Direct Tests of Models of Social Preferences and Introduction of a New Model

James C. Cox
University of Arizona and Indiana University
jcox@eller.arizona.edu

Vjollca Sadiraj
University of Arizona
vsadiraj@econlab.arizona.edu

February 2005

Direct Tests of Models of Social Preferences and Introduction of a New Model*

Abstract: *Departures from "economic man" behavior in many games in which fairness is a salient characteristic are now well-documented in the experimental literature. These data have inspired development of new models of social preferences incorporating inequality aversion and quasi-maximin preferences. We report experiments that provide direct tests of these social preference models. Data from the experiments motivate a new model of egocentric altruism that we develop and apply to data from several experiments, including data from our direct test experiments and data from experiments with proposer competition and responder competition. We discuss generalizations of the egocentric altruism model that incorporate agents' intentions and thus provide a unified approach to modeling behavior in games both with and without reciprocal motivation.*

Keywords: social preferences, fairness, experiments

JEL Classifications: A12, A13, B49, C70, C91, D63

1. Introduction

Economics has a long history of using models of preferences. The preferences are conventionally represented by utility functions and their indifference maps. Until recently, the preferences most commonly used have been self-regarding (or "economic man") preferences in which an agent cares about his own material payoffs but is indifferent about the material payoffs of others. Recently, in response to a rapidly growing literature reporting data from games in which fairness is a salient characteristic, new models of other-regarding (or "social") preferences have been developed.

Fairness game data come from experiments with different types of games, including: (a) reciprocal-motivation games, such as the ultimatum game, in which beliefs about others' possible future actions and imputations of the intentions behind their past actions can affect agents' behavior; and (b) simple distribution games, such as the dictator game, in which such beliefs and imputations are irrelevant (within the context of the experiment). Our strategy in modeling social preferences is to implement a unified approach that first develops a model that is consistent with behavior in simple distribution games and subsequently generalizes that model to incorporate intentions.

We report direct tests of the central properties of recent prominent contributions to modeling social preferences, including inequality aversion models (Fehr and Schmidt, 1999 and Bolton and Ockenfels, 2000) and a quasi-maximin model (Charness and Rabin, 2003). The distinguishing characteristic of inequality aversion models is that utility is increasing with one's own material payoff but decreasing with the difference between one's own and others' material payoffs. The distinguishing characteristic of the quasi-maximin model is that utility is increasing with an agent's own payoffs, with the lowest of all agents' payoffs, and with the total of all agents' payoffs. These characteristics of the two types of models provide the basis for simple direct tests of the models.

Data from the experiments motivate a new model of egocentric altruism that we develop and apply to data from our four experiments and data from experiments with proposer competition (Roth, et al., 1991) and responder competition (Güth, Marchand, and Rulliere, 1997). We also discuss generalizations and applications of the egocentric altruism model that incorporate agents' intentions and thus provide a unified approach to modeling behavior in games both with and without reciprocal motivation.

2. Experiment 1: A Direct Test of Inequality Aversion

Inequality aversion models are based on the assumption that an agent's utility is increasing with her own material payoff but decreasing with the difference between her own and others' material payoffs. For the special case of two agents and (scalar) money payoffs, the fundamental property of inequality aversion models is that the indifference curves have positive slopes in the part of the money payoff space in which the other's payoff is higher than one's own. This property forms the basis of a direct test of inequality aversion.

2.1 Experimental Design and Procedures

Experiment 1 involves a dictator game with the following characteristics. Subjects are randomly assigned to pairs. Each subject in a pair is given $10. The "non-dictators" have no decision to make. The dictators are told that they may keep their entire $10 endowments or divide them in any way they choose between themselves and the other person in whole dollar amounts. Each dollar that a dictator transfers to the other person is multiplied by three by the experimenters. The experimental protocol uses double-blind payoff procedures in which neither the other subjects nor the experimenters can identify the individual who has chosen any specific action. All of the features of the experiment, including the equal endowments of dictators and non-dictators, are common information given to the subjects. The experiment procedures are described in detail in Appendix 1. The subject instructions are available on an author's homepage.[1]

2.2 Predictions of the Inequality-Aversion Models

Figure 1 shows typical indifference "curves" for the Fehr and Schmidt (1999) model for the dictator's ("my") money payoff m and the other subject's ("your") money payoff y. All parameter values for this model that are consistent with its defining characteristic of inequality (or "inequity") aversion imply that the indifference "curves" have positive slope above the 45-degree line. (See Appendix 2 for derivation of these properties of the indifference "curves.") The budget constraint of a dictator in experiment 1 consists of ordered pairs of integers on the dashed line in Figure 1 extending from the point (10,10) on the 45-degree line to the point (0, 40) on the vertical axis. In this dictator game, since the budget line lies entirely (weakly) above the 45-degree line and has negative slope, the Fehr-Schmidt model makes the same predictions as does the traditional self-regarding preferences model: that the dictator will choose the point (10, 10) or, equivalently, that the dictator will give 0 dollars to the other subject.

Figure 2 shows typical graphs of the level sets or indifference curves of the Bolton and Ockenfels (2000) "motivation function" for the two-agent case with $m + y > 0$. (See Appendix 2 for derivation of these properties of the indifference curves.) This model also predicts that the dictator will give 0 dollars to the other subject for the same reason as does the Fehr-Schmidt model: above the 45-degree line, the indifference curves have positive slope whereas the budget line has negative slope.

2.3. *Subjects' Behavior in Experiment 1*

Data from experiment 1 are reported in Figure 3 with the light-colored bars. In this experiment 19 of 30 or 63% of the dictators gave positive amounts to the other person and, hence, exhibited behavior that is inconsistent with inequality aversion. The light-colored bars in Figure 4 show the distribution of relative payoffs to the subject pairs in experiment 1. The 63% of dictators who sent positive amounts of money to the other subjects were not averse to a degree of inequality that gave the other paired subjects significantly higher payoffs. Furthermore, the behavior of the 37% of subjects who did *not* give any money to the paired subject can be explained by self-regarding (or "economic man") preferences. Therefore, inequality aversion is not needed to explain the behavior of any subject in this experiment.

Other papers that report experimental tests of inequality aversion models (Charness and Rabin, 2003; Engelmann and Strobel, forthcoming) also find that the models' predictions are inconsistent with the behavior of a large proportion of the subjects. Therefore there is a moderately large body of data that is inconsistent with inequality aversion, which suggests the need for a different type of model. After reporting their tests of inequality aversion, Charness and Rabin introduce the quasi-maximin model and apply it to data from several experiments. We next report direct tests of that model.

3. Direct Tests of Quasi-Maximin Preferences

3.1 The Quasi-Maximin Model

Let x_i denote the own money payoff of an agent *i* and x_j, $j \in \{1,2,\ldots,n\} \setminus \{i\}$ as the money payoff to some other person j. Charness and Rabin's (2003) "reciprocity-free" model is based on the assumption that the utility function of agent *i* is increasing with the amount of her own money payoff (x_i), the minimum of all agents' payoffs ($\min_{j \in \{1,\ldots,n\}} \{x_j\}$), and the total of all agents' payoffs $\sum_{j=1}^{n} x_j$. The quasi-maximin utility function is assumed to be:

$$u^i(x) = (1-\gamma)x_i + \gamma\left[\delta \min_{j \in \{1,\ldots,n\}} \{x_j\} + (1-\delta)\sum_{j=1}^{n} x_j\right] \qquad (1)$$

where $\gamma \in [0,1]$ and $\delta \in (0,1)$. γ measures the relative importance of own money payoff compared to the two other arguments of the utility function. δ measures the relative importance of these other two arguments, the minimum payoff and total payoff (or "efficiency").

3.2 Is the Quasi-Maximin Model Consistent with Behavior in Experiment 1?

Indifference curves for the two-agent (*n* = 2) special case of utility function (1) are piece-wise linear, with negative slope below the 45-degree line that is flatter than the negative slope above the 45-degree line except in the limiting case where γ = 0 (the case of self-regarding preferences). In experiment 1, a dictator can choose a whole dollar amount (weakly) between 0 and 10 to send to the paired subject. If a dictator behaves as if he wants to maximize the two-agent version of utility function (1), then he will choose to give: (a) all $10 to the other subject if $\delta < 1 - 1/(3\gamma)$; or (b) nothing to the other subject if $\delta > 1 - 1/(3\gamma)$. In the event that $\delta = 1 - 1/(3\gamma)$, the model has no testable implication with these data: all feasible choices are consistent with the model. 49% of the dictators sent either \$0 or \$10 to the paired subject in experiment 1. But the behavior of the 36% of dictators who sent \$0 can be explained by self-regarding preferences; hence the quasi-maximin model adds explanatory power only for the 13% of dictators who sent all \$10 to the other subject. In addition, the quasi-maximin model has no predictive value for explaining the behavior of 51% of the subjects who sent positive amounts of money less than \$10.

The inconsistency of experiment 1 data with the quasi-maximin model is *not* a fundamental inconsistency, in contrast to the fundamental inconsistency between that data and models of inequality aversion. A nonlinear generalization of the quasi-maximin model could be consistent with experiment 1 data.

3.3 Experiments 2 and 3: Direct Tests of Quasi-Maximin Preferences

The three arguments of utility function (1) suggest the design of two experiments that provide direct tests for fundamental inconsistency between subjects' behavior and quasi-maximin preferences. In experiment 2, we offer subjects choices between alternatives in a dictator game in which the dictator's own payoff and the minimum payoff are constant but the sum of all payoffs changes. The three rows in Table 1 show the choices open to a subject in experiment 2. Because the dictator's payoff is the same in all rows and the lowest payoff is the same in all rows, the quasi-maximin model predicts that an agent will choose row 3, which has the highest total payoff to all agents (except in the limiting case in which $\gamma = 0$, where the model makes no prediction because this is the special case of self-regarding preferences). This row 3 prediction is *independent* of the specific piece-wise linear form of the Charness-Rabin (2003) utility function. Thus the experiment provides a direct test of the assumed preference for efficiency *per se*.

Whereas experiment 2 tests for a preference for efficiency, experiment 3 tests for the other defining property of the quasi-maximin model, the preference for increasing the payoff to the lowest paid agent (the maximin property). Thus, in experiment 3, we offer subjects choices in a dictator game in which the dictator's own payoff and the total payoff are constant but the minimum payoff changes. The three rows in Table 2 show the choices open to a subject in this dictator experiment. Since the dictator's payoff is the same in all rows and the total payoff to all agents is the same in all rows, the quasi-maximin model predicts that an agent will choose row 3, which has the maximin payoff (except in the limiting case in which $\gamma = 0$, where the model makes no prediction because this is the special case of self-regarding preferences). This row 3 prediction is *independent* of the specific piece-wise linear form of the Charness-Rabin (2003) utility function. Thus the experiment provides a direct test of the assumed maximin property *per se*.

3.3 Procedures in Experiments 2 and 3

Experiments 2 and 3 have the following characteristics. Subjects are randomly assigned to groups of four that consist of a dictator and three "non-dictators." The dictators are told that they must choose one row from Table 1 in experiment 2 or one row from Table 2 in experiment 3. Different subjects participated in experiments 2 and 3. The experimental protocol uses double-blind payoff procedures in which neither the other subjects nor the experimenters can identify the individual who has chosen any specific action. All of the features of the experiment are common information given to the subjects. The experiment procedures are described in detail in Appendix 1. The subject instructions are available on an author's homepage, as explained in footnote 1.

3.4 Behavior in Experiments 2 and 3

Subjects' behavior in experiments 2 and 3 is reported in Figure 5. We observe that 5 of 33 (or 15%) of the subjects chose row 3 in experiment 2, which is the unique prediction of the quasi-maximin model. Thus, the behavior of 85% of the subjects in experiment 2 is inconsistent with quasi-maximin preferences. In experiment 3, 2 of 32 (or 6%) of the subjects chose row 3, the unique prediction of the quasi-maximin model. Hence the behavior of 94% of the subjects in experiment 3 is inconsistent with the quasi-maximin model.

4. The Egocentric Altruism Model

The very high rates of inconsistency between subjects' behavior and the testable implications of the inequality-aversion and quasi-maximin models suggest the need for a model with somewhat different properties. Indeed, Charness and Rabin (2003, fn. 45) anticipated the need for a model of preferences over "…the full distribution of payoffs..." As it turns out, the behavior observed in experiments 1-3 and in similar experiments can be rationalized by a utility function that is quasi-concave (i.e., has indifference curves that are convex to the origin) and monotonically increasing in all arguments, the dictator's income x_i and others' incomes $x_j, j \in \{1,\ldots,n\} \setminus \{i\}$. We present such a model in sections 4.1 and 4.2, beginning with the two-agent special case.

Before introducing the model, we first discuss behavior in some dictator experiments with varying parameterizations. Comparison of data from experiment 1 with data from other dictator experiments provides additional insight into the properties of other-regarding preferences. In the (DB1 and DB2) double-blind dictator experiments reported by Hoffman, McCabe, Shachat, and Smith (1994), the average amount sent to the paired subjects by the dictators was \$1. In our experiment 1 dictator game, the average amount sent by the dictators was \$3.60. The price to the dictator of buying an additional \$1 of income for the paired subject was \$1 in the Hoffman, *et al.* experiment and it is \$0.33 in our experiment 1. A model of altruistic preferences in which utility is globally increasing in both the dictator's own money payoff and the other person's money payoff can explain data from our experiment 1 and the DB1 and DB2 experiments. The implied (arc) price elasticity of demand for increasing the other subject's income is -1.13, a quite reasonable figure.

Andreoni and Miller (2002) test data from many dictator games, with varying budgets and own-payoff prices for altruistic actions, for consistency with utility-maximizing behavior by testing the data for consistency with the generalized axiom of revealed preference (GARP). They report that 98 percent of their subjects make decisions that are consistent with GARP and therefore are, in that specific sense, rational altruists. Furthermore, Andreoni and Miller report that constant elasticity of substitution (CES) parametric utility functions provide a good fit to data for their subjects. We develop a model based on CES utility functions that are modified to capture salient characteristics of subjects' altruistic preferences that are revealed by our experiments.

4.1 The Two-Agent Egocentric Altruism Model

The two-agent special case of the egocentric altruism model represents other-regarding preferences with a modified constant elasticity of substitution (CES) utility function in which the weights can differ depending on whether the agent's own payoff m is higher or lower than the other's payoffs y :

(2) $$u(m, y) = [(1-\theta)m^{\alpha} + \theta y^{\alpha}]^{1/\alpha},$$

where[2]

(3) $$\theta = \theta_{-}, \text{ if } m < y$$

$$= \theta_{+}, \text{ if } m \geq y.$$

This utility function is assumed to be monotonically increasing in m and y and to have indifference curves that are negatively-sloped and convex to the origin except for the boundary values of $\theta_{-} = \theta_{+} = 0$, in which case the model is equivalent to the model of self-regarding preferences. The agent's altruistic preferences are assumed to be "egocentric," by which we mean that between two money allocations (a,b) and (b,a) the agent prefers the one that allocates the larger payoff to himself:

(4) $u(b,a) > u(a,b)$, for all a and b such that $b > a \geq 0$.

Together, these assumptions imply the following restrictions on the parameters of the utility function:

(5) $0 < \alpha < 1$; $0 \leq \theta_+ < 1$; $0 \leq \theta_- \leq \theta_+$; and $\theta_- < 1 - \theta_+$.

4.2 The Many-Agent Egocentric-Altruism Model

The generalization of the egocentric altruism utility function to $n \geq 2$ agents is:

(6) $$u^i(x) = \left[(1 - \sum_{j \neq i} \theta_j^i) x_i^\alpha + \sum_{j \neq i} \theta_j^i x_j^\alpha \right]^{1/\alpha}$$

where

(7) $\theta_j^i = \theta_-^i / (n-1)$, if $x_i < x_j$,

$= \theta_+^i / (n-1)$, if $x_i \geq x_j$.

The egocentric property generalizes to the *n*-agent case as follows. First define:

(8) $x^{ab}, x^{ba} \in \Re_+^n$ such that $x_i^{ab} = x_k^{ba} = a$, $x_k^{ab} = x_i^{ba} = b$, for some *i* and *k,* and

$x_j^{ab} = x_j^{ba}$, for all $j \neq i, k$.

Then the egocentric property is

(9) $u^i(x^{ba}) > u^i(x^{ab})$, for all x^{ab}, x^{ba} and *a*, *b* such that $b > a \geq 0$.

The parameter restrictions implied by monotonicity, egocentricity, and convexity are:

(10) $0 < \alpha < 1$; $0 \leq \theta_+^i < 1$; $0 \leq \theta_-^i \leq \theta_+^i$; and $\theta_-^i < (n-1)(1 - \theta_+^i)$.

5. Explanatory Power of the Egocentric Altruism Model

We first show that the egocentric altruism model is consistent with behavior in several dictator game experiments. Subsequently, we show that the model can rationalize data from experiments with proposer competition and responder competition. This procedure provides a strong check on how robust the model is to explaining behavior in distribution games.

5.1 Is the Egocentric Altruism Model Consistent with Behavior in Experiments 1-3?

We first ask whether the behavior of subjects in experiment 1 is consistent with the egocentric altruism model. In other words, are the data from experiment 1 consistent with the two-agent utility function given by equations (2) and (3) and the parameter restrictions in statement (5) that are implied by monotonicity, egocentricity, and convexity?

In experiment 1, a dictator can choose an amount *s* weakly between 0 and 10 to send to the paired subject. Each $1 sent decreases the dictator's money payoff by $1 and increases the paired subject's money payoff by $3. Hence, the slope of the dictator's budget line is -3 (weakly) above the 45-degree line. Utility function (2) implies that the slope of a dictator's indifference curve at some (m,y) above the 45-degree line is equal to $-\frac{1-\theta_-}{\theta_-}\left(\frac{y}{m}\right)^{1-\alpha}$. The limiting value of the slope of an indifference curve as it approaches the 45-degree line from above is $-\frac{1-\theta_-}{\theta_-}$.

Thus, if $-\frac{1-\theta_-}{\theta_-} > -3$ then the dictator's optimal amount to send is positive; otherwise, it is 0. Therefore, the egocentric altruism model predicts that a dictator will give some of her endowment to the paired subject only if $\theta_- > 0.25$ and that a dictator with $\theta_- \leq 0.25$ will not give away any money. The data reported in Figure 3 reveals that 63% of the subjects make a choice consistent with $\theta_- > 0.25$ and 37% make the choice consistent with $\theta_- \leq 0.25$.

We now ask whether the subjects' behavior in experiments 2 and 3 is consistent with the egocentric altruism model. In other words, are the data from experiments 2 and 3 consistent with the many-agent utility function given by equations (6) and (7) and the parameter restrictions in statement (10) that are implied by monotonicity, egocentricity, and convexity?

The egocentric altruism model predicts that the dictator will choose either row 2 or row 3 in experiment 2 (see Table 1), depending on the relative values of the parameters of the utility function. For experiment 2, it is clear that the egocentric altruism model ranks row 2 higher than row 1 because the utility function is positively monotonic in all payoffs and some row 2 payoffs are larger than the corresponding row 1 payoffs and no row 2 payoffs are lower than corresponding row 1 payoffs. But the utilities of rows 2 and 3 depend on the relative magnitudes of the α , θ_+ , and θ_- parameters of the utility function. There are four agents, hence $n-1$ = 3 in statement (7). Two of the other agents have higher payoffs than the dictator in row 2, which means the weight on each of their payoffs is θ_- /3, and one other agent has a lower payoff than the dictator, which means the weight on her payoff is θ_+ /3. In contrast, there are two lower payoffs in row 3, with weight θ_+ /3 on each one, and one higher payoff with weight θ_- /3. Therefore, the dictator prefers row 2 to row 3 if $\theta_- < f(\alpha)\ \theta_+$ and she prefers row 3 to row 2 if the strict inequality holds in the opposite direction, where $f(\alpha) = (10^\alpha - 2^\alpha)/(33^\alpha + 10^\alpha - 2\times 15^\alpha)$. The egocentric altruism model predicts choice of row 2, even for the linear special case with $\alpha = 1$, as follows. For the linear case, row 2 is preferred to row 3 if $\theta_- < (8/13)\,\theta_+$ and row 3 is preferred to row 2 if the strict inequality holds in the opposite direction. Furthermore, if $\theta_- > f(\alpha)\ \theta_+$ then row 3 is the most preferred row for the dictator because, from transitivity, if row 2 is preferred to row 1 and row 3 is preferred to row 2 then row 3 is preferred to row 1.

We observe from Figure 5 that the egocentric altruism model is consistent with the behavior of 28 of 33 (or 85%) of the subjects in experiment 2 who chose either row 2 or row 3. Data from this experiment reveal that the 70% of dictators who chose row 2 are consistent with $\theta_{-} < f(\alpha)\ \theta_{+}$ and the 15% of the subjects who chose row 3 are consistent with the opposite strict inequality. Furthermore, these data reveal that $\alpha \geq 0.72$ because if it were true that $\alpha < 0.72$ then it would also be true that $f(\alpha) > 1$, which in turn implies that no dictator would choose row 3. More importantly, the parameter inequalities that make the egocentric altruism model consistent with 100% of the data from experiment 1 are consistent with the parameter inequalities that make the model consistent with 85% of the data from experiment 2.

The egocentric altruism model predicts that the dictator will choose row 2 in experiment 3 (see Table 2). It is clear that the egocentric altruism model ranks row 2 (at least weakly) higher than row 1 because they differ only by a transfer of $3 from an agent with higher payoff than the dictator to an agent with lower payoff and, according to statement (10), $\theta_{-}^{i} \leq \theta_{+}^{i}$. It is straightforward to show that the model ranks row 2 higher than row 3 if $\theta_{-} < g(\alpha)\ \theta_{+}$, where $g(\alpha) = (10^{\alpha} - 5^{\alpha} - 4^{\alpha} + 3^{\alpha})/(21^{\alpha} - 12^{\alpha} - 15^{\alpha} + 10^{\alpha})$. Since $g(\alpha)$ is strictly decreasing for $\alpha \in (0,1]$, and $g(1) = 1$, it is clear that the egocentric altruism model ranks row 2 over row 3 in experiment 3. We observe from Figure 5 that 28 of 32 subjects chose row 2 in experiment 3; hence the egocentric altruism model is consistent with the behavior of 88% of the subjects in that experiment.

5.2 *Summary Comparison to Inequality Aversion and Quasi-Maximin Models*

Comparisons across the three models of their rates of consistency and inconsistency with data from experiments 1-3 reveal some of the properties that a model must have in order to maintain consistency with data from simple distribution games. We first consider the two-agent experiment 1 and, subsequently, turn our attention to the four-agent experiments 2 and 3.

Experiment 1 provides an opportunity for a dictator to pay a low cost of 33 cents for each dollar increase in the payoff to a paired subject whom she knows will have a payoff that is higher than her own payoff as a result of such a gift. A large majority of 63% of the dictators did send money to their paired subjects and thereby revealed that they did not have inequality-averse preferences. The budget constraint for dictators in experiment 1 is negatively-sloped and linear between the 45-degree line and the vertical axis, as shown in Figures 1 and 2. Therefore the negatively-sloped, piece-wise linear indifference curves for the quasi-maximin model predict a corner solution, either on the 45-degree line or on the vertical axis, except in the knife-edge case where this model is consistent with all possible choices in this experiment (and hence has no testable implication). The quasi-maximin model has power in explaining only the 13% of the choices of subjects who sent all of their $10 endowments. But this low explanatory power is not a fundamental problem for quasi-maximin preferences; a non-linear generalization of the two-agent quasi-maximin utility function could increase the explanatory power of the model for behavior in experiment 1. Indeed, the two-agent egocentric altruism model is consistent with all of the data from experiment 1, and this model looks like a non-linear version of the two-agent quasi-maximin model. The fundamental differences between the egocentric altruism and quasi-maximin models only become apparent when one considers more than two agents, as is required in order to apply the models to data from experiments 2 and 3. Experiments 2 and 3 also bring to light some further similarities and differences between the egocentric altruism model and the inequality aversion models.

In experiment 2, the many-agent quasi-maximin model predicts that a subject will choose row 3 in Table 1. In contrast, the many-agent inequality aversion models predict that a subject will choose row 1. The many-agent egocentric altruism model predicts that a subject will choose either row 2 or row 3. As reported in Figure 5, the behavior of 85% of subjects (who chose row 1 or 2) was inconsistent with the quasi-maximin model and the behavior of another 85% of subjects (who chose row 2 or 3) was inconsistent with the inequality aversion models. In contrast, the behavior of 85% of subjects (who chose row 2 or 3) was consistent with the egocentric altruism model.

In experiment 3, the many-agent quasi-maximin model predicts that a subject will choose row 3 in Table 2. The Fehr-Schmidt many-agent inequality aversion model predicts that a subject will choose row 2 and the Bolton-Ockenfels model is consistent with all possible choices and hence not testable with data from this experiment. The many-agent egocentric altruism model predicts that a subject will choose row 2. As reported in Figure 5, the behavior of 94% of subjects (who chose row 1 or 2) was inconsistent with the quasi-maximin model. In contrast, the behavior of 88% of subjects (who chose row 2) was consistent with the egocentric altruism model. In this experiment, for which it has the same prediction as the egocentric altruism model, the Fehr-Schmidt inequality aversion model has the same high rate of 88% consistency with the data.

5.3 Experiment 4: How Do Dictators Respond to the Opportunity to Take Money?

A common feature of experiments 1–3 is that the dictator cannot take money from another subject and appropriate it himself. This means that the designs of those experiments do not cause the subjects to reveal information about some characteristics of their preferences. In order to construct an experiment that will reveal more about the subjects' preferences, we expand the feasible set of experiment 1 to include opportunities to take money as well as give it away.

5.3.1 Experimental Design and Procedures

In experiment 4, subjects are randomly assigned to pairs. Each subject in a pair is given $10. The "non-dictators" have no decision to make. The dictator is asked to decide whether he wants to give part or all of his $10 endowment to the other subject or take up to $5 of the other subject's endowment or neither give nor take anything. Any amount given to the other subject is multiplied by three by the experimenter. Each dollar taken from the other subject increases the dictator's payoff by one dollar; that is, there is no multiplication by three by the experimenter. The experimental protocol uses double-blind payoff procedures. All of the features of the experiment, including the equal endowments, are common information given to the subjects. In summary, experiment 4 differs from experiment 1 *only* by introduction of the opportunity to take money from the paired subject. The experiment procedures are described in Appendix 1. The subject instructions are available on an author's homepage, as explained in footnote 1.

5.3.2 Subjects' Behavior in Experiment 4

Data from experiment 4 are reported in Figure 3 with the dark-colored bars. Note that 22 of 32 or 69% of the dictators took money from the other person and 18 of 32 or 56% took the maximum possible amount, $5. Also, 3 of 32 or 9% of the dictators gave money to the other person and 7 of 32 or 22% neither gave nor took any money (they chose $0 as the amount to send or take). The dark-colored bars in Figure 4 show the distribution of payoffs in experiment 2. Note that 22% of the dictators chose equal payoffs and 78% chose unequal payoffs.

Figures 3 and 4 show very different outcomes in experiments 1 and 4. Introduction of the opportunity to take money changes the distribution of behavior from appearing to be predominantly altruistic to appearing to be predominantly selfish. Just such behavior is consistent with a model with down-sloping indifference curves that are convex to the origin. For example, Figure 6 shows indifference curves that would cause a single subject to choose the modal observations in experiments 1 and 4.

In experiment 4, a dictator can choose an amount s (weakly) between -5 and $+10$ to "send" to the paired subject. If $s > 0$ then each $1 sent decreases the dictator's money payoff by $1 and increases the paired subject's money payoff by $3. Hence, the slope of the dictator's budget line is -3 above the 45-degree line. If $s < 0$ then each $1 taken increases the dictator's money payoff by $1 and decreases the paired subject's payoff by $1. Hence the slope of the dictator's budget line is -1 below the 45-degree line. The limiting value of the slope of a dictator's indifference curve is $-(1-\theta_-)/\theta_-$ as it approaches the 45-degree line from above and the slope is $-(1-\theta_+)/\theta_+$ as it approaches the 45-degree line from below. In the case where $-\frac{1-\theta_-}{\theta_-} > -3$, define s_-^* as the value of s at which a dictator's indifference curve is tangent to the line, $y = 40 - 3m$ above the 45-degree line. In the case where $-\frac{1-\theta_+}{\theta_+} < -1$, define s_+^* as the value of s at which a dictator's indifference curve is tangent to the line $y = 20 - m$ below the 45-degree line. Appendix 3 contains a straightforward proof of the following.

Proposition 1. Let s° be the amount sent by a dictator in experiment 4. One has:

(i) *if* $-\frac{1-\theta_-}{\theta_-} \le -3$ *and* $-\frac{1-\theta_+}{\theta_+} \ge -1$ *then* $s^\circ = 0$*;*

(ii) *if* $-\frac{1-\theta_-}{\theta_-} > -3$ *and* $-\frac{1-\theta_+}{\theta_+} \ge -1$ *then* $s^\circ = s_-^* > 0$*;*

(iii) *if* $-\frac{1-\theta_-}{\theta_-} \le -3$ *and* $-\frac{1-\theta_+}{\theta_+} < -1$ *then* $s^\circ = \max(s_+^*, -5) < 0$*; and*

(iv) *if* $-\frac{1-\theta_-}{\theta_-} > -3$ *and* $-\frac{1-\theta_+}{\theta_+} < -1$ *then* $s^\circ = s_-^* > 0$ *or*

$s^\circ = \max(s_+^*, -5) < 0$.

A distribution of preference parameters in the subject pool, all of which satisfy the inequalities in statement (5), can make the egocentric altruism model consistent with the behavior of 100% of the subjects in experiment 4 in which 22% of dictators chose according to part (*i*), 9% chose in ways that are consistent with both parts (*ii*) and (*iv*), and 69% chose in ways that are consistent with both parts (*iii*) and (*iv*) of Proposition 1. The distribution of preference parameters that makes it possible for the egocentric altruism model to rationalize all of the data from experiment 4 is shown in Figure 7. This distribution of preference parameters will be used in applications of the model to data from other types of experiments. In Figure 7: for at least 69% of the subjects, $\theta_+ \in [0, 0.5)$; for at most 9% of the subjects, $\theta_+ \in [0.5, 1)$ and $\theta_- \in (0.25, 1-\theta_+]$; and for 22% of the subjects, $\theta_+ \in [0.5, 1)$ and $\theta_- \in [0, \min(0.25, 1-\theta_+))$.

The subjects in experiments 1 and 4 were drawn from the same subject pool, hence they should have about the same distribution of preference parameters. Note that 63% of the subjects in experiment 1 revealed preference parameters consistent with both parts (*i*) and (*iii*) of Proposition 1 and 37% of those subjects revealed preference parameters consistent with both parts (*ii*) and (*iv*).

5.4 Explaining Competition

Fehr and Schmidt (1999), Bolton and Ockenfels (2000), and Charness and Rabin (2003) demonstrate that their inequality aversion and quasi-maximin models can rationalize data from several kinds of experiments. However, those models are called into question by simple direct tests with specially-designed dictator games. Experiments with dictator games designed to directly test the central defining characteristics of those models show that a large majority of subjects make choices that are inconsistent with the models. In contrast, there is a distribution of individual utility function parameter values for the egocentric altruism model that makes that model consistent with the behavior of an overwhelming majority of subjects in the four dictator game experiments. This brings up the important question of how robust is the empirical consistency of the egocentric altruism model. Does the distribution of individual utility function parameter values that rationalizes the dictator game data also make the egocentric altruism model consistent with the behavior of subjects in other experiments reported in the literature?

5.4.1 Game with Proposer Competition

Roth, Prasnikar, Okuno-Fujiwara, and Zamir (1991) report results from experiments in four countries with the following game of proposer competition. $n-1$ proposers can simultaneously propose shares $s_j \in [0,1]$, $j = 1,2,,\cdots,n-1$. The responder can accept or reject the highest share offered, $\bar{s} = \max\{s_1, s_2,\cdots s_{n-1}\}$. If the responder accepts the highest offer then the proposer that made the offer gets $1-\bar{s}$, the other proposers get 0, and the responder gets $\bar{s}$. If more then one proposer made the accepted offer then one of the proposers is randomly selected, with equal probability for all tied proposers, to get $1-\bar{s}$. If the responder rejects the highest offer then all players get 0. Results from experiments in all four countries were that the accepted (highest) proposal converged to 1 in five to six periods. The egocentric altruism model predicts this outcome, as can be seen from the following.

A formal proof is provided in Appendix 3. Here we offer an informal demonstration that conveys intuition about the implications of the egocentric altruism model for subgame perfect equilibrium in the game with proposer competition. The responder prefers the money payoffs implied by any $\bar{s} > 0$ to the outcome in which everyone gets 0 because utility function (6) is monotonically increasing in all payoffs. Thus the responder will accept all offers. Because of egocentricity, a proposer prefers the payoffs from the set of offers in which he is one of k

proposers, $1 < k \leq n-1$, who submits $\bar{s} > 0$ and has $1/k$ probability of receiving $1-\bar{s}$ to the payoffs from $s^P < \bar{s}$ where he gets 0 for sure. For a sufficiently small value of $\xi > 0$, a proposer prefers the payoffs from the set of offers in which he submits $s^P = \tilde{s} + \xi$ to the payoffs from the set of offers in which k proposers, $1 < k \leq n-1$, submit offers of $\tilde{s} = \bar{s} < 1$, because of convexity and egocentricity. Thus one has the following.

Proposition 2. Let $\bar{s}$ *be the highest offer submitted by proposers in the game of proposer competition. One has:*

(i) *The responder will accept any offer* $\bar{s} \geq 0$;

(ii) *There can be no subgame perfect equilibrium in which proposers offer* $\bar{s} < 1$; *and*

(iii) *There is a unique subgame perfect equilibrium in which at least two proposers offer* $\bar{s} = 1$ *and the responder accepts it.*

5.4.2 Game with Responder Competition

Güth, Marchand, and Rulliere (1997) report an experiment with a game in which a proposer proposes a share $s \in [0,1]$ to $n-1$ responders. A responder can accept or reject the proposal. If only one of the responders accepts the offer then she gets s, the other responders get 0, and the proposer gets $1-s$. If more then one responder accepts the proposal then one of the responders is randomly selected to get s. If all of the responders reject the proposal then all players get 0. The experiments were run with a design in which responders were asked to pre-commit to acceptance thresholds for a period before observing the proposal for that period. Results from a limited number of experiments were that the average responder threshold had declined to less than 0.05 by the fifth period, 71% of responders chose 0 thresholds, and 9% of the responders chose a threshold of 0.02. Also by period 5, on average the proposals had decreased to 0.15. The predictions of the egocentric altruism model are as follows.

Appendix 3 presents a formal proof of the egocentric altruism model's subgame perfect equilibrium for the game with responder competition. Here we present an informal demonstration that conveys intuition about the properties of the equilibrium. Consider any proposal, $s > 0$. If all responders reject the proposal, everyone gets 0. But a responder prefers to accept any $s > 0$, rather than accept the outcome where everyone gets 0, because of the positive monotonicity of utility function (6). Similarly, if k other responders, $1 \leq k \leq n-2$, accept s > 0, the utility function implies that a responder prefers also to accept s because the payoffs determined by a $1/(k+1)$ probability of receiving s are preferable to those where the responder receives 0 for sure. Since responders will accept all proposals, the proposer will propose that offer which maximizes his utility function. One responder will be randomly selected to receive s^P and the proposer will receive $1-s^P$, hence the proposer's utility implied by (6) is

(12) $$u^P = [(1-\theta_+^P)(1-s)^\alpha + \theta_+^P s^\alpha /(n-1)]^{1/\alpha}, \quad \text{if } 0 \leq s \leq 0.5$$

$$= [(1-\left((n-2)\theta_+^P - \theta_-^P\right)/(n-1))(1-s)^\alpha + \theta_-^P s^\alpha /(n-1)]^{1/\alpha}, \text{ if } 0.5 < s \leq 1.$$

Differentiation of (12) reveals that the proposer's optimal proposal s^P will be positive or 0, depending on the relative values of θ_+^P and n. Appendix 3 contains a proof of the following.

Proposition 3. Let s^P be the proposer's subgame perfect equilibrium proposal. One has:

(i) *All responders accept all proposals;*

(ii) *The proposer's offer is*

$s^P = 0$, *if* $\theta_+^P = 0$, *the limiting case of self-regarding preferences;*

$s^P = 1/(1+((n-1)(1-\theta^P{}_+)/\theta^P{}_+)^{1/(1-\alpha)})$, *if* $\theta_+^P \in (0,(n-1)/n)$*;*

$s^P = 0.5$, *otherwise.*

An immediate extension of Proposition 3 comes from noting that $s^P = 0$ for the discrete-variable case of the experiment, given that

(13) $\theta_+^P < 1/(1 + s_*^\alpha /[(1-(1-s_*)^\alpha)(n-1)])$,

where $s_* = \min S/\{0\}$. If $s_* \geq 0.01$, $\alpha \geq 0.76$ then one can calculate the entries in Table 3. Referring to the distribution of preference parameters shown in Figure 7, one has: 69% of subjects have parameters $\theta_+^P \leq 0.5$, and hence the proposal average is predicted to be not bigger than 0.155 (= 0×0.69 + 0.5×0.31). This is nearly the same value as the empirical proposal average, which was 0.15.

5.5 Robust Explanatory Power of the Egocentric Altruism Model

The preceding parts of section 5 explain that the egocentric altruism model can explain behavior in experiments 1– 4 reported in this paper and in experiments with proposer and responder competition reported by other researchers. As we have explained, the distribution of preference parameters in Figure 7 is consistent with data for: (a) 100% of subjects in experiments 1 and 4; (b) 85% of subjects in experiment 2; and (c) 88% of subjects in experiment 3. Furthermore, the egocentric altruism model is consistent with outcomes that subjects' converge to in experiments with proposer and responder competition.

6. Concluding Remarks

As shown in Charness and Rabin (2003), Engelman and Strobel (forthcoming), and section 2 above, a large majority of subjects make choices that are inconsistent with inequality aversion models in experiments designed to provide direct tests for inequality aversion. As shown in section 3, most subjects make choices that are inconsistent with the quasi-maximin model in dictator games that are designed to directly test the central defining characteristics of that model. In contrast, the egocentric altruism model is consistent with the behavior of a large majority of subjects in all four different types of dictator games. Furthermore, the egocentric altruism model is consistent with the behavior of subjects in experiments with proposer competition and responder competition. The common feature of these experiments is that they involve games that do not elicit reciprocal motives. This suggests the important question of whether the egocentric altruism model can be generalized to incorporate reciprocal motives, thus holding out the promise of an empirically-motivated and unified approach to modeling social preferences in environments with and without reciprocal motives.

Two generalizations of the two-agent egocentric altruism model are being developed. A parametric generalization of the two-agent model is reported in Cox, Friedman, and Gjerstad (2005) and a nonparametric generalization is reported in Cox, Friedman, and Sadiraj (2005). In the parametric model, the weight on the other person's payoff in the agent's CES utility function depends on the kindness or unkindness of others' choices (their revealed intentions) and on their status relative to the agent. The parametric model incorporating intentions and status is applied to data from two types of games with reciprocal motivations, Stackelberg duopoly games and mini-ultimatum games. In the nonparametric model, the indifference curves of the utility function are always convex to the origin, but the marginal rate of substitution between ones' own and another's payoff depends on the other's previous actions. Two partial orderings are introduced, an ordering of preferences by a formal representation of "more altruistic than" and an ordering of opportunity sets by a formal representation of "more generous than." These partial orderings are linked by the "reciprocity axiom" which specifies that more generous choices by a first mover induce more altruistic preferences in a second mover. This model is applied to data from Stackelberg duopoly games and to data from the Stackelberg mini-games that are introduced in the paper.

Endnotes

∗ We are grateful for research support from the National Science Foundation (grant numbers SES-9818561 and DUE-0226344).

1. Open http://uaeller.eller.arizona.edu/~jcox/index.html; then click on Subject Instructions; then click on the title of this paper.

2. Using plus and minus signs makes it easier to remember in which zone the parameters are relevant. In the zone in which the difference between "my payoff" and "the other's payoff" is negative, the parameters are indexed with $-$, and in the zone in which that difference is positive or zero the parameters are indexed with +.

References

Andreoni, James and John Miller, “Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism,” *Econometrica*, v. 70, no.2, 2002, pp. 737-753.

Bolton, Gary E. and Axel Ockenfels, “ERC: A Theory of Equity, Reciprocity, and Competition,” *American Economic Review*, 90, 2000, pp. 166-93.

Charness, Gary and Matthew Rabin, “Understanding Social Preferences with Simple Tests,” *Quarterly Journal of Economics*, 117, 2002, pp. 817-69.

Cox, James C., Daniel Friedman and Stephen Gjerstad, “A Tractable Model of Reciprocity and Fairness,” discussion paper, Department of Economics, University of Arizona, 2005.

Cox, James C., Daniel Friedman and Vjollca Sadiraj, “Revealed Altruism,” discussion paper in progress, University of Arizona and University of California at Santa Cruz, 2005.

Engelmann, Dirk and Martin Strobel, “Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments,” *American Economic Review*, forthcoming.

Fehr, Ernst and Klaus M. Schmidt, “A Theory of Fairness, Competition, and Cooperation,” *Quarterly Journal of Economics*, August 1999, pp. 817-68.

Güth, Werner, Nadège Marchand, and Jean-Louis Rulliere, “On the Reliability of Reciprocal Fairness-An Experimental Study,” Discussion Paper, Humboldt University of Berlin, 1997.

Hoffman, E., McCabe, K., Shachat, K., and Smith, V. (1994). “Preferences, Property Rights, and Anonymity in Bargaining Games.” *Games and Economic Behavior*, **7**, 346-380.

Roth, Alvin E., Vesna Prasnikar, Masahiro Okuno-Fujiwara, and Shmuel Zamir, “Bargaining and Market Behavior in Jerusalem, Ljublana, Pittsburgh, and Tokyo,” *American Economic Review*, 81, 1991, pp. 164-212.

Table 1. Feasible Choices in Experiment 2

m	y_1	y_2	y_3
10	0	6	6
10	0	15	15
10	0	2	33

Table 2. Feasible Choices in Experiment 3

m	y_1	y_2	y_3
10	0	15	15
10	3	12	15
10	4	5	21

Table 3. Rationalizing Data from Games with Responder Competition

Number of Responders	(n-1)/n	RHS of Inequality (13)
5	0.8	0.501955
6	0.833333	0.557486
7	0.857143	0.601876
8	0.875	0.638172
9	0.888889	0.668402
10	0.9	0.693971

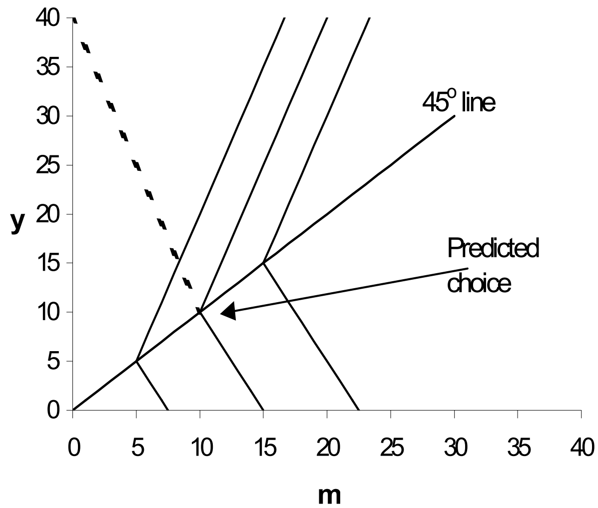

Figure 1. Budget (dashed) Line and Indifference "Curves" for the Fehr- Schmidt Model ($\alpha = 1/2$ and $\beta = 1/3$).

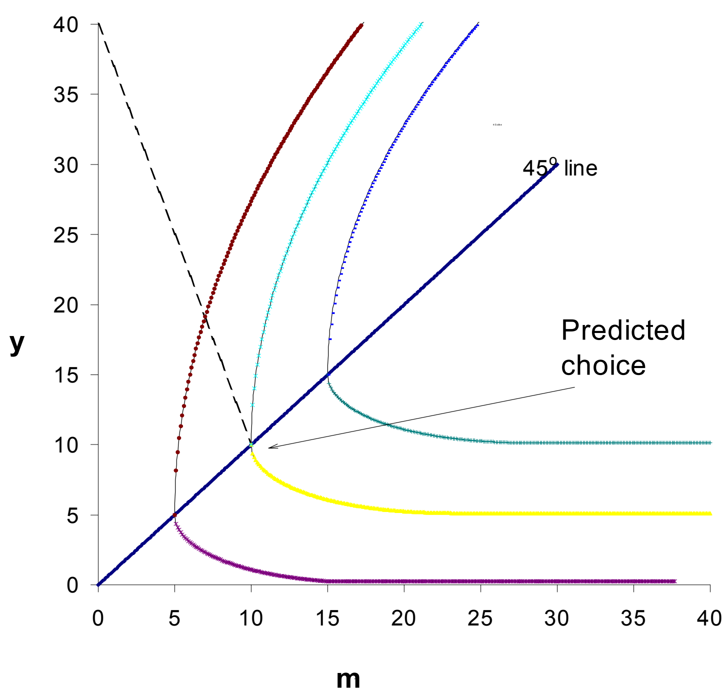

Figure 2. Budget (dashed) Line and Indifference Curves" for the Bolton-Ockenfels Model.

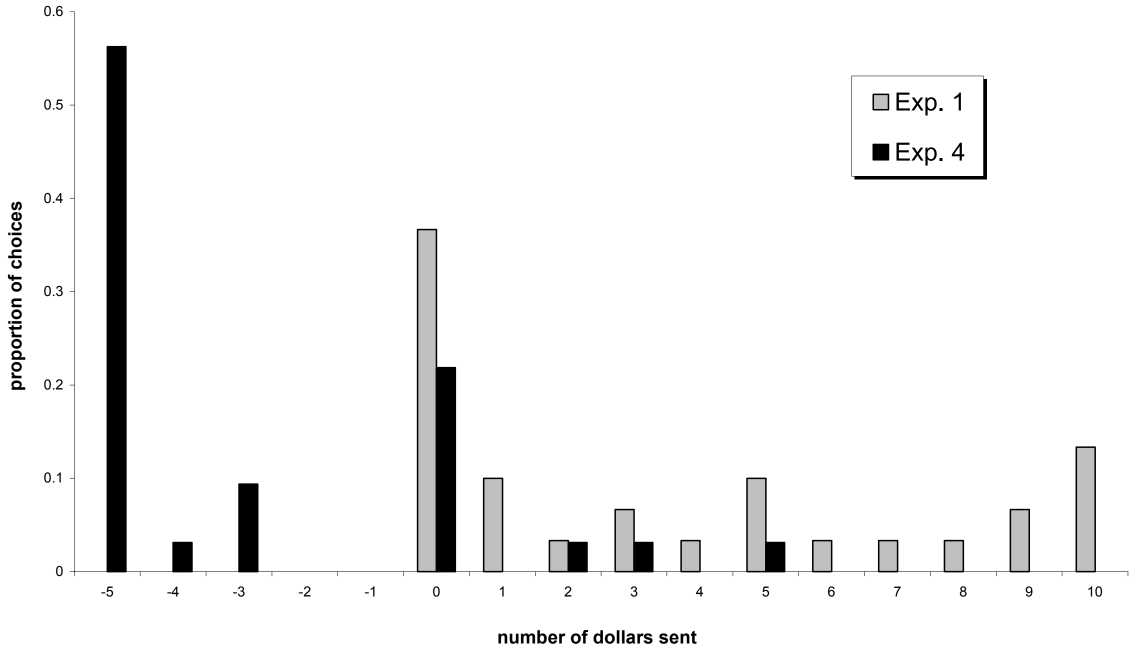

Figure 3. Dictators' Decisions in Experiments 1 and 4

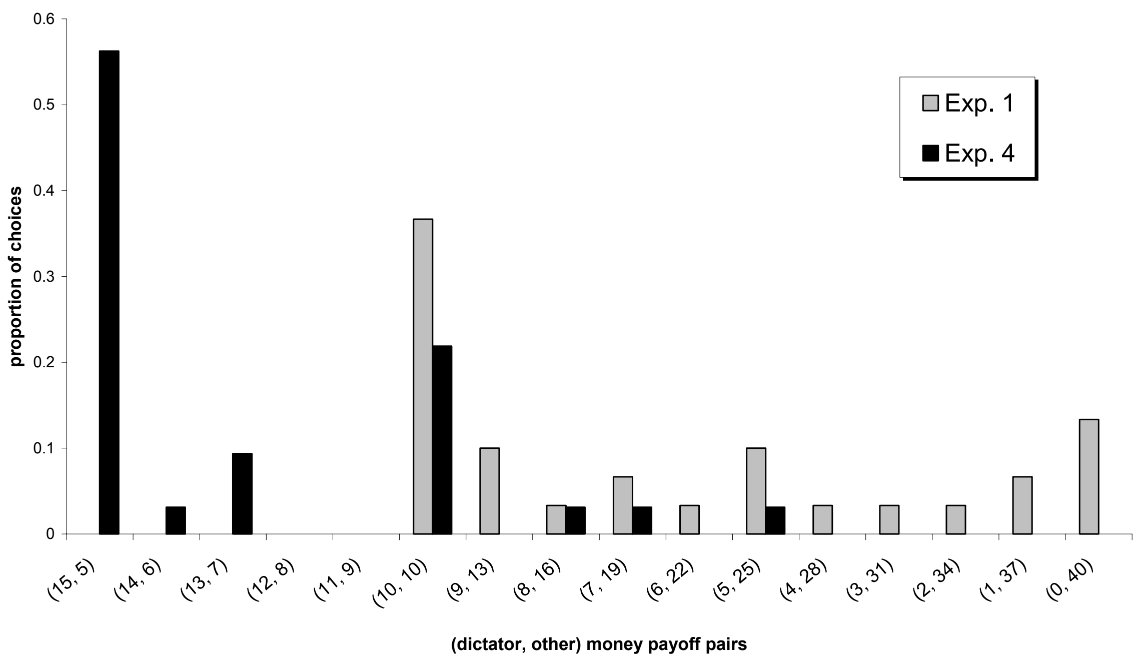

Figure 4. Relative Payoffs in Experiments 1 and 4

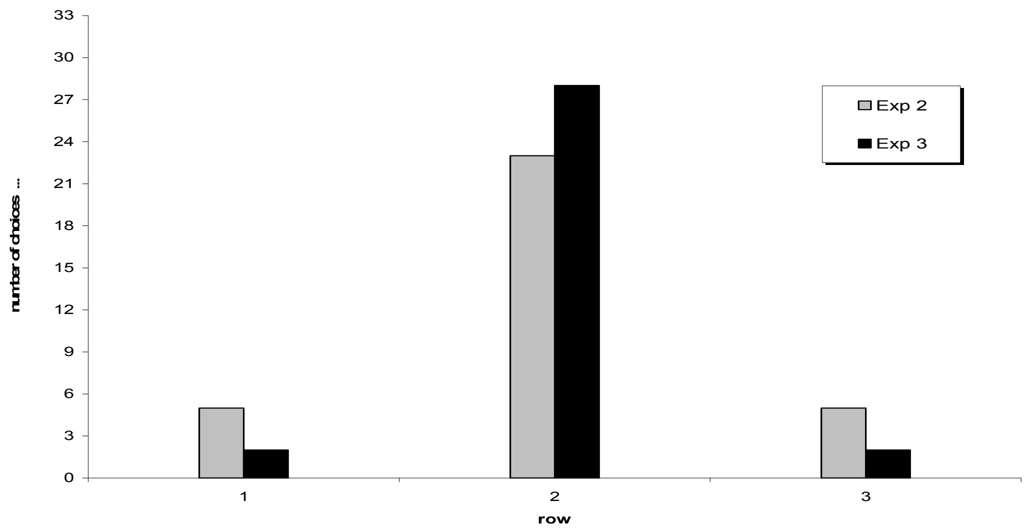

Figure 5. Dictators' Decisions in Experiments 2 and 3

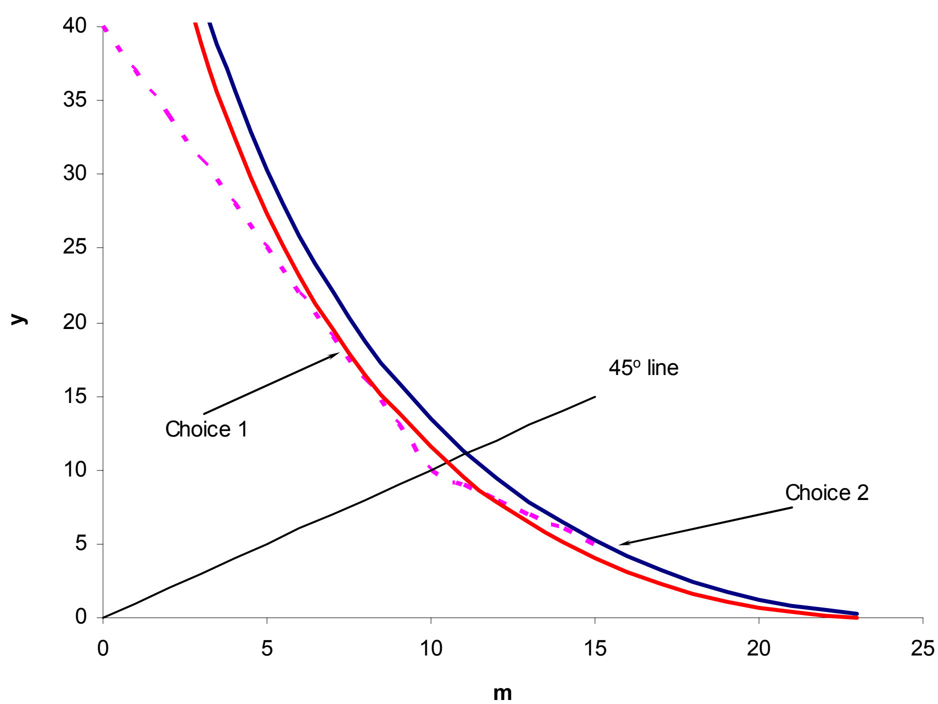

Figure 6. Rationalizing the Modal Observations

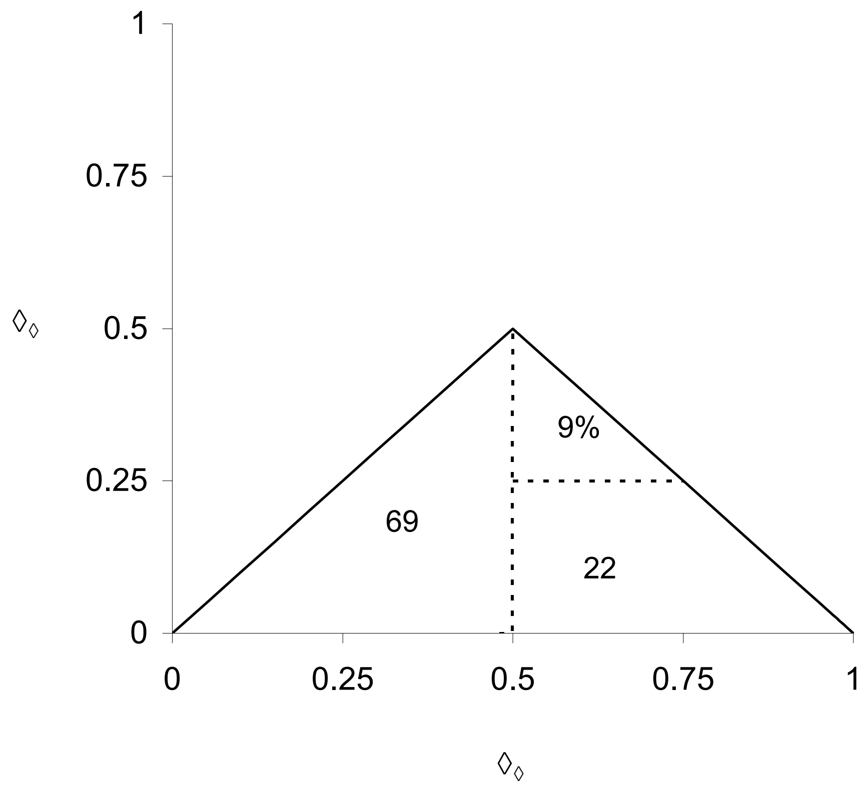

Figure 7: Distribution of Egocentric Utility Function Parameters that Rationalize Experiment Data

Appendix 1: Experiment Procedures

Procedures of Experiment 1

Experiment 1 was a manual (i.e., non-computerized) experiment with a double-blind payoff protocol. During the decision-making part of an experiment session, all interactions with the subjects were carried out by a "monitor" that had been randomly selected from the subject pool. In addition to distributing and collecting "decision envelopes," the monitor was given the responsibility of verifying that the experimenters accurately implemented subjects' decisions in calculating payoffs. The monitor did not discuss the decision task with the subjects. The monitor was paid $20. The amount paid to a monitor was not announced to the subjects so as to avoid the possible suggestion of a focal point earnings figure.

The subjects first gathered in a room adjacent to the laboratory. The monitor was randomly selected from the group of subjects by drawing a ball from a bingo cage in the presence of all of the subjects. Next, all of the rest of the subjects were randomly assigned to "group X" and "group Y." All subjects then entered the laboratory. Group X subjects were seated at widely separated computer terminals with privacy side and front partitions. (The computers were not used.) The group Y subjects were standing at the back of the room. Each subject and the monitor were given copies of the instructions. Then the experimenter read aloud the instructions. After the reading of instructions was completed, the group Y subjects were escorted back to the adjacent room. The group X subjects had no further contact with the group Y subjects. Then the group X subjects were given the opportunity to raise their hands if they had questions. If a subject raised his hand, he was approached by the experimenter and given an opportunity to ask questions and receive answers in a low voice that could not be overheard by other subjects. When there were no more questions, the experimenter left the room and the monitor took over. There was no interaction between the experimenter and the subjects during the decision-making part of an experiment session. All distribution and collection of envelopes containing subject response forms was done by the monitor. Because subject decision forms were inside envelopes when they were distributed and collected, not even the monitor could know any subject's personal decision.

The payoff procedure was double blind: (a) subject responses were identified only by letters that were private information of the subjects; and (b) money payoffs were collected in private from sealed envelopes contained in lettered mailboxes. Double blind payoffs were implemented by having each subject draw an unmarked sealed envelope containing a lettered key from a box containing many envelopes. At the end of the experiment, the subjects used their keys to open lettered mailboxes that contained their money payoffs in sealed envelopes. The experimenters were not present in the mailbox room when the subjects collected their payoff envelopes.

All of the above-described features of the experimental design and procedures were common information given to the subjects before they made their decisions. The subject instructions and response forms did *not* use evocative labels in referring to the two groups of subjects. Instead, the terms "group X" and "group Y" were used. No subject participated in more than one experiment session.

All of the experiment sessions ended with each subject being paid an additional $5 for filling out a questionnaire. Group X and group Y subjects had distinct questionnaires. The questions asked had three functions: (a) to provide additional data; (b) to provide a check for possible subject confusion about the decision tasks; and (c) to provide checks for possible recording errors by the experimenters and counting errors by the subjects. Subjects did *not* write their names or any other identity-revealing information on the questionnaires. The additional data provided by the questionnaires included the subjects' reports of their payoff key letters. Data error checks provided by the questionnaires came from asking the group X subjects to report the numbers of tokens sent. These reports, together with two distinct records kept by the experimenters, provided accuracy checks on data recording.

Procedures of Experiment 2

Experiment 2 was a manual (i.e., non-computerized) experiment with a double-blind payoff protocol. Double-blind payoffs could be, and were, implemented in this experiment without the use of a monitor and mailbox payoff procedure because the payoffs to the dictators were independent of their choices.

The subjects first gathered in a room adjacent to the laboratory. The stations in the laboratory were randomly assigned, in equal numbers, to four groups. This was done independently for each experiment session. A large manila envelope containing experiment documents was placed at each station. Subjects entered the laboratory and sat at any station they chose but without any way of knowing which of the four groups that station had been randomly assigned to. Each station had privacy side and front partitions. Procedural instructions were projected on a screen at the front of the room. The dictators were designated group W. The other subjects were designated groups X, Y and Z. The subjects in groups X, Y, and Z were given questionnaires to fill out; in this way it was not clear to other groups during the experiment which of the seats were occupied by individuals randomly assigned to be dictators. After the passage of more than enough time for decisions to be recorded, an experimenter asked from a back laboratory door for everyone who had completed his questionnaire to raise his or her hand. After all hands were raised, two messages were alternated on the projection screen at the front of the laboratory. One message instructed group X, Y and Z subjects to "wait for further instructions." The other message instructed group W subjects to put all of the experiment material except the disclaimer form and the sealed white legal-size envelope back in the large manila envelope and deposit the manila envelope in a bvx at the front of the laboratory while exiting. They were instructed that the sealed white envelope contained their payoffs ($10, as shown in Table 1 or 2, plus the show-up fee of $5). These envelopes were sealed and had labels attached on both sides with the instruction "not to open this envelope until after exiting the building." After all group

W subjects had exited, an experimenter retrieved the box and took it to the separate rear "monitor" room. The forms with the group W subjects' decisions were extracted. Then the group X, Y and Z subjects were called, one at a time to receive their show-up fees and the payoffs determined by the group W subjects' decisions. This process involved another randomization: the group W subjects' decision forms were applied in random order to determine the payoffs of group X, and Y, and Z subjects. The three screens of projected instructions and the printed instructions contained in the large manila envelopes are available on an experimenter's homepage, as explained in footnote 1.

All of the above-described features of the experimental design and procedures were common information given to the subjects. The subject instructions and response forms did *not* use evocative labels, such as "dictator" in referring to the four groups of subjects. Instead, the terms "group W," "group X," "group Y" and "group Z" were used. No subject participated in more than one experiment session.

Procedures of Experiment 3

The experiment procedures and subject instructions for experiment 3 were the same as those for experiment 2 except for the use of the choices in Table 2 rather than Table 1.

Procedures of Experiment 4

The experiment procedures and subject instructions for experiment 4 were the same as those for experiment 1 except for minimal changes necessary to introduce the opportunity for group X subjects to take money from paired group Y subjects as well as give them money.

Appendix 2: Derivation of Indifference Curves

Indifference Curves for the Fehr-Schmidt Model

The F&S model is based on the assumption that agent i, where $i = 1,2,\cdots n$, has preferences that can be represented by utility functions of the form

(1) $$u_i(x) = x_i - \alpha_i \frac{1}{n-1}\sum_{j \neq i} \max(x_j - x_i, 0) - \beta_i \frac{1}{n-1}\sum_{j \neq i} \max(x_i - x_j, 0)$$

where $\beta_i \leq \alpha_i$ and $0 \leq \beta_i < 1$. The utility function for the two-agent case can be written as

(2) $$u_m(m, y) = m - \alpha(y - m), \text{ if } m < y$$

$$= m - \beta(m - y), \text{ if } m \geq y$$

where $m = x_i$ is "my income" and $y = x_j$ is "your income."

Figure 1 shows indifference "curves" for the F&S model for $\alpha = 1/2$ and $\beta = 1/3$. All parameter values that are consistent with inequality aversion imply that the indifference "curves" have positive slope above the 45-degree line because $MRS(m, y) = -\dfrac{\partial u_m / \partial m}{\partial u_m / \partial y} > 0$ for $m < y$. The budget constraint of a dictator in experiment 1 consists of ordered pairs of integers on the dashed line in Figure 1 extending from the point (10,10) on the 45-degree line to the point (0, 40) on the vertical axis. In this dictator game, since the budget line is above the 45-degree line and has negative slope, the F&S model makes the same predictions as does the traditional self-regarding preferences model: that the dictator will choose the point (10, 10) or, equivalently, that the dictator will give 0 dollars to the other subject.

Indifference Curves for the Bolton-Ockenfels Model

The B&O model is based on a "motivation function" of the form,

(3) $$v_i = v_i(x_i, \lambda_i)$$

where

(4) $$\lambda_i = x_i / \sum_{j=1}^{n} x_j \text{ , if } \sum_{j=1}^{n} x_j > 0$$

$$= 1/n \text{ , if } \sum_{j=1}^{n} x_j = 0 .$$

Given that the sum of my income (*m*) and your income (*y*) is positive, the motivation function for the two-agent case can be written as

(5) $$v = v(m, m/(m+y)) ,$$

where $v(\cdot)$ is (B&O, pgs. 171-172) globally non-decreasing and concave in my income *m*, strictly concave in relative income $m/(m+y)$, and has a partial derivative with respect to relative income with the property

(6) $v_2(m,1/2) = 0$, for all *m*.

Figure 2 shows typical graphs of the level sets or "indifference curves" of the B&O motivation function for the two-agent case with $m + y > 0$. This model also predicts that the dictator will give 0 dollars to the other subject for the same reason as does the F&S model: above the 45-degree line, the "indifference curves" have positive slope whereas the budget line has negative slope.

Appendix 3: Derivations of Propositions 1-3

Proposition 1: Dictator Game

In the dictator game 1, a dictator can decide to give or take away money, by choosing an element s from [-5,10]. The game allows the dictator to: (i) leave both payoffs unchanged (s=0); (ii) increase the other's payoff by 3s and decrease his own by s when s>0; or (iii) decrease the other's payoff by |s| and increase his own payoff by |s| when s<0. Thus, the budget set is $\{(m,y) \mid m = 10 - s, y = 10 + s, \quad if\ s \in [-5,0]; m = 10 - s, y = 10 + 3s, \quad if\ s \in [0,10]\}$.

Referring to *u(m,y)* as given in (2) one has

(7)
$$-\frac{u_m}{u_y}(m,y) = -\frac{1-\theta_+}{\theta_+}\left(\frac{m}{y}\right)^{\alpha-1}, \quad if\ m > y,$$
$$= -\frac{1-\theta_-}{\theta_-}\left(\frac{m}{y}\right)^{\alpha-1}, \quad if\ m < y.$$

Writing it for points in the budget set and as a function of s, one has

(8)
$$-\frac{u_m}{u_y}(s) = -\frac{1-\theta_+}{\theta_+}\left(\frac{10-s}{10+s}\right)^{\alpha-1}, \quad if\ s \in [-5,0),$$
$$= -\frac{1-\theta_-}{\theta_-}\left(\frac{10-s}{10+3s}\right)^{\alpha-1}, \quad if\ s \in (0,10].$$

Note that $-\frac{u_m}{u_y}(s)$ decreases as *s* increases, that is $\forall s, s' \in (0,10]$ such that $s < s'$,

(9)
$$-\frac{u_m}{u_y}(s) > -\frac{u_m}{u_y}(s').$$

Part (i) of Proposition 1. Suppose that (*i.1*) $-\frac{1-\theta_-}{\theta_-} \leq -3$ and (*i.2*) $-\frac{1-\theta_+}{\theta_+} \geq -1$. Then the optimal choice for the dictator is $s = 0$, as follows. Let $U(s)$ denote the value of a dictator's indifference curve through $(10-s, 10+s)$, where $s \in [-5,0)$. Hence, $U(s) = u(10-s, 10+s)$, and $dU(s) = -u_m(s)ds + u_y(s)ds$. The first derivative of $U(s)$ is strictly positive in [-5,0) if for all such s from [-5,0) one has $-\frac{u_m}{u_y}(s) > -1$ (recall that $u_y > 0$). Condition (*i.2*) and inequality (9) imply that this is true. Thus $U(s)$ is a monotonically decreasing function in [-5,0) and its limit as s goes to 0 is 10, which is the value of $u(10,10)$ as well. This implies that all allocations of money payoffs in the budget set that are below the 45-degree line are on lower indifference curves than the allocation (10,10). Similarly, for all $s \in (0,10]$ one has $U(s) = u(10-s, 10+3s)$, $dU(s) = -u_m(s)ds + 3u_y(s)ds$, and the first derivative of $U(s)$ is strictly negative in (0,10] if for all s from (0,10], $-\frac{u_m}{u_y}(s) < -3$. Condition (*i.1*) and (9) imply that this is true and therefore all allocations in the budget set and above the 45-degree line are on lower indifference curves than (10,10). Hence, the best choice for the dictator is $s = 0$.

Part (ii) of Proposition 1. Suppose that (*ii.1*) $-\frac{1-\theta_-}{\theta_-} > -3$ and (*ii.2*) $-\frac{1-\theta_+}{\theta_+} \geq -1$.

Similarly to part (*i*) above, one can show that (*ii.2*) and (9) imply that all allocations (m, y) in the budget set and below the 45-degree line are on lower indifference curves than (10,10). Condition (*ii.2*) and inequality (9) on the other hand imply that either: (a) there exists some $s_-^* \in (0,10]$ such that $-\frac{u_m}{u_y}(s_-^*) = -3$; or (b) $-\frac{u_m}{u_y}(s) > -3$ for all $s \in (0,10]$. In case (a), inequality (9) tells us all allocations in the budget set and above the 45-degree line are on lower indifference curves than the one passing through (10-s_-^*,10+3s_-^*), hence $s^o = s_-^* > 0$. In case (b) the first derivative of $U(s)$ is positive for all s in (0,10] and therefore all allocations in the budget set and above the 45-degree line are on lower indifference curves than (0,40), hence $s^o = 10$. Note that $U(s_-^*)$ is an upper bound for $U(s)$, for all s in (s_-^*,0), and therefore the limit of $U(s)$ as s goes to 0 from above, which is 10 ($= u(10,10)$), is not larger than the upper bound $U(s^*_-)$; that is allocation (10,10) is not on a higher indifference curve than $(10 - s^*_-, 10 + 3s^*_-)$. Thus, the (maybe weakly) optimal choice for the dictator is s^*_- >0.

Part (iii)-(iv) of Proposition 1. Proofs are similar to parts *(i)* and *(ii).*

<u>Proposition 2: Games with Proposer Competition</u>

Let the maximum offer be $\bar{s} \geq 0$.

Part (i). Suppose that the responder rejects it. Then all players get money payoffs 0 and hence utility 0. From the monotonicity assumption on money payoffs, the responder derives a positive utility by deviating and accepting *s*. Hence, it is a dominant strategy for the responder to accept any offer $\bar{s} \geq 0$.

Part (ii). Suppose that $\bar{s} < 1$. Then, since from part (*i*) the maximum offer is accepted, the utility of a proposer *i* who offers some *s* is given by:

(a) if $s = \bar{s}$ and *(k-1)* other proposers submit $\bar{s}$, then the expected utility of proposer *i* is

$$Eu^i(\bar{s}) = \frac{1}{k} u^1(1 - \bar{s}, \bar{s}, 0, 0, \ldots, 0) + (1 - \frac{1}{k}) u^1(0, \bar{s}, 1 - \bar{s}, 0, \ldots, 0);$$

(b) if $s < \bar{s}$ then the utility of proposer *i* is

$$u^i(s) = u^1(0, \bar{s}, 1 - \bar{s}, 0, \ldots, 0).$$

To show that there are no subgame perfect equilibria with a maximum offer of $\bar{s} \in [0,1)$, we