Distinguishing Social Preferences from Preferences for Altruism*

Raymond Fisman† Columbia University

Shachar Kariv‡ UC Berkeley

Daniel Markovits§ Yale University

November 15, 2005

Abstract

We report a laboratory experiment that enables us to distinguish preferences for altruism (concerning tradeoffs between own payoffs and the payoffs of others) from social preferences (concerning tradeoffs between the payoffs of others). By using graphical representations of three-person Dictator Games that vary the relative prices of giving, we generate a very rich data set well-suited to studying behavior at the level of the individual subject. We attempt to recover subjects' underlying preferences by estimating a constant elasticity of substitution (CES) model that represents altruistic and social preferences.

*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, Syngjoo Choi, Liran Einav, Douglas Gale, Thomas Palfrey, Ben Polak, and Tim Salmon for helpful discussions. This paper has also benefited from suggestions by the participants of seminars at Berkeley, NYU, Yale and UCSD. Syngjoo Choi provided excellent research assistance. We would also 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. Kariv is grateful to the hospitality of the School of Social Science in the Institute for Advances Studies.

†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).

We find that both social preferences and preferences for altruism are highly heterogeneous, ranging from utilitarian to Rawlsian. In spite of this heterogeneity across subjects, there exists a strong positive within-subject correlation between the efficiency-equity tradeoffs made in altruistic and social preferences. (*JEL*: C79, C91, D64)

1 Introduction

Individuals often sacrifice their own payoffs in order to increase the payoffs of others. Moreover, they do so even in circumstances that do not engage reciprocity motivations or strategic behavior. This has led economists to begin the systematic study of the *distributional preferences* that govern such behavior. Distributional preferences may naturally be divided into two qualitatively different types, which we call *preferences for altruism* and *social preferences*. Preferences for altruism govern the tradeoffs that a person $self$ makes between her payoffs and the payoffs of $others$ (i.e. all persons except $self$). Social preferences govern the tradeoffs $self$ makes among the payoffs to $others$.[1]

Although the two types of distributional preferences often operate together, as when we decide both how much to give to charity and how to allocate our donations across causes, they remain conceptually distinct. Certainly there is no *a priori* reason to insist that preferences for altruism and social preferences have the same (or even a similar) form. Indeed, it is natural to suspect that while preferences for altruism will be influenced by some measure of bias in favor of $self$ over $others$, social preferences (at least over distributions to *anonymous* $others$) will exhibit no indexical preference for any particular $other$. Similarly, it seems at least plausible that attitudes towards inequality, and consequently willingness to trade equality and efficiency, will differ depending on whether or not $self$ is implicated in the inequalities at issue. It is surprising, therefore, that in spite of the massive outpouring of work on distributional preferences in recent years, little attention has been paid to the distinction between preferences for altruism and social preferences and, moreover, that there has been virtually no systematic experimental study of social preferences. The lack of positive studies of social preferences is all the more striking given the richness and prominence of the normative economic analysis of social preference in the social choice

[1] We know that the terms "distributional preferences" and "social preferences" are used interchangeably in the literature and that our usage is not quite standard. Nevertheless, the distinctions that we draw are straightforward and (as our analysis reveals) capture important differences.

literature.

This gap in economic understanding is practically important. Distributional preferences quite generally, including both preferences for altruism and social preferences, are important inputs into any broader measure of social welfare, so that correctly distinguishing social preferences from preferences for altruism, and accurately measuring both, is crucial to evaluating a range of socioeconomic polices and institutions. Finally, the empirical study of social preferences, including especially their relationship to preferences for altruism, is essential to understanding the practical influence of broader theories of justice. These theories suggest, in the spirit of Harsanyi and Rawls, that *fair-minded* people should aspire to apply unified distributive principles across both realms.

In this paper, we seek to initiate the systematic experimental study of social preferences by distinguishing them experimentally from preferences for altruism and comparing these two classes of distributional preferences. In order better to focus on behavior motivated by purely distributional preferences, we restrict attention to a dictator game and ignore the complications that strategic behavior and reciprocity introduce in response games. We use a novel graphical representation of three-person dictator games that vary the *relative prices of giving*, so that each subject faces a large and rich menu of *budget sets* representing the feasible monetary payoffs for $self$ and two $others$. This environment is richer and more flexible than the one in the existing literature. Most importantly, it generates a very rich data set well-suited to studying behavior at the level of the individual subject.

With these data, we can thoroughly address three types of questions concerning distributional preferences. First, and most narrowly, how does increasing the number of $others$ affect preferences for altruism? Second, how can social preferences be characterized experimentally? And third, what is the relationship between preferences for altruism and social preferences? We emphasize that we investigate behavior at the level of the individual subject and thus also thoroughly address other sorts of questions concerning behavior, such as whether behavior is consistent with the utility maximization model and how distributional preferences differ across subjects.

Our results can be summarized as follows. First, the extent of altruism changes surprisingly little when there are two potential beneficiaries of altruism rather than just one. We compare this study with an earlier study of the otherwise identical two-person dictator experiment of Fisman, Kariv, and Markovits (2005) (hereafter, FKM) and find that although our current subjects did on balance give more away in the presence of two $others$ than the subjects in the two-person experiment gave away in the presence of one *other*, the addition of a second *other* fell far short of generating a proportional increase in the overall level of giving. We also extend the conclusions of FKM that classical demand theory can explain altruistic preferences and that although individual preferences for altruism are highly heterogeneous, and range from Rawlsian to utilitarian to perfectly selfish, subjects display a pronounced (although far from monolithic) emphasis on increasing aggregate payoffs of $self$ and $others$ rather than reducing the differences in payoffs between $self$ and $others$.

Second, we take up social preferences and provide a comforting confirmation of the strong (indeed almost irresistible) intuition that social preferences should accord equal weight to payouts given to anonymous $others$. Moreover, and more substantially, we find that classical demand theory can also explain social preferences and that although individual social preferences are again highly heterogeneous, and range from Rawlsian to utilitarian, they also display a pronounced (although far from monolithic) emphasis on increasing aggregate payoffs rather than reducing the differences in payoffs between $others$.

Third, and most importantly, we compare preferences for altruism and social preferences and find (although with a few interesting exceptions) that subjects display a strong positive correlation between the efficiency-equity tradeoffs that they make in their altruistic and social preferences. Thus, although there is considerable heterogeneity in preferences for altruism and social preferences *across* subjects, there is a strong association between preferences for altruism and social preferences *within* subjects. This finding decides a genuinely open question rather than just confirming rigorously what was already intuitively clear. Inequality between $self$ and $others$ and inequality across $others$ are entirely distinct phenomena; no more closely connected conceptually than $self$ - *other* and *other* - *other* authority relations, for example. There is therefore no a priori reason why attitudes to the two types of inequality should be related.

Our paper thus contributes to the vast body of research on distributional preferences, including Loewenstein, Bazerman, and Thompson (1989), Bolton (1991), Rabin (1993), Levine (1998), Fehr and Schmidt (1999), Bolton and Ockenfels (1998, 2000), Charness and Rabin (2002), and Andreoni and Miller (2002) among others. Camerer (2003) provides a comprehensive discussion of experimental and theoretical work in economics focusing on dictator, ultimatum and trust games. Charness and Rabin (2002) test a few simple three-person dictator and ultimatum games. They conclude that, contrary to assumptions made by Bolton and Ockenfels (1998, 2000) and elsewhere, subjects are not indifferent to the distribution of payoffs among other individuals.

The rest of the paper is organized as follows. Section 2 introduces the template for our analysis. Section 3 describes the experimental design and procedures. Section 4 summarizes some important features of the data. Section 5 describes the consistency of the data with the maximization hypothesis. Section 6 discusses the results and Section 7 contains some concluding remarks. The experimental instructions are reproduced in Section 8.

2 Template for Analysis

We investigate choices made by a person $self$ that have consequences for her own payoff and the payoffs of two *anonymous* $others$. Throughout, we denote persons $self$ and $others$ by S and $O = \{A, B\}$, respectively, and the associated monetary payoffs by π_S and a *profile* $\pi_O = (\pi_A, \pi_B)$. Specifically, we study a three-person dictator game in which $self$ must allocate an *endowment* m across $\pi = (\pi_S, \pi_O)$ at *prices* $p = (p_S, p_O)$, such that $p_S\pi_S + p_O\pi_O = m$. This configuration creates *budget sets* over π_S and π_O. An example of one such budget set is illustrated in Figure 1 below. This game is a generalization of the game employed in FKM to study individual preferences for altruism, and the present study therefore incorporates the earlier study's methodological advances in analyzing behavior at the level of the individual subject.

[Figure 1 here]

Given observations on individual-level data (p^t, π^t) (i.e. the t^{th} observation of prices and associated quantities), a *nondegenerate* utility function $U_S = u_S(\pi_S, \pi_O)$ that captures the possibility of giving is said to *rationalize* the behavior of $self$ if $u_S(\pi^t) \geq u_S(\pi)$ for all π such that $p^t\pi^t \geq p^t\pi$ (i.e. u_S achieves the maximum on the *budget set* at the chosen bundle). If a well-behaved utility function $u_S(\pi_S, \pi_O)$ that the choices maximize exists, it becomes natural to explore the structure of the utility functions that rationalize the observed data. This is of particular interest insofar as it facilitates the analysis of the two types of *distributional preferences* that our experiment engages – *preferences for altruism* and *social preferences*. Once again, preferences for altruism address tradeoffs between the payoffs to $self$ and the payoffs to $others$. Person $self$ is perfectly *selfish* when $u_S(\pi) \geq u_S(\pi')$ if and only if $\pi_S \geq \pi'_S$ and otherwise displays some form of *altruism*. In contrast, social preferences address tradeoffs between the payoffs to $others$ (i.e. all persons except $self$).

A common assumption used in demand analysis allows for a clear demarcation between social preferences and preferences for altruism:

Independence *For any* π_S, π'_S, π_O *and* π'_O, $u_S(\pi_S, \pi_O) > u_S(\pi_S, \pi'_O)$ *if and only if* $u_S(\pi'_S, \pi_O) > u_S(\pi'_S, \pi'_O)$.

The independence property entails that if π_O is preferred to π'_O for some π_S, then π_O is preferred to π'_O for all π_S. That is, the preferences of $self$ over the payoffs of $others$ are independent of her $self$-interestedness. If this independence property is satisfied, then the utility function $u_S(\pi_S, \pi_O)$ is (weakly) *separable* in the sense that we can find a *subutility* function $w_S(\pi_O)$ and a *macro* function $v_S(\pi_S, w_S)$ with v_S strictly increasing in w_S such that

$$u_S(\pi_S, \pi_O) \equiv v_S(\pi_S, w_S(\pi_O)).$$

This formulation makes it possible to represent distributional preferences in a particularly convenient manner, because the macro utility function $v_S(\pi_S, w_S)$ represents preferences for altruism (i.e. $self$ versus $others$), whereas the subutility function $w_S(\pi_O)$ represents social preferences (i.e. *other* versus *other*).[2] Moreover, separability imposes convenient (if restrictive) patterns on demand behavior. First, separability entails that the substitutability between the payoffs for $others$ is independent of the payoff for $self$. Separability also entails that the payoff for any *other* person is a function only of the prices p_O and the total expenditure on $others$. The price p_S is relevant only insofar as it affects the total expenditure on $others$.

Although a separable utility function is very convenient for distinguishing preferences for altruism from social preferences and commonly employed in demand analysis, it should not necessarily be given any psychological interpretation. This approach is useful in interpreting the data, but our analysis does not stand or fall on the literal truth of separability. Rather, we use separable utility as an *as if* methodology and confront this formulation with the experimental data. We are not dogmatic about this approach; it just seems a natural starting point. Finally, we note that the starting point of several theories, such as Fehr and Schmidt (1999) and Charness and Rabin (2002), is to make rather specific and quit distinct assumptions on the form of the utility function in order to yield empirically testable restrictions on observed behavior.

[2] Karni and Safra (2000) introduce an axiomatic model of choice among random social allocation procedures. Their utility representation is also decomposed in a similar way, and they also provide conditions under which the representation is additively separable.

3 Design and Procedures

The experiment was conducted at the Experimental Social Science Laboratory (X-Lab) at UC Berkeley under the X-Lab Master Human Subjects Protocol. The 65 subjects in the experiment were recruited from all undergraduate classes and staff at UC Berkeley and had no previous experience in experiments of dictator, ultimatum, or trust games. After subjects read the instructions (see Section 8), the instructions were read aloud by an experimenter. No subject reported any difficulty understanding the procedures or using the computer program. Each experimental session lasted for about one and a half hours. A \$5 participation fee and subsequent earnings, which averaged about \$15, were paid in private at the end of the session. Throughout the experiment we ensured anonymity and effective isolation of subjects in order to minimize any interpersonal influences that could stimulate other-regarding behavior.

The procedures described below are identical to those used by FKM to study two-person dictator games. Each experimental session consisted of 50 independent decision-problems. In each decision problem, each subject was asked to allocate tokens between $self$ and two anonymous $others$, indexed by A and B, that were 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 π_S and $\pi_O = (\pi_A, \pi_B)$. Each decision problem $t = 1, ..., 50$ started by having the computer select a budget set randomly from the set

$$p_S^t \pi_S + p_A^t \pi_A + p_B^t \pi_B = m^t$$

where $m^t/p_i^t \leq 100$ for all persons and $m^t/p_i^t \geq 50$ for at least one person (so that the budget sets intersected with at least one of the axes at 50 or more tokens, but with no intercept exceeding 100 tokens). The budget 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. In contrast with the two-person games reported in FKM, choices were restricted to allocations on the budget constraint, so that subjects could not violate *budget balancedness*, $p^t \pi^t = m^t$.

The π_S-axis, π_A-axis and π_B-axis were labeled Hold, Pass A and Pass B 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 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 randomly repositioned on the budget set at the beginning of each round. 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. The computer program dialog window is shown in Attachment 3 in Section 8.

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 subjects with whom he was matched received the tokens that he passed π_A and π_B. Thus, each subject received three groups of tokens, one based on her own decision to hold tokens and two based on the decisions of two other random subjects to pass tokens. The computer program ensured that the same two subjects were not paired twice. At the end of the experiment, the tokens were converted into money. Each token was worth \$0.25. Subjects received their payments privately as they left the experiment.

Experimental research has been very fruitful in both establishing the empirical reliability of distributional preferences and directing theoretical attention to such preferences. At the same time, existing work has typically collected only a few decisions from each subject and offered subjects a binary choice in extreme rather than typical decision problems designed to encourage violations of specific theories. Although this is understandable given the purposes for which the experiments were designed, it limits the usefulness of data generated for other purposes. Most importantly, the small data sets in the existing literature force experimenters to pool data and to ignore individual heterogeneity.

This has led us to develop an experimental design that is quite different from those used in the literature. The choice from a budget set provides more information about preferences than a discrete choice would reveal and allows us to apply powerful techniques from demand analysis to determine whether the observed behavior is consistent with utility maximization. Additionally, our experiments employ decision problems that are representative (both in the statistical sense and in the economic sense) of broad classes of distributional choices rather than being narrowly tailored to capture a particular phenomenon. Finally, the rich data sets generated by this design allow us to analyze behavior at the level of the individual subject. There is no need to pool data or to assume that subjects are homogeneous.

4 Descriptive Statistics

We begin with an overview of some basic features of the experimental data. For comparative purposes, we present our results alongside the results of the two-person experiment reported in FKM. The histograms in Figure 2 show the distributions of the fraction given to *others*, defined in a couple of ways, and compare them with the analogous distributions of the fraction given to *other* reported in FKM. Figure 2A depicts the distribution of the expenditure on tokens given to *others* as a fraction of total expenditure

$$\frac{p_A^t \pi_A^t + p_B^t \pi_B^t}{p_S^t \pi_S^t + p_A^t \pi_A^t + p_B^t \pi_B^t},$$

which captures the presence of price changes, and Figure 2B depicts the distribution of the tokens given to *others* as a fraction of the sum of the tokens kept and given

$$\frac{\pi_A^t + \pi_B^t}{\pi_S^t + \pi_A^t + \pi_B^t}.$$

The horizontal axis identifies the fractions for different intervals and the vertical axis reports the percentage of decisions corresponding to each interval.

[Figure 2 here]

In Figure 2A, the distributions in the two- and three-person experiments are quite similar, although, most interestingly, there is a larger fraction of *selfish* allocations of 0.05 or less of the total expenditure on tokens for *others* in the three-person case. The patterns in Figure 2B are even more similar. Additionally, perhaps as expected, in the three-person case, subjects gave more than half of the tokens to *others* with much greater frequency than in the two-person case. Overall, our subjects gave approximately 26 percent of the tokens to *others*, accounting for 25 percent of total expenditure, which is only marginally higher than the 19 percent and 21 percent, respectively, in the two-person experiment reported in FKM. In the studies of standard split-the-pie two-person dictator games reported in Camerer (2003), the typical mean allocations are of about 20 percent.

Figure 2 potentially obscures the presence of individual concerns *on average* for *others*. For example, a person who gives everything to *others* 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 3 shows the distribution of the expenditure on tokens given to *others* as a fraction of total expenditure averaged at the subject level and compares it with the analogous distribution in the two-person experiment reported in FKM. Since this takes an average over all prices, the distribution should be similar for the tokens given to *others* as a fraction of the sum of the tokens kept and given (in practice, the histograms are identical). The horizontal axis identifies the fractions for different intervals and the vertical axis reports the percentage of subjects corresponding to each interval.

[Figure 3 here]

Both distributions in Figure 3 show a pattern with a mode around the midpoint (i.e. same total expenditure on $self$ and *others*). Interestingly, the mode around the midpoint is more pronounced in the three-person than in the two-person case, in spite of the increased number of *others*. Moreover, only seven subjects (10.8 percent) in the three-person experiment spent, on average, more than half of their endowment on tokens given to *others*. We consider this to be surprisingly low, although no subjects in the two-person experiment spent more than half of their endowment on *others* on average. Finally, and perhaps also surprisingly, the three-person experiment found a larger fraction of selfish subjects (who, on average, spent less than 0.05 of their endowment on the tokens given to *others*) than the two-person experiment, although this difference is not statistically significant (p-value=0.14).

Since persons A and B are two anonymous *others* and thus indistinguishable from the perspective of $self$, it would be natural for them to receive approximately equal total allocations, aside from the heterogeneity generated by differences in the (random) prices. In other words, preferences over π_O should not depend on the identity of *others*, only the levels of payoffs involved. To investigate how $self$ trades off the payoff of person A against that of person B, Figure 4 depicts the distribution of the expenditure on tokens given to person A as a fraction of total expenditure on tokens given to *others*

$$\frac{p_A^t \pi_A^t}{p_A^t \pi_A^t + p_B^t \pi_B^t}.$$

After screening the data for selfish allocations that spend 0.05 or less of the total expenditure to *others* (which account for 50.2 percent of all allocations), we present the distribution based on the full sample, as well as distributions with the sample divided into three *relative distributional price* terciles: intermediate relative prices of around 1 ($0.70 \leq p_A/p_B \leq 1.43$), steep prices ($p_A/p_B > 1.43$) and symmetric flat prices ($p_A/p_B < 0.70$).

[Figure 4 here]

For the full sample, the distribution is nearly symmetric around the midpoint of 0.5 (i.e. same expenditure on persons A and B) indicating that $others$ are treated identically on average. For the distributions by tercile, we find that there is also a very pronounced mode at the midpoint of 0.5 for the middle tercile. Most interestingly, the distribution for the steep tercile is bimodal with modes at $0.95 - 1$ and $0.35 - 0.45$. For the flat tercile, the pattern is the mirror image. Thus, subjects respond symmetrically to changes in the relative price p_A/p_B. We obtain similar patterns if we look at the distribution of the tokens given to person A as a fraction of the tokens given to $others$ $\pi^t_A/(\pi^t_A + \pi^t_B)$. The only difference is that the intermediate mode is at $0.45 - 0.55$ rather than $0.35 - 0.45$.

In Figure 5, we averaged at the subject level the distribution of the expenditure on tokens given to person A as a fraction of total expenditure on tokens given to $others$ presented in Figure 4, with the sample limited to the 41 subjects (63.1 percent) that did not make exclusively selfish allocations. For 29 of these subjects (70.7 percent), the fraction of expenditure on tokens given to person A as a fraction of total expenditure on tokens given $others$ is between 0.45 and 0.55. This increases to a total of 38 subjects (92.7 percent) if we consider the bounds $0.35 - 0.65$. Thus, $others$ are treated symmetrically by $self$. This is a natural result of the anonymity of $others$.

[Figure 5 here]

5 Testing for Consistency

Before postulating a parametric family of functional forms for the utility function and fitting the derived demand functions to the data, we test wether choices can be utility-generated. Let (p^t, π^t) for $t = 1, ..., 50$ be some observed individual data (i.e. p^t denotes the t^{th} observation of the prices and π^t denotes the associated allocation). Throughout this section it will be convenient to normalize the prices by the endowment at each observation so that $p^t\pi^t = 1$ for all t.

Following *Afriat's theorem*, we employ the *Generalized Axiom of Revealed Preference* (GARP) to test whether the finite set of observed price and quantity data that our experiment generated may be rationalized by a utility function $u_S(\pi_S, \pi_O)$. GARP (which is a generalization of various other revealed preference tests) requires that if π^t is indirectly revealed preferred to π^s, then π^s is not *strictly* directly revealed preferred (i.e. $p^s \pi^t \geq p^s \pi^s$) to π^t. The theory tells us that if the data satisfy GARP, then a utility function that rationalizes the observed allocations exists and, moreover, may be chosen to be *increasing*, *continuous* and *concave*.

The broad range of budget sets that our experiment involves provides a rigorous test of GARP. 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, 1983), Bronars (1987) and Andreoni and Harbaugh (2005)). Our experiment is therefore sufficiently powerful to detect whether or not utility maximization explains behavior in the laboratory. We refer the interested reader to the Appendix for details on testing for consistency with 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 assess how nearly the data complies with GARP by calculating Afriat's (1972) *Critical Cost Efficiency Index* (CCEI). This measures the amount by which each budget constraint must be relaxed in order to remove all violations of GARP. The CCEI is bounded between zero and one. The closer it is to one, the smaller the perturbation of budget sets required to remove all violations and thus the closer the data are to satisfying GARP.

Over all subjects, the CCEI scores averaged 0.924 which is close enough to passing GARP to suggest that our subjects' choices are indeed consistent with utility maximization. To make this suggestion more precise, we generate a benchmark against which to compare our CCEI scores using the test designed by Bronars (1987) which builds on Becker (1962) and employs the choices of a hypothetical subject who randomizes uniformly among all allocations on each budget set as a point of comparison. Figure 6 shows the distribution of CCEI scores generated by a random sample of 25,000 hypothetical subjects and the actual distribution. We allow for a narrow confidence interval of one token to account for small mistakes resulting from the slight imprecision of subjects' handling of the mouse (i.e. for any t and $s \neq t$, if $d(\pi^t, \pi^s) \leq 1$ then π^t and π^s are treated as the same allocation). The horizontal axis identifies intervals of CCEI scores and the vertical axis reports the percentage of subjects corresponding to each interval.

[Figure 6 here]

The histograms in Figure 6 show that the distribution of CCEI scores shifts considerably to the right when calculated using our actual data as compared to randomly generated allocations. This makes plain that the significant majority of our subjects came much nearer to consistency with utility maximization than random choosers would have done and that their CCEI scores were only slightly worse than the score of one of the perfect utility maximizers. We therefore conclude 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 distributional preferences or constructing appropriate utility functions. Bronars' test (i.e. the probability that a random subject violates GARP) has also been applied in other experimental papers. The setup used in this study has the highest Bronar power of one (i.e. all random subjects had violations). As a practical note, these results strongly suggest that subjects did not have any difficulties in understanding the procedures or using the computer program.

6 Individual Preferences

6.1 Prototypical Distributional Preferences

The aggregate distributions above tell us little about the particular allocations chosen by individual subjects. In select cases, it is possible readily to identify subjects whose choices correspond to *prototypical* distributional preferences simply from the scatterplots of their choices. Figure 7A depicts the choices of a *selfish* subject (ID 101) $u_S(\pi_S, \pi_O) = \pi_S$, Figure 7B shows the choices of a subject with *utilitarian* preferences (ID 105) $u_S(\pi_S, \pi_O) = \pi_S + \pi_A + \pi_B$, and Figure 7C depicts the choices of a *Rawlsian* subject (ID 124) $u_S(\pi_S, \pi_O) = \min\{\pi_s, \pi_A, \pi_B\}$. For each subject, the choices are depicted as points in a sequence of scatterplots.

[Figure 7 here]

Of our 65 subjects, 24 subjects (36.9 percent) behaved perfectly selfishly. Additionally, three subjects (4.6 percent) displayed utilitarian distributional preferences (allocating all their tokens to person i for whom $p_i < p_j$ for any $j \neq i$), and one subject made nearly equal allocations indicating Rawlsian distributional preferences. By comparison, in FKM, we report that, of the 76 subjects, 20 of them (26.3 percent) behaved perfectly selfishly, two (2.6 percent) fit with utilitarian preferences, and two (2.6 percent) were consistent with Rawlsian preferences. We also find many intermediate cases, but these are difficult to see directly on a scatterplot, because both p and m shift in each new allocation. In order to recover the underlying distributional preferences and to assess any possible relationship between preferences for altruism and social preferences we must impose further structure on the data, which we now proceed to do in our econometric analysis.

6.2 Econometric Specification

Our subjects' CCEI scores are sufficiently near one to justify treating the data as utility-generated, and Afriat's theorem tells us that the underlying utility function $u_S(\pi_S, \pi_O)$ that rationalizes the data can be chosen to be increasing, continuous and concave. Additionally, we assume a separable utility function, which may be expressed in terms of a subutility function $w_S(\pi_O)$ and macro utility function $v_S(\pi_S, w_S)$ with v_S strictly increasing in w_S. Finally we suppose that $w_S(\pi_O)$ and $v_S(\pi_S, w_S)$ are members of the constant elasticity of substitution (CES) family commonly employed in demand analysis.

We therefore write:

$$w_S(\pi_O) = [\alpha' (\pi_A)^{\rho'} + (1-\alpha')(\pi_B)^{\rho'}]^{1/\rho'}$$

and

$$v_S(\pi_S, w_S) = [\alpha (\pi_S)^{\rho} + (1-\alpha) [w_s (\pi_O)]^{\rho}]^{1/\rho}$$

We thus generate a family of CES functions that embed preferences for altruism and social preferences in a particularly convenient manner as

$$U_S = [\alpha(\pi_S)^{\rho} + (1-\alpha)[\alpha'(\pi_A)^{\rho'} + (1-\alpha')(\pi_B)^{\rho'}]^{\rho/\rho'})]^{1/\rho}$$

where $0 \neq \rho, \rho' < 1$.

This CES formulation is very flexible since it "spans" a range of well-behaved utility functions by means of the parameters α, α', ρ and ρ'. Specifically, α represents the relative weight on $self$ versus $others$ and ρ expresses the curvature of the altruistic indifference curves. Analogously, α' represents the relative weight on person A versus person B, and ρ' expresses the curvature of the social indifference curves. Similarly $\sigma = 1/(\rho - 1)$ and $\sigma' = 1/(\rho' - 1)$ are, respectively, the (constant) elasticities of altruistic substitution between $self$ and $others$, and of social substitution between $others$. Clearly, when $\alpha = 1/3$ and $\alpha' = 1/2$, $U_S \to \pi_S + \pi_A + \pi_B$ (the purely utilitarian case) as $\rho, \rho' \to 1$ $(\sigma, \sigma' \to \infty)$, and $U_S \to \min\{\pi_S, \pi_A, \pi_B\}$ (the Rawlsian case) as $\rho, \rho' \to -\infty$ $(\sigma, \sigma' \to 0)$. As $\rho, \rho' \to 0$ $(\sigma, \sigma' \to 1)$, the indifference curves approach those of a Cobb-Douglas function. Further, any

$0 < \rho, \rho' \leq 1$ indicate distributional preference weighted towards increasing total payoffs, whereas any $\rho, \rho' < 0$ indicate distributional preference weighted towards reducing differences in payoffs.

For our purposes, the advantages of the CES formulation are therefore flexibility, tractability and straightforward interpretation. The CES is also the parametric form chosen by FKM and Andreoni and Miller (2002) for recovering preferences for altruism in two-person dictator games. Additionally, the additively separable structure of the CES formulation imposes *two-stage budgeting*: in the first stage $self$ considers how much to keep according to the macro utility maximization, and in the second stage how much to give to each of $others$ according to subutility maximization.

Put precisely, by direct calculation, the solution to the subutility maximization problem is given by

$$\pi_A(p_O, m_O) = \left[\frac{g'}{(p_B/p_A)^{r'} + g'}\right]\frac{m_O}{p_A}$$

where $r' = -\rho'/(1-\rho')$, $g' = [\alpha'/(1-\alpha')]^{1/(1-\rho')}$ and $m_O = p_O\pi_O$ is the total expenditure on tokens given to $others$. The solution to the macro utility maximization problem is then given by

$$\pi_S(p, m) = \left[\frac{g}{q^r + g}\right]\frac{m}{p_S}$$

where $r = -\rho/(1-\rho)$, $g = [\alpha/(1-\alpha)]^{1/(1-\rho)}$ and q is a *weighted relative price of giving* defined by

$$q = \frac{(p_A/p_S) + (p_B/p_S)\left[(\alpha'/(1-\alpha'))(p_B/p_A)\right]^{1/(\rho'-1)}}{\left[\alpha' + (1-\alpha')\left[(\alpha'/(1-\alpha'))(p_B/p_A)\right]^{\rho'/(\rho'-1)}\right]^{1/\rho'}}.$$

This generates the following individual-level two-stage econometric specification for each subject n:

$$\frac{\pi^t_{A,n}}{m^t_{O,n}/p^t_{A,n}} = \frac{g'_n}{\left(p^t_{B,n}/p^t_{A,n}\right)^{r'_n} + g'_n} + \epsilon'^t_n \tag{1}$$

and

$$\frac{\pi^t_{S,n}}{m^t_n/p^t_s} = \frac{g_n}{(q^t_n)^{r_n} + g_n} + \epsilon^t_n \tag{2}$$

where ϵ_n^t and $\epsilon_n'^t$ are assumed to be distributed normally with mean zero and variance σ_n^2 and $\sigma_n'^2$ respectively. Note that the demands (1) and (2) are estimated as budget shares, which are bounded between zero and one, with an *i.i.d.* error term. Using nonlinear tobit maximum likelihood estimation, we first generate estimates of $\hat{g}_n'$ and $\hat{r}_n'$ using (1) and use this to infer the values of the underlying subutility parameters, $\hat{\alpha}_n'$ and $\hat{\rho}_n'$, and the elasticity of social substitution $\hat{\sigma}_n'$. Then, the estimated parameters for the subutility function are employed in estimating the parameters $\hat{g}_n$ and $\hat{r}_n$ using (2), which are then used to infer the values of the parameters of the macro utility function $\hat{\alpha}_n$ and $\hat{\rho}_n$ and the elasticity of altruistic substitution $\hat{\sigma}_n$.

Before proceeding to the estimations, we omit the eight subjects with a CCEI score below 0.80 (ID 103, 106, 107, 109, 110, 119, 204 and 208) as their choices are not sufficiently consistent to be considered utility-generated. We also screen subjects with readily identifiable preferences for whom the CES function is not well defined. These include the 24 subjects with uniformly selfish allocations (those with average $p_s\pi_s/m \geq 0.95$), as well as the two pure utilitarians (ID 105 and 120) and one pure Rawlsian (ID 124). One final subject (ID 326) perfectly implemented utilitarian social preferences and implemented utilitarian preferences for altruism with slight imperfections. Throughout this section, we will also classify this subject as utilitarian. This leaves a set of 29 subjects (44.6 percent) for whom we need to recover the underlying distributional preferences by estimating the CES model.

Table 1 presents the results of the estimations $\hat{a}_n$, $\hat{\rho}_n$, $\hat{\sigma}_n$, $\hat{a}_n'$, $\hat{\rho}_n'$ and $\hat{\sigma}_n'$ sorted according to ascending values of $\hat{\rho}_n$. The additional columns list the CCEI scores. We emphasize again that our estimations will be done for each subject n separately. Throughout this section, whenever we list the number and percentages of subjects with particular properties, we will be considering the 33 subjects with consistent non-selfish preferences. That is, the 29 subjects listed in Table 1 plus the four subjects whose choices correspond precisely to utilitarian or Rawlsian distributional preferences.

[Table 1 here]

6.3 Social Preferences

The estimated parameters for the subutility function $w_S(\pi_O)$, α' and ρ', reflect social preferences (i.e. *other* versus *other*). The coefficient α' expresses the weight that $self$ accords payouts to a particular $other$: $\alpha_n' > 1/2$ ($\alpha_n' < 1/2$) indicates that subject n is biased toward person A (B). Of the 33 subjects with non-selfish consistent preferences, 24 subjects (72.7 percent) have $0.45 \leq \hat{\alpha}' \leq 0.55$, and this increases to a total of 31 subjects (93.9 percent) if we consider $0.4 \leq \hat{\alpha}' \leq 0.6$. We cannot reject the hypothesis that $\hat{\alpha}'_n = 1/2$ for all but four subjects at the 95 percent significance level. This provides a strong support for the inference that subjects do not have any bias towards a particular person, A or B. Thus, we conclude that $others$ are treated symmetrically by $self$, which is a natural result of the anonymity of $others$.

Figure 8 presents the distribution of $\hat{\rho}'_n$, which parameterizes attitudes towards efficiency-equity tradeoff concerning $others$, rounded to a single decimal. Of the 33 subjects with consistent, non-selfish preferences, 14 subjects (42.4 percent) have social preferences that are cleanly classifiable through direct observation of their scatterplots, or through econometric estimation: five subjects (15.2 percent) have perfect substitutes social preferences ($\hat{\rho}' \approx 1$), three subjects (9.1 percent) exhibit Cobb-Douglas social preferences ($\hat{\rho}' \approx 0$), and six subjects (17.2 percent) exhibit extreme aversion to inequality (low $\hat{\rho}'$-values) or Leontief social preferences. Since $others$ are treated symmetrically by $self$, we conclude that both utilitarian and Rawlsian social preferences are well represented among our subjects.

[Figure 8 here]

Moreover, there is considerably heterogeneity in subjects' social preferences among those that cannot be cleanly categorized: 17 subjects (51.5 percent) have $0.1 \leq \hat{\rho}' \leq 0.9$ so that the expenditure on tokens given to person A as a fraction of total expenditure on $others$, $p_A\hat{\pi}_A/m_O$, increases with the relative price p_B/p_A; these subjects thus show a preference for increasing the total payoffs of $others$. On the other hand, only two subjects (6.1 percent) have negative values of $\hat{\rho}$ that are not 'too low' $-0.9 \leq \hat{\rho}' \leq -0.1$ so that $p_A\hat{\pi}_A/m'_O$, decreases with the relative price p_B/p_A; these subjects thus show aversion to inequality between $others$. Overall, we conclude that a significant majority of subjects are concerned with increasing the aggregate payoffs of $others$ rather than reducing differences in payoffs between $others$.

6.4 Preferences for Altruism

The estimates of the two relevant parameters for the macro function $v_S(\pi_O, w_S)$, α and ρ, reflect preferences for altruism (i.e. $self$ versus $others$). The coefficient α represents the relative weight on the payoff for $self$ and ρ parameterizes attitudes towards efficiency-equity tradeoff between $self$ and $others$. As a preview, Figure 9 shows a scatterplot of $\hat{a}_n$ and $\hat{\rho}_n$ (with subjects ID

124 and 306 excluded because they have very low $\hat{\rho}$-values), and compares the estimated parameters with the analogous parameters in the CES model for the two-person study reported in FKM. Note that in both the three- and two-person games there is considerable heterogeneity in both parameters, $\hat{a}_n$ and $\hat{\rho}_n$, and that their values are negatively correlated ($r^2 = -0.43$ and $r^2 = -0.35$ respectively). Perhaps not surprisingly, $\hat{a}_n > 1/2$ for all n in the two-person case, whereas in the three-person case $\hat{a}_n > 1/3$ for all n.

[Figure 9 here]

Of the 33 subjects with consistent, non-selfish preferences, 8 subjects (24.2 percent) have preferences for altruism that are cleanly classifiable through econometric analysis (if not directly from the scatterplots of their decisions): four subjects (12.1 percent) have perfect substitutes preferences for altruism ($\hat{\rho} \approx 1$), three subjects (9.1 percent) exhibit Leontief preferences (low $\hat{\rho}$-values) and one subject exhibited Cobb-Douglas preferences ($\hat{\rho} \approx 0$). There are additionally many subjects with intermediate values of $\hat{\rho}$: 18 subjects (54.5 percent) have $0.1 \leq \hat{\rho} \leq 0.9$ so that the expenditure on tokens kept as a fraction of total expenditure, $p_s \hat{\pi}_s / m$, increases with the weighted relative price of giving q; these subjects thus show a preference for increasing total payoffs of $self$ and $others$. The seven other subjects (21.2 percent) have negative values of $\hat{\rho}$ that are not 'too low' $-0.9 \leq \hat{\rho} \leq -0.1$ so that $p_s \hat{\pi}_s / m$ decreases with the price of giving q; these subjects thus show a preference for reducing differences in payoffs between $self$ and $others$. Figure 10 presents the distribution of $\hat{\rho}_n$ for the 33 subjects with consistent, non-selfish preferences, rounded to a single decimal and and compares it with the analogous distribution in the two-person experiment reported in FKM. The distributions are very similar and skewed to the right so that, as in FKM, our results lean overall toward a social welfare conception of preferences for altruism.

[Figure 10 here]

6.5 Preferences for Altruism versus Social Preferences

While the comparisons we have so far drawn between the three- and two-person experiments are based on different subject pools, the primary innovation of our experimental design is that it allows for a within-subject comparison of preferences for altruism and social preferences. Specifically, we can make within-subject comparisons of the estimated CES parameter of the macro utility function $\hat{\rho}$ (preferences for altruism) and the parameter of the subutility function $\hat{\rho}'$ (social preferences). In other words, each subject's efficiency-equity tradeoff for $self$ versus $other$ embodied in the $\hat{\rho}$ estimator may be compared directly to her efficiency-equity tradeoff between $others$ embodied in the the $\hat{\rho}'$ estimator. Figure 11 shows a scatterplot of $\hat{\rho}_n$ and $\hat{\rho}'_n$. Subjects with very low values for $\hat{\rho}_n$ or $\hat{\rho}'_n$ (ID 114, 124, 201, 304, 306, 318, and 324) are omitted to facilitate presentation of the data.

[Figure 11 here]

The data are concentrated in the upper right quadrant ($0 < \hat{\rho}_n, \hat{\rho}'_n \leq 1$). Of the 33 subjects with consistent, non-selfish preferences, 21 subjects (63.6 percent) have positive values for both $\hat{\rho}_n$ and $\hat{\rho}'_n$, so that for a majority of subjects, both preferences for altruism and social preferences emphasize increasing aggregate payoffs rather than reducing differences in payoffs. Two of the remaining subjects on the graph and six of the seven subjects omitted from the graph because of low $\hat{\rho}_n$ or $\hat{\rho}'_n$ values are located in the lower left quadrant ($\hat{\rho}_n, \hat{\rho}'_n < 0$). Hence, a total of eight subjects (24.2 percent) emphasize reducing difference in payoffs for both altruistic and social preferences.

Interestingly, four subjects exhibit opposite tradeoffs between efficiency and equity in their altruistic and social preferences. Two subjects (ID 123 and 320), who fall in the lower right quadrant ($0 < \hat{\rho}_n \leq 1$ and $\hat{\rho}'_n < 0$), show a preference for increasing total payoffs of $self$ and $others$ while reducing differences in payoffs between $others$. In contrast, two subjects (ID 114 who is omitted from the graph because of a low $\hat{\rho}_n$-value and ID 312) who fall in the top left quadrant ($\hat{\rho}_n < 0$ and $0 < \hat{\rho}'_n \leq 1$) show a preference for reducing differences in payoffs between $self$ and $others$ while increasing total payoffs of $others$. Note, however, that in only two of these four cases both $\hat{\rho}_n$ and $\hat{\rho}'_n$ are significantly different from zero.

Perhaps most interestingly, there is a strong similarity between the efficiency-equity tradeoffs subjects make when allocating between $self$ and $others$ and when allocating across $others$. Specifically, for the subjects shown in Figure 11, the correlation between $\hat{\rho}_n$ and $\hat{\rho}'_n$ is positive ($r^2 = 0.48$). Accordingly, although we find considerable heterogeneity of attitudes towards the efficiency-equity tradeoff *across* subjects, there is a strong association between preferences for altruism and social preferences *within* subjects.

7 Conclusion

A new experimental design - employing graphical representations of three-person dictator games - enables us systematically to distinguish preferences for altruism experimentally from social preferences. Moreover, our experimental method enables us to collect many observations per subject, and we can therefore analyze both types of distributional preferences at the individual level. Most importantly, the broad range of budget sets that our experiment employs provides a serious test of the ability of the theory, and a structural econometric model based on the theory, to interpret the data. In this way, we present the first systematic experimental study of individual social preferences and compare these preferences to individual preferences for altruism.

We conclude by re-emphasizing that the strong correlation between the equality-efficiency tradeoffs subjects make in their altruistic and social preferences is anything but expected. Individuals behave differently when their own payoffs are at stake than when they are not and there is therefore no conceptual reason to expect that preferences concerning the tradeoff between equality and efficiency should be stable over the two scenarios. Indeed it might even seem intuitive to think that individuals who are plainly more inclined to sacrifice efficiency to secure their own payoffs than to secure the payoffs of other individuals will also be more inclined, for example, to sacrifice efficiency to combat inequality that leaves them with less.

The strong correlation between our subjects' altruistic and social preferences concerning efficiency-equity tradeoffs suggests that this intuition is mistaken, or at least captures only a modest effect. Subjects' special concern for themselves seems not to distort impartiality with respect to efficiency-equity tradeoffs nearly as much as it does with respect to the indexical weights that they place on $self$ versus $others$ payoffs. And insofar as this is so, it suggests that at least with respect to preferences concerning efficiency versus equity, subjects actually act on the unified distributive principles that fair-minded people, proceeding in the spirit of Harsanyi and Rawls, would aspire to apply.

8 Experimental Instructions

Introduction This is an experiment in decision-making. Research foundations have provided funds for conducting this research. Your payoffs will depend partly on your decisions and the decisions of the other participants and partly on chance. Please pay careful attention to the instructions as a considerable amount of money is at stake.

The entire experiment should be complete within an hour and a half. At the end of the experiment you will be paid privately. At this time, you will receive $5 as a participation fee (simply for showing up on time). Details of how you will make decisions and receive payments will be provided below.

During the experiment we will speak in terms of experimental tokens instead of dollars. Your payoffs will be calculated in terms of tokens and then translated at the end of the experiment into dollars at the following rate: 4 Tokens = 1 Dollar.

A decision problem In this experiment, you will participate in 50 independent decision problems that share a common form. This section describes in detail the process that will be repeated in all decision problems and the computer program that you will use to make your decisions.

In each decision problem you will be asked to allocate tokens between yourself (Hold) and two other persons, A (Pass A) and B (Pass B) who will be chosen at random from the group of participants in the experiment. The other persons will not be told of your identity. Note that the persons will be chosen at around in each problem. For each allocation, you and the two other persons will each receive tokens.

Each choice will involve choosing a point on a three-dimensional graph representing possible token allocations, Hold / Pass A / Pass B. In each choice, you may choose any combination of Hold / Pass A / Pass B that is on the plane that is shaded in gray. Examples of planes that you might face appear in Attachment 1.

[Attachment 1 here]

Each decision problem will start by having the computer select such a plane randomly from the set of planes that intersect with at least one of the axes (Hold-axis, Pass A-axis or Pass B-axis) at 50 tokens or more but with no intercept exceeding 100 tokens. The planes selected for you in different decision problems are independent of each other and independent of the planes selected for any of the other participants in their decision problems.

For example, as illustrated in Attachment 2, choice 1 represents an allocation in which you hold approximately 20 tokens (Hold), pass 40 tokens to person A (Pass A) and 10 tokens to person B (Pass B). Thus, if you choose this allocation, you will receive 20 tokens, the participant with whom you are matched as person A in that round will receive 40 tokens and the participant with whom you are matched as person B in that round will receive

10 tokens. Another possible allocation is choice 2, in which you receive approximately 30 tokens (Hold), the participant with whom you are matched as person A receives 10 tokens (Pass A) and the participant with whom you are matched as person B receives 20 tokens (Pass B).

[Attachment 2 here]

To choose an allocation, use the mouse to move the pointer on the computer screen to the allocation that you desire. On the right hand side of the program dialog window, you will be informed of the exact allocation that the pointer is located. When you are ready to make your decision, left-click to enter your chosen allocation. After that, confirm your decision by clicking on the Submit button. Note that you can choose only Hold / Pass A / Pass B combinations that are on the gray plane. To move on to the next round, press the OK button. The computer program dialog window is shown in Attachment 3.

[Attachment 3 here]

Next, you will be asked to make an allocation in another independent decision problem. This process will be repeated until all 50 rounds are completed. At the end of the last round, you will be informed the experiment has ended.

Earnings Your payoffs are determined as follows. At the end of the experiment, the computer will randomly select one decision round (that is, 1 out of 50) from each participant to carry out. That participant will then receive the tokens that she allocated to Hold in this round, the participant with whom she was matched as person A will receive the tokens that she allocated to Pass A and the participant with whom she was matched as person B will receive the tokens that she allocated to Pass B. The round selected depends solely upon chance. For each participant, it is equally likely that any round will be chosen.

Each participant will therefore receive three groups of tokens, one based on her own decision to hold tokens, one based on the decision of another random participant to pass tokens to her as person A and one based on the decision of another random participant to pass tokens to her as person B. The computer will ensure that the same two participants are not matched more than once.

The round selected, your choice and your payment will be shown in the large window that appears at the center of the program dialog window. At the end of the experiment, the tokens will be converted into money. Each token will be worth 0.25 Dollars. Your final earnings in the experiment will be your earnings in the round selected plus the $5 show-up fee. You will receive your payment as you leave the experiment.

Rules Your participation in the experiment and any information about your payoffs will be kept strictly confidential. Your payment-receipt and participant form are the only places in which your name and social security number are recorded.

You will never be asked to reveal your identity to anyone during the course of the experiment. Neither the experimenters nor the other participants will be able to link you to any of your decisions. In order to keep your decisions private, please do not reveal your choices to any other participant.

Please do not talk with anyone during the experiment. We ask everyone to remain silent until the end of the last round. If there are no further questions, you are ready to start. An instructor will approach your desk and activate your program.

9 Appendix

To better understand Afriat's (1972) Critical Cost Efficiency Index (CCEI), it is instructive to describe the basic idea underlying the algorithm. Consider a *directed graph* G with a set of *nodes*

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

and set of *edges*

$$E = \cup_{t=1}^{T} \{tj : p^t \pi^t \geq p^t \pi^s\}.$$

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 t to s does not imply the existence of an edge from s to t (in fact, this would imply a GARP violation if one of the inequalities were strict).

For any nodes t and s, an $t - s$ *path* is a finite sequence $t_1, ..., t_K$ such that $t_1 = t$, $t_K = s$ and $p^k \pi^k \geq p^k \pi^{k+1}$ for $k = 1, ..., K-1$ (i.e. a sequence of nodes $t_1, ..., t_K$ linked by E). Note that a path represents a revealed preferred relation in the data (i.e. π^t is revealed preferred to π^s if and only if there exists an $t - s$ path). A cyclic sequence of nodes that creates an $t - t$ path called a *cycle*. The length of a cycle is its number of edges, and