Direct Tests of Individual Preferences for Efficiency and Equity

By James C. Cox and Vjollca Sadiraj

Experimental Economics Center
Georgia State University

April 2010

Direct Tests of Individual Preferences for Efficiency and Equity

By James C. Cox and Vjollca Sadiraj*

Abstract: Departures from "economic man" behavior in many games in which fairness is a salient characteristic are now well documented in the experimental economics literature. These data have inspired development of models of social preferences that assume agents have preferences for equity and efficiency as well as their own material payoffs. Empirical failure of the economic man model comes from experiments that provide direct tests of its distinguishing characteristic: indifference to the payoffs of others. This paper reports an experiment that subjects popular social preferences models to the same type of empirical challenge. We report direct tests of the distinguishing characteristics of these models: preference for allocations that have higher efficiency and greater equity.

1. Introduction

Economics has a long history of using models of preferences to explain choices. 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. There is now a large literature that supports the conclusion that self-regarding preferences models are mostly inconsistent with observed behavior in experiments in which fairness is a salient characteristic of the decision tasks. In response, various models of social preferences have been developed and applied to data from experiments.

Two prominent types of models of social preferences are inequality (or inequity) aversion models (Fehr and Schmidt, 1999 and Bolton and Ockenfels, 2000) and the quasi-maximin model (Charness and Rabin, 2002). These models have been widely applied to data in the past and continue to be applied in current literature (Chen and Li, 2009; Fehr, Klein, and Schmidt, 2007). In the present paper, we focus on testing the distinguishing characteristics of the models rather than fitting their parameters to data. The distinguishing characteristic of inequality aversion models is that utility is decreasing with the absolute (value of the) difference between one's own and others' material payoffs as well as being increasing with one's own payoff. The

distinguishing characteristic of the quasi-maximin model is that utility is increasing with the lowest of all agents' payoffs (the maximin property) and the total of all agents' payoffs (the efficiency property) as well as being increasing with one's own payoff.

Inequality aversion, efficiency, and maximin have been described as "motives for behavior" and controversy has developed about the relative importance of these motives for explaining behavior (Engelmann and Strobel, 2004, 2006; Bolton and Ockenfels, 2006; Fehr, Naef, and Schmidt, 2006). We here apply a somewhat different approach with an experiment that includes three treatments. Each treatment implements an experiment design that identifies and tests an observable consequence of a single one of these possible properties of distributional preferences; we refer to this type of test as a "direct test." Treatment one implements a direct test for inequality aversion: a subject whose preferences include an aversion to inequality in payoffs favoring another person will make one specific choice in this treatment while other feasible choices are inconsistent with inequality aversion. Treatment two implements a direct test in which one feasible choice is consistent with preferences that are monotonically increasing in the total of all agents' payoffs while other feasible choices are inconsistent with such preference for efficiency. Treatment three implements a similar direct test for preferences that include the maximin property.

Choices made by large majorities of subjects in the three direct-test experiment treatments reported herein are inconsistent with preferences characterized by inequality aversion, efficiency, or maximin. This might seem surprising, given the many applications of the inequality aversion and quasi-maximin models that seem to show that the models fit data from various experiments pretty well. But in many experiments the implications of inequality aversion, efficiency, and maximin are confounded with the implications of the conventional convexity property of preferences. We will explain that most data from our three experiment treatments that are inconsistent with the inequality aversion and quasi-maximin models are, in contrast, consistent with preference models that include conventional properties such as convexity

and positive monotonicity for all agents' payoffs. One such model is the egocentric altruism model (Cox and Sadiraj, 2007).[1]

2. Treatment 1: A Direct Test for Inequality Aversion

Utility functions for inequality aversion models are increasing with an agent's own material payoff but decreasing with the absolute (value of the) difference between her own payoff and others' material payoffs. For the special case of two agents and 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 for inequality aversion.

2.1 Experiment Design and Procedures

Treatment 1 involves a dictator game with the following characteristics. Subjects are randomly assigned to pairs. In addition to a show-up fee of $5, each of the two subjects in a pair is given an endowment of $10. The "non-dictator" in a pair of subjects has no decision to make. A dictator has (only) one decision to make. A dictator is told that he/she can send zero or a positive amount (in whole dollar units), up to $10, from his/her endowment to the other person. Each dollar that a dictator transfers to the other person is multiplied by three by the experimenters.

After the dictator makes his/her decision, both subjects in the pair are informed of their final money payoffs that are determined by the dictator's decision. Each pair of subjects is informed only about their own two amounts of money payoffs; they are not informed of the payoffs of other pairs of subjects. The experiment protocol uses a double-blind payoff procedure in which neither the other subjects nor the experimenters can identify the individual who has made any specific decision. This protocol is implemented by having each subject draw an envelope from a box of identical envelopes containing uniquely numbered mailbox keys. The number on a subject's mailbox key is private information of the subject. This mailbox key

number is the only way that a subject's responses are identified in the experiment and data record. Subjects use their mailbox keys to collect their payoffs (contained in sealed envelopes) from their mailboxes in private. All of the features of the experiment, including the equal endowments of dictators and non-dictators, are common information given to the subjects. The subject instructions are available on the journal's website.[2]

2.2 Predictions of the Inequality-Aversion Models

Figure 1 shows typical indifference curves for the two-agent version of the Fehr and Schmidt (1999) model for the dictator's ("my") money payoff m and the other person'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. Including the $5 show-up fees in payoffs, the budget constraint of a dictator in treatment 1 consists of ordered pairs of integers on the dashed line in Figure 1 extending from the point (15,15) on the 45-degree line to the point (5, 45) near the vertical axis. In this dictator game, the Fehr-Schmidt model predicts that a dictator will give 0 dollars to the other subject. (See appendix section A.1 for a formal derivation.)

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 $m + y > 0$. 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. (See appendix section A.2 for a formal derivation.)

Treatment 1 provides a general test for inequality aversion. The test does *not* depend on the specific parametric utility functions conventionally used to represent the Fehr-Schmidt and Bolton- Ockenfels models.

`

2.3 Subjects' Behavior in Treatment 1

As we explained, both inequality aversion models predict that a dictator will send 0, which results in the payoffs $(m,y) = (15,15)$.   Data from treatment 1 are reported in Figure 3.   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 63% of dictators who sent positive amounts of money to the other subjects imposed significant costs on themselves to *increase* inequality favoring others. This behavior is inconsistent with the central distinguishing characteristic of inequality aversion models. The average amount given away by the dictators was $3.60, which gave the average recipient a payoff of $25.80 (= $5 + $10 + 3×$3.60). This left the dictators with an average payoff of $11.40 (= $5 + $10 - $3.60), and produced an allocation with much higher inequality favoring non-dictators than the initial money allocation. The disadvantageous difference in average payoffs increased from $0 to $14.40 whereas the average proportion of total payoffs received by dictators decreased from 0.5 (the most favored ratio by inequality aversion) to 0.31 (= 11.40/37.20).   Finally, note that 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 even one subject in treatment 1.

2.4 Questions about the Experiment Design and Protocol

It is natural to ask what might account for the generosity exhibited by the subjects in this treatment. We discuss four possible explanations, stated in the form of questions that might be asked: (a) Were the subjects confused? (b) Was there an experimenter "demand effect"? (c) Does generosity vary with the price of giving? (d) Do the subjects exhibit a preference for efficiency?

It is difficult to believe that subjects could be confused about the simple dictator game explained in our instructions (see the instructions on the journal website). Furthermore, after

completing their decision forms, subjects filled out a questionnaire including questions about their reasons for making the decisions they recorded in the experiment. No responses to these questions showed confusion about the decision task or other conditions of the experiment. Instead, in explaining reasons for giving a typical response from subjects who gave positive amounts to others was wording similar to: "It only cost me $1 to give $3 to the other guy." The dictators were clearly informed that the other (non-dictator) subjects had each been given the same $10 endowment and $5 show-up fee as the dictators.

It is hard to construct a story about a demand effect in this experiment. Since the payoff procedure was double blind, it was impossible for a subject to acquire a reputation for generosity either with other subjects or the experimenter. Furthermore, subjects would undertake the same physical action in recording a 0 on their response forms as in recording some positive integer.[3]

Comparison of data from treatment 1 with data from another dictator experiment 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 treatment 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 was $0.33 in our treatment 1. The implied (arc) price elasticity of demand for increasing the other subject's payoff is –1.12, a quite reasonable figure. In this way, a preference model with conventional properties of convexity and monotonicity and positive income effects (or normal goods) can account for the data from these dictator experiments.

An off-heard interpretation of behavioral inconsistency with inequality aversion is based on the view that people have a preference for "efficiency" of final outcomes. In treatment 1, equality of payoffs is free: it can be implemented at zero cost to the dictator by a choice of zero as the amount to send to the other. In this way, the inconsistency with inequality aversion is a strong result. Choosing larger total payoffs ("efficiency") is costly to the dictator: each one dollar sent

to the other person increases total payoff to the pair of subjects by $2 but costs the dictator $1. In this way, it may seem unlikely that a preference for efficiency can explain subjects' choices in treatment 1, but a dictator with a sufficiently strong preference for efficiency may be willing to pay a high price to attain it. Charness and Rabin (2002) offer a social preferences model that includes a preference for efficiency. Their model of quasi-maximin preferences preserves the inequality aversion property only with respect to the poorest individual, and it incorporates a preference for efficiency. We next report direct tests of the quasi-maximin model.

3. Treatments 2 and 3: Direct Tests for Quasi-Maximin Preferences

The quasi-maximin model, as introduced by Charness and Rabin (2002), is based on the assumption that an agent's utility is increasing in: (a) his or her own money payoff; (b) the total of all individuals' money payoffs (efficiency); and (c) the minimum money payoff across individuals (maximin). An agent with quasi-maximin preferences will prefer an option with a higher amount of total money payoff to an option with a lower amount of total money payoff when the other two outcome measures are the same in the two options. An agent with quasi-maximin preferences will prefer an option with a higher payoff to the lowest paid individual to an option with a lower payoff to the lowest paid individual when the other two outcome measures are the same in the two options. These properties form the bases for direct tests for quasi-maximin preferences.

3.1 Experiment Design

In treatment 2, we offer subjects choices between alternatives in a dictator game in which the dictator's own payoff and the lowest individual payoff are constant but the sum of all payoffs changes. Hence in treatment 2 the dictator's price (in own payoff foregone) is 0 for increasing efficiency. The feasible set in this treatment includes three alternative allocations of money to four individuals. Each of the three possible allocations pays the dictator $10 and pays the lowest-

paid individual $0. Payoffs to the other two individuals are ($6,$6) or ($15,$15) or ($2,$33) in the three alternative allocations. The top part of Table 1 shows the choices (including the show-up fee of $5) available to a subject in treatment 2. Note that the total payoff varies from $42 to $60 to $65 while the dictators own payoff (m) remains constant at $15 and the minimum payoff remains constant at $5. Therefore, the quasi-maximin model which includes preferences that are monotonically increasing in own payoff, minimum payoff, and total payoff (or efficiency) predicts that subjects will choose the row with the maximum total payoff of $65. (See appendix section A.3 for a formal derivation.) This treatment provides a direct test for a preference for efficiency.

Whereas treatment 2 tests for a preference for efficiency, treatment 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). In treatment 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. Hence in treatment 3 the dictator's price (in own payoff foregone) is 0 for increasing the payoff of the lowest paid individual.

The bottom part of Table 1 shows the choices (including the show-up fee of $5) available to a dictator in treatment 3. Note that the dictator's payoff is the same in all three of these allocations, and so is the total payoff to all agents. But the minimum payoff varies from $5 to $8 to $9 and is highest in the bottom row. Choice of the bottom row is the unique prediction of the quasi-maximin model. (See appendix section A.3 for a formal derivation.) This treatment provides a direct test of the maximin property of the model.

3.3 Procedures in Treatments 2 and 3

Treatments 2 and 3 have the following characteristics. Subjects are randomly assigned to groups of size four that consist of a dictator and three "non-dictators." A dictator is asked to choose one of three available allocations of payoffs for the four individuals in his/her group. Each dictator

makes (only) one decision. Each non-dictator has no decision to make. Different subjects participate in treatments 2 and 3. The experiment 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. At the end of a treatment, each group of subjects is informed of their own four payoffs but is not informed of the payoffs of other groups. All of the features of the experiment are common information given to the subjects. The subject instructions are available on the journal's website.

3.4 Behavior in Treatments 2 and 3

Subjects' behavior in treatments 2 and 3 is reported in Figure 4. We observe that only 5 of 33 (or 15%) of the subjects chose the efficient allocation (15,5,7,38) in treatment 2, which is the unique prediction of the quasi-maximin model. Hence the behavior of 85% of the subjects in treatment 2 is inconsistent with preferences for efficiency. Recall that the price (in decreased own money payoff) of buying the most efficient allocation is zero, hence an agent with quasi-maximin preferences will do so.

In treatment 3, only 2 of 32 (or 6%) of the subjects chose the maximin allocation (15,9,10,26), which is the unique prediction of the quasi-maximin model. Hence the behavior of 94% of the subjects in treatment 3 is inconsistent with maximin preferences. Recall that the price (in decreased own money payoff) of buying the maximin allocation is zero in treatment 3, hence an agent with quasi-maximin preferences will do so.

4. Can Data from All Three Treatments be Rationalized by Any One Model?

We next consider data from all three treatments and ask whether any model can rationalize all of the data.

4.1 The Inequality Aversion and Quasi-Maximin Models Fail for At Least One Treatment

The quasi-maximin model of Charness and Rabin (hereafter CR) can explain only 15% of the choices observed in treatment 2 and only 6% of the choices in treatment 3 as shown in the two rows containing CR in the Predictions column of Table 1. The quasi-maximin model can potentially rationalize treatment 1 data. As shown in appendix section A.3, this model predicts choice of (15,15) if $\gamma(1-\delta) < 1/3$ and choice of (5,45) if $\gamma(1-\delta) > 1/3$. With $\gamma(1-\delta) = 1/3$ the model predicts indifference among all feasible allocations. Data from treatment 1 show that half of the subjects made choices that correspond to neither the (15,15) nor the (5,45) allocation, which according to the quasi-maximin model leaves only parameter values such that $\gamma(1-\delta) = 1/3$.

We next ask whether the inequality aversion models can rationalize treatment 2 and 3 data. Utility in the Fehr and Schmidt (hereafter FS) model is increasing in my payoff and decreasing in both payoff differences that favor me and payoff differences that favor others. The dictator's payoff is the same in all allocations in treatment 2. The allocation (15,5,11,11) in the second row of Table 1 has 18 favorable payoff differences and zero unfavorable payoff differences. This is clearly better than the 18 favorable differences and 23 unfavorable payoff differences in allocation (15,5,7,38) in the first row of the table. It is also better than the 10 favorable differences and 10 unfavorable differences for allocation (15,5,20,20) in the third row of Table 1 because unfavorable differences have a weakly higher weight than favorable differences. Therefore, as reported in the Predictions column of Table 1, the FS model predicts choice of (15,5,11,11) in treatment 2. (See appendix section A.1 for a formal derivation.) The FS model can explain only 15% of the treatment 2 data. In treatment 3, both favorable and unfavorable payoff differences are lowest (at 7) for allocation (15,8,17,20). Therefore, the FS model predicts choice of this allocation. (See appendix section A.1 for a formal derivation.) As reported in Table 1, the FS model can explain choices by 88% of the subjects in treatment 3.

Utility in the Bolton and Ockenfels (hereafter BO) model is increasing in my payoff m and decreasing in the absolute difference between 0.5 and the ratio of my payoff to total payoff of all subjects ($m/(m + y_1 + y_2 + y_3)$). It can be easily seen in the top part of Table 1 that the BO model predicts allocation (15,5,11,11) in treatment 2 because this allocation has the payoff ratio 0.36 that is closest to 0.5 and all allocations pay the same amount 15 to the dictator. (See appendix section A.2 for a formal derivation.) Therefore, as reported in the right-most two columns of Table 1, the BO model can explain only 15% of the treatment 2 data. The BO model predicts indifference among all three allocations in treatment 3 because they all have the same ratio and the same payoff to the dictator. (See appendix section A.2 for a formal derivation.) Since we do not observe choices to be randomly distributed among the three allocations in treatment 3, we conclude that these data are also inconsistent with the BO model.

The very high rates of inconsistency between subjects' behavior and predictions for treatments that test the defining characteristics of the models are quite striking because the own-payoff price of buying the supposedly preferred outcome – equality of payoffs or efficiency or maximin – was zero. This suggests the importance of the question of whether the behavior observed in treatments 1-3 can be rationalized by some other type of model.

4.2 The Egocentric Altruism Model Can Rationalize the Data

A model with conventional indifference curves that are downward sloping and convex to the origin can potentially explain 100% of treatment 1 data, 85% of treatment 2 data, and 88% of treatment 3 data. One such model is the egocentric altruism model (Cox and Sadiraj, 2007). In addition to convexity, the agent's altruistic preferences are assumed to be "egocentric." Egocentricity means that, for payoffs that are unequal, I prefer that I get the larger payoff rather than someone else get it. For the special case of my utility $u(m, y)$ defined over my payoff m and your payoff y, egocentricity means $u(b,a) > u(a,b)$ for all a and b such that $b > a \geq 0$.

The statement of the egocentricity property is somewhat less transparent when there are more than two agents. Let $y^k(z) \equiv (y_1, \cdots, y_{k-1}, z, y_{k+1}, \cdots, y_n)$ and let $u(m, y^k(z))$ denote my utility for my own and others' payoffs. The egocentricity property means:

(1) $u(b, y^k(a)) > u(a, y^k(b))$, for $k = 1, 2, \cdots, n$, for all a and b such that $b > a \geq 0$.

Egocentricity implies that my willingness to pay to increase someone else's payoff is less than one when our payoffs are equal. Other properties of the model are that my willingness to pay to increase another's payoff: (a) is everywhere positive (monotonicity); and (b) increases as my payoff increases and another's decreases, given constant utility (convexity).

It is natural to assume one other property of altruistic preferences for environments with anonymity (such as experiments with single-blind or double-blind payoffs). If I don't know *who* is "agent k" and *who* is "agent j" then my preferences do not discriminate between their payoffs. For the special case of my utility $u(m, y_1, y_2)$ defined over my payoff m and the payoff of the first anonymous other person y_1 and the second anonymous other person y_2, "nondiscrimination" means $u(m, a, b) = u(m, b, a)$ for all positive m, a, and b. The statement of the nondiscrimination property is somewhat less transparent when there are more than two other agents. Let my utility $u(m, y)$ depend on my payoff m and the n-vector of payoffs y to the n other agents. The nondiscrimination property is:

(2) $u(m, \tilde{y}) = u(m, \hat{y})$ for all m and all $\tilde{y}$ and $\hat{y}$ that are permutations of each other.

A special case, parametric version of the egocentric altruism model is give by the CES functional form reported in appendix section A.4. This parametric form exhibits the four qualitative properties discussed above for the nonparametric model: monotonicity, convexity, egocentricity, and nondiscrimination.

We next ask whether the behavior of subjects in treatments 1, 2, and 3 is consistent with the egocentric altruism model. Appendix section A.4 presents formal derivations for the parametric

specification of the egocentric altruism model. Here we discuss results that follow from the four qualitative properties of the model as well as the sharper results that follow from the parametric specification.

Recall that, in treatment 1, the dictator's budget line is negatively sloped and the price of giving is 1/3. By the monotonicity assumption, indifference curves of the egocentric altruism model are downward sloping everywhere. Furthermore, willingness to pay above the 45-degree line (where your payoff is higher than mine) is lower than 1 by convexity and egocentricity. This is consistent with my sending you a positive amount of money or sending you zero, depending on the curvature of indifference curves above the 45-degree line. In this way, the egocentric altruism model is consistent with all of the data from treatment 1. The parametric form of the model predicts the dictator will give the other person zero if the value of the parameter θ, the weight on the other's payoff, is $\theta \leq 1/3$ and will give a positive amount if $\theta > 1/3$.

For treatment 2, the egocentric altruism model ranks allocation (15,5,20,20) higher than (15,5,11,11) because of positive monotonicity in all payoffs. So the model predicts that no agent would choose allocation (15,5,11,11). Ranking of allocations (15,5,20,20) and (15,5,7,38) depends on the curvature of indifference curves. We observe from Figure 4 and Table 1 that the egocentric altruism model is consistent with the behavior of 28 out of 33 (or 85%) of the subjects in treatment 2 who chose either (15,5,20,20) or (15,5,7,38). The 15% of dictators who chose allocation (15, 5, 11, 11) violate the positive monotonicity in others' money payoff property of the model.

In treatment 3, allocation (15,8,17,20) is preferred to allocation (15,5,20,20) since the utility gain from an increase of the payoff of (the lower-paid) agent 2 by 3 units is larger than the decrease in the utility from a reduction of the payoff of (the higher-paid) agent 3 by 3 units because of convexity and nondiscrimination. As shown in appendix section A.4, the parametric form of the egocentric altruism model has the additional implication that allocation (15,8,17,20) is preferred to allocation (15,9,10,26). In this way, the parametric form of the egocentric altruism

model predicts that the dictator will choose allocation (15,8,17,20) in treatment 3. We observe from Figure 4 that 28 out of 32 subjects chose this allocation in treatment 3; hence the egocentric altruism model is consistent with the behavior of 88% of the subjects in that treatment.

5. Concluding Remarks

A large majority (63%) of subjects made choices that are inconsistent with inequality aversion models in our treatment 1 dictator game designed to provide a direct test for inequality aversion. This is consistent with findings from other types of experiments reported by Charness and Rabin (2002), Deck (2001), and Engelman and Strobel (2004). Most subjects (85% and 94%, respectively) made choices that are inconsistent with the quasi-maximin model in our treatment 2 and treatment 3 dictator games designed to provide direct tests of that model's efficiency and maximin properties. In contrast, the egocentric altruism model is consistent with the behavior of most of the subjects in all three dictator games (100%, 85%, and 88%, respectively, in treatments 1, 2, and 3). Hence, the model is inconsistent with only 15% and 12% of the data, respectively, in treatments 2 and 3.

Andreoni and Miller (2002) report tests of data from many dictator games, with varying budgets and own-payoff prices of 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. They conclude that a CES utility function similar to the egocentric altruism model can be used to represent preferences revealed by subjects in their experiment.

The inequality aversion models, quasi-maximin model, and egocentric altruism model all exhibit a fundamental property of neoclassical preference theory (Hicks, 1946; Samuelson, 1947), which is that preferences over allocations are an invariant characteristic of an agent that is independent of others' actions. Extensions of the egocentric altruism model to incorporate reciprocity constitute a departure from neoclassical theory in which my willingness to pay to

increase or decrease your material payoff depends on your prior actions that help or harm me. Extensions of the egocentric altruism provide a theory of reciprocity. In the parametric model of reciprocity reported by Cox, Friedman, and Gjerstad (2007), 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. In the nonparametric model reported by Cox, Friedman, and Sadiraj (2008), 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 choices and whether those choices were acts of commission or acts of omission. These models have been successfully applied to data from several different types of experiments including the dictator game (with and without earned endowments), ultimatum game, min-ultimatum game, investment game, moonlighting game, Stackelberg duopoly game, Stackelberg mini-game, and carrot and stick games.

Some recent experiments provide difficulties for all of the models discussed above including both the unconditional preference models such as the inequality aversion, quasi-maximin, and egocentric altruism models and the models of reciprocal preferences such as revealed altruism theory. For example, Cherry, Frykblom, and Shogren (2002) report an experiment in which dictators behave differently with earned and unearned endowments. Such dependence is not modeled by any of the theories we have discussed. Bardsley (2008) and List (2007) report experiments with dictator games that produce another interesting challenge to existing theory. One of the treatments in List's paper enlarges the feasible set from opportunities (only) to give money to opportunities to both give and take. The most striking result reported by List is that introduction of a symmetric take opportunity into a gift-only dictator game shifts median behavior from gift-giving generosity to confiscatory selfishness. Bardsley reports experiments in which asymmetric take opportunities are introduced into give-only dictator games and a third experiment with treatments that involve mirror-image give and take opportunities. The Bardsley and List experiments challenge all of the above preference theories, and many other

theories (Neilson, 2009), because they elicit behavior that is linked not only to individuals' preferences but also to properties of the situation (or context) in which behavioral responses are elicited. These data have motivated an extension of the egocentric altruism model based on an axiomatic characterization of power and legitimacy of power that rationalizes the Bardsley, List and Cherry, et al. data (Cox and Sadiraj, 2009). This model also incorporates the axioms of revealed altruism theory (Cox, Friedman, and Sadiraj, 2008), and in that way provides an integrated theoretical treatment of both unconditional altruism and of reciprocity, power, and property rights legitimacy.

Endnotes

1. The egocentric altruism model is also more consistent with subjects' end-period choices in voluntary contributions public goods experiments than is the Fehr-Schmidt inequality aversion model (Cox and Sadiraj, 2007).

2. Data from this treatment have been previously reported, as treatment B in Cox (2004), and used therein to test other hypotheses.

3. Recent discussion about possible determinants of behavior in dictator games (Bardsley, 2008) includes the "Hawthorne effect" in which subjects' knowledge that they are being observed is said to change their behavior. Careful examinations of the original Hawthorne plant data (Jones, 1992; Levitt and List, 2009) find little empirical support for the existence of the "Hawthorne effect" asserted in much subsequent literature. Nevertheless, data from experiments that motivated development of social preferences models came from experiments equally subject to, or not subject to, possible Hawthorne effects as the experiment reported herein.

References

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

Bardsley, N. "Dictator Game Giving: Altruism or Artefact?" *Experimental Economics*, 11, 2008, 122-33.

Bolton, G. E., and A. Ockenfels. "ERC: A Theory of Equity, Reciprocity, and Competition." *American Economic Review*, 90, 2000, 166-93.

--------------. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments: Comment." *American Economic Review*, 96, 2006, 1906-11.

Charness, G., and M. Rabin. "Understanding Social Preferences with Simple Tests." *Quarterly Journal of Economics*, 117, 2002, 817-69.

Chen, Y. and S. X. Li. "Group Identity and Social Preferences." *American Economic Review,* 99, 2009, 431-57.

Cox, J. C. "How to Identify Trust and Reciprocity." *Games and Economic Behavior*, 46, 2004, 260-81.

Cox, J. C., D. Friedman, and S. Gjerstad. "A Tractable Model of Reciprocity and Fairness." *Games and Economic Behavior*, 59, 2007, 17-45.

Cox, J. C., D. Friedman, and V. Sadiraj. "Revealed Altruism." *Econometrica*, 76, 2008, 31–69.

Cox, J. C., and V. Sadiraj. "On Modeling Voluntary Contributions to Public Goods." *Public Finance Review*, 35, 2007, 311-32.

--------------. "Revealed Altruism without Independence of Irrelevant Alternatives." Experimental Economics Center Working Paper, Georgia State University, 2009.

Deck, C. A. "A Test of Behavioral and Game Theoretic Models of Play in Exchange and Insurance Environments." *American Economic Review*, 91, 2001, 1546-55.

Engelmann, D., and M. Strobel. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments." *American Economic Review*, 94, 2004, 857-69.

---------------. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments: Reply." *American Economic Review*, 96, 2006, 1918-23.

Fehr, E., A. Klein, and K. M. Schmidt. "Fairness and Contract Design." *Econometrica*, 75, 2007, 121-54.

Fehr, E., M. Naef, and K. M. Schmidt. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments: Comment." *American Economic Review*, 96, 2006, 1912-17.

Fehr, E., and K. M. Schmidt. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics*, 1999, 114, 817-68.

Hicks, J. R., *Value and Capital*. Oxford: Clarendon Press, second edition, 1946.

Hoffman, E., K. McCabe, K. Shachat, and V. L. Smith. "Preferences, Property Rights, and Anonymity in Bargaining Games." *Games and Economic Behavior*, 7, 1994, 346-80.

Jones, S. R. G. "Was There a Hawthorne Effect?" *American Journal of Sociology*, 98, 1992, 451-68.

Levitt, Steven D. and John A. List, "Was There Really a Hawthorne Effect at the Hawthorne Plant? An Analysis of the Original Illumination Experiments." NBER Working Paper No. 15016, 2009.

List, John A., "On the Interpretation of Giving in Dictator Games." *Journal of Political Economy*, 115, 2007, 482-93.

Neilson, W. S. "A Theory of Kindness, Reluctance, and Shame for Social Preferences." *Games and Economic Behavior*, 66, 2009, 394-403.

Samuelson, P. A. *Foundations of Economic Analysis*. Cambridge, MA: Harvard University Press, 1947.

Table 1. Budget sets, Model Predictions, and Data Distributions for Two Treatments

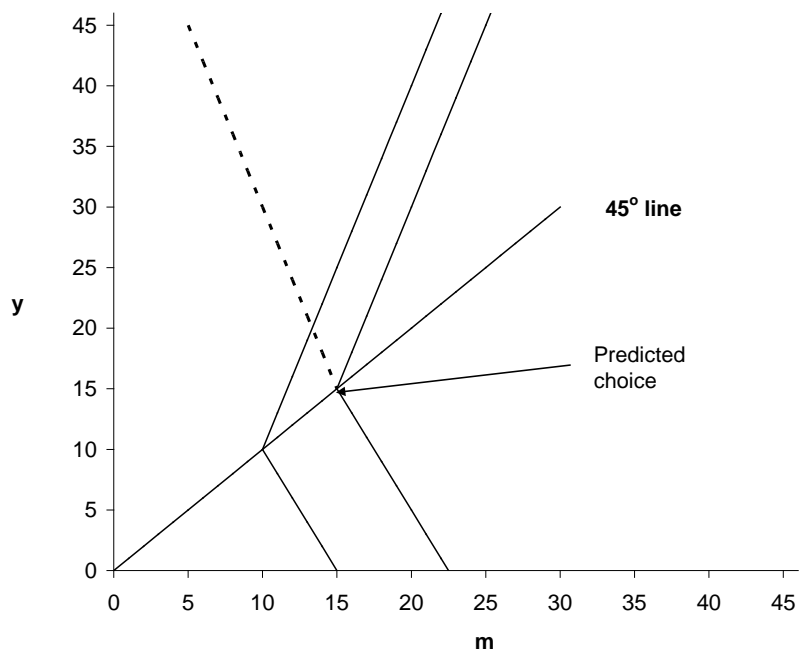| Treatment (nobs)	Budget Sets				Properties of feasible money allocations				Predictions	Observed Choices (percent)
	m	y_1	y_2	y_3	Total Payoff	Own/ Total	Sum of Payoff Differences		Model	
							Favorable	Unfavorable		
2 (33)	15	5	7	38	65	0.23	18	23	CR, EA	15%
	15	5	11	11	42	0.36	18	0	BO, FS	15%
	15	5	20	20	60	0.25	10	10	EA	70%
3 (32)	15	5	20	20	60	0.25	10	10		6%
	15	8	17	20	60	0.25	7	7	FS, EA	88%
	15	9	10	26	60	0.25	11	11	CR	6%

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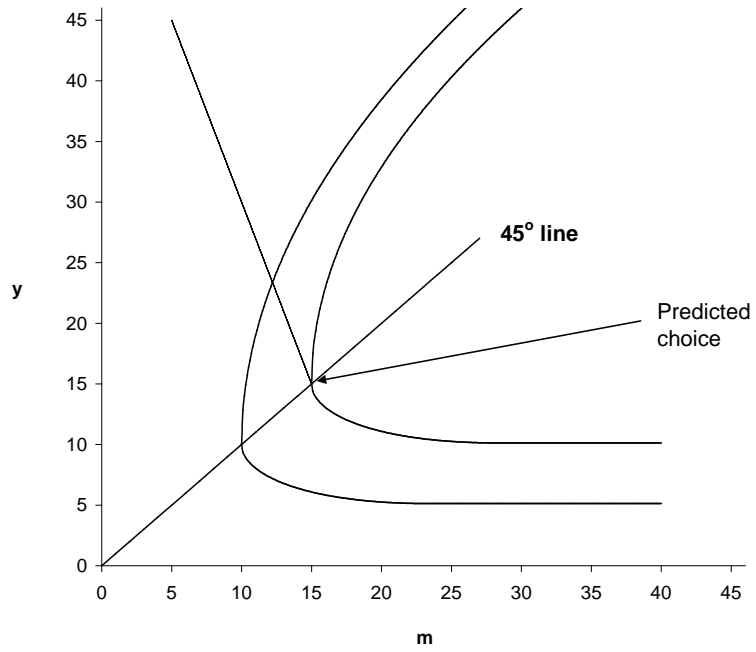($\alpha = 1/2$ and $\beta = 1/3$)

F**igure 2. Budget (dashed) Line and Indifference Curves for the Bolton-Ockenfels Model**
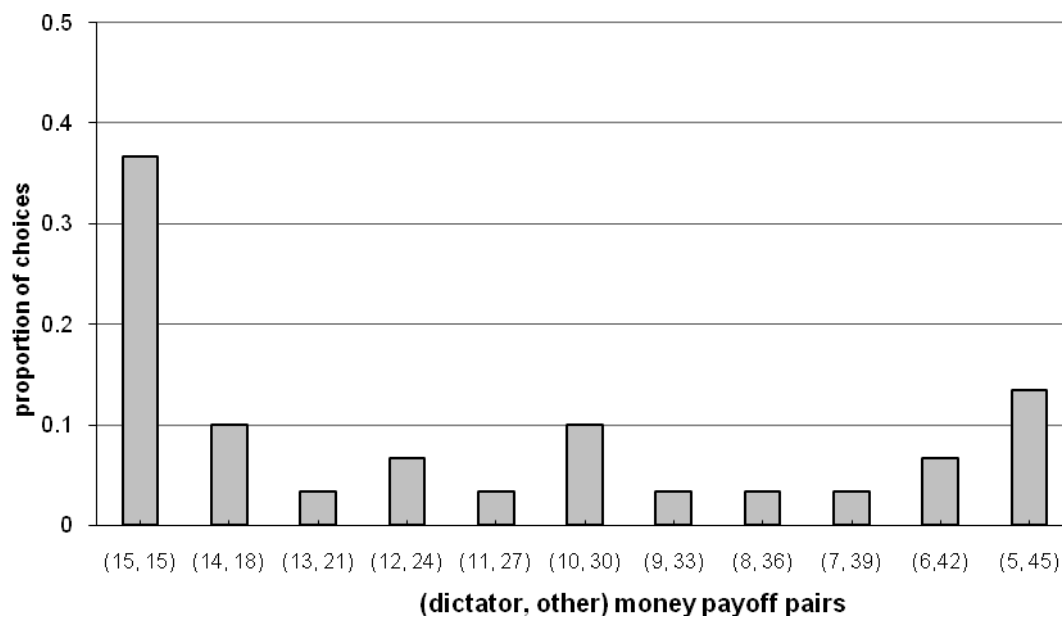
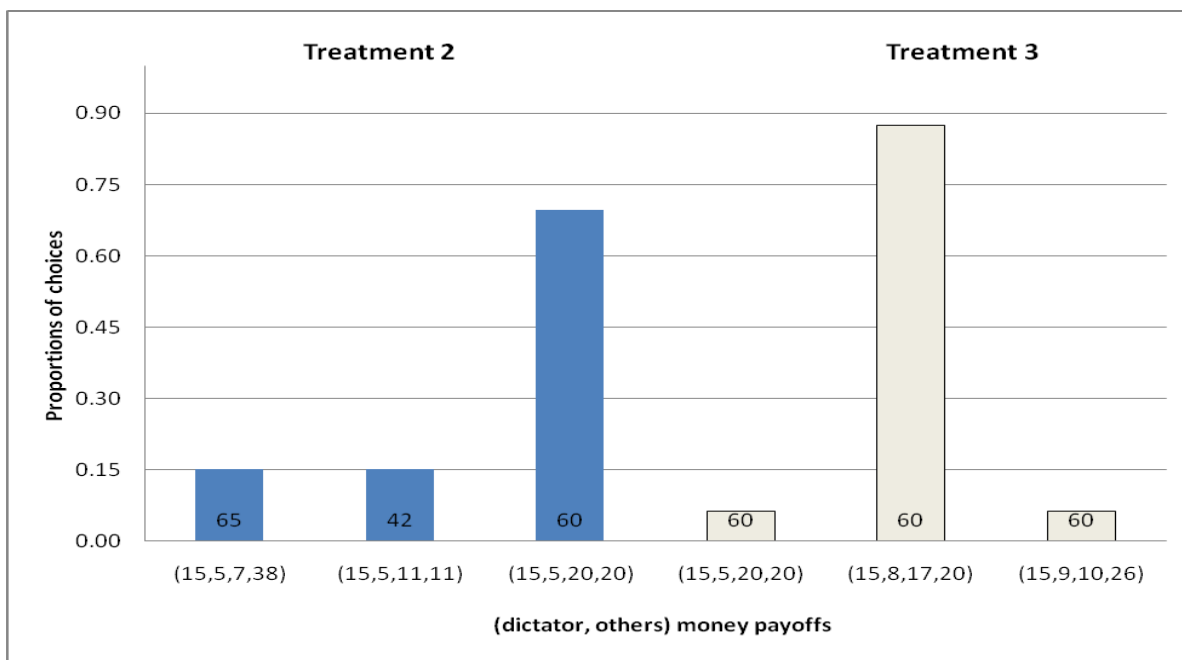**Figure 3. Dictators' Choices in Treatment 1**

Figure 4. Dictators' Choices in Treatments 2 and 3 (figures within bars are total payoffs)

Appendix 1: Predictions of the Models

A.1. The Fehr-Schmidt Model

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

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

where $\beta_i \leq \alpha_i$ and $0 \leq \beta_i < 1$.

The FS predictions. The FS model predicts that $\forall \alpha_i \geq \beta_i > 0$ the chosen set, C in:

1. Treatment 1 is   C = {(15,15)},
2. Treatment 2 is C = {(15,5,11,11)},
3. Treatment 3 is C= {(15,8,17,20)},

Derivations are straightforward. For treatment 1, it can be easily verified that $u_i(15 - s, 15 + 3s) = 15 - (1 + 4\alpha_i)s, \forall s \geq 0$. Note that $u_{is} = -(1 + 4\alpha_i)$ is negative for all $\alpha_i \geq 0$. The last inequality is true by assumptions since $\alpha_i \geq \beta_i > 0$. Therefore, $u_i(15,15) > u_i(15 - s, 15 + 3s), \forall s > 0$; hence (15, 15) is the most preferred feasible allocation.

For treatment 2, one has

$$u_i(15,5,11,11) = 15 - 6\beta_i$$
$$u_i(15,5,20,20) = 15 - 10(\alpha_i + \beta_i)/3 < 15 - 10 \times 2\beta_i/3 < u_i(15,5,11,11)$$
$$u_i(15,5,7,38) = 15 - 6\beta_i - 23\alpha_i/3 < u_i(15,5,11,11)$$

where the inequalities follow from $\alpha_i \geq \beta_i > 0$. For treatment 3, one has

$$u_i(15,5,20,20) = 15 - 10(\alpha_i + \beta_i)/3 < u_i(15,8,17,20)$$
$$u_i(15,8,17,20) = 15 - 7(\alpha_i + \beta_i)/3$$
$$u_i(15,9,10,26) = 15 - 11(\alpha_i + \beta_i)/3 < u_i(15,8,17,20)$$

where the inequalities follow from $\alpha_i + \beta_i > 0$.

A.2. The Bolton-Ockenfels Model

The BO model is based on a "motivation function" of the form,

(A. 2) $\qquad v_i = v_i(x_i, \sigma_i)$

where

(A. 3) $\qquad \begin{aligned} \sigma_i &= x_i / \sum_{j=1}^{n} x_j, \quad \text{if } \sum_j x_j \neq 0 \\ &= 1/n, \qquad \text{if } \sum_j x_j = 0 \end{aligned}$,

where $v(\cdot)$ is (BO, pgs. 171-172) continuous and twice differentiable on (x_i, σ_i), non-decreasing and concave in my income x_i,

(A. 4) $\qquad (v_{i1}(x_i, \sigma_i) \geq 0, \quad and \quad v_{i11}(x_i, \sigma_i) \leq 0)$,

strictly concave in relative income σ_i, and has a partial derivative with respect to relative income with the property

(A. 5) $\qquad v_{i2}(x_i, \sigma_i) = 0, \quad for \quad \sigma_i = 1/n; \quad and \quad v_{i22}(x_i, \sigma_i) < 0$.

Note that from (A.5) one has

(A. 6) $\qquad \begin{aligned} v_{i2}(x_i, \sigma_i) &> 0, \ \text{for } \sigma_i \in [0, 1/n) \\ &< 0, \ \text{for } \sigma_i \in (1/n, 1] \end{aligned}$

The BO predictions. The BO model predicts that the chosen set in:

1. Treatment 1 is $\{(15,15)\}$
2. Treatment 2 is $\{(15,5,11,11)\}$
3. Treatment 3 is $\{(15,5,20,20),(15,8,17,20),(15,9,10,26)\}$

Derivations are straightforward. For treatment 1, one has

$$v_i(15 - s, 15 + 3s) = v(15 - s, \frac{1}{2} - \frac{s}{15 + s})$$

and

$$\frac{dv}{ds} = -v_1(15 - s, \frac{1}{2} - \frac{s}{15 + s}) - v_2(15 - s, \frac{1}{2} - \frac{s}{15 + s}) \frac{15}{(15 + s)^2}.$$

The last expression and statements (A.4) and (A.6) imply that $\dfrac{dv}{ds} < 0, \forall s = 0..10$. Therefore

$v_i(15,15) > v_i(15 - s, 15 + 3s), \forall s = 0..10$. Thus, allocation (15, 15) is the BO most preferred

feasible allocation in treatment 1.   For treatment 2, one has

$$v_i(15,5,11,11) = v(15,5/14) = v(15,1/2 - 1/7)$$
$$v_i(15,5,20,20) = v(15,1/4) = v(15,1/2 - 1/4) < v_i(15,5,11,11)$$
$$v_i(15,5,7,38) = v(15,3/13) = v(15,1/2 - 7/26) < v_i(15,5,11,11)$$

where inequalities follows from (A.6).   For treatment 3, the BO utility of all three allocations is

$v_i(15,1/4)$. Therefore, a BO agent chooses randomly one of the three feasible allocations.

A.3. The Quasi-maximin Model

Let x denote a vector of money payoffs of n agents and x_i denote the payoff of agent i. Charness

and Rabin's (2002) "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 ($\min\limits_{j \in \{1,...,n\}} \{x_j\}$), and the total of all agents' payoffs ($\sum\limits_{j=1}^{n} x_j$). The quasi-maximin

model's utility function is:

(A. 7) $$u_i(x) = (1 - \gamma)x_i + \gamma \left[\delta \min_{j \in \{1, \text{fi}., n\}} \{x_j\} + (1 - \delta) \sum_{j=1}^{n} x_j \right],$$

where

(A. 8) $\gamma \in [0,1]$ and $\delta \in (0,1)$.

The γ parameter measures the relative importance of own money payoff compared to the two

other arguments of the utility function. The δ parameter measures the relative importance of

these other two arguments, the minimum payoff and total payoff (or "efficiency").

The CR predictions. The CR model predicts that the chosen set, C in:

$$
\begin{array}{lll}
 & C = \{(15,15)\}, & \text{if } \gamma(1\text{-}\delta)<1/3; \\
\text{1. Treatment 1 is} & = \{(5,45)\}, & \text{if } \gamma(1\text{-}\delta)>1/3; \\
 & = \{(15\text{-s},15\text{+3s})\,|\,s \in \{0,...,10\}\}, & \text{if } \gamma(1\text{-}\delta) = 1/3.
\end{array}
$$

2. Treatment 2 is $C = \{(15,5,7,38)\}$

3. Treatment 3 is $C = \{15,9,10,26\}$

Derivations are straightforward. For treatment 1 one has

$$u_i(15-s, 15+3s) = 15 - s + \gamma(1-\delta)(15+3s), \forall s > 0,$$

Hence, $du_i / ds = -1 + 3\gamma(1-\delta)$, which is positive if and only if $\gamma(1-\delta) > 1/3$. So,

if $\gamma(1-\delta) > 1/3$ then $u_i(5,45) > u_i(15-s, 15+3s), \forall s > 0$; if $\gamma(1-\delta) < 1/3$ then

$u_i(15,15) > u_i(15-s, 15+3s), \forall s > 0$. Indifference holds for the special case of $\gamma(1-\delta) = 1/3$.

For treatment 2, one has

$u_i(15,5,11,11) = 15 + \gamma(27 - 37\delta)$
$u_i(15,5,20,20) = 15 + \gamma(45 - 55\delta) = u_i(15,5,11,11) + 18\gamma(1-\delta) > u_i(15,5,11,11)$
$u_i(15,5,7,38) = 15 + \gamma(50 - 60\delta) = u_i(15,5,20,20) + 5\gamma(1-\delta) > u_i(15,5,20,20)$

where inequalities follow from (A.8). By transitivity then $(15,5,7,38)$ is the CR most preferred feasible money allocation in treatment 2.

For treatment 3, one has

$u_i(15,5,20,20) = 15 + \gamma(45 - 55\delta)$
$u_i(15,8,17,20) = 15 + \gamma(45 - 52\delta) = u_i(15,5,20,20) + 3\gamma\delta > u_i(15,5,20,20)$
$u_i(15,9,10,26) = 15 + \gamma(45 - 51\delta) = u_i(15,8,17,20) + \gamma\delta > u_i(15,8,17,20)$

where inequalities follow from (A.8). By transitivity then $(15,9,10,26)$ is the CR most preferred feasible money allocation in treatment 3.

A.4. The Egocentric Altruism Model

The egocentric altruism (EA) parametric utility function is:

(A. 9)
$$u_i(x) = \frac{1}{\alpha}\left[x_i^\alpha + \theta \sum_{j \neq i} x_j^\alpha \right], \quad \alpha \in (-\infty, 1) \setminus \{0\},$$

$$= x_i \left(\Pi x_j \right)^\theta, \qquad \alpha = 0.$$

Cox, Friedman, and Gjerstad (2007, appendix A) shows that CES indifference curves for $\alpha \neq 0$ converge point wise as $\alpha \to 0$ to indifference curves for Cobb-Douglas preferences with $\alpha = 0$. The parameter restrictions implied by monotonicity, egocentricity, and convexity are

(A. 10) $\qquad \alpha \in (-\infty, 1), \theta \in [0,1)$.

The EA predictions. The EA model predicts that the chosen set C in:

1. In Treatment 1 is

$$
\begin{aligned}
C &= \{(15,15)\} && \text{if } 0 \le \theta < 1/3; \\
&= \{(15\text{-}s^*, 15+3s^*)\} && \text{if } 1/3 \le \theta < \min\{3^{1-2\alpha}, 1\} \\
&= \{(5,45)\} && \text{otherwise}
\end{aligned}
$$

2. In Treatment 2 is

$$
\begin{aligned}
C &= \{(15,5,7,38)\}, && \text{if } \alpha \ge 0.594, \\
&= \{(15,5,20,20))\}, && \text{if } \alpha < 0.594.
\end{aligned}
$$

3. In Treatment 3 is $\quad C = \{(15,8,17,20)\}$

Derivations are straightforward. The budget set in treatment 1 is given by $\{(15-s, 15+3s) \mid s \in \{0,...,10\}\}$.

The dictator's utility as a function of argument s is:

(A. 11) $\quad U(s) = u_1(15-s, 15+3s), \quad s \in [0,10]$

with $u^1(m, y)$ is

$$
\begin{aligned}
u_i(15-s, 15+3s) &= \frac{1}{\alpha}\left[(15-s)^\alpha + \theta(15+3s)^\alpha\right], && \alpha \in (-\infty, 1)\setminus\{0\}, \\
&= (15-s)(15+3s)^\theta, && \alpha = 0.
\end{aligned}
$$

The slope dy/dm of an indifference curve through $(15\text{-}s, 15+3s)$ is given by

(A. 12) $\quad \dfrac{dy}{dm} = -\dfrac{1}{\theta}\left(\dfrac{15+3s}{15-s}\right)^{1-\alpha}$

The preferred allocation then is $(10-s^*, 10+3s^*)$ where

$$
\begin{aligned}
s^* &= 0 && \text{if } \theta \in [0, 1/3) \\
(A.\,13) \qquad &= 15\frac{(3\theta)^{1/(1-\alpha)} - 1}{(3\theta)^{1/(1-\alpha)} + 3} && \text{if } \theta \in [1/3, \min\{3^{1-2\alpha}, 1\}), \\
&= 10 && \text{otherwise}
\end{aligned}
$$

This follows from solving equation $U'(s) = 0$ for $s > 0$ (or $\dfrac{dy}{dm}(10-s^*, 10+3s^*) = -3$)).

Note that $s^*(\alpha, \theta)$ takes values from 1 to 10 and therefore sending any integer from 1 to 10 which is observed in the data can be explained by this model.

For treatment 2, one has

$$u_i(15,5,11,11) = [15^\alpha + \theta(5^\alpha + 11^\alpha + 11^\alpha)]/\alpha$$

$$u_i(15,5,20,20) = [15^\alpha + \theta(5^\alpha + 20^\alpha + 20^\alpha)]/\alpha > u_i(15,5,11,11)$$

$$u_i(15,5,7,38) = [15^\alpha + \theta(5^\alpha + 7^\alpha + 38^\alpha)]/\alpha$$

if $\alpha = 0$ then

$$u_i(15,5,11,11) = 15(5 \times 11 \times 11)^\theta$$

$$u_i(15,5,20,20) = 15(5 \times 20 \times 20)^\theta > u_i(15,5,11,11)$$

$$u_i(15,5,7,38) = 15(5 \times 7 \times 38)^\theta < u_i(15,5,20,20)$$

Note that $(15,5,11,11)$ is always dominated, so it would never be chosen by an EA dictator. An EA dictator with $\alpha = 0$ will choose $(15,5,20,20)$ in treatment 2. An EA dictator with $\alpha \neq 0$ and $\theta \neq 0$ prefers $(15,5,20,20)$ to $(15,5,7,38)$ if $\alpha < .593$ and she prefers $(15,5,7,38)$ to $(15,5,20,20)$ if $\alpha > .594$.

For treatment 3, it can be verified that for all $0 \neq \alpha < 1$

$$u_i(15,5,20,20) = [15^\alpha + \theta(5^\alpha + 20^\alpha + 20^\alpha)]/\alpha < u_i(15,8,17,20)$$

$$u_i(15,8,17,20) = [15^\alpha + \theta(8^\alpha + 17^\alpha + 20^\alpha)]/\alpha$$

$$u_i(15,9,10,26) = [15^\alpha + \theta(9^\alpha + 10^\alpha + 26^\alpha)]/\alpha < u_i(15,8,17,20)$$

if $\alpha = 0$ then for all $\theta > 0$

$$u_i(15,5,20,20) = 15(5 \times 20 \times 20)^\theta$$

$$u_i(15,8,17,20) = 15(8 \times 17 \times 20)^\theta > u_i(15,5,20,20)$$

$$u_i(15,9,10,26) = 15(9 \times 10 \times 26)^\theta < u_i(15,8,17,20)$$

In treatment 3, the prediction is unique: $(15,8,17,20)$ is the best choice for all EA agents with positive θ.