

Fairness and the Roads Not Taken:
An Experimental Test of Non-Reciprocal Set-Dependence in
Distributive Preferences

Martin Eiliv Sandbu

December 14, 2006

Department of Legal Studies and Business Ethics

Wharton School of Business

University of Pennsylvania

3730 Walnut Street, Philadelphia, PA 19104-6340

Email: sandbu@post.harvard.edu

Phone: +1-215-573 4864

Fax: +1-215-573 2006

This version completed March 2006

The author is grateful for very useful help and comments from James Alt, Christopher Avery, Gary Bolton, Urs Fischbacher, Jerry Green, Macartan Humphreys, Emir Kamenica, David Laibson, Nien-hê Hsieh, Xaq Pitkow, Alvin Roth, Amartya Sen, Jeremy Tobacman, Richard Tuck, seminar participants at Harvard University and Pennsylvania State University, the staff at the Harvard Business School's Computer Lab for Experimental Research, the associate editor of this journal and three anonymous referees. The research reported in this paper was made possible with financial support from the Research Council of Norway and the Center for Basic Research in the Social Sciences at Harvard University.

Abstract

Experimental investigations of preferences for fairness reveal systematic set-dependence in people's allocative choices: Choices over identical options can be reversed by strategically irrelevant changes in the set of possible outcomes of the larger game. Set-dependent choice behavior is often thought to reflect *reciprocity motives*, the desire to reward good intentions and punish bad ones. This paper investigates set-dependence in unilateral decision problems, where reciprocity motives cannot be at work. By examining *truncated dictator games*—dictator games with constrained budgets—I document choice reversals that are systematically related to changes in the available set, yet cannot be explained by the desire to reciprocate intentions.

JEL classification: D01, D64

Keywords: Fairness, set-dependence, reciprocity, dictator games, truncated games, independence of irrelevant alternatives, reference-dependent utility functions, state-dependent utility functions

1 Introduction

The analysis of other-regarding preferences and fairness motivations in behavioral economics can be conveniently divided into two approaches. The first sees fairness as “outcome fairness” and models it with utility functions defined over distributive outcomes, *i.e.* payoff allocations (Fehr and Schmidt 1999, Bolton and Ockenfels 2000, Andreoni and Miller 2002 are important examples). These models depart from the conventional assumptions that economic agents only care about their own payoffs, but retain a “consequentialist” framework with preferences defined over ultimate consequences, without reference to the process that produced the outcomes. The other approach can be called one of “contextual fairness.” In this approach, the same payoff allocation can be judged differently depending on the context in which it arose.¹ The most successful theory of this type is arguably *reciprocity theory* (Rabin 1993, Falk and Fischbacher 2000, Dufwenberg and Kirchsteiger 2004), which assumes that agents are motivated to reward those who are well-intentioned towards them, and hurt those who are ill-intentioned. Reciprocity theory has the appealing feature that it can rationalize some types of *set-dependence*. Set-dependence obtains when choices over the *same* payoff allocations vary systematically with the set of possible outcomes *off* the path of actual play—the allocations that *could* have occurred—even when those unchosen paths are strategically irrelevant (such as in the last stage of a game). In the words of one study: “Not only is the actual allocation producing fairness, but the road to that allocation and the roads not taken along the way are also inputs into the production of fairness” (Andreoni et al. 2002). Set-dependence is clearly anomalous from an outcome-based point of view, as it violates the axiom of Independence of Irrelevant Alternatives (IIA) when preferences are defined over ultimate outcomes. Yet the phenomenon has been documented in a large number of studies (Prasnikar and Roth 1992, Blount 1995, Brandts and Sola 2001, Güth et al. 2001, Andreoni et al. 2002, Falk et al. 2003). I discuss some examples in section 2.

Reciprocity theory rationalizes some forms of set-dependent behavior, because what actions *could* be taken affects which intentions can be inferred from the actions *actually* performed. (In the ultimatum game, for example, an unequal offer from a proposer who *could* have chosen an equal offer may be a sign of unkind or unfair intentions.) The question addressed here is whether allowing for reciprocity in our models is sufficient to accommodate the choice reversals that are observed in experiments. To the extent that set-dependence is caused by reciprocity, it should disappear in unilateral decision problems such as dictator games, where one player is passive. This paper examines empirically if this is the case, by measuring the extent to which set-dependence occurs in an experiment involving *truncated dictator games*—that is, dictator games in which the set of available divisions is continuous but limited. The results reveal that set-dependence is statistically and quantitatively significant even in the absence of reciprocity motives. This suggests

¹Kahneman et al. (1986a,b) provide classic examples of how people judge the fairness of payoff allocations differently depending on the context or frame within which the outcome arises. In the main text, I focus on reciprocity theory, since it gives an explanation of set-dependent behavior, as explained below. Other recent experimental and theoretical studies have investigated other contextual features that might affect distributive preferences, as pointed out by Dufwenberg and Kirchsteiger (2004, footnote 1). Examples include promise-keeping (Charness and Dufwenberg 2004), truth-telling (Brandts and Charness 1999), “status” (Cox et al. forthcoming), gender (Andreoni and Petrie 2005), and property rights (Gächter and Riedl 2005).

that findings of set-dependence must be interpreted with caution, and that a more general theory of set-dependent behavior may be needed.

In section 3 I present a simple set-dependent utility function that can be calibrated to experimental data to measure whether set-dependence occurs. Section 4 describes the dictator game experiment and the econometric approach and reports the results. In section 5 I revisit the well-known study by Charness and Rabin (2002, hereafter CR), and show that the same type of set-dependence shows up in their data even after controlling for their measure of reciprocity motives. The final section concludes.

2 Intuitions: The relevance of irrelevant alternatives

A series of studies have documented that people's choices over payoff allocations are influenced by the available outcomes on the *unchosen* paths of the game trees, even in the last stage of the games. One simple but compelling example is given by Falk et al. (2003), who studied four different "mini-ultimatum games." In each game, player 1 could offer player 2 one of two allocations, which player 2 could then accept or reject. If he rejected, both players would get a payoff of zero. In each of the four games, one of the allocations was (8,2) in player 1's favor. The alternative allocation in each respective game was (10,0), (8,2), (5,5) and (2,8). Outcome-based theories predict that rejection rates for the (8,2)-allocation should be the same across all games. Instead, the researchers found substantially higher rejection rates in the game where the alternative offer was (5,5) than in the other games. A similar result was found by Güth et al. (2001) who used three mini-ultimatum games, where one offer was always (17,3) in player 1's favor. The alternative offer was (11,9), (10,10) or (9,11) in the three games. When player 2 had to choose between (17,3) and (0,0), the offer on the unchosen path mattered considerably for the choice, with rejection rates for the (17,3)-allocation being the highest when the alternative offer was the egalitarian (10,10).² In Brandts and Sola's (2001) study of five mini-ultimatum games, variations in the offer on the unchosen path again influenced actual choices between alternatives that were identical across the five games. Finally, Andreoni et al. (2002), who ran an experiment with a much larger menu of first and second moves, concluded that "if models of fairness are to predict the observed difference across... games, then they must allow the evaluation of actions to depend on the actions not chosen."³

If preferences are only defined over distributive outcomes, the choices observed in the studies just cited violate IIA. This evidence therefore lends credence to non-outcome-based theories in general, and to reciprocity theory in particular. A comprehensive survey of behavioral game theory (Camerer 2003) assesses the issue as follows:

²Güth et al. (2001) note that the effect disappears when they use the strategy method instead of sequential play; indeed a conclusion of their study is the sensitivity of experimental behavior to methodological changes. (I thank an anonymous referee for making this observation.) Falk et al. (2003), however, obtain their result using the strategy method, showing that set-dependence does not only occur in sequentially played games.

³In addition to experimental studies, there are several theoretical investigations of set-dependence. Sen (1997) provides a formal treatment of "menu-dependence" as well as "chooser-dependence." A philosophical discussion on how the available set of outcomes may matter is Elster (1983). In bargaining theory, the Kalai-Smorodinsky solution is set-dependent (Kalai and Smorodinsky 1975).

Most data support either inequality-aversion or the reciprocal approaches [as against “pure altruism”-approaches]. . . Inequality-aversion theories are very promising but predict a kind of separability—utilities of terminal-node allocations are independent of how those allocations arose, and of allocations from unchosen alternatives—that is psychologically suspect and violated in several experiments. (p. 113)

The reciprocity-based view is surely more psychologically correct because players do care about the intentions of other players and unchosen paths. (p. 116)

Set-dependence (“non-separability” in Camerer’s terminology) is often taken as evidence in favor of reciprocity theory. This is natural in light of its many appealing features. Reciprocity theory makes intuitive psychological sense, lends itself well to formal game-theoretic modelling, and can account for such behavioral regularities as rejections in ultimatum games (Roth 1995) and the willingness to punish people who do not contribute to public goods (Fehr and Gächter 2002).

An unfortunate aspect of reciprocity theory’s success, however, has been a relative penury of experiments in which set-dependence is not observationally equivalent with reciprocal behavior. (One exception is CR’s data, which I investigate in section 5.) This is unfortunate because of the obvious fact that set-dependent behavior *by itself* reveals nothing about intentions. It is theoretically possible for set-dependence to be caused by other psychological factors that have nothing to do with assessing the intentions of others. Put differently, there could be both “reciprocal set-dependence” and “non-reciprocal set-dependence,” but we do not have much data with which to determine how much to attribute to one or the other. It may be that all or most set-dependent behavior is indeed caused by reciprocity motives, in which case this theoretical worry can be dismissed. To do so, however, we must first empirically investigate the hypothesis that agents exhibit non-reciprocal set-dependence, which is what the present paper does by analyzing data that allow reciprocal and non-reciprocal set-dependence to be identified separately. This investigation demonstrates that, contrary to what is sometimes implicitly assumed, experimental subjects exhibit set-dependent behavior that cannot be attributed to reciprocity motives.

3 Formalizations

This study follows earlier work on other-regarding preferences in fitting a utility function to experimental data (Fehr and Schmidt 1999, Andreoni and Miller 2002, Charness and Rabin 2002). One advantage of this method is that it can easily capture preference heterogeneity (by allowing shifts in the estimates depending on the type of agent), which has been found to be an important feature of other-regarding preferences (Fehr and Schmidt 1999, Andreoni and Miller 2002). I use a simple but versatile diagnostic utility function for the purpose of measuring set-dependence. I assume that people have *state-dependent preferences* over payoff vectors $\mathbf{x} \in \mathbb{R}_+^n$ (for n individuals), that is, their ordering

of payoff vectors can be different depending on the state of the world. Formally, this is usually represented as a family of utility functions over payoff vectors $U_s(\mathbf{x})$ indexed by the state of the world s . Since we are investigating set-dependence, the relevant description of the state of the world is the *available payoff set* $\mathbf{X} \subseteq \mathbb{R}_+^n$, which I define as the set of payoff allocations that appear at some end node of the game in question. The state-dependent utility function $U_s(\mathbf{x})$ is more precisely expressed as a *set-dependent* utility function $U_{\mathbf{X}}(\mathbf{x})$. In the two-person case, the utility function is defined over nonnegative payoff pairs $(x_{self}, x_{other}) \in \mathbb{R}_+^2$ and indexed by the set of available payoff pairs $\mathbf{X} \in \mathbb{R}_+^2$ from which an allocation is selected. We shall use the constant elasticity of substitution (CES) functional form:⁴

$$U_{\mathbf{X}}(x_{self}, x_{other}) = \left[x_{self}^{\rho} + \omega(\mathbf{X}) x_{other}^{\rho} \right]^{\frac{1}{\rho}}. \quad (1)$$

The “altruism weight” $\omega(\mathbf{X})$ is the marginal rate of substitution (MRS) between own and other’s payoff at equality. A perfectly selfish person would have $\omega(\mathbf{X}) = 0$. Here the relative weight on the other person’s payoff is not fixed as in outcome-based models, but can vary with the set of available payoff allocations. The curvature parameter ρ measures the degree of inequality-aversion.

For the purposes of simply distinguishing non-reciprocal set-dependence from reciprocity motives, the functional form of $\omega(\mathbf{X})$ is of little import. Evidence for *any* kind of set-dependence in situations where reciprocity motives cannot be a factor—such as the dictator games studied in the next section—would establish the difference. To operationalize the model, however, a functional form must be specified. To do so, start by noting one way in which the set of available outcomes could make a difference. As Güth et al. (2001) point out, some allocations—for example equal splits—are more *salient* for fairness than others. Thus the availability of some allocations rather than others might “trigger” fairness motives more or less strongly. In a manner analogous to the theory of reference-dependence in consumer choice (Tversky and Kahneman 1991, Bateman et al. 1997, Munro and Sugden 2003), one can model this as the altruism weight depending on a *reference allocation* $\mathbf{r} \in \mathbb{R}_+^n$ (with $\mathbf{r} \equiv (r_{self}, r_{other})$ in the two-person case), where the reference point in turn is a function of the available set: $\mathbf{r} = \mathcal{R}(\mathbf{X})$. That is, $\omega(\mathbf{X}) \equiv A(\mathcal{R}(\mathbf{X}))$, where $A(\mathbf{r})$ is a function of vectors in \mathbb{R}_+^n . The reference point $\mathcal{R}(\mathbf{X})$ should be interpreted as the agent’s notion of what each individual “could have received” given the set of available outcomes in the game, and the assumption is that these reference payoffs affect the weight she places on the actual payoffs.⁵

⁴One shortcoming of the CES function is that it does not allow for a negative expression inside the brackets when its exponent $1/\rho$ is non-integer. From a theoretical perspective, this should be allowed, since the weight on other people’s payoff could be negative. A different normalization would be $U_{\mathbf{X}}(x_{self}, x_{other}) = \text{sign}(\rho) \left[x_{self}^{\rho} + A(\mathbf{X}) x_{other}^{\rho} \right]$, which is ordinally equivalent to the CES function but leaves preferences well-defined over the entire positive quadrant in payoff space. The estimating equation introduced below (equation 4) would not be affected by using this other normalization of the utility function.

⁵Earlier uses of reference points in formal representations of other-regarding preferences include Rabin (1993) (who uses an “equitable payoff” in the specification of the kindness of the players), Falk and Fischbacher (2000) (“reference payoff”) and Cox et al. (forthcoming). Brandts and Sola (2001) use the same term, although they do not model it. One appealing feature of formalizing fairness preferences as reference-dependent is that it allows the agent to recognize some allocation as a fair reference without necessarily *choosing* exactly that allocation. (She chooses the $\mathbf{x}^* \in \mathbf{X}$ that maximizes the utility function, and there is no requirement that $\mathbf{r} = \mathbf{x}^*$.) This seems psychologically plausible: Fairness influences people’s

Thus we write

$$U_{\mathbf{X}}(x_{self}, x_{other}) = \left[x_{self}^\rho + A(\mathcal{R}(\mathbf{X})) x_{other}^\rho \right]^{\frac{1}{\rho}}. \quad (2)$$

In the econometric specification below, I make the altruism weight A a simple linear function of the ratio of the reference payoffs:

$$A(\mathbf{r}) \equiv A(r_{self}, r_{other}) = a + cr, \text{ with } r \equiv \frac{r_{other}}{r_{self}} = \frac{\mathcal{R}_{other}(\mathbf{X})}{\mathcal{R}_{self}(\mathbf{X})}. \quad (3)$$

I specify $\mathcal{R}(\mathbf{X})$ as the most egalitarian Pareto-efficient allocation in the available set, subject to not putting the decision-maker behind.⁶ This specification strikes a reasonable compromise between the salience of the equal split and a natural concern to avoid Pareto-inferior outcomes and excessive sacrifices by decision-makers. While people are often willing to sacrifice to benefit others, the evidence shows that they are much less willing to sacrifice when they are behind than when they are ahead (see CR for an example). Several of the studies mentioned above, moreover, find that responders in mini-ultimatum games reject unequal offers less frequently when the alternative for the proposer is to offer an allocation very unfavorable to herself. Both of these findings can be interpreted as an aversion to “unreasonable” sacrifices. The idea that fairness does not require an unreasonable sacrifice is also used by Falk and Fischbacher (2000) in their model of what decision-makers consider fair.

4 Measuring set-dependence

4.1 Modifying Andreoni and Miller (2002)

This section reports a dictator game experiment constructed to identify set-dependent behavior that cannot be attributed to reciprocity motives. In a dictator game, the recipient takes no action and can display no intentions; consequently, reciprocity motives cannot be the cause of any variation in outcomes across treatments. The design was based on an experiment by Andreoni and Miller (2002, hereafter AM). Like them, I offer subjects a unilateral choice of how to preferences somewhat, but does not determine them completely—self-interest remains a powerful motive.

⁶If all available payoff vectors in \mathbf{X} are unequal in the other person’s favour, the reference payoff vector is the most egalitarian Pareto-efficient allocation in \mathbf{X} . Formally, denote by $\mathbf{X} \equiv \left\{ \mathbf{x} \in \mathbb{R}_+^2 \mid \mathbf{x} \text{ is available} \right\}$ the set of available payoff distributions, where a distribution is available if it is achievable through some combination of actions by the agents involved. Denote by $\mathbf{X}^{advantaged}$ the subset of \mathbf{X} whose elements give the decision-maker at least as much as the other person, that is, $\mathbf{X}^{advantaged} \equiv \{ \mathbf{x} \in \mathbf{X} \mid x_{other} \leq x_{self} \}$. The fair reference allocation is the function $\mathcal{R}(\mathbf{X}) \equiv (\mathcal{R}_{self}(\mathbf{X}), \mathcal{R}_{other}(\mathbf{X}))$ characterized by:

- If $\mathbf{X}^{advantaged}$ is non-empty, $\mathcal{R}(\mathbf{X})$ is the payoff vector \mathbf{r} in $\mathbf{X}^{advantaged}$ that satisfies:
 - (a) $r_{other} \geq x'_{other}$, for all $\mathbf{x}' \in \mathbf{X}^{advantaged}$, and
 - (b) $r_{self} \geq x''_{self}$, for all $\mathbf{x}'' \in \mathbf{X}^{advantaged}$ satisfying $x''_{other} = r_{other}$.
- If $\mathbf{X}^{advantaged}$ is empty, $\mathcal{R}(\mathbf{X})$ is the payoff vector \mathbf{r} in \mathbf{X} that satisfies:
 - (c) $r_{self} \geq x'_{self}$ for all $\mathbf{x}' \in \mathbf{X}$, and
 - (d) $r_{other} \geq x''_{other}$ for all $\mathbf{x}'' \in \mathbf{X}$ satisfying $x''_{self} = r_{self}$.

divide a series of budgets between themselves and another participant. I test for non-reciprocal set-dependence by adding an important feature to the AM design, which is to truncate the budget sets so that the decision problems have different reference points.

Each subject had to make a total of 36 decisions.⁷ The opportunity cost of giving could take on one of four values. In 27 decision problems the dictator was faced with a budget of 120 “tokens” that could be worth 20 cents each to the dictator and 10 each to the recipient (nine decisions); 10 cents to both players (nine decisions); or 10 cents to the dictator and 20 to the recipients (nine decisions). In the remaining nine decisions the budget was 140 tokens that were worth 10 cents each to the dictator and 30 cents each to the recipient. Thus the smallest possible pie across all decisions was \$12 and the largest possible \$42 (only achievable if the opportunity cost was 1/3 and the divider let the recipient have everything). This variation facilitated the estimation of the utility function’s curvature, a measure of inequality-aversion. The four opportunity costs were matched with nine truncation schemes, to create 36 unique decision problems. The nine types of truncation are schematically illustrated in figure 1. For each opportunity cost, three decisions had budgets truncated from above, three decisions had budgets truncated from below, and three decisions had symmetric budgets. The asymmetrically truncated budgets ensured that either the dictator (with the truncations illustrated in row 1) or the recipient (with the truncations illustrated in row 2) could get at most 1/3, 1/2, or 2/3 of the aggregate dollar earnings in the division, respectively. The symmetric budgets (illustrated in row 3) included one unrestricted budget (the divider could give any amount between zero and everything), one which required that both participants should get at least 1/3 of the total dollar earnings, thus limiting the maximum amount of inequality, and one which required that one of the two participants (the divider’s choice) should get at least 2/3 of the total dollar earnings, thus making an equal division unavailable. Figure 2 illustrates this last truncation with all four corresponding budget sets, one for each exchange rate. Across all truncations, the reference payoff ratio r could be 1/2, 1 or 2; the average across all decisions was $\bar{r} = 1$.

The instructions given to the subjects specified which kinds of divisions were allowed in each case. I followed AM’s terminology of “passing tokens” to avoid connotations of giving and charity, and to make my results more directly comparable with theirs. Two typical questions read:

“Divide 120 tokens. Each token is worth 10 cents to you, 20 cents to the other person. You must pass at least 20 tokens and at most 60 tokens.”

“Divide 120 tokens. Each token is worth 10 cents to you, 10 cents to the other person. You must pass less than 40 tokens or more than 80 tokens.”

⁷This design means that most of the coefficient estimates are driven by within-subject rather than between-subject variation. This might be a source of concern if subjects exaggerate the variation in their behavior when they know there is a large number of slightly different choices to be made. However, set-dependence has also been found to be a robust phenomenon in between-subject studies, such as those cited at the beginning of section 2. The findings in this section are further supported by the results in section 5, where I calibrate the utility function to an entirely different data set with mostly between-subject variation. Still, future research should investigate to what extent set-dependence occurs in truncated dictator games like the ones presented here when there is only between-subject variation. (I thank an anonymous referee for raising this concern.)

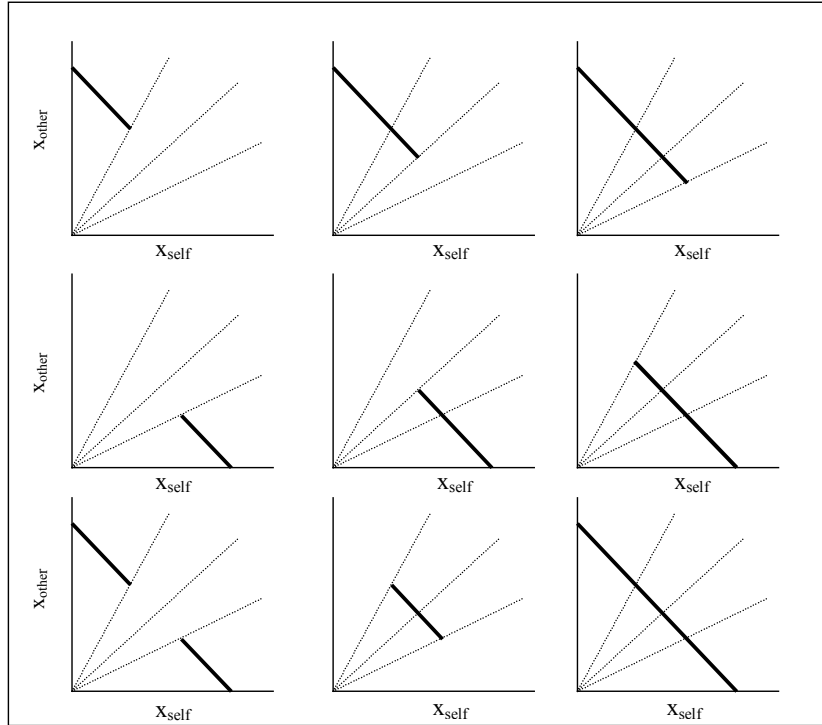


Figure 1: Truncated budget sets in the dictator game experiment

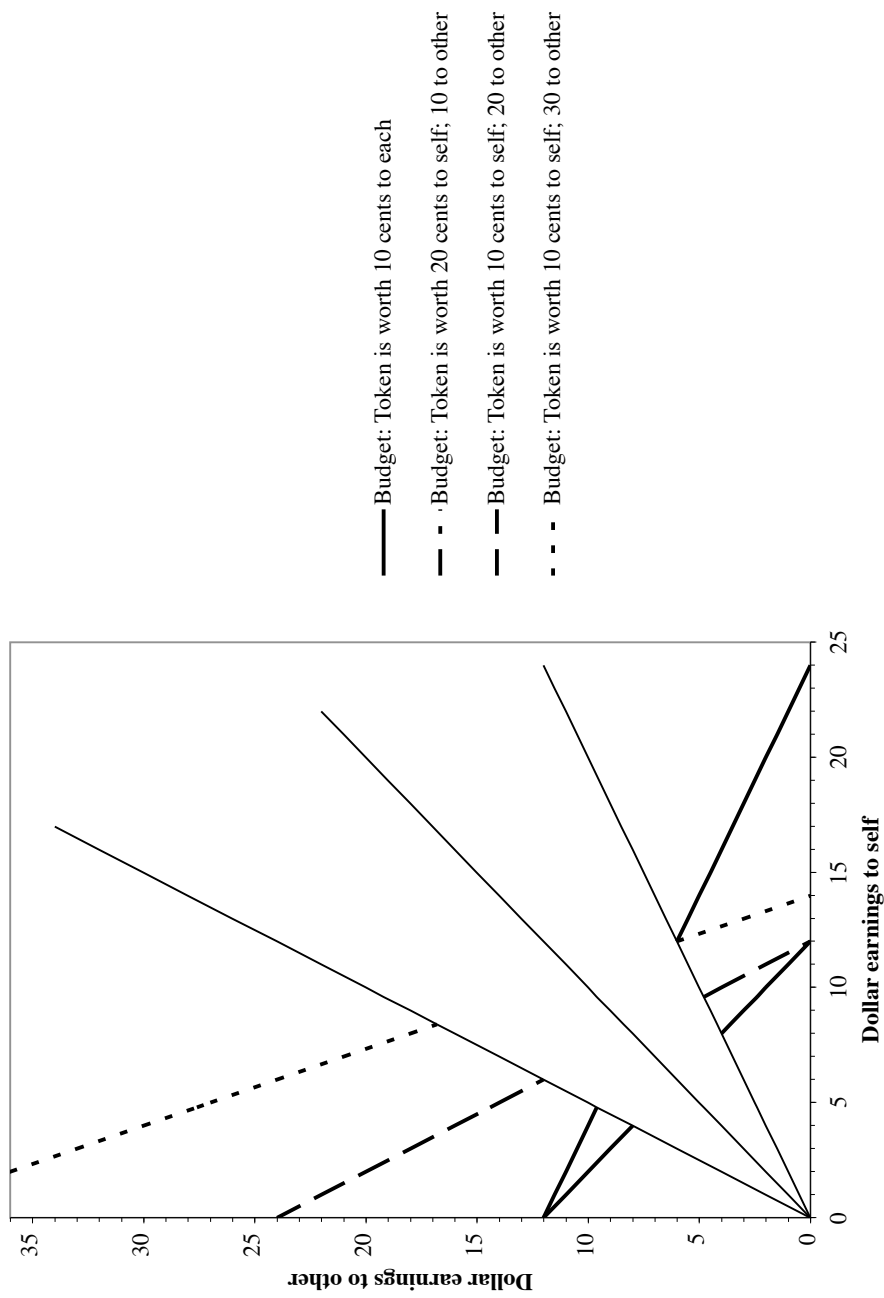


Figure 2: One truncation scheme with four budget sets

For each decision, the subjects entered the number of tokens they passed to the other person, after which the computer showed the monetary distribution implied by the decision and gave the subjects a chance to change their minds. The subjects had an unlimited opportunity to go back and change the number of tokens before finalizing each decision. This procedure minimized the risk that confusion and mathematical complexity prevented the subjects from acting on their preferences. Each subject was presented with the same 36 decision problems, but in different (random) orders. This was common knowledge. The subjects were told that one of their decisions would be picked at random at the end of the session.⁸ They made money from what they had decided to keep in that decision, and from being a recipient from somebody else’s division. The recipients were also paid a show-up fee of \$10. The experiment was programmed and conducted using the zTree software for economic experiments (Fischbacher 1999). It was carried out at the Harvard Business School’s Computer Lab for Experimental Research in the course of three sessions in the spring of 2003, and included a total of 96 subjects. The detailed instructions are available from the author.

4.2 Estimation method

As stated above, a convenient and easily interpreted formulation of the altruism weight in the two-person case is a simple linear function of the *ratio* of the two reference payoffs: $A(\mathbf{r}) = a + cr$, with $r \equiv r_{other}/r_{self}$. In everything that follows, the reference payoffs are defined by the function of the available payoff set posited above, so $\mathbf{r} \equiv (r_{self}, r_{other}) = (\mathcal{R}_{self}(\mathbf{X}), \mathcal{R}_{other}(\mathbf{X}))$. Then equations (2) and (3) give us three parameters to estimate: ρ , a and c , each of which characterizes an intuitive aspect of the agent’s preferences. The curvature parameter ρ (and the elasticity of substitution $(\rho - 1)^{-1}$) measures the sensitivity to inequality. The parameters a and c measure the set-independent and the set-dependent components of the decision-maker’s altruism, respectively. If $c = 0$, the weight on the other person’s payoff is exactly a , which can then be seen as a measure of “pure” altruism.

As mentioned above, experimental investigations have found considerable heterogeneity in other-regarding preferences (Fehr and Schmidt 1999, Andreoni and Miller 2002). Following AM, I define three “ideal-types” of preferences. I label individuals with $a = c = 0$ *strongly selfish*; those with $\rho = -\infty$ *strongly egalitarian*; and those with $\rho = 1$ *strongly utilitarian* or *perfect substitute types*. Those individuals that do not conform exactly to one of these ideal types can be sorted as “weak” versions of the three strong types depending on which type is closest to their actual choices (measured by the Euclidian distance in payoff space). The utility function can then be calibrated to the full sample of “weak” types, allowing for parameter shifts across types to capture preference heterogeneity. Specifically,

⁸Earlier readers of this paper have expressed worries about this way of rewarding the subject. Since the recipient cannot in any way affect the payoffs of the divider he or she is matched with, there is of course no direct reciprocity here. But the fact that everyone is both a divider and a recipient (but matched with different people in the two cases) might create *indirect* reciprocity motives. That is, a subject in her recipient role might expect to be treated generously or stingily by her divider, and will “indirectly reciprocate” in the divider role. But it should be noted that even if there are indirect reciprocity motives, they do not vary across the various decisions the subjects have to make. A statistically significant effect of variations in the reference point, therefore, cannot be explained by indirect reciprocity motives.

the following estimating equation is derived from the first-order condition of the utility function (equation 2):

$$\frac{x_{other,ij}}{x_{self,ij} + x_{other,ij}} = \frac{A_{ij}^{\frac{1}{1-\rho_i}}}{p_j^{\frac{1}{1-\rho_i}} + A_{ij}^{\frac{1}{1-\rho_i}}} + \varepsilon_{ij}, \quad (4)$$

with

$$A_{ij} = a + b_E WEAKEGAL_i + b_{PS} WEAKPSUB_i + cr_j \quad (5)$$

$$\text{and } \rho_i = g + g_E WEAKEGAL_i + g_{PS} WEAKPSUB_i \quad (6)$$

where i indexes the subject and j indexes the decision problem. The dependent variable in equation (4) is the share of the monetary pie allocated by subject i to the recipient in decision problem j . Subject i 's altruism weight in decision problem j is measured by A_{ij} , and the curvature of her utility function by ρ_i . The altruism weight and the curvature can differ across subject types, with the shifts estimated by the coefficients on indicator variables WEAKEGAL and WEAKPSUB for weak egalitarians and weak perfect substitute types. The opportunity cost of giving in game j is p_j , and ε_{ij} is a normally distributed observation disturbance. I use maximum likelihood to estimate the equation as a tobit model, which treats corner choices in the truncated budgets as censored observations. This prevents the coefficient c from picking up effects that simply reflect the removal of preferred allocations from the choice set. The estimates reported below are adjusted for clustering within subjects.

4.3 Results

Out of the 96 subjects, 33 subjects (35%) conform perfectly to one of the three ideal-types (32 are strongly selfish, and one is strongly utilitarian). Their behavior is exactly captured with the relevant outcome-based utility function, so these subjects clearly exhibit no set-dependence. The majority of the remaining 63 subjects are weakly selfish (45 subjects or 71%), a large minority are weakly egalitarian (17 or 27%), and only one subject is weakly utilitarian.⁹ With 36 decisions made by each person, this gives 2268 observations to which I calibrate the utility function.

Before presenting the coefficient estimates, it is instructive to observe a particularly strong example of set-dependence, which can be found in one subject's choices in the three symmetric budgets when the price of giving was 1/3. The three choice situations are depicted in figure 3. In these situations, the subject was asked to divide 140 tokens with a recipient, each token being worth 10 cents to the divider and 30 cents to the recipient. When the division was unconstrained (this implies $r = 1$), the subject chose a \$18-\$8 division in the recipient's favour. When the number of tokens

⁹The proportion of weak types (65% of all subjects) is higher than the roughly 50% that AM found in their experiment. This is not surprising, since my subjects faced many more choice situations and thus had more opportunity to make "imperfect" choices. Among the weak subjects, my pool had somewhat more selfish types and fewer perfect substitute-types than AM.

passed was constrained to an egalitarian range (between 20 and 56, implying $r = 1$), the subject chose the allocation closest to the unconstrained choice, which led to a \$16.80-\$8.40 division in the recipient's favour. In the third decision, the available allocations were limited to the *inegalitarian* ranges (the number of tokens passed had to be no more than 20 or no fewer than 56 tokens, so $r = 1/2$), which include the allocation the subject had preferred in the unconstrained choice. Yet from this budget, the subject opted for a \$12-\$6 division in his or her *own* favour! If preferences were only defined over distributive outcomes, these choices would clearly violate IIA. The only rationalization consistent with outcome-based models would be that the subject were indifferent between all the available allocations. Such complete indifference is implausible; moreover, it is unlikely that an indifferent subject would have chosen corner solutions in two out of three cases. This choice pattern, therefore, is not well explained by outcome-based theories. Nor can it be explained by reciprocity theory, since no action is chosen and no intention is displayed by the recipient.

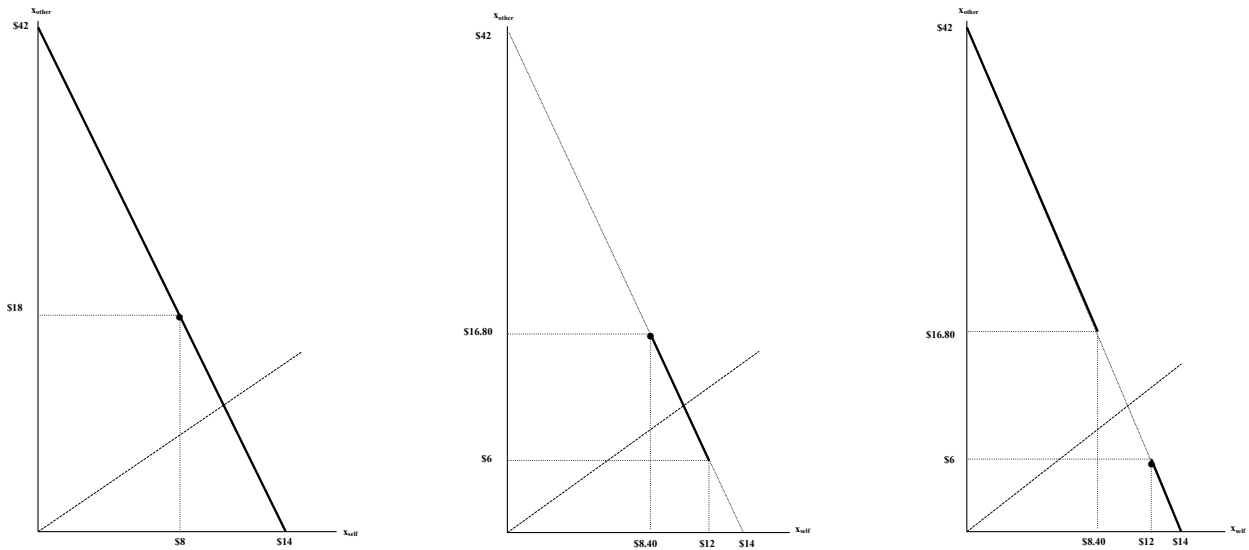


Figure 3: Example of set-dependence in truncated dictator games

To see if these results hold more generally, I turn to the estimated parameters of the calibrated utility function, reported in table 1. The first two columns show the results from fitting equation (4) to the dictator game data set, with and without the restriction $c = 0$, respectively. If outcome-based theories are correct, the coefficient c in the unconstrained model should be zero. A nonzero coefficient estimate constitutes evidence of non-reciprocal set-dependence, since reciprocity motives are absent in these individual decision problems.

The parameters in the set-independent model (SI, column 1) are unsurprising (and consistent with AM's results). At egalitarian payoff allocations ($x_{other}/x_{self} = 1$) the weakly selfish subjects value a marginal dollar to the other person as equivalent to a marginal 16 cents to themselves, while the weak utilitarians value a marginal dollar to the

other at 77 cents to themselves ($\$0.16 + \$0.61 = \$0.77$). The weak egalitarians are not altruistic when both subjects get the same amount, but the strong curvature ($\hat{\rho} = 0.62 - 5.05 = -4.43$, implying an elasticity of substitution of -0.18) means that the marginal rate of substitution changes dramatically as soon as one of the subjects gets less than the other. The weakly selfish and weakly utilitarian types, in contrast, have relatively flat indifference curves ($\hat{\rho} = 0.62$ and $\hat{\rho} = 0.62 + 0.01 = .063$, respectively).

Column 2 reports the estimates for the set-dependent model (SD), which show that the specific example above is indeed representative of a general pattern of set-dependence. The estimated coefficient \hat{c} on the reference payoff ratio r is 0.13 and strongly significant—both the t -statistic and a likelihood ratio test comparing models SD and SI reject the restriction $c = 0$ at any conventional significance level. To assess the magnitude of this effect, we can calculate the estimated altruism weight $\hat{A} = \hat{a} + \hat{c}r$ at various values of r .¹⁰ At the sample average level of $r = 1$, the estimates imply that weakly selfish dividers value a marginal dollar to the recipient at 15 cents to themselves (at egalitarian allocations). When the reference payoff ratio is at its sample maximum ($r = 2$), however, the estimates imply an altruism weight of 0.28 for weakly selfish subjects. That is almost 70% more than the altruism weight estimated in the set-independent model (0.16), and more than *three times* the implied 0.08 altruism weight implied by a reference payoff ratio at the sample minimum ($r = .5$). In comparison, changing the *actual* payoff ratio from $x_{other}/x_{self} = 2$ to $x_{other}/x_{self} = .5$ leads to less than a doubling of the marginal rate of substitution for weakly selfish subjects. These are economically significant magnitudes, and the comparison shows that at sample averages, set-dependence is at least as important as “outcome fairness” in determining the weight placed on the recipient’s payoff. In unit-independent terms, we may compare the elasticity of the MRS with respect to the reference payoff ratio and the elasticity of the MRS with respect to the actual payoff ratio (the inverse of the elasticity of substitution). These elasticities are reported in the table, which shows that the MRS with respect to the reference payoff ratio, calculated at sample averages, is about twice (the absolute value of) the inverse of the elasticity of substitution for weakly selfish subjects.

In sum, the data reveal a strong pattern of set-dependent behavior that cannot be attributed to reciprocity. This set-dependence is a statistically and quantitatively important phenomenon. The estimates imply that changes in the available set of outcomes can substantially affect the weight a decision-maker places on other people’s payoffs. This effect is comparable in magnitude—if not larger—to the effect of changes in distributive shares on allocative decision (“outcome fairness”).

¹⁰The text focuses on the weakly selfish individuals. For the other two weak types, the complete altruism weight is adjusted by the relevant intercept shift, so that $\hat{A} = \hat{a} + \hat{b}_E + \hat{c}r$ for the weak egalitarians, and $\hat{A} = \hat{a} + \hat{b}_{PS} + \hat{c}r$ for the weak perfect substitute type.

5 Revisiting Charness and Rabin (2002)

5.1 Data and comparative framework

The presence of set-dependent behavior in unilateral decision problems means that not all set-dependence is due to reciprocity motives. Can we separate reciprocal from non-reciprocal set-dependence—and assess their relative magnitude—in games that might include both? A clean separation may be impossible as long as we cannot measure intentions directly, but we can give a first approximation of the two effects by adding non-reciprocal set-dependence (in the form of the variable r used above) to earlier empirical assessments of reciprocity. This section analyzes the data set from CR, who report the results of 27 two-player games designed to compare different models of other-regarding preferences. Seven of these were unilateral decision problems where one player chose between two allocations of payoffs. The remaining twenty were sequential-move games, where the first mover decided between ending the game and implementing a certain allocation, versus letting the second player decide between two other allocations. Together, the games give 903 observations of binary choices.

CR fit utility functions to these data to show that reciprocity theories have better explanatory power than purely outcome-based theories.¹¹ They measure reciprocity motives by including in the altruism weight of the utility function an indicator variable q which is set to 1 if the other player has “misbehaved” and zero otherwise. The average value of q in the data set is 0.36. Misbehavior is defined as occurring when the first mover makes a choice by which the second mover must end up with a lower payoff than if the first mover had chosen differently (so q is always zero in the dictatorial decision problems). Conditional on the second mover actually having to make a decision, therefore, q is actually a function of the available set of outcomes, just like my fair reference allocation $\mathcal{R}(\mathbf{X})$ (although it is of course defined differently). For consistency, we shall write $q = \mathcal{Q}(\mathbf{X})$ as a single-valued function of the set of available payoffs, where $\mathcal{Q}(\mathbf{X})$ captures CR’s definition of misbehavior. This illustrates the important fact that CR do not actually model intentions explicitly, although the misbehavior variable is intuitively linked to intentionality. This limits the kind of separation of reciprocal from non-reciprocal set-dependence that will be feasible—theoretically speaking, CR’s q might not reflect intentions. The comparison of the respective effects of r and q should therefore be seen only as a rough approximation.

With this caveat, I proceed to calibrating the set-dependent constant elasticity of substitution (CES) function from before, now enhanced with the reciprocity parameter:

$$U_{\mathbf{X}}(x_{self}, x_{other}) = \left[x_{self}^{\rho} + A(\mathcal{R}(\mathbf{X}), \mathcal{Q}(\mathbf{X})) x_{other}^{\rho} \right]^{\frac{1}{\rho}} \quad (7)$$

¹¹CR use a piecewise linear utility function. The results using the CES function given in equation 7 are fully consistent with CR’s own results. Additionally, the CES formulation fits CR’s data much better than the piecewise linear function, judging by the likelihood values and by CR’s own precision measure γ (defined below).

now with

$$A(\mathbf{r}, q) \equiv A((r_{self}, r_{other}), q) = a + cr + \theta q, \text{ with } r \equiv \frac{r_{other}}{r_{self}} \text{ as before.} \quad (8)$$

In the absence of reciprocity (imposing the restriction $\theta = 0$), this utility function is the same as that defined by equations 2 and 3. Like CR I calibrate the utility function with a logit model estimated by maximum likelihood, using the following likelihood function:

$$\Pr(x'_{self}, x'_{other}) = \frac{e^{\gamma U_{\mathbf{X}}(x'_{self}, x'_{other})}}{e^{\gamma U_{\mathbf{X}}(x'_{self}, x'_{other})} + e^{\gamma U_{\mathbf{X}}(x''_{self}, x''_{other})}} \quad (9)$$

where (x'_{self}, x'_{other}) and $(x''_{self}, x''_{other})$ are the two payoff vectors the decision-maker is choosing between (a dictatorial decision or a second move, as described above). The parameter γ measures the precision of the model. The parameters are estimated with robust standard errors.

5.2 Set-dependence versus reciprocity: Results

I estimate four models with the CR data. Columns 3 and 4 report the results for the set-independent (SI) and set-dependent (SD) models without reciprocity, the same models that were calibrated to the dictator game data (they impose the restriction $\theta = 0$, and SI also imposes $c = 0$). The last two models include reciprocity effects, as defined by CR. A “set-independent reciprocal” model SI-R ($c = 0$, but θ is freely estimated) is reported in column 5, and the last column reports the results from the full model, which allows both non-reciprocal set-dependence and reciprocity effects (model SD-R).

The results for model SI are mostly interesting for comparing the estimates of altruism across data sets, which I postpone until the next subsection. The results for model SD (column 4) show that the point estimate of \hat{c} is 0.72 and highly significant. In other words, CR’s subjects also exhibit a strong sensitivity to changes in the reference payoff ratio. The estimated “pure altruism” coefficient \hat{a} is *negative* (at -0.49), meaning that at low reference payoff ratios, the decision-maker would be willing to forego payoffs to himself in order to reduce the payoff of the other player. At sample averages, however, the subjects put positive weight on the other player’s payoff. The reference payoff ratio in this sample varies from 0.33 to 1, with an average value of 0.91. The corresponding value of $\hat{A}(r)$ —the estimated MRS at equal allocations—varies from -0.25 to $+0.23$, with an average value of 0.16. The elasticity of the MRS with respect to the reference payoff ratio is 4.06, an order of magnitude greater than the elasticity of the MRS with respect to the actual payoff ratio (the inverse of the elasticity of substitution). Clearly behavior is motivated by more than selfishness and concerns for outcome fairness only.

Can this set-dependence be attributed to the effect of CR’s “misbehavior” variable? Column 5 corroborates CR’s finding that misbehavior has a very strong effect on choices. The estimate $\hat{\theta}$ is statistically and quantitatively significant

at -0.45 . Together with the estimate of pure altruism \hat{a} at 0.25 , we see that reciprocity motives can turn generosity into spite. At equal allocations, the average subject values a marginal dollar to the other player at 25 cents to himself, but not if the other player has misbehaved, in which case he is willing to *give up* 20 cents ($0.25 - 0.45$) to reduce her payoff by a marginal dollar. When I include both variables in the model (column 6), however, we see that misbehavior is not the source of the set-dependence I found in model SD (column 4). When θ and c are estimated freely, they both remain significant, and the magnitude of the coefficients hardly changes from the restricted models SI-R (column 5) and SD (column 4).¹² (Nor does allowing for set-dependence of either type affect the estimate of inequality-aversion, as measured by the curvature parameter.)

The main result of this section is that non-reciprocal set-dependence and reciprocity have separate, significant effects on choices. We finish by reporting what the data can say about the relative importance of the two effects. The increase in the log-likelihood from including reciprocity in the model (9.2 from model SI to model SI-R; 9.7 from model SD to model SD-R) is considerably larger than the increase from including set-dependence (5.5 from model SI to model SD; 6.0 from model SI-R to model SD-R), even though the latter is also highly significant. The same pattern can be seen in the change of the estimated precision parameter $\hat{\gamma}$. Since these statistics are measures of the models' goodness of fit, this suggests that reciprocity explains more of the *variation* in behavior in this data set than does non-reciprocal set-dependence. The data is less helpful for comparing the *magnitudes* of the two effects, which for purposes of comparison are best measured by the elasticity of the MRS with respect to r and q . This is because by construction of the model, these elasticities are different for different values of the variables (the estimated MRS elasticity with respect to r is $\hat{c}r/\hat{A}$ and the estimated MRS elasticity with respect to q is $\hat{\theta}q/\hat{A}$). Which is larger will therefore depend on where they are calculated. For illustrative purposes, I calculate the elasticities at sample averages ($\bar{r} = 0.91$ and $\bar{q} = 0.36$). At those values, non-reciprocal set-dependence turns out to have a much larger effect than reciprocity, with an elasticity of 5.10 as against -1.86 (a Wald test that their absolute values are equal is rejected with $p = 0.056$). This comparison cannot, however, be generalized.

5.3 Cross-study comparison

One important difference between the two data sets is that CR's subjects make very few decisions (sometimes only one). This makes it impossible to allow for preference heterogeneity and assign subjects to the same categories as in section 4. Comparisons of coefficient estimates across the two data set must therefore be between the full sample from the CR data and subsamples from the dictator game data. I focus here on the weakly selfish individuals from section 4, this being the single largest group and accounting for close to half of the subjects in that experiment. (Moreover,

¹²A closer examination of CR's 27 games shows why the estimates of the two effects are unchanged across the different models. The majority of the games either have no misbehavior or they have a reference payoff ratio of 1. The non-reciprocal set-dependence results are driven by six games (one of which is a dictator-type game) that have $r < 1$, and the reciprocity results are caused by ten games in which the misbehavior dummy is set to 1. It turns out that the two subsets of games only have one game in common, making the estimators almost orthogonal to each other.

the weak egalitarians have almost the same altruism weight, and the weak utilitarian has almost the same curvature, as the weakly selfish subjects.) Allowing for this difference, start by noting the remarkable similarity of the estimated altruism weight in the outcome-based models across the two data sets (model SI; columns 1 and 3). CR's subjects value a marginal dollar to the other person at 18 cents to themselves (at equal allocations); that is just 2 cents more than the weakly selfish subjects in the dictator game experiments. They are also similarly inequality-averse, as is evinced by a curvature parameter of 0.51 for CR's subjects against 0.62 for my weakly selfish subjects.¹³

If we compare the coefficients in the set-dependent models without reciprocity (model SD; columns 2 and 4), we again find very similar estimates, with an altruism weight of 0.15 for my weakly selfish subjects versus 0.16 for CR's subjects (at sample averages of r), and a curvature parameter of 0.56 versus 0.50, respectively. These estimates do not change much when I include CR's reciprocity parameter. The main difference between the two data sets is found in the estimate of \hat{c} , the coefficient on the reference payoff ratio, which is much larger in the CR data. The difference could be due to the discreteness of the choice sets in the CR games, which requires discontinuous choices, unlike the continuous budget sets in the dictator games. Notwithstanding the difference in magnitude, the sign and statistical significance of the estimated effect of the reference point in the CR data corroborate the findings in section 4.

6 Conclusion

Many studies of distributive choices have now documented behavior that would amount to systematic violations of IIA if purely outcome-based models were correct. These findings pose the question of how we should explain such set-dependence. Since it is difficult to measure intentions directly, set-dependent behavior is often taken as evidence for reciprocity motives. This paper has examined whether set-dependence occurs even in situations where reciprocity motives cannot possibly be the cause. By calibrating a utility function to data from truncated dictator games, I showed that set-dependence is a statistically and quantitatively significant phenomenon even when the other player cannot display any intentions. I fitted the same model to a different data set in which Charness and Rabin (2002) had found evidence of reciprocal behavior, and showed that there, too, subjects exhibit set-dependence that cannot be reduced to CR's measure of reciprocity. The immediate lesson to draw from these experimental results is that researchers cannot take reciprocity as the only explanation of set-dependence. This points to two important matters for future research. Firstly, one should devise better experimental tests to isolate what reciprocity theory is really about, namely intentions. One approach is to adapt the technique presented here to suitably chosen unilateral choice problems, and use the resulting estimate of *non*-reciprocal set-dependence to correct measures of reciprocity motives. Secondly, the results here have documented set-dependence in unilateral choice problems, but we have not attempted to explain the

¹³The minor differences are consistent with the conjecture that the CR subjects contain all the "weak" types of subjects in roughly the same proportion as the dictator game data. Note that the CR coefficient estimates are comparable to weighted averages of the type-specific estimates from column 1.

phenomenon. A task for the research program on other-regarding preferences should be to develop a coherent theory that can predict and explain such non-reciprocal set-dependence in a way consistent with what we already understand about pro-social behavior.

A Appendix: Table of estimation results

Data	Dictator games		Charness and Rabin (2002)			
	1	2	3	4	5	6
Model	SI	SD	SI	SD	SI-R	SD-R
Non-reciprocal set-dependence included?	No	Yes	No	Yes	No	Yes
Reciprocity included?	No	No	No	No	Yes	Yes
Altruism weight (A):						
Constant (a)	0.16 (5.63)	0.02 (0.30)	0.18 (3.96)	-0.49 (2.85)	0.25 (7.71)	-0.18 (1.51)
Set-dependence (c)		0.13 (2.97)		0.72 (3.85)		0.46 (3.70)
CR's misbehavior (θ)					-0.45 (5.74)	-0.42 (5.70)
$A(\bar{r})$ (columns 1–2) or $A(\bar{r}, 0)$ (column 3–6):	0.16	0.15	0.18	0.16	0.25	0.23
WEAKEGAL (b_E)	0.00 (0.02)	0.02 (0.15)				
WEAKPSUB (b_{PS})	0.61 (21.16)	0.63 (20.69)				
Curvature (ρ):						
Constant (g)	0.62 (5.25)	0.56 (5.45)	0.51 (8.41)	0.50 (8.70)	0.45 (9.06)	0.46 (10.37)
WEAKEGAL (g_E)	-5.05 (1.81)	-5.01 (1.92)				
WEAKPSUB (g_{PS})	0.01 (0.11)	0.06 (0.58)				
Implied elasticities: (at sample averages)						
Elasticity of MRS with respect to r		0.87		4.06		5.10
Elasticity of MRS with respect to q					-1.83	-1.86
Inverse of elasticity of substitution (weakly selfish subjects in columns 1–2)	-0.38	-0.44	-0.49	-0.50	-0.55	-0.54
Model statistics						
N	2268	2268	903	903	903	903
Precision: σ (DG data), γ (CR data)	0.238	0.232	0.175	0.179	0.370	0.396
Log-likelihood	-670.5	-658.0	-558.6	-553.1	-549.4	-543.4
Likelihood-ratio test:						
Comparison model		SI		SI	SI	SI-R
χ^2		24.98		11.0	18.5	12.0
p-value		0.000		0.001	0.000	0.001

Robust t-statistics in parentheses. All elasticities calculated at sample averages. σ is the standard deviation of the observation disturbance (columns 1–2); γ is the precision parameter in the logit model (columns 3–6).

Table 1: Estimation results

References

- Andreoni, J., Brown, P. M., Vesterlund, L., 2002. What makes an allocation fair? some experimental evidence. *Games and Economic Behavior* 40, 1–24.
- Andreoni, J., Miller, J., 2002. Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica* 70 (2), 737–753.
- Andreoni, J., Petrie, R., 2005. Beauty, gender and stereotypes: Evidence from laboratory experiments.
- Bateman, I., Munro, A., Rhodes, B., Starmer, C., Sugden, R., 1997. A test of the theory of reference-dependent preferences. *The Quarterly Journal of Economics* 112, 479–505.
- Blount, S., 1995. When social outcomes aren't fair: The effect of causal attribution on preferences. *Organizational Behavior and Human Decision Processes* 63, 131–144.
- Bolton, G. E., Ockenfels, A., 2000. Erc: A theory of equity, reciprocity, and competition. *The American Economic Review* 90 (1), 166–193.
- Brandts, J., Charness, G., 1999. Retribution in a cheap-talk experiment.
- Brandts, J., Sola, C., 2001. Reference points and negative reciprocity in simple sequential games. *Games and Economic Behavior* 36 (2), 138.
- Camerer, C. F., 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. The Roundtable Series in Behavioral Economics. Princeton University Press and Russell Sage Foundation, New York and Princeton.
- Charness, G., Dufwenberg, M., 2004. Promises and partnership.
- Charness, G., Rabin, M., 2002. Understanding social preferences with simple tests. *The Quarterly Journal of Economics* 117, 817–869.
- Cox, J. C., Friedman, D., Gjerstad, S., forthcoming. A tractable model of reciprocity and fairness. *Games and Economic Behavior*.
- Dufwenberg, M., Kirchsteiger, G., 2004. A theory of sequential reciprocity. *Games and Economic Behavior* 47 (2), 268–298.
- Elster, J., 1983. *Sour Grapes: Studies in the subversion of rationality*. Maison des Sciences de l'Homme and Cambridge University Press, Cambridge, UK and Paris.
- Falk, A., Fehr, E., Fischbacher, U., 2003. On the nature of fair behavior. *Economic Inquiry* 41 (1), 20–26.
- Falk, A., Fischbacher, U., 2000. A theory of reciprocity. *Games and Economic Behavior* 54 (2), 293–315.
- Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. *Nature* 415 (10 January 2002), 137–140.
- Fehr, E., Schmidt, K. M., 1999. A theory of fairness, competition and cooperation. *The Quarterly Journal of Economics* 114, 817–868.
- Fischbacher, U., 1999. z-tree: Zurich toolbox for readymade economic experiments. experimenter's manual.
- Gächter, S., Riedl, A., 2005. Moral property rights in bargaining with infeasible claims. *Management Science* 51 (2), 249–263.
- Güth, W., Huck, S., Müller, W., 2001. The relevance of equal splits in ultimatum games. *Games and Economic Behavior* 37, 161–169.
- Kahneman, D., Knetsch, J. L., Thaler, R. H., 1986a. Fairness and the assumptions of economics. *Journal of Business* 59 (4), S285–S300.

- Kahneman, D., Knetsch, J. L., Thaler, R. H., 1986b. Fairness as a constraint on profit seeking: Entitlements in the market. *The American Economic Review* 76 (4), 728–41.
- Kalai, E., Smorodinsky, M., 1975. Other solutions to nash's bargaining problem. *Econometrica* 43, 513–18.
- Munro, A., Sugden, R., 2003. On the theory of reference-dependent preferences. *Journal of Economic Behavior and Organization* 50 (4), 407–428.
- Prasnikar, V., Roth, A. E., 1992. Considerations of fairness and strategy: Experimental data from sequential games. *The Quarterly Journal of Economics* 107, 865–888.
- Rabin, M., 1993. Incorporating fairness into game theory and economics. *The American Economic Review* 83 (5), 1281–1302.
- Roth, A. E., 1995. Bargaining experiments. In: Kagel, J., Roth, A. E. (Eds.), *Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ, pp. 253–348.
- Sen, A., 1997. Maximization and the act of choice. *Econometrica* 65 (4), 745–779.
- Tversky, A., Kahneman, D., 1991. Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics* 106 (4), 1039–1061.