Individual Preferences for Giving*

Raymond Fisman† Columbia University
Shachar Kariv‡ UC Berkeley
Daniel Markovits§ Yale University

February 14, 2005

Abstract

This paper reports an experimental test of individual preferences for giving. We use graphical representations of modified Dictator Games that vary the price of giving. This generates a very rich data set well-suited to studying behavior at the level of the individual subject. We test the data for consistency with preference maximization, and we recover underlying preferences and forecast behavior using both nonparametric and parametric methods. Our results emphasize that classical demand theory can account very well for behaviors observed in the laboratory and that individual preferences for giving are highly heterogeneous, ranging from utilitarian to Rawlsian to perfectly selfish. (*JEL*: C79, C91, D64)

*This research was supported by the Experimental Social Science Laboratory (X-Lab) at the University of California, Berkeley. We are grateful to Jim Andreoni, Colin Camerer, Gary Charness, Ken Chay, Stefano DellaVigna, Steve Goldman, Botond Koszegi, John List, Ben Polak, Jim Powell, Matthew Rabin, Al Roth, Ariel Rubinstein, Chris Shannon, Andrew Schotter and Hal Varian for helpful discussions. This paper has also benefited from suggestions by the participants of seminars at several universities. Syngjoo Choi and Benjamin Schneer provided excellent research assistance. In addition, we would like to thank Brenda Naputi and Lawrence Sweet from the X-Lab for their valuable assistance, and Roi Zemmer for writing the experimental computer program. For financial support, Fisman thanks the Columbia University Graduate School of Business; Kariv acknowledges University of California, Berkeley; Markovits thanks Yale Law School and Deans Anthony Kronman and Harold Koh.

†Graduate School of Business, Columbia University, Uris 823, New York, NY 10027 (E-mail: rf250@columbia.edu, URL: http://www-1.gsb.columbia.edu/faculty/rfisman/).

‡Department of Economics, University of California, Berkeley, 549 Evans Hall # 3880, Berkeley, CA 94720 (E-mail: kariv@berkeley.edu, URL: http://socrates.berkeley.edu/~kariv/).

§Yale Law School, P.O. Box 208215, New Haven, CT 06520. (E-mail: daniel.markovits@yale.edu, URL: http://www.law.yale.edu/outside/html/faculty/ntuser93/profile.htm)

1 Introduction

We study *individual* preferences for giving. We emphasize the term individual to highlight that we will investigate behavior at the level of the individual subject. We utilize a modified dictator game that expands upon an experiment conducted by Andreoni and Miller (2002) (hereafter AM) in a number of fundamental ways. The modified dictator game varies the endowments and the prices of giving, so that each subject faces a menu of budget sets over his own payoff and the payoff of one other subject. We introduce a graphical computer interface that allows for the efficient collection of many observations per subject. Given our rich data set, we can thoroughly address four types of questions concerning preferences for giving (Varian 1982): (*i*) *Consistency.* Is behavior consistent with the utility maximization model? (*ii*) *Recoverability.* Can underlying preferences be recovered from observed choices? (*iii*) *Extrapolation.* Can behavior on previously unobserved budget sets be forecasted? (*iv*) *Heterogeneity.* To what degree do preferences differ across subjects?

The importance of studying individual heterogeneity was emphasized by AM, who also test observed behavior for consistency with the utility maximization model by applying the axiom of revealed preference. AM report that classical demand theory is indeed relevant to the interpretation of the experimental data and that preferences are very heterogeneous. Because of this heterogeneity, AM conclude that it is necessary to investigate behavior at an individual level. However, their experimental design involves few budgets sets and allows for only a small number of decisions per subject, which necessitates the pooling of observations across subjects. Our experimental design allows subjects to make numerous choices over a wide range of budget sets, and this yields a rich dataset that is well-suited to analysis at the level of the individual subject.

We begin our analysis of the experimental data by testing for consistency with utility maximization using revealed preference axioms. Afriat (1967) shows that if finite choice data satisfy the Generalized Axiom of Revealed Preference (GARP), there exists a well-behaved utility function that the choices maximize. Varian (1982) modifies Afriat's (1967) results and describes the most efficient and general techniques for testing the extent to which choices satisfy GARP. Since GARP offers an exact test (i.e. either the data satisfy GARP or they do not) and choice data almost always contain at least some violations, we calculate a variety of goodness-of-fit indices described by Varian (1991) to quantify the extent of violation. We find that most subjects exhibit behavior that appears to be *almost* optimizing in the sense that their choices nearly satisfy GARP, so that the violations are minor enough to ignore for the purposes of recovering preferences or constructing appropriate utility functions.

We therefore move to recovering underlying preferences and forecasting behavior in new situations. We begin by using the revealed preference techniques developed by Varian (1982). This approach is purely nonparametric and compares arbitrary allocations or budgets assuming only consistency with previously observed behavior and making no assumptions about the parametric form of the underlying utility function. Because of our rich data set, we are able to generate fairly tight bounds on these measures. We conduct case studies of several subjects that almost satisfy GARP and who serve to illustrate ideal types whose choices fit with prototypical preferences.

Motivated by patterns observed in the nonparametric approach, we estimate constant elasticity of substitution (CES) demand functions for giving at the individual level. The CES form is very useful in many applications since it spans a range of well-behaved utility functions by means of a single parameter. We find that the CES demand function provides a good fit at the individual level. Moreover, the parameter estimates vary dramatically across subjects, implying that individual preferences for giving are very heterogeneous, ranging from perfect substitutes to Leontief. However, like Charness and Rabin (2002), we do find that a significant majority of subjects are concerned with increasing aggregate payoffs rather than reducing differences in payoffs.

Our paper contributes to the large and growing body of work that seeks to explain departures from self-interest in economic experiments. Social preferences theories include altruism, relative income and envy, inequality aversion, and altruism and spitefulness (see, respectively, Charness and Rabin (2002), Bolton (1991), Levine (1998) and Fehr and Schmidt (1999), among others). Our paper provides a couple of important innovations over previous work. Our primary methodological contribution is an experimental technique that enables us to collect richer data about preferences for giving than has heretofore been possible. This enables us to analyze preferences for giving at the individual level. Further, previous experimental studies have inferred preferences from a small number of individual decisions and hence have been forced to set up relatively extreme choice scenarios. In contrast, we do not force subjects into choices that suggest discrete and extreme prototypical preference types. The experimental techniques and our results provide promising tools for future work in this area.

The rest of the paper is organized as follows. The next section describes the experimental design and procedures. Section 3 summarizes some important features of the data. Section 4 provides the nonparametric analysis, and Section 5 provides the parametric analysis. Section 6 concludes.

2 Experimental design and procedures

2.1 Design

We begin by discussing a very general formulation of the dictator game, before outlining the linear version that we utilize in our experiments. We adopt AM's notation and model a generalized dictator game by a set of two persons s and o (s for $self$ and o for *other*), a compact subset of the plane Π representing the feasible monetary payoff choices of person s (i.e. each point $\pi = (\pi_s, \pi_o)$ in Π corresponds to the payoffs to persons $self$ and *other*, respectively), and a preference ordering $\succsim_s$for person s. For any $\pi, \pi' \in \Pi$, $\pi \succsim_s \pi'$ if and only if person s prefers π to π' or is indifferent between them. The preference ordering $\succsim_s$can be represented by a utility function $U_s = u_s(\pi_s, \pi_o)$ that captures the possibility of giving if $u_s(\pi) \geq u_s(\pi')$ whenever $\pi \succsim_s \pi'$. Person s is said to be *selfish*, or an own-payoff maximizer, when $\pi \succsim_s \pi'$ if and only if $\pi_s \geq \pi'_s$. The objects s, Π and $\succsim_s$ define the dictator game.

The set Π of feasible payoff pairs may take many forms and one should ultimately put no restrictions directly on Π. Yet, in a typical dictator experiment, first introduced by Forsythe, et al. (1994), subject s divides some *endowment* m between $self$ and $other$ in any way he wishes such that

$$\Pi = \{(\pi_s, \pi_o) : \pi_s + \pi_o = m\}.$$

One respect in which this framework is restrictive is that the set Π is always the line with a slope of -1 so the problem faced by person s is simply allocating a fixed total income between $self$ and $other$.

A full generalization of the dictator game would allow for nonlinear budget sets. But because a great deal of classical theory is built on the assumption of a simple linear budget constraint, the simplest and perhaps most important generalization of the dictator game, developed by AM, allows for an endowment which is to be spent on π_s and π_o at given fixed corresponding *price* levels p_s and p_o. By choosing p_s as the numeraire price, the set of normalized possible payoffs for the game can then be written

$$\Pi = \{(\pi_s, \pi_o) : \pi_s + p\pi_o = m'\}$$

where $p = p_o/p_s$ is the *relative price of giving* and $m' = m/p_s$. This configuration creates *budget sets* over π_s and π_o that allow for the thorough testing of observed dictator behavior for consistency with utility maximization.[1]

2.2 Procedures

The experiment was conducted at the Experimental Social Science Laboratory (X-Lab) at the University of California, Berkeley under the X-Lab Master Human Subjects Protocol.[2] The 76 subjects in the experiment were recruited from all undergraduate classes at UC Berkeley and had no previous experience in experiments of dictator, ultimatum, or trust games. After subjects read the instructions, the instructions were read aloud by an experimenter.[3] The experimental instructions build on AM with minor changes to reflect employment of computers and are available upon request. Each experimental session lasted for about one and a half hours. A \$5 participation fee and subsequent earnings, which averaged \$16.0, were paid in private at the end of the session.[4] Throughout the experiment we ensured anonymity and effective isolation of subjects

[1] Andreoni and Vesterlund (2001) utilize the same design as AM to study whether preferences for giving differ by gender. They find that men are more responsive to price change and likely to be either perfectly selfish or utilitarian, whereas women tend to be Rawlsian.

[2] See http://xlab.berkeley.edu/cphs/master_protocol.pdf.

[3] At the end of the instructional period subjects were asked if they had any questions or any difficulties understanding the experiment. No subject reported any difficulty understanding the procedures or using the computer program.

[4] Payoffs were calculated in terms of tokens and then translated at the end of the experiment into dollars at the rate of 3 tokens = 1 dollar.

in order to minimize any interpersonal influences that could stimulate other-regarding behavior.[5]

Each session consisted of 50 independent decision-problems. In each decision problem, each subject was asked to allocate tokens between himself π_s and an anonymous subject π_o, where the anonymous subject was chosen at random from the group of subjects in the experiment. Each choice involved choosing a point on a graph representing a budget set over possible token allocations. Each decision problem $i = 1, ..., 50$ started by having the computer select a budget set randomly from the set

$$\{p_s^i \pi_s + p_o^i \pi_o \leq m^i : \quad m^i/p_n \leq 100 \text{ for both } n = s, o \\ \text{and } m^i/p_n \geq 50 \text{ for some } n = s, o\}$$

(i.e. budget sets that intersect with at least one of the axes at 50 or more tokens, but with no intercept exceeding 100 tokens). For comparison, in AM, the full menu of budgets offered included only 8 budget constraints in four sessions and 11 in one session with $m = \{40, 60, 75, 100\}$ and $p = \{1/4, 1/3, 1/2, 1, 2, 3, 4\}$. The sets selected for each subject in different decision problems were independent of each other and of the sets selected for any of the other subjects in their decision problems. Notice that choices were not restricted to allocations on the budget constraint so that, in theory, subjects could costlessly dispose payoffs and violate *budget balancedness* (i.e. $p^i \pi^i = m^i$).

To choose an allocation, subjects used the mouse or the arrows on the keyboard to move the pointer on the computer screen to the desired allocation. Subjects could either left-click or press the Enter key to make their allocation. At any point, subjects could either right-click or press the Space key to find out the allocation at the pointer's current position. The π_s-axis and π_o-axis were labeled Hold and Pass respectively and scaled from 0 to 100 tokens. The resolution compatibility of the budget sets was 0.2 tokens; the sets were colored in light grey; and the frontiers were not emphasized. The graphical representation of budget sets also enabled us to avoid emphasizing any particular allocation. At the beginning of each decision round, the experimental program dialog window went blank and the entire setup reappeared. The appearance and behavior of the pointer were set to the Windows mouse default and the pointer was automatically repositioned at the origin $\pi = (0, 0)$ at the beginning of each round.

This process was repeated until all 50 rounds were completed. At the end of the experiment, payoffs were determined in the following way. The experimental program first randomly selected one decision round from each subject to carry out. That subject then received the tokens that he held in this round π_s, and the subject with whom he was matched received the tokens that he passed π_o. Thus, as in AM, each subject received two groups of tokens, one based on his own decision to hold tokens and one based on the decision of another random subject to pass tokens. The computer program ensured that the same two subjects were not paired twice. Subjects received their payment privately as

[5] Subjects' work-stations were isolated by cubicles making it impossible for subjects to observe other screens or to communicate.

they left the experiment.

The experiments provide us with a very rich dataset. We have observations on $76 \times 50 = 3800$ individual decisions over a variety of different budget sets. Most importantly, the broad range of budget sets provides a serious test of the ability of classical theory, and a structural econometric model based on the theory, to interpret the data. In particular, the changes in endowments and relative prices are such that budget lines cross frequently. This means that our data lead to high power tests of revealed preference conditions (see Varian (1982), Bronars (1987), Russell (1992) and Blundell, Browning and Crawford (2003)).

Aside from pure technicalities, our paper provides a couple of important innovations over previous work. First it allows us to test a wider range of budget sets than can be tested using the pencil-and-paper experimental questionnaire method of AM. Second, it generates a rich dataset that provides the opportunity to address recoverability and extrapolation at the level of the individual subject. This is particularly important. As AM emphasize, individual heterogeneity requires behavior to be examined at an individual level, before preferences can properly be understood.

3 Overview of experimental data

In this section, we provide an overview of some important features of the experimental data, which we summarize by reporting the distribution of allocations in a number of ways. The histograms in Figure 1 show the distributions of the fraction given to *other*, defined in a couple of ways. Figure 1A depicts the distribution of the expenditure on tokens given to *other* $p_o^i \pi_o^i$ as a fraction of total expenditure $p_s^i \pi_s^i + p_o^i \pi_o^i$, or simply

$$\frac{p^i \pi_o^i}{\pi_s^i + p^i \pi_o^i},$$

and Figure 1B depicts the distribution of the tokens given to *other* as a fraction of the sum of the tokens kept and given

$$\frac{\pi_o^i}{\pi_s^i + \pi_o^i}.$$

Clearly, this distinction is only relevant in the presence of price changes. Further, in each case we present the distribution for *all* allocations, as well as the distributions by three price terciles: intermediate prices of around 1 ($0.70 \leq p \leq 1.43$), steep prices ($p > 1.43$) and symmetric flat prices ($p < 0.70$). The horizontal axis measures the fractions for different intervals and the vertical axis measures the percentage of decisions corresponding to each interval.

[Figure 1 here]

In Figure 1A, for the full sample there is a mode at the midpoint of 0.5 (i.e. same expenditure on *self* and *other*, $p_s^i \pi_s^i = p_o^i \pi_o^i$). The number of allocations then decreases as we move to the left, before increasing rapidly to selfish allocations of 0.05 or less of the total expenditure on tokens for *other*, which account for 40.5 percent of all allocations. This masks some heterogeneity by price. For the middle tercile, the pattern is somewhat more pronounced, while for the flat tercile, there is no peak at the midpoint. Not surprisingly, the distribution is generally further to the left for steeper-sloped budgets. Figure 1B shows a similar pattern to Figure 1A, though it is somewhat more skewed to the left. Compared with studies of split-the-pie dictator games (See, Camerer (2003) for a comprehensive discussion), the mode at the midpoint is relatively less pronounced and the distribution is much smoother, even for the intermediate tercile allocations. Over all prices, our subjects gave to *other* about 19 percent of the tokens, accounting for 21 percent of total expenditure; this is very similar to typical mean allocations of about 20 percent in the studies reported in Camerer (2003).

The decision-level graphs in Figure 1 potentially obscure the presence of individual concerns *on average* for others. For example, a person who gives everything to *other* half of the time and keeps everything for *self* the other half would generate extreme giving values, when in fact such a person keeps an intermediate fraction on average. Hence, Figure 2 shows the distribution of $p^i \pi_o^i / (\pi_s^i + p^i \pi_o^i)$ presented in Figure 1A aggregated to the subject level. Since this takes an average over all prices, the distribution should be similar for $\pi^i / (\pi_s^i + \pi_o^i)$ (in practice, the histograms are identical once we aggregate by decile). The horizontal axis measures the fractions for different intervals and the vertical axis measures the percentage of subjects corresponding to each interval. Perhaps as expected, the distributions show a pattern with a larger mode around the midpoint and no observations below the midpoint.

[Figure 2 here]

These aggregate distributions tell us little about the particular allocations chosen by individual subjects. As a preview, Figure 3 shows the scatterplot of the choices made by subjects whose choices correspond to prototypical preferences. Figure 3A depicts the choices of a selfish subject (ID 4) $U(\pi_s, \pi_o) = \pi_s$, Figure 3B depicts the allocations of a subject with utilitarian preferences (ID 54) $U(\pi_s, \pi_o) = \pi_s + \pi_o$, and Figure 3C depicts the other extreme, that of a Rawlsian subject (ID 46) $U(\pi_s, \pi_o) = \min\{\pi_s, \pi_o\}$. Of our 76 subjects, 20 (26.3 percent) behaved perfectly selfishly. Only two (2.6 percent) subjects allocated all their tokens to *self* if $p_s < p_o$ and to *other* if $p_s > p_o$ implying utilitarian preferences, and two (2.6 percent) subjects made nearly equal expenditure on *self* and *other* indicating Rawlsian preferences. By comparison, AM report that 40 subjects (22.7 percent) behaved perfectly selfishly, 25 subjects (14.2 percent) could fit with utilitarian preferences, and 11 subjects (6.2 percent) were consistent with Rawlsian preferences. We also find many intermediate cases, but these are difficult to see directly on a scatterplot due to the fact that both p

and m shift in each new allocation. This is the purpose of our individual-level analysis below.

[Figure 3 here]

4 Nonparametric analysis

4.1 Consistency

Varian (1982) describes a set of techniques, based on revealed preference axioms, through which to test a *finite* set of data for consistency with utility maximization, to recover preferences from the data, and to forecast demand behavior on budget sets which have not been previously observed.

Let (p^i, π^i) for $i = 1, ..., n$ be some observed individual data (i.e. p^i denotes the i^{th} observation of the prices and π^i denotes the associated allocation). Throughout this section it will be convenient to normalize the prices, p_s^i and p_o^i, by the endowment m^i at each observation so that $p^i\pi^i = 1$ for all i.

First, we recall that π^i is *directly revealed preferred* to π if $p^i\pi^i \geq p^i\pi$ and *strictly directly revealed preferred* if the inequality is strict. The relation *indirectly revealed preferred* is the transitive closure of the directly revealed preferred relation. We then define the following:

Generalized Axiom of Revealed Preference (GARP) If π^i is indirectly revealed preferred to π^j, then π^j is not strictly directly revealed preferred (i.e. $p^j\pi^i \geq p^j\pi^j$) to π^i.

A utility function $u_s(\pi)$ *rationalizes* the observed behavior if $u_s(\pi^i) \geq u_s(\pi)$ for all π such that $p^i\pi^i \geq p^i\pi$ (i.e. u_s achieves the maximum on the budget set at the chosen bundle). GARP is tied to utility representation through the following theorem, which was first proved by Afriat (1967):

Afriat's Theorem The following conditions are equivalent: (i) The data satisfy GARP. (ii) There exists a non-satiated utility function that rationalizes the data. (iii) There exists a concave, monotonic, continuous, non-satiated utility function that rationalizes the data.[6]

The equivalence of (i) and (ii) establishes GARP as a direct test for whether a finite data set may be rationalized by a utility function, and the equivalence of (ii) and (iii) tells us that when a rationalizing utility function exists, it can be chosen to be well-behaved.

In order to verify that the data satisfies GARP it is only necessary to have an efficient way to compute the transitive closure of a binary relation. A variety of easily implementable graph theoretic algorithms, such as that described in Varian (1982), provide an efficient way to check GARP by computing the transitive closure of the directly revealed preference relation.

[6] This statement of the theorem follows Varian (1982), who replaced the condition Afriat called *cyclical consistency* with GARP.

To better understand our measures of consistency, it is instructive to describe the basic idea underlying the algorithm (the computer program and details are available from the authors upon request). Consider a *directed graph* G with a set of *nodes*

$$V = \{1, ..., n\}$$

and set of *edges*

$$E = \cup_{i=1}^{n} \{ij : p^i \pi^i \geq p^i \pi^j\}.$$

That is, the graph is a pair $G = (V, E)$ of sets satisfying $E \subseteq [V]^2$ with nodes representing individual decisions and edges representing directly revealed preferred relations. Note that the edges need not be symmetric: the existence of an edge directed from i to j does not imply the existence of an edge from j to i (in fact, this would imply a GARP violation if one of the inequalities were strict).

For any nodes i and j, an $i - j$ *path* is a finite sequence $i_1, ..., i_K$ such that $i_1 = i$, $i_K = j$ and $p^k \pi^k \geq p^k \pi^{k+1}$ for $k = 1, ..., K - 1$ (i.e. a sequence of nodes $i_1, ..., i_K$ linked by E). Note that a path represents a revealed preferred relation in the data (i.e. π^i is revealed preferred to π^j if and only if there exists an $i - j$ path). A cyclic sequence of nodes that creates an $i - i$ path called a *cycle*. The length of a cycle is its number of edges, and a cycle of length k is called a k-*cycle*. It follows directly from the definition that if G contains a cycle with at least one strict inequality, then we have a violation of GARP. The number of cycles in G is the number of GARP violations.

4.2 Goodness-of-fit

As formulated above, GARP offers an exact test: either the data satisfy GARP and are therefore consistent with the utility maximization model, or they do not. Unfortunately, choice data almost always contain at least some violations. Hence, it is useful to measure the *extent* of GARP violation. We report measures of GARP violation based on three indices: Afriat (1972), Houtman and Maks (1985), and Varian (1991) which we refer to as the E-*index*, V-*index* and EV-*index*, respectively, according to whether they involve conditions on edges, vertices, or both.

The first index is what Afriat (1972) calls the *critical cost efficiency index* (CCEI); it measures the amount by which each budget constraint must be relaxed in order to remove all violations of GARP. More precisely, let (e^i) be a vector of numbers with $0 \leq e^i \leq 1$. Define G' to be a *spanning subgraph* of G (i.e. $G' = (V', E')$ with $V' = V$ and $E' \subseteq E$) with

$$E' = \cup_{i=1}^{n} \{ij : e^i p^i \pi^i \geq p^i \pi^j\}.$$

Then the CCEI is the largest number e^i such that the subgraph G' does not contain any cycle, with at least one of the inequalities holding strictly. We call this number the E-index. Note that the E-index provides a summary statistic of the overall consistency of the data with GARP but gives no hint as to which of the observations are causing violations. Figure 4 illustrates the construction of the E-index for a simple violation of GARP.[7] Note that a small perturbation $A/B < C/D$ of the budget constraint through observation π' removes the violation.

[Figure 4 here]

The second test, proposed by Houtman and Maks (1985), finds the largest subset of choices that is consistent with GARP. More formally, let G'' be an *induced subgraph* of G (i.e. $G'' = (V'', E'')$, with $V'' \subseteq V$ and E'' containing all edges $ij \in E$ with $i, j \in V''$) and find the graph G'' with the largest order $|G''|$ (i.e. number of nodes) that does not contain any cycle. We call the ratio $|G''| / |G|$ the V-index; note that $|G''| / |G| \in [0, 1]$, with higher numbers reflecting a larger subset of consistent observations. This method has a couple of drawbacks. First, observations may be discarded even if the associated GARP violations could be removed by small perturbations of the budget constraint. Further, it is computationally very intensive and thus impractical if, roughly speaking, cycles often overlap. As a result, we were unable to calculate the V-indices for a small number of subjects who often violated GARP, and we therefore report only lower bounds.

The third test follows Varian (1991), who refined the E-index (Afriat's CCEI) by providing a more disaggregated measure that indicates the minimal amount required to perturb *each* observation in order to remove all violations of GARP. With a slight abuse of notation, let (v^i) be a vector of numbers with $0 \leq v^i \leq 1$. Let G''' be a *spanning subgraph* of G with

$$E''' = \cup_{i=1}^{n} \{ij : v^i p^i \pi^i \geq p^i \pi^j\}.$$

Then, v^i is the largest number such that the subgraph G''' does not contain any $i - i$ cycles, with at least one of the inequalities holding strictly. Thus, $v^i = 1$ if and only if G does not contain an $i - i$ cycle. Note that (v^i) need not be the smallest perturbation of the data that is consistent with GARP because an $i - i$ cycle may often be broken by perturbing an edge which joins two other nodes of the cycle. To describe efficiency, Varian (1991) uses the smallest (v^i), which we call EV-index. Note that the EV-index is a lower bound on the E-index. The reasons for this discrepancy are discussed in Varian (1991).

Table 1 lists, by subject, the number of violations of budget balancedness and GARP, and also reports the three indices according to descending E-values. We need to allow for small mistakes resulting from the slight imprecision of subjects' handling of the mouse. The results presented in Table 1 allow for a narrow confidence interval of one token (i.e. for any i and $j \neq i$, if $d(\pi^i, \pi^j) \leq 1$ then π^i and π^j are treated as the same allocation).

[Table 1 here]

[7] Here we have a violation of the Weak Axiom of Revealed Preference (WARP) (i.e. π^i is directly revealed preferred to $\pi^j \neq \pi^i$ and π^j is directly revealed preferred to π^i).

Half of our subjects have no violations of budget balancedness, even with the narrow one token confidence interval; if we allow for a five token confidence interval, 64 subjects (84.2 percent) have no violations of budget balancedness. The second column of Table 1 reports the confidence interval required to remove *all* budget balancedness violations.[8]

Turning now to GARP violations, out of the 76 subjects, only 8 (10.5 percent) have no violations of GARP, but all of these are subjects that always chose selfish allocations $\pi^i = (m^i/p_s^i, 0)$ for all i. However, 54 subjects, (71.1 percent) had E-indices above the 0.90 threshold, and of those 41 subjects (53.9 percent) were above the 0.95 threshold. In contrast, AM report that only 18 of their 176 subjects (10.2 percent) violated GARP, and of those only 3 had E-indices below the 0.95 threshold. This is as expected, as our subjects were given a larger and richer menu of budget sets which provides more opportunities to violate GARP and thus improves the power of nonparametric tests of revealed preference theory. Figure 5 shows the distributions of the three indices. The horizontal axis measures the indices for different intervals and the vertical axis measures the percentage of subjects corresponding to each interval.

[Figure 5 here]

Nonetheless, we argue that most of our subjects are close enough to passing GARP so that we may not want to reject that their choices are consistent with utility maximization. To give greater precision to this notion, we wish to generate a benchmark levels of consistency with which we may compare our E-index and EV-index. As in AM, we use the test designed by Bronars (1987) that uses the choices of a hypothetical subject who randomizes uniformly among all allocations on each budget line as a benchmark. To this end, we generated a random sample of 25,000 subjects and found that all of them violated GARP. Their E-index and EV-index values averaged 0.60 and 0.25 respectively.[9] If we choose the 0.9 efficiency level as our critical value, we find that only 12 of the random subjects' E-index values were above the threshold and none of the subjects' EV-index values were above this threshold. Using the same approach on only 8 random budget sets, which is the number of decisions made by most of AM's subjects, we find that random decisions generate average E-index and EV-index values of 0.91 and 0.78 respectively.[10]

In order to determine the extent to which subjects are consistent with GARP relative to random allocations, for each subject we take the difference between his actual index and the average random-choice index as a fraction of the difference between consistent-choice index of one and the average random-choice

[8] We note that there are a few subjects that required large confidence intervals, but that these subjects also have many GARP violations even if the choices that violate budget balancedness are removed.

[9] Note that we cannot generate an average V-index for random subjects because of the computational complexity required.

[10] Bronars' test (i.e. the probability that a random subject violates GARP) has also been applied to experimental data by Cox (1987), Sippel (1997), Mattei (2001) and Harbaugh, Krause and Berry (2001). Our study has the highest Bronar power of one (i.e. all random subjects had violations).

index. We do this for both the E-index and EV-index. Subjects with no GARP violations have a consistency of one and the average random subject has a consistency of zero; obviously, negative values are possible if a subject performs worse than randomness. The results are presented in the final two columns of Table 1 above, and illustrated in the histograms of consistency values in Figure 6. As Figure 6 shows, the distribution of consistency values is skewed to the right and almost uniformly positive; this provides a clear graphical illustration of the extent to which subjects did worse than choosing consistently and the extent to which they did better than choosing randomly.

[Figure 6 here]

4.3 Recovering preferences and forecasting behavior

If choice data satisfy GARP we would ideally like to extract a rationalizing utility function through which to recover preferences and forecast choices on out-of-sample allocations. Varian (1982) uses GARP and assumptions of convexity and monotonicity to generate an algorithm that serves as a partial solution to this so-called *recoverability problem*. This algorithm can also recover preferences from choices, such as the ones in our experiment, that are almost consistent with GARP. Further, since we observe many choices over a wide range of budget sets, we can describe preferences with some precision.

We give a brief outline of Varian's algorithm, which provides the tightest possible bounds on indifference curves through an allocation π^0 that has not been observed in the previous data (p^i, π^i) for $i = 1, ..., n$. First, we consider the set of prices at which π^0 could be chosen and be consistent (i.e. does not add violations of GARP) with the previously observed data. This set of prices is the solution to the system of linear inequalities constructed from the data and revealed preference relations. Call this set $S(\pi^0)$. Second, we use $S(\pi^0)$ to generate the set of observations, $RP(\pi^0)$, revealed preferred to π^0 and the set of observations, $RW(\pi^0)$, revealed worse than π^0.

It is not difficult to show that $RP(\pi^0)$ is simply the convex monotonic hull of all allocations revealed preferred to π^0. To understand the construction of $RW(\pi^0)$, note that if π^0 is directly revealed preferred to some observation π^i for all prices $p^0 \in S(\pi^0)$ (i.e. $p^0\pi^0 \geq p^0\pi^i$), then it is indirectly revealed preferred to any allocation in the budget set $(p^i, 1)$ on which π^i was chosen. Similarly, it is indirectly revealed preferred to all observations that π^i is revealed preferred to and so on. Hence, the two sets $RP(\pi^0)$ and the complement of $RW(\pi^0)$ form the tightest inner and outer bounds on the set of allocations preferred to π^0. Similarly, $RW(\pi^0)$ and the complement of $RP(\pi^0)$ form the tightest inner and outer bounds on the set of allocations worse than π^0.

Figure 7 depicts the construction of the bounds described above through some allocation π^0 for several subjects who almost satisfy GARP and who illustrate ideal types whose choices fit with prototypical utility functions.[11] Since

[11] Our computational experience with this technique reveals that if the data are not very close to satisfying GARP then $RP(\pi^0)$ and $RW(\pi^0)$ overlap.

the data is clustered in very different areas on the graphs for different subjects, we look at indifference curves through the average choices of each subject. In addition to the $RW(\pi^0)$ and $RP(\pi^0)$ sets, Figure 7 also shows the subjects' choices $(\pi^1, ...\pi^{50})$ as well as the budget sets used to construct $RW(\pi^0)$. Figure 7A (top left) shows the bounds on the indifference curve through $(40, 25)$ for a utilitarian subject (ID 54) where the revealed worse and preferred sets closely bound a linear indifference curve with slope of about -1.[12] Figure 7B (top right) shows the bounds for a Leontief subject (ID 46) through $(23, 20)$ where the bounds suggest a near-right angled indifference curve. Finally, Figure 7C (bottom) shows the bounds for an intermediate-case subject (ID 30) through $(41, 14)$ where the bounds imply indifference curves with some degree of curvature, but with greater weight on π_s. While these cases generate a particularly close fit, we may generally provide reasonably precise bounds for subjects that nearly satisfy GARP, as the variation in p and m ensure that budget sets intersect frequently.

[Figure 7 here]

We next turn our attention to forecasting behavior. We ask, that is, what set of allocations could be chosen on a previously unobserved budget set. Varian (1982) shows that this is the set of all allocations that support the budget set in the sense of consistency with GARP. We skip the development of the algorithm in the interests of brevity, and because it builds on the same algorithm used to recover preferences above; we note that this set is the tightest estimate of the underlying inverse demand correspondence. Figure 8 gives examples of this set for the same group of subjects as in Figure 7 for the budget sets $(p, m) = \{(1, 40), (3/2, 40), (2, 40)(2/3, 60), (1/2, 80)\}$. Again, note the tightness of some of the sets and the differences among subjects. In the figure, the dotted lines show hypothetical budget sets faced by the subjects, and the solid portions of these lines show the set of points that could be chosen on these lines. The dotted lines in Figure 8 show hypothetical budget sets faced by the subjects, and the solid portions of these lines show the set of points that could be chosen on these lines.

[Figure 8 here]

5 Parametric analysis

Afriat's theorem tells us that if a rationalizable utility function exists, it can be chosen to be increasing, continuous, and concave. Additionally, the patterns observed in the nonparametric approach suggest that it is appropriate to estimate a CES demand function. The CES is useful because attitudes towards giving

[12] The difference in the average values of π_s and π_o for subject ID 54 might look surprising, given that he apparently has utilitarian preferences. This is due to the potential asymmetry in the (randomly chosen) budget sets faced by each subject. In the case of subject ID 54, the randomly selected budget sets had a relatively large fraction of steep slopes.

can be adjusted by means of a single parameter. This is also the parametric form chosen by AM.

The CES utility function is given by

$$U_s = [\alpha(\pi_s)^\rho + (1-\alpha)(\pi_o)^\rho]^{1/\rho}$$

where ρ represents the curvature of the indifference curves, α represents the relative weight on the payoff for $self$, and $\sigma = 1/(\rho - 1)$ is the (constant) elasticity of substitution. The CES approaches a perfect substitutes utility function as $\rho \to 1$ and the Leontief form as $\rho \to -\infty$. As $\rho \to 0$, the indifference curves approach those of a Cobb-Douglas function.

The CES demand function is given by

$$\pi_s(p, m') = \frac{A}{p^r + A} m'$$

where

$$r = -\rho / (\rho - 1)$$

and

$$A = [\alpha / (1-\alpha)]^{1/(1-\rho)},$$

which is homogeneous in degree zero in (p, m).[13] This generates the following individual-level econometric specification for each subject n:

$$\frac{\pi_{sn}^i}{m_n'^i} = \frac{A_n}{(p_n^i)^{r_n} + A_n} + \epsilon_n^i$$

where ϵ_n^i is assumed to be distributed normally with mean zero and variance σ_n^2. Note that, as in AM, the demands are estimated as budget shares, which are bounded between zero and one, with an *i.i.d.* error term. We generate estimates of $\hat{A}_n$ and $\hat{r}_n$ using non-linear tobit maximum likelihood, and use this to infer the values of the underlying CES parameters $\hat{\alpha}_n$ and $\hat{\rho}_n$ and the elasticity of substitution $\hat{\sigma}_n$. We emphasize again that our estimations will be done for each subject n separately, generating separate estimates $\hat{A}_n$ and $\hat{r}_n$.

Before proceeding to the estimations, we screen the data for subjects with an E-index below 0.80, as the choices of subjects with E-indices not sufficiently close to one cannot be utility-generated; we also screen out subjects with uniformly selfish allocations (average $\pi_s/m' \geq 0.95$). This left a total of 45 subjects (59.2 percent) for whom we estimated parameters. Table 2 presents the results of the estimations $\hat{A}_n$, $\hat{r}_n$, $\hat{a}_n$, $\hat{\rho}_n$ and $\hat{\sigma}_n$ sorted according to ascending values of $\hat{\rho}_n$.[14]

[Table 2 here]

[13] In the case of two goods, WARP and budget balancedness imply that demand functions must be homogeneous of degree zero in (p, m).

[14] We generate virtually identical parameter values using non-linear least squares.

Figure 9 shows a scatterplot of our estimates of the two relevant parameters in the CES model, $\hat{\rho}_n$ and $\hat{a}_n$ (subjects ID 3, 46, 55 and 73 are excluded because they have very low $\hat{\rho}$-values). We distinguish the between the estimates for subjects with an E-index above 0.80, 0.90 and 0.95. We also note that there is considerable heterogeneity in both parameters, and that their values are negatively correlated ($r^2 = -0.35$).

[Figure 9 here]

Of the 45 subjects listed in Table 2, nine subjects (20.0 percent) have cleanly classifiable preferences: two subjects (4.4 percent) have perfect substitutes preferences ($\hat{\rho} \approx 1$), five subjects (11.1 percent) exhibit Cobb-Douglas preferences ($\hat{\rho} \approx 0$), and two subjects (4.4 percent) exhibits Leontief preferences (very low $\hat{\rho}$-values). More interestingly, there is considerably heterogeneity in subjects' preferences among those that cannot be cleanly categorized: 22 subjects (48.9 percent) have $0.1 \leq \hat{\rho} \leq 0.9$ so that the fraction kept, $\hat{\pi}_s/m'$, increases with the price of giving p; these subjects thus show a preference for increasing total payoffs. The 14 other subjects (31.1 percent) have negative values of $\hat{\rho}$ that are not 'too low' so that the fraction kept, $\hat{\pi}_s/m'$, decreases with the price of giving p; these subjects thus show a preference for reducing differences in payoffs. Figure 10 presents the distribution of $\hat{\rho}_n$ rounded to a single decimal (subjects ID 3, 46, 55 and 73 are again excluded). We present the distribution for all subjects, as well as the distributions for subjects with an E-index above 0.90 and 0.95.

[Figure 10 here]

The distributions shown in Figure 10 inform the debate about whether preferences for giving are best thought of as preferences for increasing total payoffs or for reducing differences in payoffs. Previous experimental studies have inferred preferences from a small number of individual decisions and hence have been forced to set up relatively extreme choice scenarios. Our method enables us to confront subjects with a wide range of prices for giving, so that the specification of choice sets is less likely to influence subjects' decisions. In particular, we do not force subjects into choices that suggest discrete and extreme prototypical preference types.

Figure 10 emphasizes the heterogeneity of preferences that we find. Nevertheless the distribution of types is skewed toward preferences for increasing total payoffs rather than reducing differences in payoffs. Also, interestingly, there is a significant portion of subjects for whom $\hat{\rho} \approx 0$ (whose preferences are Cobb-Douglas), and hence whose fraction kept $\hat{\pi}_s/m'$ is insensitive to the price of giving. Thus, our results lean overall toward a social welfare conception of preferences.

Figure 11 shows the relationship between $\log(p)$ and $\hat{\pi}_s(p, m')/m'$ for the three representative subjects that we followed in the non-parametric section and also for some intermediate cases. Figure 11A shows the relationship for a subject (ID 54) with $\hat{\rho} = 0.99$ and $\hat{\alpha} = 0.50$, who is one of two subjects who very precisely implemented utilitarian preferences. Figure 10B shows the subject (ID 46) who most closely approximated Leontief preferences with $\hat{\rho} = -5.37$ and $\hat{\alpha} = 0.66$. Figure 11C shows a subject (ID 40) with Cobb-Douglas preferences, again very precisely implemented with $\hat{\rho} = -0.01$ and $\hat{\alpha} = 0.51$. More interestingly, some subjects made choices that may reflect preferences with intermediate parameter values. Figures 11D ($\hat{\rho} = 0.43$) and Figure 11E ($\hat{\rho} = -0.51$) show such subjects (ID 30 and 18 respectively).

[Figure 11 here]

Finally, the good fit between observed behavior and classical utility theory makes it appropriate to bring standard tools to bear on preferences for giving. Figure 12 depicts the estimated indirect utility function $\hat{v}_s(p,m) = u_s(\hat{\pi}_s(p,m))$ by showing a three-dimensional mapping of $\log(p)$ and m onto $\hat{v}_s$ for the five subjects that we examine in Figure 11. More interestingly, Figure 13 shows the effect of p on $\hat{v}_s$ holding m constant, and Figure 14 shows the typical (p,m) indifference curves for these subjects.

[Figure 12 here]
[Figure 13 here]
[Figure 14 here]

6 Concluding remarks

A new experimental design - employing graphical representations of modified dictator games - enables us to collect richer data about preferences for giving than has heretofore been possible. This enables us to analyze preferences for giving at the individual level.

Our results are summarized as follows. First, a significant majority of our subjects exhibit behavior that appears to be almost optimizing in the sense that their choices are close to satisfying both budget balancedness and GARP. Second, the CES utility function provides a good fit at the individual level. Third, individual preferences are very heterogeneous, ranging from utilitarian to Rawlsian to perfectly selfish. A significant majority of subjects have estimated parameters that indicate a preference for increasing total payoffs rather than reducing differences in payoffs.

Many open questions remain about preferences for giving. For example, one promising direction is to study decisions over non-linear budgets. Our setup enables us to enrich the dictator experiment paradigm by testing any set, theoretically yielding richer data about choices than in the case of a simple linear budget set. Implicit in the linear budget constraint are the assumptions of constant and smooth substitution between payoffs to $self$ and $other$ and negligible transfer costs. However, it is clear that even in the simple case of spending a fixed total on $self$ and $other$, nonlinearities may arise. For example, an important nonlinearity arises when the dictator has some initial endowment and faces a different price ratio in each direction.

A different set of important nonlinearities arise when we move to questions of preferences for giving over strictly convex budget sets. Such a setting would be the obvious way to study the behavior of, for example, the circumstances faced by head-of-family farmers in developing countries. Finally, nonlinearities become even more pervasive when we move to non-convex sets, in which choices that reduce differences in payoffs may make both players worse off. The experimental techniques that we have developed provide some promising tools for future work in these areas.

References

[1] Afriat, S. (1967) "The Construction of a Utility Function from Expenditure Data." *Econometrica*, *6*, pp. 67-77.

[2] Afriat, S. (1972) "Efficiency Estimates of Production Functions." *International Economic Review*, *8*, pp. 568-598.

[3] Andreoni, J. and J. Miller (2002) "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism." *Econometrica*, *70*, pp. 737-753.

[4] Andreoni, J. and L. Vesterlund (2001) "Which Is The Fair Sex? Gender Differences In Altruism." *The Quarterly Journal of Economics*, *116*, pp. 293-312.

[5] Blundell R., M. Browning, and I. Crawford (2003) "Nonparametric Engel Curves and Revealed Preference." *Econometrica*, *71*, pp. 205.240.

[6] Bolton, G. (1991) "A Comparative Model of Bargaining: Theory and Evidence." *American Economic review*, *81*, pp. 1096-1136.

[7] Bronars, S. (1987) "The power of nonparametric tests of preference maximization." *Econometrica*, *55*, pp. 693-698.

[8] Camerer, C. (2003) "Behavioral Game Theory: Experiments in Strategic Interaction." Princeton University Press.

[9] Charness, G. and M. Rabin (2002) "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics*, *117*, pp. 817-869.

[10] Cox, J. (1997) "On Testing the Utility Hypothesis." *The Economic Journal*, *107*, pp. 1054-1078.

[11] Fehr, E. and K. Schmidt (1999) "A Theory of Fairness, Competition and Co-operation." *Quarterly Journal of Economics*, *114*, pp. 817-868.

[12] Forsythe, R., J. Horowitz, N. Savin, and M. Sefton (1994) "Fairness in Simple Bargaining Games." *Games and Economic Behavior*, *6*, pp. 347-369.

[13] Harbaugh, W., K. Krause, and T. Berry (2001) "GARP for Kids: On the Development of Rational Choice Behavior." *American Economic Review*, *91*, pp. 1539-1545.

[14] Houtman, M. and J. Maks (1985) "Determining all Maximial Data Subsets Consistent with Revealed Preference." *Kwantitatieve Methoden*, *19*, pp. 89-104.

[15] Levine, D. (1998) "Modeling Altruism and Spitefulness in Experiments." *Review of Economic Dynamics*, *1*, pp. 593-622.

[16] Russell, R. (1992)"Remarks on the Power of Nonparametric tests of Consumer-Theory Hypotheses." In L. Phlips and L.D.Taylor (eds), Aggregation, Consumption and Trade, Dordrecht: Kluwer Academic Publishes, pp. 121-135.

[17] Sippel, R. (1997) "An Experiment on the Pure Theory of Consumer's Behavior." *The Economic Journal*, *107*, pp. 1431-1444.

[18] Varian, H. (1981) "The Nonparametric Approach to Demand Analysis." *Econometrica*, *50*, pp. 945-972.

[19] Varian, H. (1991) "Goodness-of-Fit for Revealed Preference Tests." University of Michigan CREST Working Paper # 13.

Table 1. Budget balancedness and GARP violations and the three indices by subject
(sorted according to descending *E*-values)

Subject ID	budget balancedness		# GARP violations	Violation indices			Consistency measures**	
	# violations	CI*		*E*	*V*	*EV*	*E*	*EV*
4	0	1	0	1.000	1.000	1.000	1.000	1.000
17	0	1	0	1.000	1.000	1.000	1.000	1.000
31	0	1	0	1.000	1.000	1.000	1.000	1.000
42	0	1	0	1.000	1.000	1.000	1.000	1.000
57	0	1	0	1.000	1.000	1.000	1.000	1.000
58	0	1	0	1.000	1.000	1.000	1.000	1.000
60	0	1	0	1.000	1.000	1.000	1.000	1.000
65	0	1	0	1.000	1.000	1.000	1.000	1.000
40	0	1	4	0.998	0.920	0.978	0.994	0.971
22	0	1	2	0.998	0.980	0.980	0.994	0.973
20	0	1	2	0.996	0.960	0.974	0.991	0.966
32	0	1	4	0.991	0.940	0.982	0.979	0.976
41	0	1	5	0.990	0.940	0.984	0.976	0.978
50	2	2	9	0.990	0.840	0.916	0.975	0.887
33	0	1	3	0.990	0.980	0.973	0.975	0.964
9	0	1	2	0.989	0.960	0.960	0.973	0.946
27	0	1	2	0.989	0.960	0.969	0.972	0.958
35	0	1	3	0.985	0.980	0.948	0.962	0.931
24	0	1	5	0.985	0.920	0.967	0.962	0.956
15	1	3	9	0.983	0.900	0.965	0.958	0.953
67	0	1	3	0.983	0.940	0.960	0.957	0.947
25	0	1	3	0.981	0.940	0.963	0.953	0.951
68	3	3	9	0.980	0.920	0.948	0.949	0.930
18	1	2	7	0.978	0.880	0.937	0.946	0.915
23	0	1	3	0.978	0.980	0.931	0.944	0.908
8	0	1	1	0.977	0.980	0.971	0.943	0.961
38	0	1	14	0.977	0.940	0.947	0.942	0.929

Table 1 (cont.)

Subject ID	budget balancedness		# GARP violations	Violation indices			Consistency measures**	
	# violations	CI*		*E*	*V*	*EV*	*E*	*EV*
54	0	1	2	0.975	0.960	0.952	0.938	0.935
44	3	3	15	0.972	0.840	0.938	0.930	0.918
30	0	1	15	0.971	0.860	0.933	0.927	0.911
55	4	3	58	0.970	0.780	0.896	0.925	0.861
56	0	1	9	0.968	0.900	0.894	0.919	0.858
10	3	2	55	0.966	0.840	0.836	0.915	0.781
5	2	2	20	0.965	0.880	0.901	0.912	0.868
49	0	1	19	0.965	0.840	0.911	0.912	0.881
59	0	1	30	0.959	0.860	0.909	0.897	0.878
61	9	2	89	0.957	0.760	0.889	0.893	0.852
28	1	3	34	0.957	0.820	0.886	0.892	0.848
62	0	1	41	0.956	0.900	0.905	0.891	0.874
13	1	2	20	0.954	0.800	0.828	0.885	0.770
72	2	2	14	0.952	0.900	0.884	0.879	0.845
39	3	6	76	0.948	0.820	0.822	0.870	0.762
6	1	2	16	0.946	0.940	0.832	0.864	0.776
69	2	2	100	0.939	0.800	0.824	0.848	0.766
12	0	1	22	0.935	0.960	0.593	0.838	0.457
52	9	2	60	0.933	0.700	0.789	0.833	0.718
45	0	1	191	0.931	0.780	0.707	0.827	0.609
37	12	4	480	0.930	0.760	0.590	0.824	0.453
7	7	3	70	0.928	0.680	0.754	0.820	0.671
34	0	1	26	0.928	0.860	0.716	0.819	0.621
51	0	1	54	0.926	0.840	0.774	0.816	0.698
36	0	1	181	0.916	0.840	0.795	0.789	0.726
46	3	5	57	0.902	0.820	0.802	0.756	0.735
29	1	2	63	0.900	0.860	0.812	0.751	0.749

Table 1 (cont.)

Subject ID	budget balancedness		# GARP violations	Violation indices			Consistency measures**	
	# violations	CI*		*E*	*V*	*EV*	*E*	*EV*
73	3	5	221	0.899	0.680	0.676	0.748	0.567
70	0	1	24	0.892	0.840	0.877	0.731	0.835
66	2	2	541	0.865	0.800	0.518	0.663	0.356
64	6	5	132	0.848	0.720	0.693	0.619	0.590
21	0	1	539	0.845	0.820	0.486	0.612	0.313
1	0	1	376	0.844	0.780	0.464	0.610	0.284
11	1	2	209	0.834	0.840	0.658	0.586	0.543
3	4	6	332	0.817	0.700	0.390	0.544	0.185
43	0	1	248	0.811	0.740	0.510	0.529	0.346
14	5	2	19	0.806	0.840	0.741	0.514	0.654
47	17	12	359	0.798	0.600	0.533	0.494	0.377
75	26	6	446	0.792	0.760	0.540	0.479	0.386
63	1	2	73	0.716	0.940	0.507	0.290	0.342
19	15	9	497	0.710	0.660	0.256	0.276	0.006
74	21	15	521	0.697	0.800	0.402	0.242	0.202
53	2	4	942	0.619	0.840	0.196	0.048	-0.074
16	42	15	1005	0.606	0.840	0.205	0.016	-0.062
71	31	23	528	0.582	0.760	0.364	-0.046	0.151
2	8	12	1089	0.517	0.840	0.244	-0.207	-0.010
48	6	20	1037	0.500	0.860	0.069	-0.251	-0.243
26	10	32	797	0.272	0.840	0.185	-0.821	-0.088
76	49	41	1216	0.211	0.860	0.066	-0.973	-0.248

* The confidence interval required to remove all budget balancedness violations.

** The difference between the actual index and the average random-choice index as a fraction of the difference between index of one and the average random-choice index.

(shaded) Lower bounds.

Table 2. Results of individual-level CES demand function estimation

ID	ρ	α	σ	A	r	Standard errors A	Standard errors r	F-value
73	-14.813	1.000	-0.063	9.980	0.937	0.708	0.055	-87.794
46	-5.369	0.658	-0.157	1.108	0.843	0.031	0.040	-86.608
55	-2.786	0.993	-0.264	3.678	0.736	0.178	0.057	-79.227
3	-2.324	0.679	-0.301	1.253	0.699	0.089	0.085	-39.736
43	-0.672	0.961	-0.598	6.814	0.402	1.120	0.134	-39.690
36	-0.511	0.866	-0.662	3.430	0.338	0.270	0.067	-52.577
18	-0.507	0.554	-0.664	1.153	0.336	0.043	0.036	-69.368
37	-0.489	0.685	-0.672	1.686	0.328	0.099	0.071	-48.821
10	-0.476	0.971	-0.678	10.749	0.322	0.974	0.066	-86.621
1	-0.439	0.827	-0.695	2.970	0.305	0.332	0.150	-27.705
64	-0.435	0.962	-0.697	9.446	0.303	1.965	0.182	-23.831
45	-0.326	0.940	-0.754	7.924	0.246	1.046	0.133	-53.673
13	-0.132	0.922	-0.883	8.912	0.117	0.792	0.092	-75.707
61	-0.115	0.959	-0.897	16.978	0.103	1.654	0.126	-98.711
70	-0.111	0.864	-0.900	5.277	0.100	0.448	0.104	-58.064
21	-0.097	0.669	-0.912	1.897	0.088	0.167	0.100	-27.197
56	-0.083	0.517	-0.924	1.065	0.076	0.045	0.043	-60.128
6	-0.060	0.503	-0.943	1.012	0.057	0.020	0.020	-99.715
32	-0.040	0.516	-0.962	1.064	0.038	0.012	0.010	-128.900
40	-0.011	0.505	-0.989	1.020	0.011	0.014	0.012	-115.900
49	0.027	0.530	-1.028	1.133	-0.028	0.061	0.059	-53.317
52	0.114	0.767	-1.129	3.834	-0.129	0.281	0.059	-57.055

Table 2 (cont.)

ID	ρ	α	σ	A	r	Standard errors		F-value
						A	r	
11	0.144	0.927	-1.168	19.575	-0.168	6.190	0.290	-36.330
72	0.170	0.670	-1.205	2.349	-0.205	0.155	0.073	-48.241
38	0.219	0.787	-1.280	5.323	-0.280	0.430	0.084	-61.866
69	0.231	0.638	-1.300	2.092	-0.300	0.197	0.108	-28.393
29	0.260	0.922	-1.351	28.095	-0.351	17.856	0.443	-20.957
34	0.291	0.532	-1.410	1.200	-0.410	0.099	0.077	-30.197
28	0.315	0.893	-1.460	22.187	-0.460	9.393	0.278	-16.375
66	0.329	0.581	-1.490	1.631	-0.490	0.181	0.123	-18.562
15	0.334	0.699	-1.501	3.549	-0.501	0.183	0.057	-72.855
50	0.337	0.579	-1.509	1.614	-0.509	0.084	0.064	-55.396
30	0.428	0.657	-1.747	3.123	-0.747	0.287	0.096	-46.677
9	0.493	0.748	-1.972	8.575	-0.972	0.905	0.110	-60.649
23	0.497	0.746	-1.987	8.508	-0.987	1.295	0.129	-32.981
7	0.543	0.706	-2.187	6.775	-1.187	1.193	0.167	-43.891
41	0.577	0.812	-2.362	31.464	-1.362	8.512	0.151	-75.664
39	0.583	0.649	-2.399	4.393	-1.399	0.709	0.152	-30.944
14	0.615	0.771	-2.598	23.339	-1.598	8.231	0.204	-38.780
5	0.658	0.576	-2.923	2.463	-1.923	0.293	0.230	-31.056
59	0.674	0.522	-3.065	1.317	-2.065	0.146	0.235	-31.640
51	0.702	0.580	-3.351	2.970	-2.351	0.766	0.373	1.957
12	0.823	0.639	-5.659	25.376	-4.659	25.809	1.452	26.995
8	0.942	0.532	-17.346	9.204	-16.346	5.941	3.734	-13.037
54	0.992	0.504	-117.970	6.345	-116.970	12.567	36.202	-96.376

Figure 1A. Decision-level distribution of expenditure on tokens given to others as a fraction of total expenditure on tokens

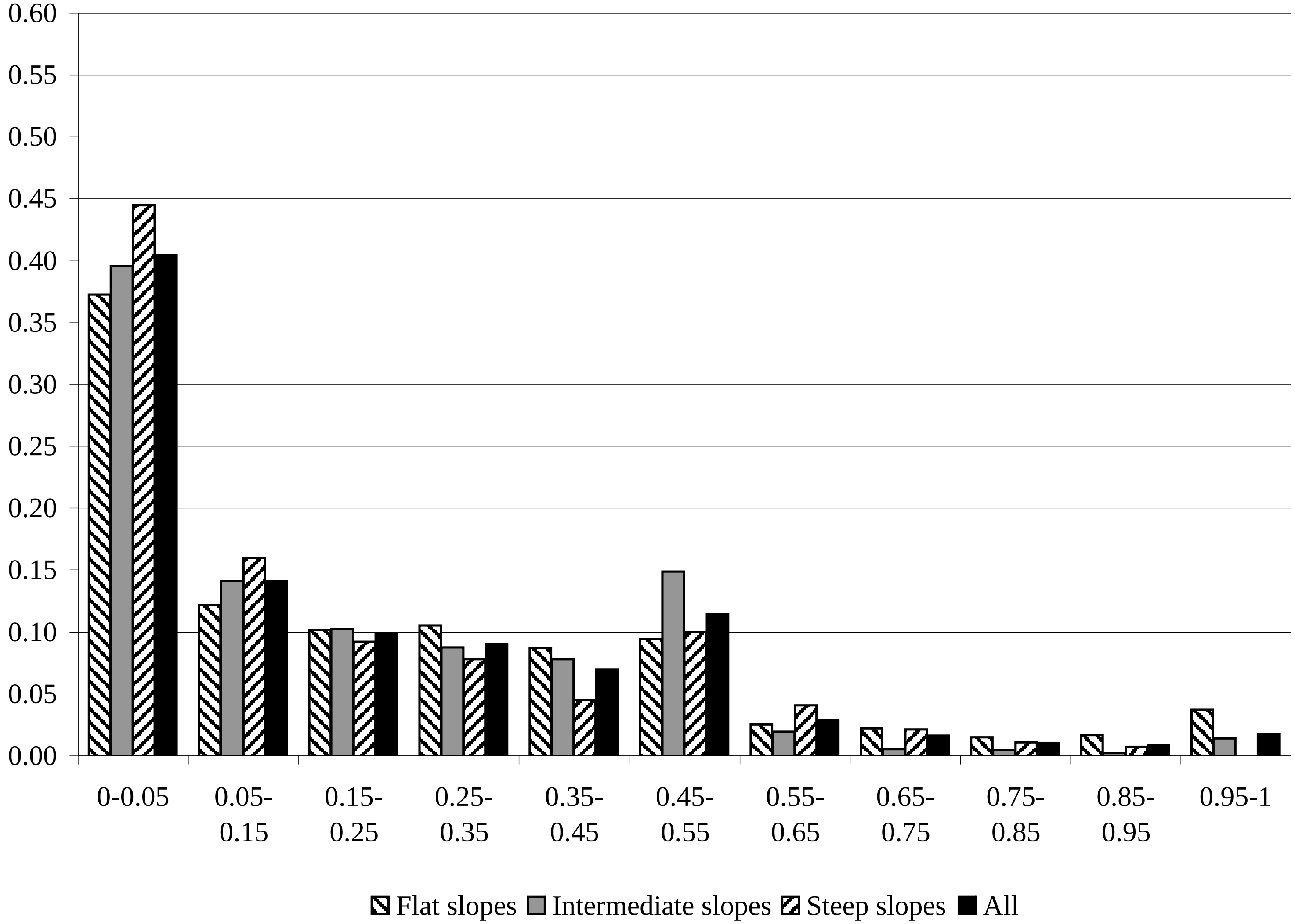

Figure 1B. Decision-level distribution of tokens given to others as a fraction of total tokens kept and given

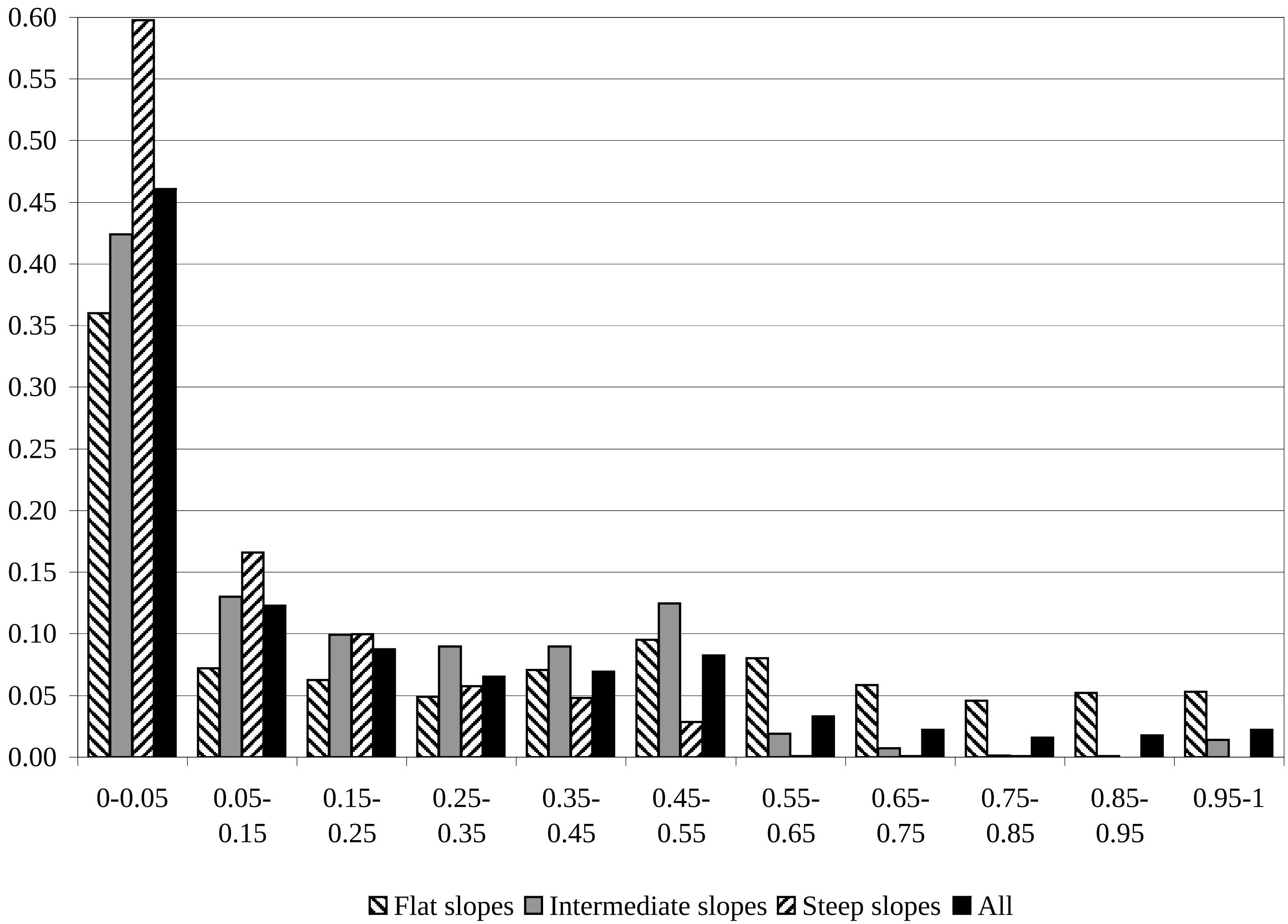

Figure 2. The distribution of the expenditure on tokens given to others a fraction of total expenditure aggregated to the subject level

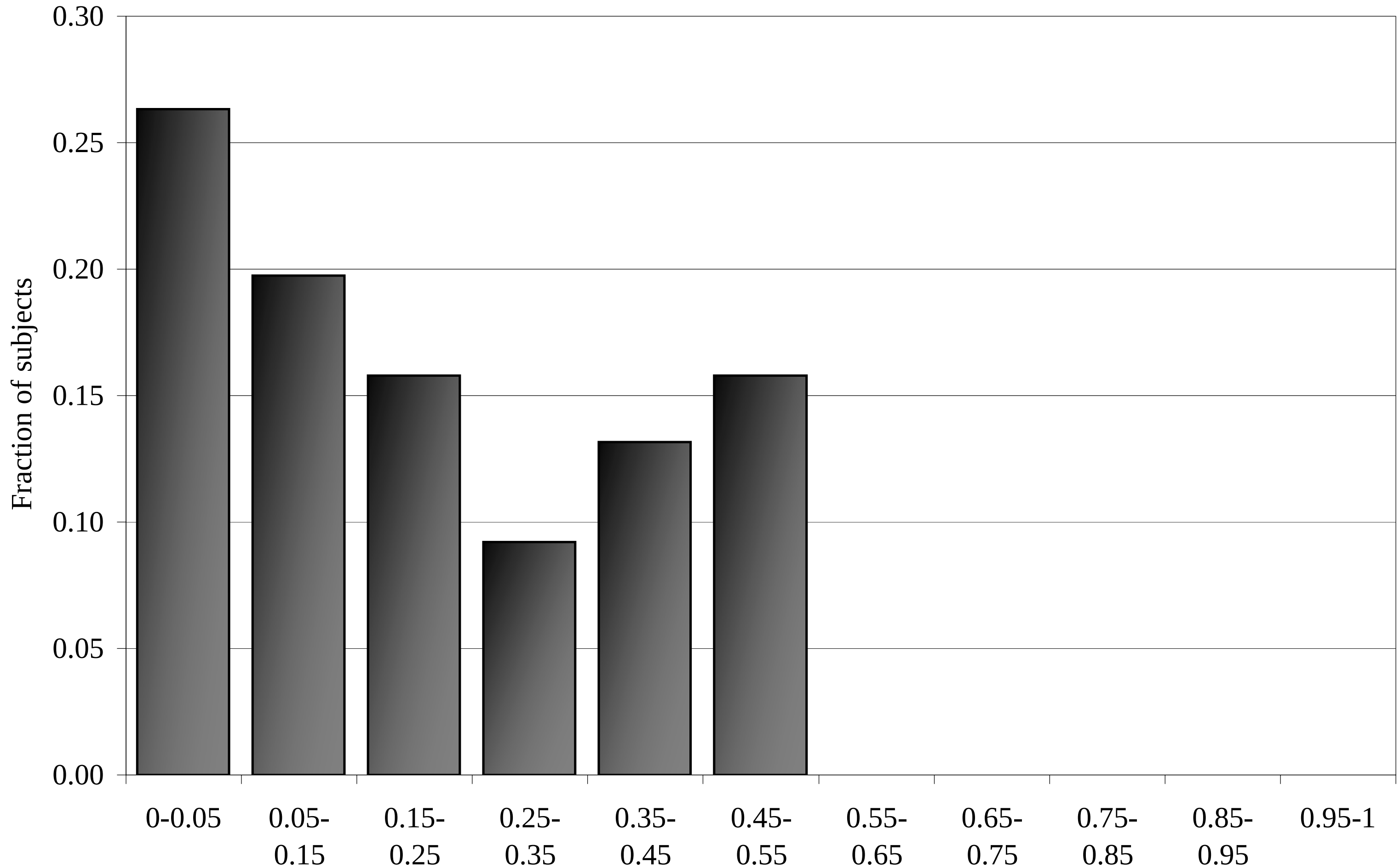

Figure 3. The allocations of subjects with prototypical references

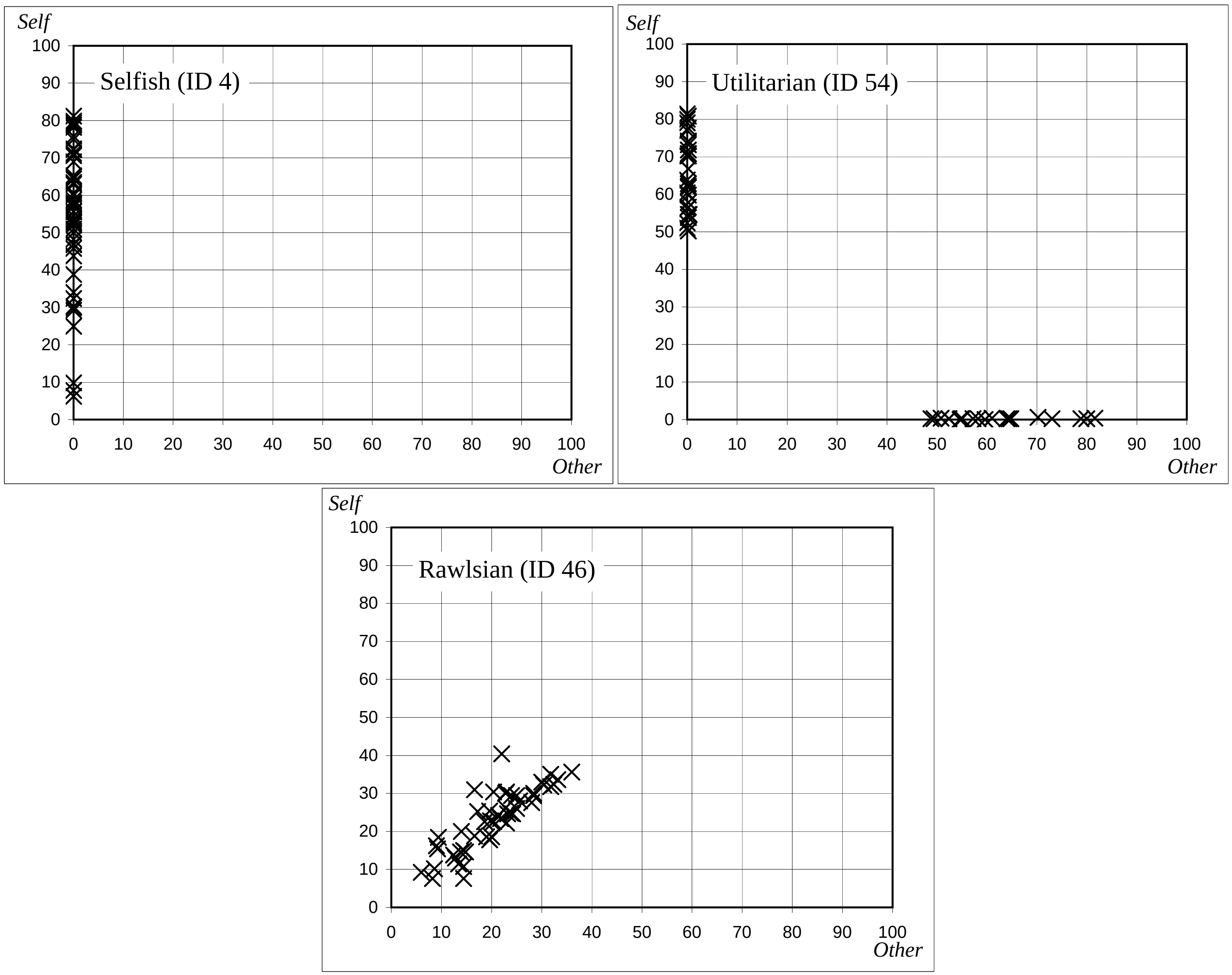

Figure 4. The construction of the E-index for a simple violation of GARP

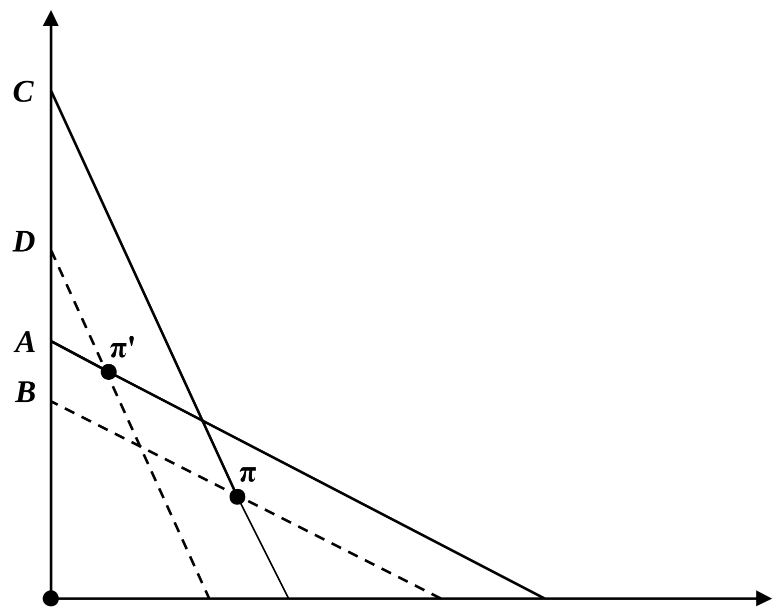

Figure 5. Distributions of GARP violation indices

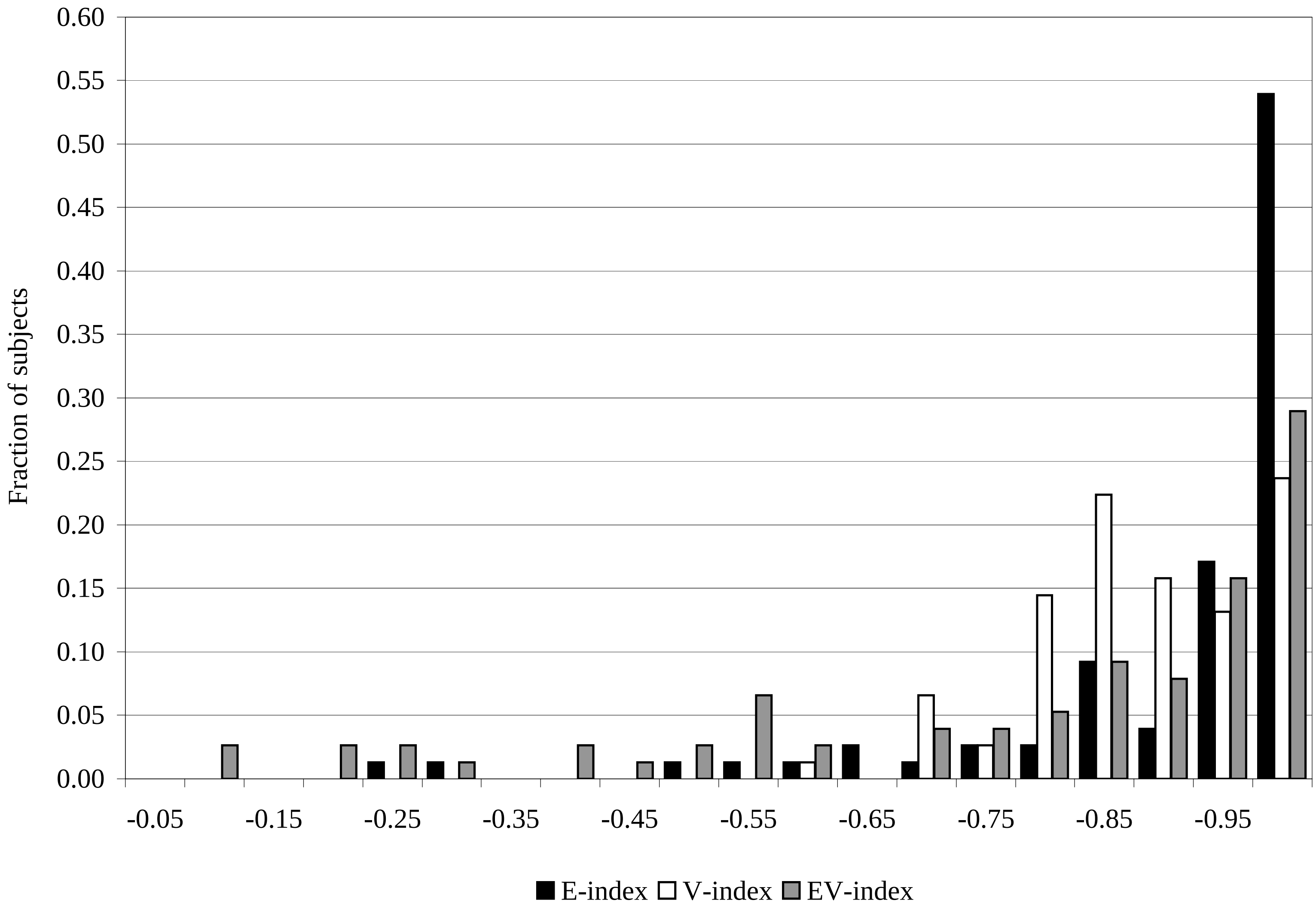

Figure 6. Distribution of (normalized) consistency values

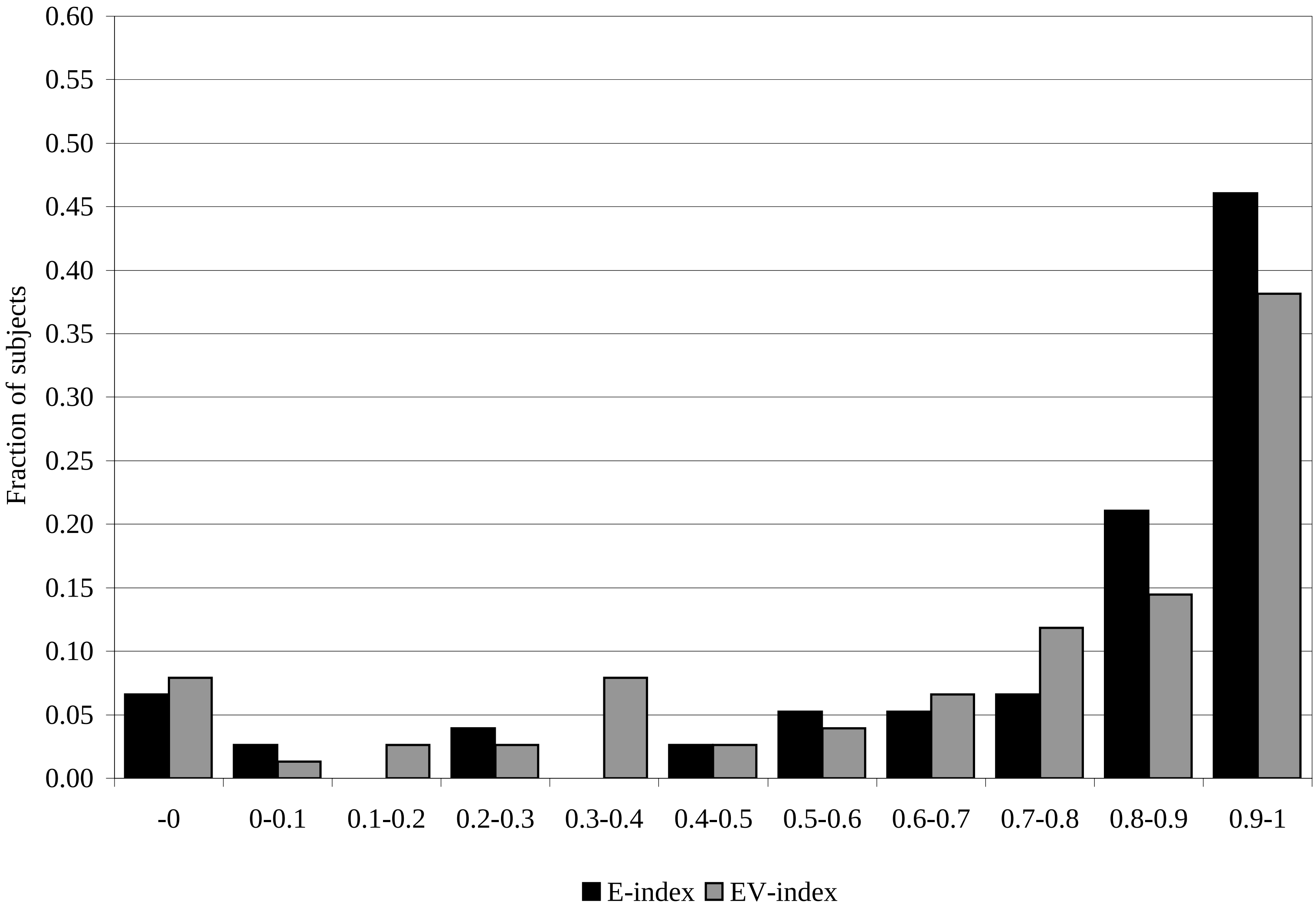

Figure 7. Illustration of recoverability for subjects whose choices fit with prototypical utility functions

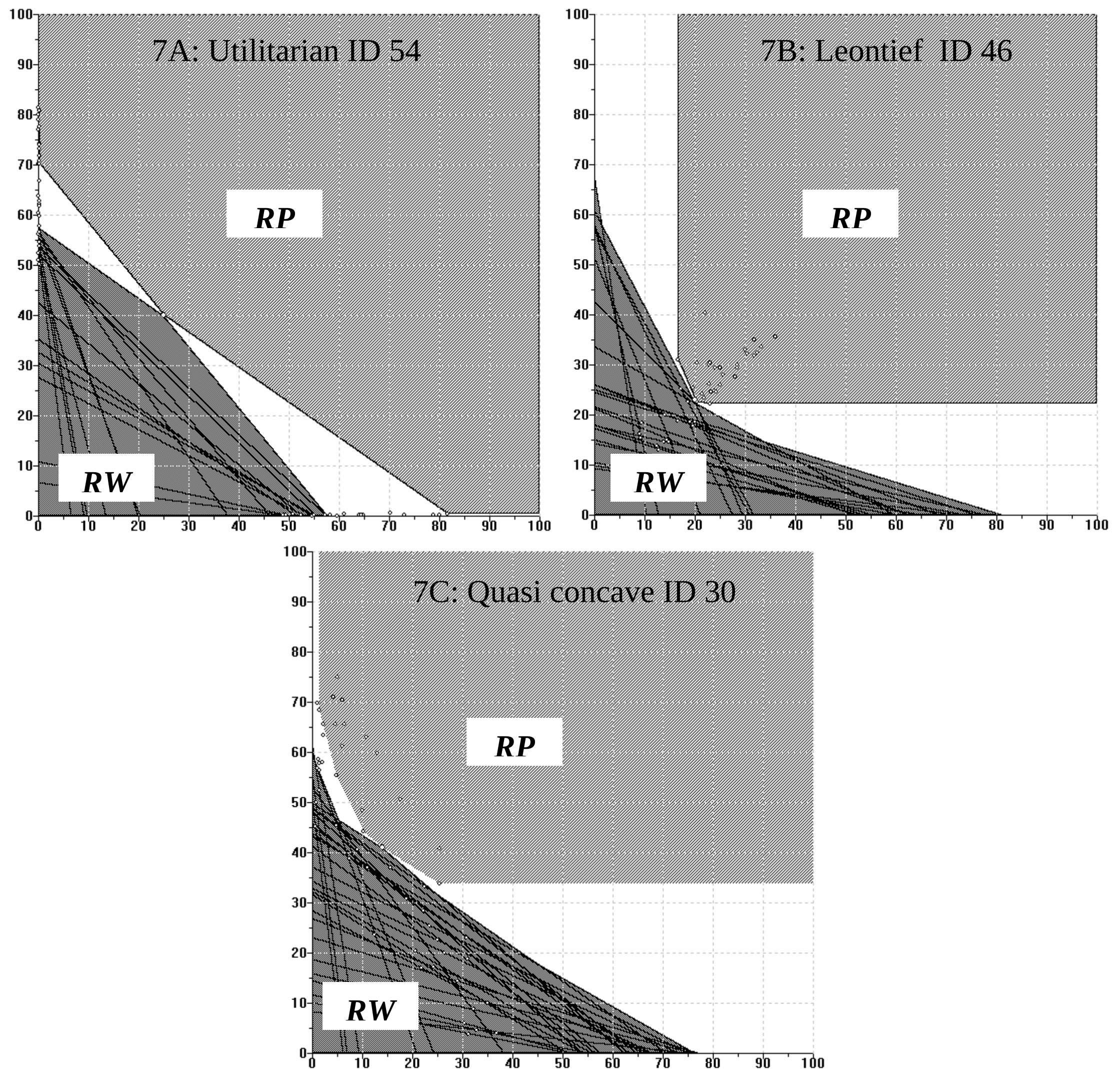

Figure 8. Illustration of forecasting behavior for subjects whose choices fit with prototypical utility functions

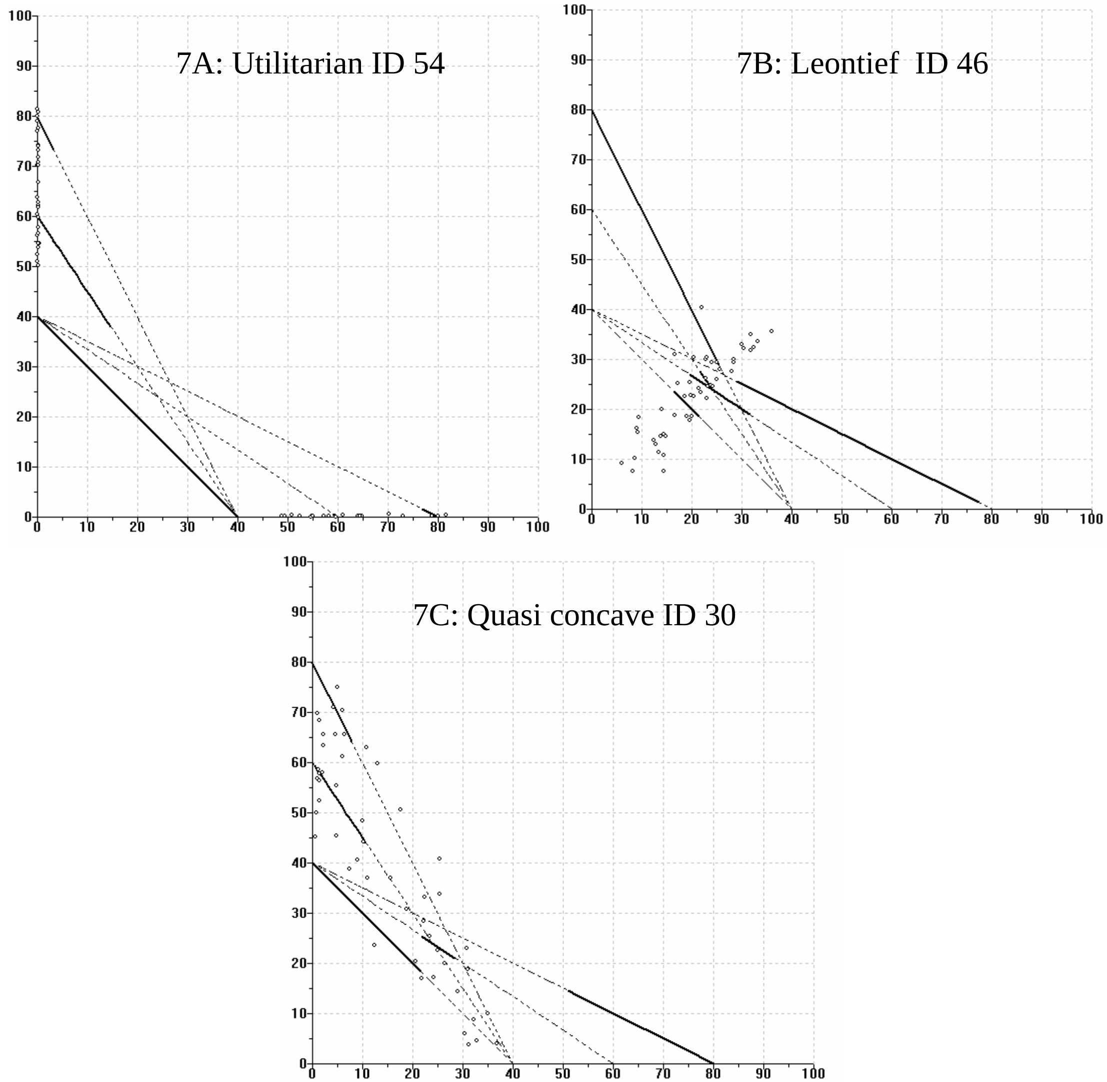

Figure 9. Scatterplot of the CES estimates

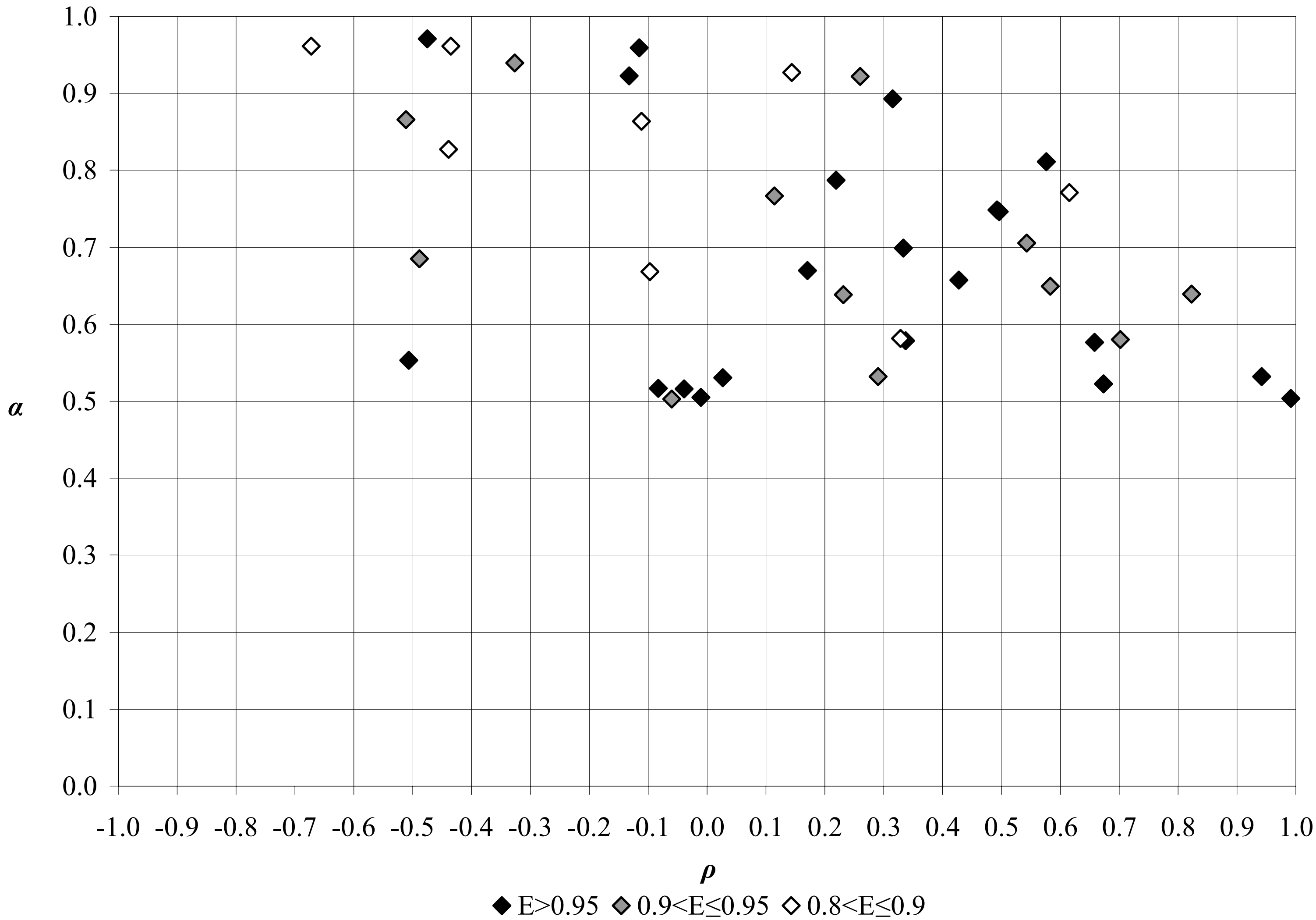

Figure 10. The distribution of the CES parameter
(rounded to a single decimal place)

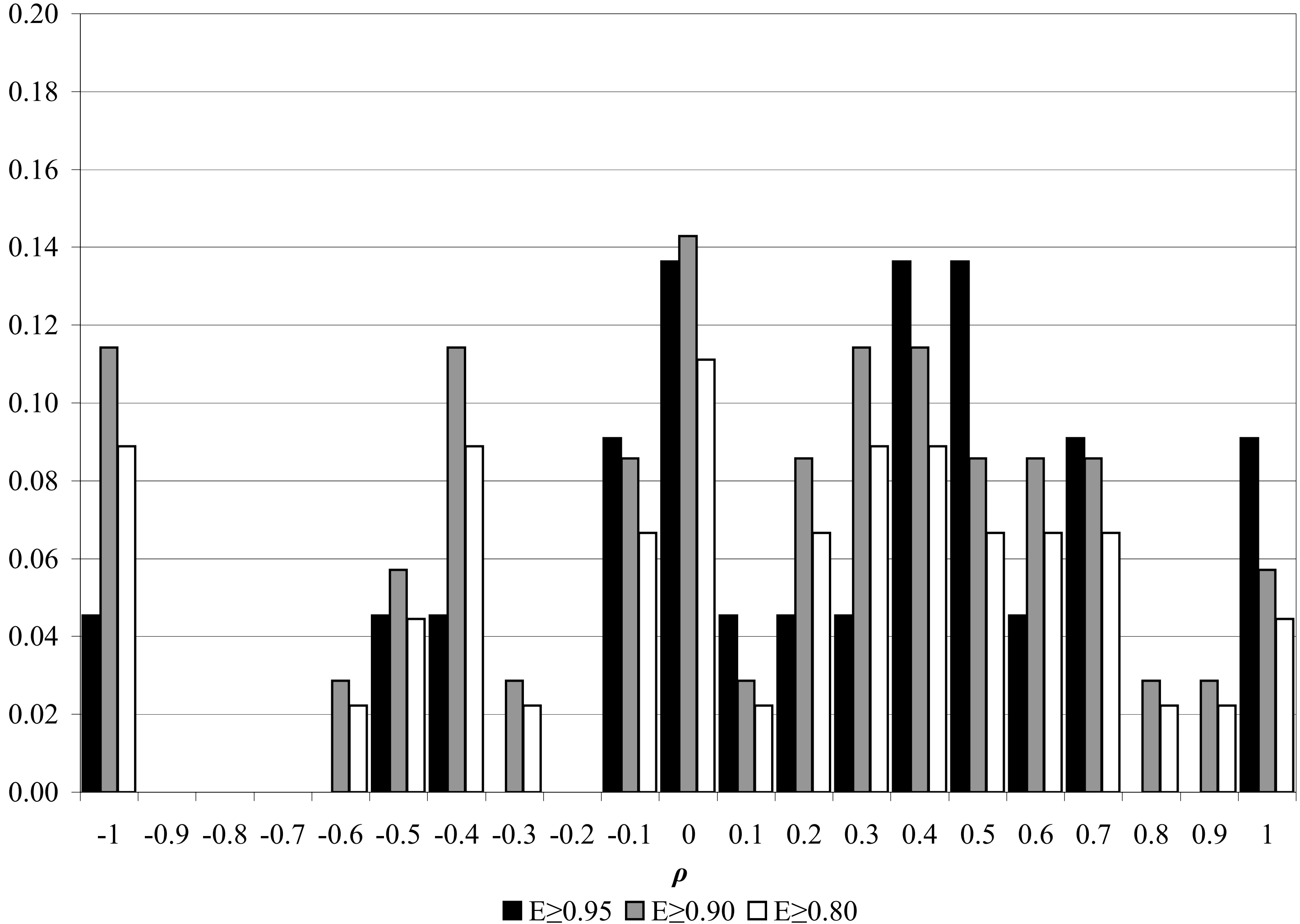

Figure 11. The estimated demand function for giving for selected subjects

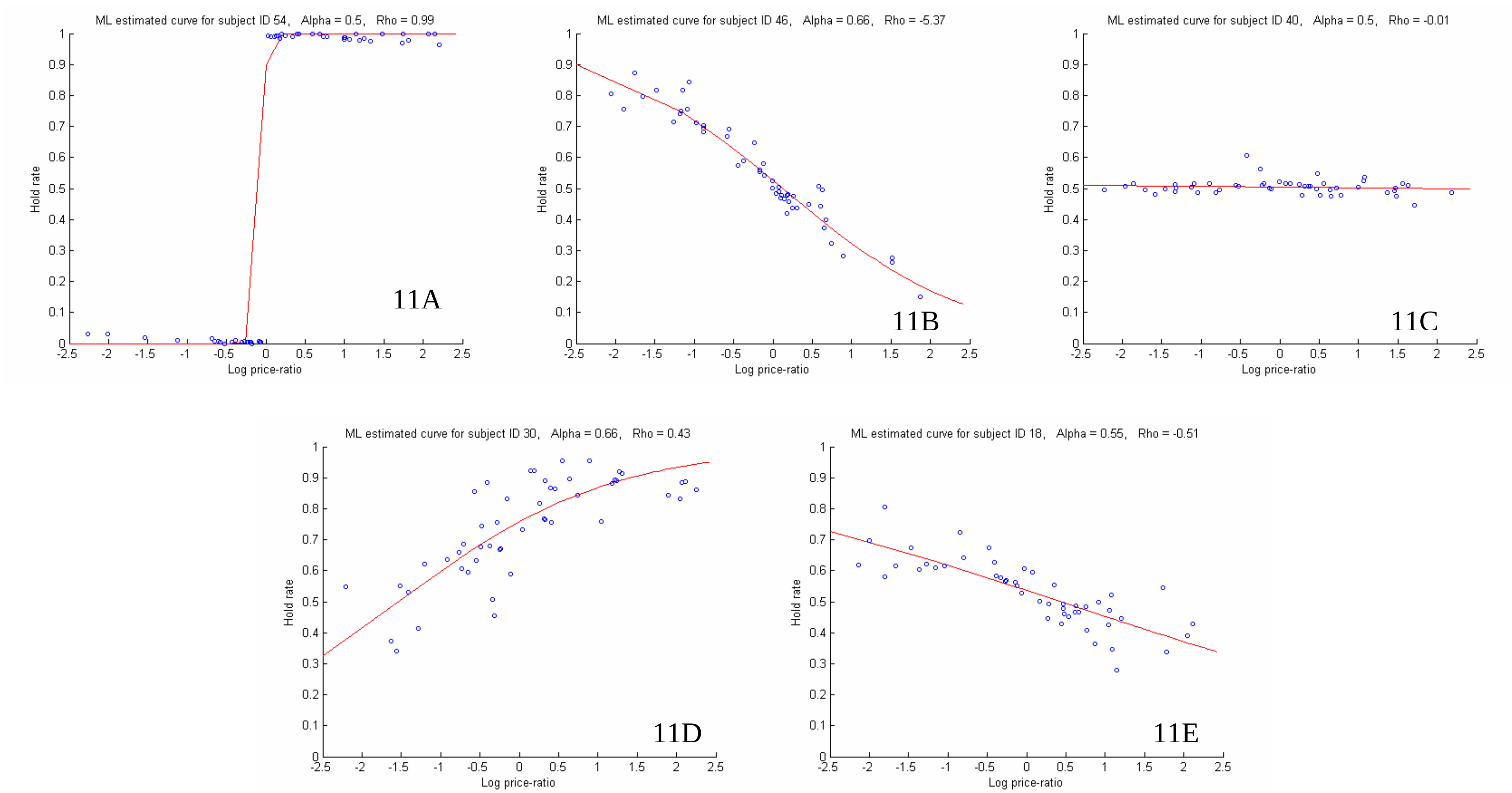

Figure 12. The estimated indirect utility function for selected subjects

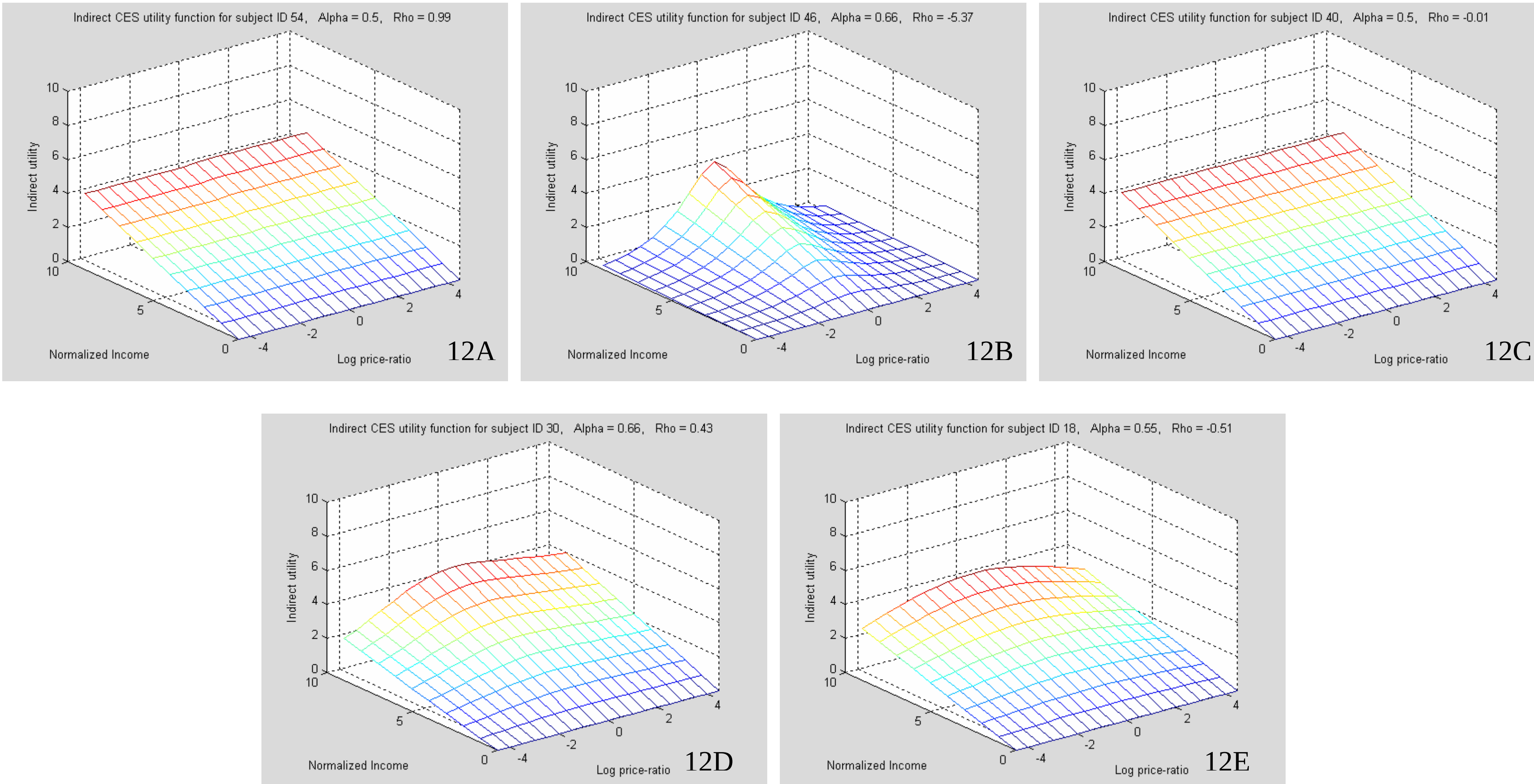

Figure 13. The effect of change in *log*(*p*) on the indirect utility (holding *m* constant) for selected subjects

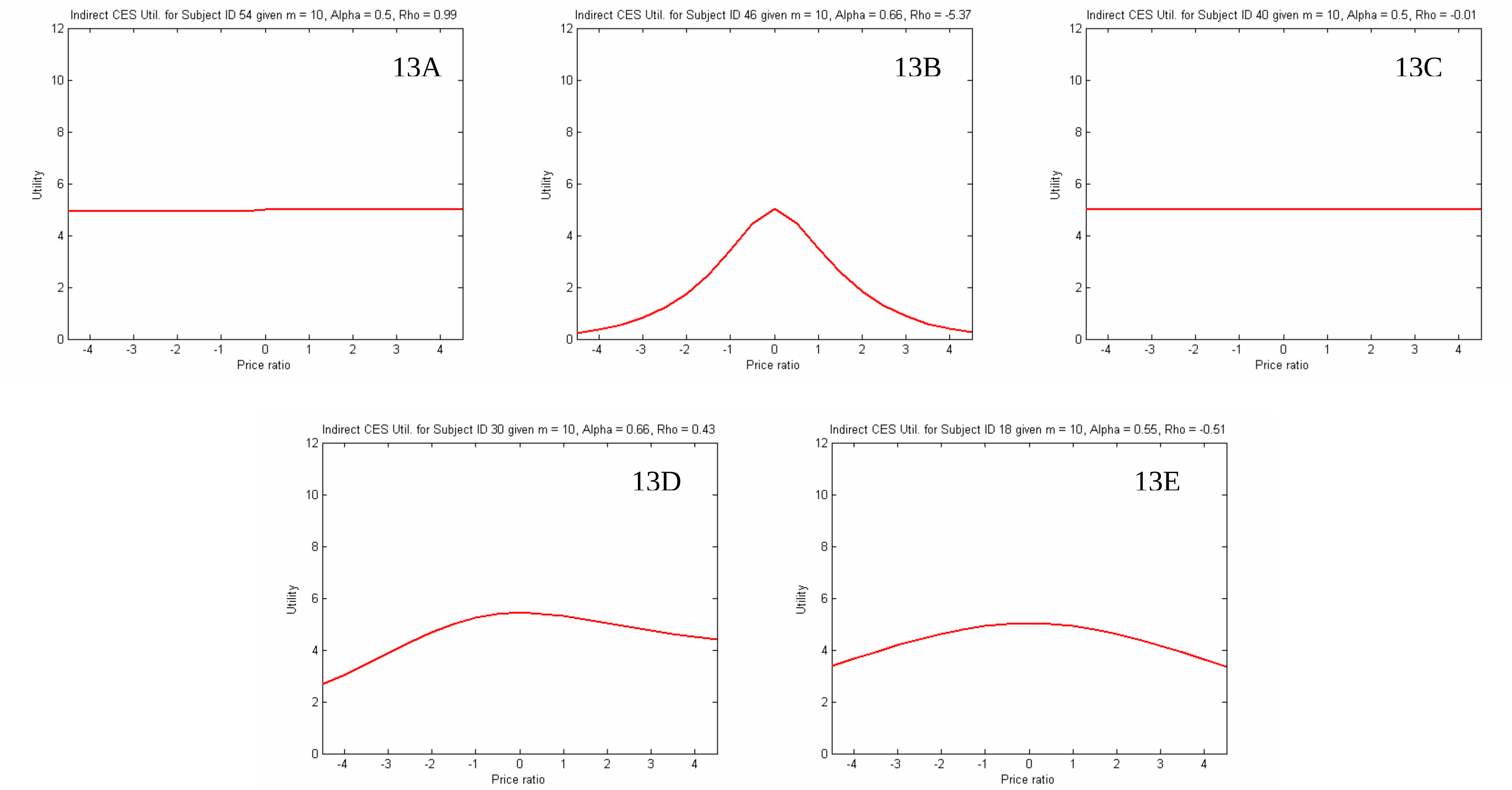

Figure 14. A typical (p,m) indifference curves for selected subjects