility functions only depend on and are increasing and concave in x_i and y_i). However, since the prices and endowments vary across the choice scenarios, the consumption $x^t_{-i} = \frac{w^t_i - k}{p^t_x}$ of the Cohort-A subject is different across the choice scenarios. Therefore, in this case, relative consumption preferences are not identical to warm glow or pure altruism and will thus, have potentially more explanatory power than relative donation preferences. Of course, in the opposite example where x^t_{-i} does not vary, relative donation will be more general. In our experiment design, the variation of both x^t_{-i} and y^t_{-i} over t ensures that the preferences are distinct with neither nesting the other.

We end this section by observing that this experiment design ensures that our tests isolate the intrinsic motivations of the subjects. This is because any extrinsic factors such as beliefs about the quality of the charity, total donations of other Cohort-B subjects or external donors will not differentially impact choices made by a Cohort-B subject across the 20 scenarios. The only thing that changes in each choice scenario as far as a Cohort-B subject is concerned is their budget set and the corresponding action of the Cohort-A subject in that scenario. By presenting all the choice scenarios on the same sheet, any signaling effect about charity quality present in the choices of Cohort A will average out and not differ across the choice scenarios for Cohort-B. Thus by conducting our test separately on the 20 choices of each Cohort-B subject, our focus is precisely on how the intrinsic motivation is affected by Cohort-A choices.

4. Experiment results

We focus analysis of the results on how different models of charitable giving rationalize the choices of our subjects. As discussed previously, we consider standard models of privately provided public goods, along with novel models that account for social comparisons. We also conduct statistical tests to evaluate the relative performance among models. 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 experiment 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

Table 3
Fixed-effects model of demand for donations.

	Coefficients	Elasticities
Price of donation (p_y^t)	-25.762^{***}	-0.419^{***}
	(3.363)	(0.055)
Price of private consumption (p_x^t)	61.222^{***}	0.995^{***}
	(9.758)	(0.159)
Token endowment (w^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.

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 is consistent with the corresponding model, rather than the proportion that is inconsistent. These results are based on 50,000 replications. The percentage of subjects rationalized by each model is based on 119 Cohort-B subjects.

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 experiment 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% of our subjects, compared to the identical percentage and 25.2% for Andreoni and Miller (2002) and Korenok et al. (2011), respectively.[17] 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% of our subjects played this strategy compared to 6.2 and 1.6% 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 2), the token pass rate of our subjects was 20.6% compared to 23.0 and 22.9% 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:

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

where y_{it} is the contribution of subject i in scenario t measured in points; p_y^t is the price of contribution points in terms of tokens (i.e., the inverse of points per passed in Table 2); p_x^t is the price of private consumption points in terms of tokens (i.e., the inverse of points per held in Table 2); w^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 fail the revealed-preference test.

Table 4 lists the different models that we consider, by name and utility function, along with the 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 Bronars 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% are consistent with warm-glow preferences, meaning that 99.96% are inconsistent with the model. With more general utility functions, the power declines, but even for relative consumption + relative donation, only 20.37% 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.[18] These percentages are derived by implementation of the tests based on Theorem 1. We carried out the analysis in MATLAB using a linear program 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 that the models meaningfully explain subject behavior. For example, nearly 50% of the subjects made choices consistent with warm-glow preferences, and relative consumption + relative donation rationalizes the choices of more than 94% of the subjects. These numbers compare to 0.04 and 20.37% for the Bronars 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

[17] 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.

[18] 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.

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 performs 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 for the differing power of tests. Beatty and Crawford (2011) provide 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 (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. [19]

This raises the question, however, about whether the differences in model performance are statistically significant. To make such inferences, it is straightforward to carry out tests of differences in the adjusted percentages rationalized. Indeed, tests between any two models are simple paired t-tests that account for the fact that different models are being tested on the same sample. The tests are two-tailed for distinct models and one-tailed for nested models.

Fig. 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 Fig. 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 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% of the data for warm glow and pure altruism, respectively (Table 4). After making the power adjustments (Fig. 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 are statistically significant ($t = 2.66$, $p < .01$; $t = 2.70$, $p < .01$) and 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.[20]

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 significantly, increasing the fit by 14.6 and 19.2 percentage points beyond pure altruism with and without the "selfish" subjects ($t = 4.32$, $p < .01$; $t = 4.53$, $p < .01$), respectively. The contrast with warm glow is even more substantial, at 22.8 and 29.8 percentage points ($t = 5.75$, $p < .01$; $t = 6.13$, $p < .01$), 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% of the data before making the power adjustment. After the adjustment, the numbers are 80.1 and 75.5% 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

$$\widetilde{U}_i\left[x_i, y_i + Y_{-i}, \alpha(y_i + Y_{-i}) + \beta\left(y_i - \frac{Y_{-i}}{N-1}\right)\right],$$

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 $\widetilde{U}_i = \widetilde{U}_i[x_i, y_i + Y_{-i}, y_i]$. 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 statistically significant ($t = 2.85$, $p < .01$; $t = 3.07$, $p < .01$), 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% of the data. After making the power adjustment, the numbers are 81.1 and 76.9% 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

[19] 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 is 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, implying that the CCEI approach is more ad hoc than employing standard thresholds (e.g., the conventional 95-percent rule).

[20] Throughout the text, we report the results of paired t-tests for cases of particular interest. The complete set of all paired results are available upon request.

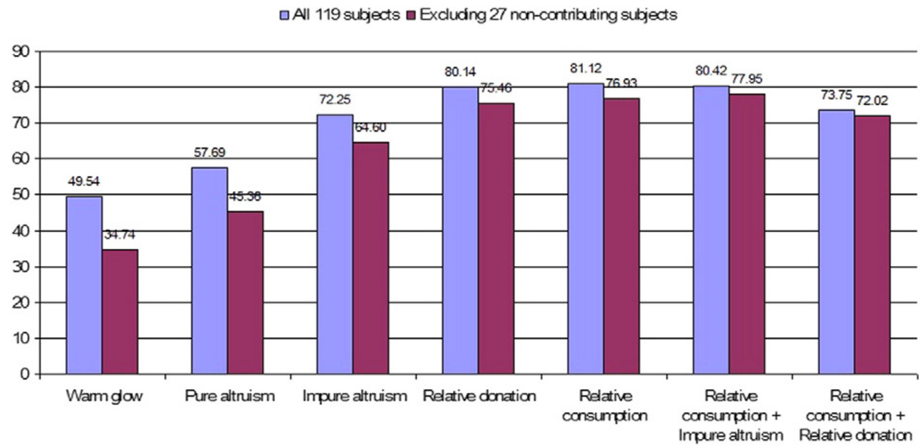

Fig. 2. Power adjusted (*r-a*) percent of Cohort-B subjects that are consistent with each model, excluding non-contributing subjects.

differences of 8.9 and 12.3 percentage points with and without the "selfish" subjects ($t = 2.99, p < .01$; $t = 3.27, p < .01$). We interpret these results as significant evidence that how one's level of philanthropy compares to others, scaled by income, helps to explain decisions about charitable giving. This is because both relative giving and endowments enter the model implicitly via y^t_{-i} and x^t_{-i} that each agent i observes when making allocation choices.[21]

The other results shown in Table 4 and Fig. 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% of the data (Table 4). But after making the power adjustments, these models have less explanatory power than relative consumption on its own (Fig. 2). We conclude, 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, tests between 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. We find little variation in the explanatory power within models and no change in the comparisons across models based on the mean results (details are in the working paper version). 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 experiment design, while they were reported (without deception) as the result of a previous subject's choices to Cohort-B.[22] 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.

The power adjusted results for Cohort-A are as follows (details are in the working paper version). 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 of models that account for relative donation, relative consumption, or both. We find that after making the power 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 other-regarding preferences – 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

[21] 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 consistent with Fehr and Schmidt (1999) of the form $U_i\left[x_i, -(x_i - \frac{X_{-i}}{N-1})^2, Y\right]$ 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 (which yields very similar results), and this underscores the generality of our framework for testing many types of models.

[22] 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.

the characteristics approach to specifying utility functions and externalities imposed by other agents—two features that are common to most models of charitable giving, as well as models with other-regarding preferences. 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, this further structure is both reasonable and intuitive in the context of charitable giving, and it leads to nontrivial testable restrictions enabling meaningful 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 that the techniques demonstrated here open the door to greater use of revealed-preference analysis in future research on charitable giving. Toward 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 focus of our experiment is on the second player in a two-player sequential game, the theory applies equally to multi-agent simultaneous games.

Our experiment 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 other-regarding preferences for understanding decisions about the level of one's charitable contributions. 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 social comparisons of either private consumption or levels of donation yield statistically significant differences in the explanatory power over and above the standard model of impure altruism. Specific features of our experiment 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 as the charitable cause and choice environment are held constant, leaving social comparisons as the only variable other than standard parameters affecting choices.

Finally, we conclude with reasons why one might expect social comparisons to play an even more important role on charitable behavior outside of a laboratory setting. Features of our experiment bias against the 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 that 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 our 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, and the same argument can be extended to all other agents.

A.1. Proof that (1) ⇒ (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

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

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 U_{ij} denotes the partial derivative of $U_i[x_i^t, c_i(x_i^t, \mathbf{x}_{-i}^t), y_i^t + Y_{-i}^t, d_i(y_i^t, \mathbf{y}_{-i}^t)]$ with respect to the jth argument, and μ^t is the Lagrangian 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: $U_i^t := U_i[x_i^t, c_i(x_i^t, \mathbf{x}_{-i}^t), y_i^t + Y_{-i}^t, d_i(y_i^t, \mathbf{y}_{-i}^t)]$, $\kappa_i^t := U_{i1}^t$, $\pi_i^t := U_{i2}^t$, $\gamma_i^t := U_{i3}^t$, $\eta_i^t := U_{i4}^t$, 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, note that for any t and t', concavity of U_i implies that

$$\begin{aligned} U_i^{t'} &\leq U_i^t + U_{i1}^t\left[x_i^{t'} - x_i^t\right] + U_{i2}^t\left[c_i^{t'} - c_i^t\right] + U_{i3}^t\left[Y^{t'} - Y^t\right] + U_{i4}^t\left[d_i^{t'} - d_i^t\right] \\ &= U_i^t + \kappa_i^t\left[x_i^{t'} - x_i^t\right] + \pi_i^t\left[c_i^{t'} - c_i^t\right] + \gamma_i^t\left[Y^{t'} - Y^t\right] + \eta_i^t\left[d_i^{t'} - d_i^t\right], \end{aligned}$$

which completes this direction of the proof.

A.2. Proof that (2) ⇒ (1)

We define agent i's utility function as follows:

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

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

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

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''):

$$c_i\left(x_i'', \mathbf{x}_{-i}\right) - c_i\left(x_i', \mathbf{x}_{-i}\right) \le \frac{\partial c_i\left(x_i', \mathbf{x}_{-i}\right)}{\partial x_i}\left[x_i'' - x_i'\right],$$
$$d_i\left(y_i'', \mathbf{y}_{-i}\right) - d_i\left(y_i', \mathbf{y}_{-i}\right) \le \frac{\partial d_i\left(y_i', \mathbf{y}_{-i}\right)}{\partial y_i}\left[y_i'' - y_i'\right].$$

We define $c_{i1}^t := \frac{\partial c_i(x_i^t, \mathbf{x}_{-i}^t)}{\partial x_i}$ and $d_{i1}^t := \frac{\partial d_i(y_i^t, \mathbf{y}_{-i}^t)}{\partial y_i}$.

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

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

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

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

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, which completes the proof.

Appendix B. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jpubeco.2014.09.009.

References

Afriat, S., 1967. The construction of a utility function from expenditure data. Int. Econ. Rev. 8, 67–77.
Afriat, S., 1972. Efficiency estimates of production functions. Int. Econ. Rev. 13, 568–598.
Andreoni, J., 1988. Privately provided public goods in a large economy: the limits of altruism. J. Public Econ. 35, 57–73.
Andreoni, J., 1989. Giving with impure altruism: applications to charity and Ricardian equivalence. J. Polit. Econ. 97, 1447–1458.
Andreoni, J., 1990. Impure altruism and donations to public goods: a theory of warm-glow giving. Econ. J. 100, 464–477.
Andreoni, J., Miller, J., 2002. Giving according to GARP: an experimental test of the consistency of preferences for altruism. Econometrica 70, 737–753.
Ariely, D., Bracha, A., Meier, S., 2009. Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. Am. Econ. Rev. 99, 544–555.
Aronsson, T., Johansson-Stenman, O., 2008. When the Joneses' consumption hurts: optimal public good provision and nonlinear income taxation. J. Public Econ. 92, 986–997.
Beatty, T.K.M., Crawford, I.A., 2011. How demanding is the revealed preference approach to demand? Am. Econ. Rev. 101, 2782–2795.
Benabou, R., Tirole, J., 2006. Incentives and prosocial behavior. Am. Econ. Rev. 96, 1652–1678.
Bergstrom, T.C., Blume, L.E., Varian, H.R., 1986. On the private provision of public goods. J. Public Econ. 29, 25–49.
Blow, L., Browning, M., Crawford, I., 2008. Revealed preference analysis of characteristics models. Rev. Econ. Stud. 75, 371–389.
Bolton, G.E., Ockenfels, A., 2000. ERC: a theory of equity, reciprocity, and competition. Am. Econ. Rev. 90, 166–193.
Boskin, M.J., Sheshinski, E., 1978. Optimal redistributive taxation when individual welfare depends upon relative income. Q. J. Econ. 92, 589–601.
Bronars, S.G., 1987. The power of nonparametric tests of preference maximization [the nonparametric approach to demand analysis]. Econometrica 55, 693–698.
Carvajal, A., 2010. The testable implications of competitive equilibrium in economies with externalities. Economic Theory 45, 349–378.
Charness, G., Cheung, T., 2013. A restaurant field experiment in charitable contributions. Econ. Lett. 119, 48–49.
Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. Q. J. Econ. 117, 817–869.
Cornes, R., Sandler, T., 1984. Easy riders, joint production, and public goods. Econ. J. 94, 580–598.
Cornes, R., Sandler, T., 1996. The Theory of Externalities, Public Goods and Club Goods, 2nd edition. Cambridge University Press, New York.
Deb, R., 2009. A testable model of consumption with externalities. J. Econ. Theory 144, 1804–1816.
DellaVigna, S., List, J.A., Malmendier, U., 2012. Testing for altruism and social pressure in charitable giving. Q. J. Econ. 127, 1–56.
Duesenberry, J., 1949. Income, Savings and the Theory of Consumer Behavior. Harvard University Press, Cambridge.
Eckel, C., Grossman, P., 1996. Altruism in anonymous dictator games. Games Econ. Behav. 16, 181–191.
Fehr, E., Schmidt, K.M., 1999. A theory of fairness, competition, and cooperation. Q. J. Econ. 114, 817–868.
Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exp. Econ. 10, 171–178.
Frank, R., 1985. Choosing the Right Pond: Human Behavior and the Quest for Status. Oxford University Press, New York.
Frank, R., 1999. Luxury Fever: Money and Happiness in a Era of Excess. The Free Press, New York.
Frey, B.S., Meier, S., 2004. Social comparisons and pro-social behavior. Am. Econ. Rev. 94 (1717), 22.
Friedman, D., Sunder, S., 1994. Experimental Methods: A Primer for Economists. Cambridge University Press, Cambridge.
Glazer, A., Konrad, K.A., 1986. A Signalling Explanation for Charity. Am. Econ. Rev. 86, 1019–1028.
Greiner, B., 2004. An online recruitment system for economic experiments. In: Kremer, Kurt, Macho, Volker (Eds.), Forschung und wissenschaftliches Rechnen 2003GWDG-Bericht, Göttingen vol. 63. Gesellschaft für wissenschaftliche Datenverarbeitung mbh, Germany, pp. 79–93.
Harbaugh, W.T., 1998. What do donations buy? A model of philanthropy based on prestige and warm glow. J. Public Econ. 67, 269–284.
Hoffman, E., McCabe, K.A., Smith, V.L., 1996. Social distance and other-regarding behavior in dictator games. Am. Econ. Rev. 86, 653–660.
Hollander, H., 1990. A social exchange approach to voluntary cooperation. Am. Econ. Rev. 80, 1157–1167.
Jamison, J., Karlan, D., Schechter, L., 2008. To deceive or not to deceive: the effect of deception on behavior in future laboratory experiments. J. Econ. Behav. Organ. 68, 477–488.
Korenok, O., Millner, E.L., Razzolini, L., 2011. Impure altruism in dictators' giving. Working Paper. Department of Economics, Virginia Commonwealth University.
Lancaster, K., 1971. Consumer Demand: A New Approach. Columbia University Press, New York.
Layard, R., 1980. Human satisfactions and public policy. Econ. J. 90, 737–750.
Ljungqvist, L., Uhlig, H., 2000. Tax policy and aggregate demand management under catching up with the Joneses. Am. Econ. Rev. 90, 356–366.
Luttmer, E., 2005. Neighbors as negatives: relative earnings and well-being. Q. J. Econ. 120, 963–1002.
Romano, R., Yildirim, H., 2001. Why charities announce donations: a positive perspective. J. Public Econ. 81, 423–447.
Rosen, J.B., 1965. Existence and uniqueness of equilibrium points for concave n-person games. Econometrica 33, 520–534.
Selten, R., 1991. Properties of a measure of predictive success. Math. Soc. Sci. 21, 153–167.
Shang, J., Croson, R., 2006. The impact of social comparisons on nonprofit fund raising. In: Mark Isaac, R., Davis, Douglas D. (Eds.), Experiments investigating fundraising and charitable contributorsResearch in Experimental Economics vol. 11. Emerald Group Publishing Limited, Bingley, pp. 143–156.
Shang, J., Croson, R., 2009. A field experiment in charitable contribution: the impact of social information on the voluntary provision of public goods. Econ. J. 119, 1422–1439.
Varian, H.R., 1983. Non-parametric tests of consumer behavior. Rev. Econ. Stud. 50, 99–110.
Veblen, Thorstein B., 1899. The Theory of the Leisure Class: An Economic Study in the Evolution of Institutions. Macmillan, New York.
Vesterlund, L., 2006. Why do people give?, In: Powell, W., Steinberg, R.S. (Eds.), The Non-profit Sector: A Research Handbook, 2nd edition Yale University Press, New Haven, pp. 568–587.
Vesterlund, L., 2012. Voluntary giving to public goods: moving beyond the linear VCM. Working Paper. Department of Economics, University of Pittsburgh.