1-1-2006

REVEALED ALTRUISM

James Cox
Georgia State University

Daniel Friedman
University of California

VJOLLCA SADIRAJ
Georgia State University

REVEALED ALTRUISM

BY JAMES C. COX, DANIEL FRIEDMAN, AND VJOLLCA SADIRAJ

ABSTRACT. This paper develops a nonparametric theory of preferences over one's own and others' monetary payoffs. We introduce "more altruistic than" (**MAT**), a partial ordering over preferences, and interpret it with known parametric models. We also introduce and illustrate "more generous than" (**MGT**), a partial ordering over opportunity sets. Several recent studies focus on two player extensive form games of complete information in which the first mover (FM) chooses a more or less generous opportunity set for the second mover (SM). Here reciprocity can be formalized as the assertion that an **MGT** choice by the FM will elicit **MAT** preferences in the SM. A further assertion is that the effect on preferences is stronger for acts of commission by FM than for acts of omission. We state and prove propositions on the observable consequences of these assertions. Finally, empirical support for the propositions is found in existing data from investment games and from Stackelberg games and in new data from Stackelberg mini-games.

J. C. Cox, Department of Economics, Andrew Young School of Policy Studies, Georgia State University, P.O. Box 3992, Atlanta, GA 30302-3992, USA. E-mail: jccox@gsu.edu.

D. Friedman, Professor of Economics, 417 Building E2, University of California, Santa Cruz, CA 95064, USA. E-mail: dan@ucsc.edu.

V. Sadiraj, Department of Economics, Andrew Young School of Policy Studies, Georgia State University, P.O. Box 3992, Atlanta, GA 30302-3992, USA. E-mail: vsadiraj@gsu.edu.

KEYWORDS: Classical Preferences, Social Preferences, Convexity, Reciprocity, Experiments.

1. Introduction

What are the contents of preferences? People surely care about their own material well-being, e.g., as proxied by income. In some contexts people also may care about others' well-being. Abstract theory and common sense have long recognized that possibility but until recently it has been neglected in applied work. Evidence from the laboratory and field (as surveyed in Fehr and Gächter (2000), for example) has begun to persuade economists to develop specific models of how and when a person's preferences depend on others' material payoffs (Sobel, 2005).

Andreoni and Miller (2002) report "dictator" experiments in which a human subject decides on an allocation for himself and for some anonymous other subject while facing a linear budget constraint. Their analysis confirms consistency with the generalized axiom of revealed preference (GARP) for a large majority of subjects. They conclude that altruism can be modeled as utility maximizing behavior.

In this paper we take three further steps down the same path. First, we create non-linear opportunity sets. Such sets allow the subject to reveal more about the tradeoff between her own and another's income, e.g., whether her indifference curves have positive or negative slope, and whether the curvature is zero or negative. Second, we give another subject an initial move that can be more or less generous. This allows us to distinguish conditional altruism—positive and negative reciprocity—from unconditional altruism. It also allows us to clarify the observable consequences of convex preferences and of reciprocal preferences. Third, we distinguish active from passive initial moves; i.e., we distinguish among acts of omission, acts of commission, and absence of opportunity to act, and examine their impacts on reciprocity.

We begin in Section 2 by developing representations of preferences over own and others' income, and formalize the idea that one preference ordering is "more altruistic than" (**MAT**) another. We include the possibility of negative regard for the other's income; in this case **MAT** really means "less malevolent than." Special cases include the main parametric models of other regarding preferences that have appeared in the literature.

[0]For helpful comments, we thank Steve Gjerstad, Stephen Leider, Stefan Traub, and Joel Sobel as well as participants in the International Meeting of the Economic Science Association (ESA) 2004, the North American Regional ESA Meeting 2004, and at Economics Department seminars at UCSC, Harvard and University College London.

Section 3 introduces opportunities and choices, and illustrates concepts with several two player games of complete information from the recent literature. Section 4 formalizes the idea that one opportunity set is more generous than (**MGT**) another, and then uses it to formalize reciprocity. Axiom **R** asserts that more generous choices by the first mover induce more altruistic preferences in the second mover. An interpretation urged in a previous paper (Cox, Friedman, and Gjerstad, forthcoming) is that preferences are emotional state-dependent, and the first mover's generosity induces a more benevolent (or less malevolent) emotional state in the second mover. Axiom **S** asserts that the reciprocity effect is stronger following an act of commission (upsetting the status-quo) than following an act of omission (upholding the status-quo), and that the effect is weaker when the first mover is unable to alter the status quo.

Section 5 reports some general theoretical propositions on the consequences of convex preferences and Axioms **R** and **S**. To illustrate the empirical content, we derive testable predictions for the well-known investment game, which features a complete **MGT** ordering of linear opportunity sets. We also derive testable predictions for Stackelberg duopoly games. These games are especially useful because a smaller output by the Stackelberg Leader induces a more generous opportunity set for the Follower, and because the opportunity sets have a parabolic shape that enables the Follower to reveal a wide range of positive and negative tradeoffs between own income and Leader's income. Some key predictions involve a variant game, called the Stackelberg mini-game, in which the Leader has only two alternative output choices, one of which is clearly more generous than the other.

The last three sections report tests of the predictions on existing investment game data, on existing Stackelberg data, and on new Stackeberg mini-game data. Within the limitations of the data, the test results are consistent with predictions. All formal proofs and other mathematical details are collected in Appendix A. Instructions to subjects in the mini-Stackelberg game appear in Appendix B.

2. Preferences

Let $Y = (Y_1, Y_2, ..., Y_N) \in \mathfrak{R}^N_+$ represent the payoff vector in a game that pays each of $N \geq 2$ players a non-negative income. Admissible preferences for each player i are smooth and convex orderings on the positive orthant $\Re^N_+$ that are strictly increasing in own income Y_i. The set of all admissible preferences is denoted $\mathfrak{P}$. Any particular preference $\mathcal{P} \in \mathfrak{P}$ can be represented by a smooth utility function $u : \Re^N_+ \to \Re$ with positive i^{th} partial derivative $\frac{\partial u}{\partial Y_i} = u_{Y_i} > 0$. The other first partial derivatives are zero for standard selfish preferences, but we allow for the possibility that they are positive in some regions (where the agent is "benevolent") and negative in others (where she is "malevolent").

We shall focus on two-player extensive form games of complete information, and to streamline notation we shall denote own ("my") income by $Y_i = m$ and the other player's ("your") income by $Y_{-i} = y$. Thus preferences are defined on the positive quadrant $\Re^2_+ = \{(m, y) : m, y \geq 0\}$. The player's marginal rate of substitution $\mathbf{MRS}(m, y) = u_m/u_y$ is, of course, the negative of the slope of the indifference curve through the given point. Unfortunately, the **MRS** is not well defined at points where the agent is selfish, and diverges to $+\infty$ and back from $-\infty$ when we pass from slight benevolence to slight malevolence. Therefore it is convenient to work with willingness to pay, $\mathbf{WTP} = 1/\mathbf{MRS}$, the amount of own income the agent is willing to give up in order to increase the other agent's income by a unit. **WTP** moves from slightly positive through zero to slightly negative when the agent goes from slight benevolence to slight malevolence.

What sort of factors might affect **WTP**? A prime candidate is relative income, as measured for non-zero allocations in $\Re^2_+$ by the ratio of other's income to own income, $d = y/m$. It is easily shown (see Appendix A.1) that only relative income d matters for homothetic preferences, i.e., **WTP** is constant along each ray $R_d = \{(t, td) : t > 0\} \subset \Re^2_+$ when preferences are homothetic.

It is intuitive that **WTP** decreases in d, that is, I'm willing to pay less to increase your income when I'm relatively poor. The intuition is formalized in the convexity assumption imposed earlier. Recall that preferences are convex if each upper contour set (i.e., the set of allocations preferred to any given allocation) is a convex subset of $\Re^2_+$. A quantitative measure of convexity is provided by curvature of the indifference curves. In terms of the first and second partial derivatives of a utility function u representing the preferences, curvature is

$$\mathbf{K} \ = \frac{u_{mm}u_y^2 - 2u_{my}u_m u_y + u_{yy}u_m^2}{(u_m^2 + u_y^2)^{3/2}} \tag{2.1}$$

The convexity assumption is that $\mathbf{K} \ \leq 0$ at every point $\Re^2_+$.[1] Negative **K** reflects decreasing **MRS** and increasing **WTP** as a benevolent player's own income

[1]This statement is true since we have chosen to work with outward-pointing normal vectors, that is the normal vector points outside the upper contour set. When the curve lies on one side of the tangent line and the normal vector on the other side then the sign of the curvature is negative.

increases along an indifference curve. Of course $\mathbf{K} = 0$ in a region where indifference curves are straight lines, and more negative $\mathbf{K}$ means that the $\mathbf{WTP}$ changes more quickly with changes in relative income d along an indifference curve. See Appendix A.2 for a derivation of the formula, and alternate expressions.

Note that both $\mathbf{WTP}$ and $\mathbf{K}$ are intrinsic for preferences. That is, if we choose another utility function $v = h \circ u$ to represent the same preferences (so $h'(t) > 0$ $\forall t \in Range(u)$), then using v in the computations for $\mathbf{WTP}$ and $\mathbf{K}$ gives us the same values that we get using u.

We are now prepared to formalize the idea that one preference ordering on $\Re^2_+$ is more altruistic than another. Two different preference orderings $\mathcal{A}, \mathcal{B} \in \mathfrak{P}$ over income allocation vectors might represent the preferences of two different players, or might represent the preferences of the same player in two different emotional states (Cox, Friedman, and Gjerstad, forthcoming).

Definition 1. For a given domain $\mathfrak{D} \subset \Re^2_+$ we say that $\mathcal{A}$ $\mathbf{MAT}$ $\mathcal{B}$ on $\mathfrak{D}$ if $\mathbf{WTP}_{\mathcal{A}}(m, y) \geq \mathbf{WTP}_{\mathcal{B}}(m, y)$, for all $(m, y) \in \mathfrak{D}$.

The idea is straightforward. In the benevolence case (where utility is monotone increasing in y) more altruistic than ($\mathbf{MAT}$) means that $\mathcal{A}$ has shallower indifference curves than $\mathcal{B}$ in (m, y) space, so $\mathcal{A}$ indicates a willingness to pay more m than $\mathcal{B}$ for a unit increase in y. In the malevolence case, $\mathbf{WTP}$ is less negative for $\mathcal{A}$, so it indicates a lesser willingness to pay for a unit decrease in y.

Of course, $\mathbf{MAT}$ is a partial ordering on $\mathfrak{P}$ (see Appendix A.3), not a total ordering for nontrivial domains $\mathfrak{D}$ that contain more than a single point.[2] When preferences are homothetic then it suffices to check for $\mathbf{MAT}$ on a thin subset of $\mathfrak{D}$, typically a single indifference curve. When no particular domain $\mathfrak{D}$ is indicated, the $\mathbf{MAT}$ ordering is understood to refer to the entire positive orthant $\mathfrak{D} = \Re^2_+$.

Example 2.1. Linear Inequality-averse Preferences (for $N = 2$ only; Fehr and Schmidt, 1999). Let preferences $\mathcal{J} = \mathcal{A}, \mathcal{B}$ be represented by $u_{\mathcal{J}}(m, y) = (1 + \theta_{\mathcal{J}})m - \theta_{\mathcal{J}} y$, where

$$\begin{aligned} \theta_{\mathcal{J}} &= \alpha_{\mathcal{J}}, \ \texttt{if } m < y \\ &= -\beta_{\mathcal{J}}, \ \texttt{if } m \geq y, \end{aligned}$$

with $\beta_{\mathcal{J}} \leq \alpha_{\mathcal{J}}$ and $0 < \beta_{\mathcal{J}} < 1$. Straightforwardly, $\mathcal{A}$ $\mathbf{MAT}$ $\mathcal{B}$ iff $\theta_{\mathcal{A}} \leq \theta_{\mathcal{B}}$.

[2]The single crossing property imposes the same partial ordering in a different context, but it is usually restricted to one-dimensional families of preferences.

Example 2.2. Nonlinear Inequality-averse Preferences (for $N = 2$, Bolton and Ockenfels, 2000). Let preferences $\mathcal{J} = \mathcal{A}, \mathcal{B}$ be represented by $u_{\mathcal{J}}(m, y) = \nu_{\mathcal{J}}(m, \sigma)$, where

$$\begin{aligned} \sigma &= m/(m+y), \ \texttt{if} \ m+y>0 \\ &= 1/2, \ \texttt{if} \ m+y=0 \end{aligned}$$

It can be easily verified that $\mathcal{A}$ **MAT** $\mathcal{B}$ iff $\nu_{\mathcal{A}1}/\nu_{\mathcal{A}2} \leq \nu_{\mathcal{B}1}/\nu_{\mathcal{B}2}$.

Example 2.3. Quasi-maximin Preferences (for $N = 2$, Charness and Rabin, 2002). Let preferences $\mathcal{J} = \mathcal{A}, \mathcal{B}$ be represented by

$$\begin{aligned} u_{\mathcal{J}}(m, y) &= m + \gamma_{\mathcal{J}}(1-\delta_{\mathcal{J}})y, \ \texttt{if} \ m<y, \\ &= (1-\delta_{\mathcal{J}}\gamma_{\mathcal{J}})m + \gamma_{\mathcal{J}}y, \ \texttt{if} \ m \geq y \end{aligned}$$

and $\gamma_{\mathcal{J}} \in [0, 1]$, $\delta_{\mathcal{J}} \in (0, 1)$. It is straightforward (although a bit tedious) to verify that $\mathcal{A}$ **MAT** $\mathcal{B}$ iff

$$\gamma_{\mathcal{A}} \geq \gamma_{\mathcal{B}} \max\left\{\frac{1}{1+(\delta_{\mathcal{A}}-\delta_{\mathcal{B}})\gamma_{\mathcal{B}}}, \frac{1-\delta_{\mathcal{B}}}{1-\delta_{\mathcal{A}}}\right\}.$$

Example 2.4. Egocentric Altruism (CES) Preferences (Cox and Sadiraj, 2004). Let preferences $\mathcal{J}$=$\mathcal{A}, \mathcal{B}$ be represented by

$$\begin{aligned} u_{\mathcal{J}}(m, y) &= \frac{1}{\alpha}(m^{\alpha} + \theta_{\mathcal{J}} y^{\alpha}), \ \texttt{if} \ \alpha \in (-\infty, 1)\diagdown\{0\} \\ &= m y^{\theta_{\mathcal{J}}}, \ \texttt{if} \ \alpha = 0. \end{aligned}$$

If $0 < \theta_{\mathcal{B}} \leq \theta_{\mathcal{A}}$ then $\mathcal{A}$ **MAT** $\mathcal{B}$. Verification is straightforward: $\mathbf{WTP}_{\mathcal{J}} = \theta_{\mathcal{J}}(m/y)^{1-\alpha}$, $\mathcal{J} = \mathcal{A}, \mathcal{B}$ imply $\mathbf{WPT}_{\mathcal{A}}/\mathbf{WTP}_{\mathcal{B}} = \theta_{\mathcal{A}}/\theta_{\mathcal{B}} \geq 1$. "Egocentricity" means that $u_{\mathcal{J}}(x+\epsilon, x-\epsilon) > u_{\mathcal{J}}(x-\epsilon, x+\epsilon)$ for any $\epsilon \in (0, x)$ which implies $\mathbf{WTP}(m, m) \leq 1$.

The exponent $\alpha \leq 1$ in the last example determines the curvature and hence the convexity of preferences. Straightforward algebra yields

$$\mathbf{K} \ = \frac{\theta(\alpha-1)m^{\alpha+1}y^{\alpha+1}(m^{\alpha}+\theta y^{\alpha})}{((m^{\alpha}y)^2 + (\theta m y^{\alpha})^2)^{\frac{3}{2}}} \leq 0.$$

On a ray $R_d = \{(m, md) : m > 0\}$ we have

$$\mathbf{K} \ = \frac{(\alpha-1)d^{\alpha+1}\theta(d^{\alpha}\theta+1)}{m(d^2 + (d^{\alpha}\theta)^2)^{\frac{3}{2}}} \leq 0.$$

Thus the curvature decreases (in absolute value) along a given ray proportionally to $1/m$, i.e., $m\mathbf{K}$ is constant along the ray. Appendix A.4 shows that the same is true for any homothetic preferences. Hence relative curvature $m\mathbf{K}$ sometimes is more useful than curvature $\mathbf{K}$.

Much of the theoretical literature on social preferences relies on special assumptions that appear to be inconsistent with the classical approach to preference (and demand) theory (Hicks, 1939; Samuelson, 1947). The preceding examples help to clarify the issues. The linear and nonlinear inequality aversion models, quasi-maximin model, and egocentric altruism all assume convex preferences, since their upper contour sets are all convex. Monotonicity is violated by strict inequality aversion: **WTP** is positive when y is below own income m, but **WTP** is negative in the region where $y > m$. For $N = 2$, the maximin property is implied by convexity and positive monotonicity, and therefore is a property of altruistic preferences with these classical properties. A preference for efficiency (i.e., for a larger total of all agents' payoffs) is consistent with a limiting case of the quasi-maximin model. In our notation, efficiency is implied by preferences with **WTP**=1. We shall now see that for more general preferences, the efficiency of choices depends on the shape of the opportunity set.

3. Opportunities

Define an opportunity set F (or synonymously, a feasible set or budget set) as a convex compact subset of $\Re^2_+$. It is convenient and harmless (given preferences monotone in own income m) to assume free disposal for own income, i.e., if $(m, y) \in F$ then $(am, y) \in F$ for all $a \in [0, 1]$. Every opportunity set F is closed (as a compact set in $\Re^2_+$) and therefore it contains its $\Re^2_+$-boundary, denoted ∂F; indeed F is the convex hull of ∂F. Convexity of F means that each boundary point has a supporting hyperplane (i.e., tangent line) defined by an inward-pointing normal vector, and F is contained in a closed positive halfspace; see for example Rockafellar, 1970, p. 100. At some boundary points X (informally called corners or kinks) the supporting hyperplane is not unique; examples will be noted later.

At regular boundary points there is a unique supporting hyperplane and the implicit function theorem guarantees a smooth function f whose zero isoquant defines the boundary locally. The marginal rate of transformation **MRT** can be expressed as a ratio of the first partial derivatives, except when the tangent is vertical. We often need to work near vertical tangents, so we use the need to pay, **NTP**, defined as $\mathbf{NTP}(X) = 1/\mathbf{MRT}(X) = f_y/f_m$. The **NTP** is single-valued except at kinks and corners of the boundary, where its values lie in some interval.

Curvature can also be defined at regular boundary points, using the same formula 2.1 for **K** with u replaced by f. Again **K** and **NTP** are intrinsic, independent of the choice f used to represent the boundary segment.

Some examples may help fix ideas.

Example 3.1. Standard budget set. Let $F = \left\{(m, y) \in \Re^2_+ : m + py \leq I\right\}$ for given $p, I > 0$. Then ∂F consists of portions of the axes together with the line segment $\left\{(m, y) \in \Re^2_+ : m + py = I\right\}$, as shown in Figure 1. The **NTP** is p along the budget line, is 0 along the y axis and is ∞ along the m-axis. **NTP** assumes all values outside the interval $(0, p)$ at the corner $(0, I/p)$, and takes all values in the interval $[-\infty, p]$ at the corner $(I, 0) \in \partial F$. As usual, curvature is not defined at the corners, while at all regular boundary points $\mathbf{K} = 0$.

Example 3.2. Ring test (Liebrand, 1984; see also Sonnemans, van Dijk and van Winden, 2005). Let $F = \left\{(m, y) \in \Re^2_+ : m^2 + y^2 \leq R^2\right\}$ for given $R > 0$, as shown in Figure 2.[3] On the circular part of the boundary, **NTP** is y/m and the curvature is $1/R$.

Example 3.3. Ultimatum game (Güth, Schmittberger, and Schwarze, 1982). The responder's opportunities in the \$10 ultimatum game consist of the origin $(0, 0)$ and (due to our free disposal assumption) the horizontal line segment from $(0, 10 - x)$ to $(x, 10 - x)$. This set is not convex so it doesn't qualify as an opportunity set by our definition. Its convex hull, however, is the opportunity set in the Convex Ultimatum game (Andreoni, Castillo and Petrie, 2003), which is identical to that of the Power to Take game in the following example.

Example 3.4. Power to take game (Bosman and van Winden, 2002). The "take authority" player chooses a take rate $b \in [0, 1]$. Then the responder with income I chooses a destruction rate $1 - \delta$. The resulting payoffs are $m = (1 - b)\delta I$ for the responder and $y = b\delta I$ for the take authority. Thus, with free disposal the responder's opportunity set is the convex hull of three points $(m, y) = (0, 0), (0, bI)$ and $((1 - b)I, bI)$. Along the Eastern boundary **NTP** is constant at $(b - 1)/b$ and the curvature is 0. Figure 3 shows opportunity sets for $b = 0.3, 0.7$ and $I = 10$.

Example 3.5. Investment game (Berg, Dickhaut, and McCabe, 1995). The First Mover (FM) and Second Mover (SM) each have an initial endowment of I. The FM sends an amount $s \in [0, I]$ to SM, who receives ks. Then the SM returns an amount $r \in [0, ks]$ to the FM, resulting in payoffs $m = I + ks - r$ for SM

[3]In the original studies ∂F is the entire circle $\left\{(m, y) \in \Re^2 : m^2 + y^2 = R^2\right\}$, not just the portion in the positive quadrant.

and $y = I - s + r$ for FM. The FM's choice of s selects the SM's opportunity set $F_s = \{(m, y) \in \Re^2_+ : y \in [I - s, I + (k-1)s], m \in [0, 2I + (k-1)s - y]\}$. Curvature $\mathbf{K} = 0$ on each segment of the boundary, and $\mathbf{NTP} = 1$ on the segment of the boundary corresponding to $m + y = 2I + (k-1)s$. Figure 4 shows F_s for $s = 3$ and 9 when $I = 10$ and $k = 3$.

Example 3.6. Moonlighting game (Abbink, Irlenbusch, and Renner, 2000). In this variant on the investment game, the FM sends $s \in [-I/2, I]$ to SM, who receives $g(s) = ks$ for positive s and $g(s) = s$ for negative s. Then the second mover transfers $t \in [(-I + s)/k, I + g(s)]$ resulting in non-negative payoffs $m = I + g(s) - |t|$ and $y = I - s + t$ for positive t and $y = I - s + kt$ for negative t. The second mover's opportunity set is the convex hull of the points $(m, y) = (0, 0), (I + g(s) - (I - s)/k, 0), (I + g(s), I - s)$, and $(0, 2I + g(s) - |s|)$. The $\mathbf{NTP}$ along the boundary of the opportunity set is 1 above and $-1/k$ below the $t = 0$ locus, is 0 along the y axis, and is ∞ along the m-axis. Again, curvature at all regular boundary points is $\mathbf{K} = 0$. Figure 5 shows SM opportunity sets for $s = -5$ and 4 when $I = 10$ and $k = 3$.

Example 3.7. Gift exchange labor markets (Fehr, Kirchsteiger, and Riedl, 1993). The employer with initial endowment I offers a wage $w \in [0, I]$ and the worker then chooses an effort level $e \in [0, 1]$ with a quadratic cost function $c(e)$. The final payoffs are $m = w - c(e)$ for the worker and $y = I + ke - w$ for the employer, where the productivity parameter $k = 10$ in a typical game. Figure 6 shows opportunity sets for $w = 3, 7$ and $I = 10$, $c(e) = 10e^2$. The worker's opportunity set is similar to the second mover's in the investment game, except that the Northeastern boundary is a parabolic arc instead of a straight line of slope -1. Along this Eastern boundary $\mathbf{NTP}$ is $2e$ and the curvature is $-1/5(4e^2 + 1)^{3/2}$. Also, if the employer offers a wage in excess of his endowment I then the opportunity set includes part of the quadrant $[m > 0 > y]$. It is straightforward but a bit messy to extend the definition of opportunity set to include such possibilities.

Example 3.8. Sequential VCM public good game with two players (Varian, 1994). Each player has initial endowment I. FM contributes $c_1 \in [0, I]$ to the public good. SM observes c_1 and then chooses his contribution $c_2 \in [0, I]$. Each unit contributed has a return of $a \in (0.5, 1]$, so the final payoffs are $m = I + ac_1 - (1 - a)c_2$ for SM and $y = I + ac_2 - (1 - a)c_1$ for FM. SM's opportunity set is the convex hull of the four points $(m, y) = (0, I - (1 - a)c_1), (I + ac_1, I - (1 - a)c_1), (aI + ac_1, (1 + a)I - (1 - a)c_1)$ and $(0, (1 + a)I - (1 - a)c_1)$. Along the

Pareto frontier, **NTP** is constant at $(1-a)/a$. Figure 7 shows opportunity sets for $c_1 = 2, 8$ and $I = 10, a = 0.75$.

Example 3.9. Stackelberg duopoly game (e.g., Varian, 1992, p. 295-298). Consider a duopoly with zero fixed cost, constant and equal marginal cost, and nontrivial linear demand. Without further loss of generality one can normalize so that the profit margin (price minus marginal cost) is $M = T - q_{_L} - q_{_F}$, where $q_{_L} \in [0, T]$ is the Leader's output choice and $q_{_F} \in [0, T - q_{_L}]$ is the Follower's output to be chosen. Thus payoffs are $m = Mq_{_F}$ and $y = Mq_{_L}$. The Follower's opportunity set therefore is bounded by a parabolic arc opening towards the y-axis, as shown in Figure 8 for $T = 24$ and $q_{_L} = 6, 8$ and 11. A calculation of **NTP** and curvature appears in Appendix A.5. Unlike the earlier examples, the **NTP** varies smoothly from negative to positive values along the boundary of the opportunity set as one moves counterclockwise.

4. Reciprocity

Reciprocity is key to our analysis. The idea is that more generous choices by one player induce more altruistic preferences in a second player. To formalize, consider a two person extensive form game of complete information in which the first mover (FM) chooses an opportunity set $C \in \mathcal{C}$, and the second mover (SM) chooses the payoff vector $(m, y) \in C$. It is natural to regard opportunity set G as more generous than (**MGT**) opportunity set F if it is obtained by stretching along the m axis (SM's potential payoffs), and shrinking (or stretching less) along the y axis (FM's potential payoffs).

Definition 2. Opportunity set $G \subset \Re^2_+$ is more generous than opportunity set F if $G = \mathbb{T}F$ for some smooth transformation $\mathbb{T} : \Re^2_+ \longrightarrow \Re^2_+$ such that $(m, y) \longmapsto (\rho(m), \vartheta(y))$ with $\rho(0) \geq 0$, $\rho'(z) \geq 1$ and $\rho(z) \geq \vartheta(z)$ for all $z \geq 0$. In this case we write G **MGT** F.

Initially, the second mover knows the collection $\mathcal{C}$ of possible opportunity sets. Prior to her actual choice she learns the actual opportunity set $C \in \mathcal{C}$, and acquires preferences $\mathcal{A}_C$. Reciprocity is captured formally in

> **Axiom R:** *Let the first mover choose the actual opportunity set for the second mover from the collection $\mathcal{C}$. If $F, G \in \mathcal{C}$ and G* **MGT** *F, then $\mathcal{A}_G$* **MAT** *$\mathcal{A}_F$.*

There is a traditional distinction between sins of commission (active choice) and sins of omission (retaining the status quo). Of course, sometimes there is no choice at all and the status quo cannot be altered. Intuitively, the second mover will respond more strongly to generous (or ungenerous) choices that overturn the status quo than to those that uphold it, or that involve no real choice by the first mover.[4] Compared to no choice, upholding the status quo should provoke the stronger response, at least when the status quo is the best or worst possible opportunity.

The following notation will help formalize the intuition. Let the first mover choose the actual opportunity set C from a collection $\mathcal{C}$. If $\mathcal{C}$ contains at least two elements, then one of them is denoted the status quo. Let $\mathcal{A}_{C^*}$ and $\mathcal{A}_{C^c}$ respectively denote the second mover's acquired preferences when C is the status quo and when it differs from the status quo. If $\mathcal{C}$ is a singleton, then the first mover has no choice and we write $\mathcal{C} = \{C^o\}$ with corresponding second mover preferences $\mathcal{A}_{C^o}$.

> **Axiom S:** *Let the first mover choose the actual opportunity set for the second mover from the collection $\mathcal{C}$. If the status quo is either F or G and G **MGT** F then*

(1) $\mathcal{A}_{G^c}$ **MAT** $\mathcal{A}_{G^*}, \mathcal{A}_{G^o}$ and $\mathcal{A}_{F^*}$, $\mathcal{A}_{F^o}$ **MAT** $\mathcal{A}_{F^c}$,
(2) $\mathcal{A}_{G^*}$ **MAT** $\mathcal{A}_{G^o}$ if G **MGT** C, for all $C \in \mathcal{C}$ and $\mathcal{A}_{F^o}$ **MAT** $\mathcal{A}_{F^*}$ if C **MGT** F for all $C \in \mathcal{C}$.

We will say that either Axiom *holds strictly* if the inequality in the **MAT** Definition 1 is strict.

It should be emphasized that the recent preference models noted in Examples 2.1 - 2.4 have no room for Axioms **R** and **S**. In those models preferences are assumed fixed, unaffected by more or less generous opportunity sets chosen by the first mover. Actual choices by a first mover are not central even in the "reciprocity" models of Charness and Rabin (2002, Appendix), Falk and Fischbacher (2006), and Dufwenberg and Kirchsteiger (2004). Those models focus on higher-order beliefs regarding other players' intentions (or, in Levine (1998), regarding other players' types). Cox, Friedman, and Gjerstad (forthcoming) implicitly consider Axiom **R**, but only within the particular parametric family of CES utility functions noted in Example 2.4.

Natural **MGT** orderings are fairly common. For example, with the standard budget set in Example 3.1, an increase in own income I or a real increase in the price of transfers (so I/p decreases) leads to a more generous budget set, as illustrated by the solid budget lines in Figure 9. Indeed, $\rho' > 1$ is simply the

[4]This intuition goes back at least to Adam Smith's *Theory of Moral Sentiments*, 1759, p. 181.

income ratio and $\vartheta' < 1$ reflects the decrease in I/p. Likewise, it is clear from Figures 3-8 that smaller take rate τ in the Power to Take game, larger s in the Investment and Moonlighting games, larger I in the Gift Exchange Labor Market, a larger contribution c_1 in the VCM public goods game and a smaller output q_L in the Stackelebrg duopoly game all create **MGT** opportunities for the second mover (Appendix A.7 verifies this for the Investment game and the Stackelberg duopoly game.)

But a few minutes study of those figures reveals that reciprocity and convexity will be difficult to disentangle. In the Investment game, for example, larger s moves the second mover's altered endowment down along the dashed line in Figure 4, increasing his relative income. Indeed, the ray through the altered endowment point for $s = 3$ has slope $d = 7/19$, compared to $d = 1/37$ for $s = 9$. Hence if the second mover's preferences are strictly convex and homothetic then a larger s implies greater **WTP** at the altered endowment even when those preferences are not at all affected by the first mover's more generous choice. The other games also conflate convexity and reciprocity. The underlying problem is that more generous choices by definition create better relative opportunities, hence lower d and (by convexity) greater **WTP**.

5. Choice

As in standard preference theory, our maintained assumption is that every player always chooses a most preferred point in the opportunity set F. By positive monotonicity in own income such points are from ∂F, and by convexity such points must form a connected subset of ∂F. If either preferences $\mathcal{A}$ or opportunities F are strictly convex then that subset is a singleton, i.e., there is a unique choice $X \in \partial F$. In this case all points in $F \setminus \{X\}$ are revealed to be on lower $\mathcal{A}$-indifference curves than X.

Not all boundary points are candidates for choice in our set up. The first result is that, due to strict monotonicity in own payoff m, only "eastern" points will be chosen, since they have larger own payoff. To formalize, define the Eastern boundary as $\partial_E F = \{(m, y) \in F : \forall x > m, (x, y) \notin F\}$. The North point N_F and the South point S_F are the points in $\partial_E F$ with respectively the largest and the smallest y component.

Proposition 1. *Suppose that either preferences $\mathcal{A}$ or the opportunity set F are strictly convex, and let X be the $\mathcal{A}-$chosen point in F. Then $X \in \partial_E F$.*

All proofs are collected in Appendix A.

The next result shows that, as admissible preferences go from maximally malevolent through neutral to maximally benevolent under the **MAT** ordering, the player's choices trace out the Eastern boundary of the budget set from South to North. To put it another way, consider the ray of slope d. As d increases from 0 to ∞, the intersection of the ray with the Eastern boundary traces out the chosen points. The notation d_X indicates the slope of the ray through X, i.e., $d_X = y/m$ if $X = (m, y)$.

Proposition 2. *Suppose that either preferences $\mathcal{A}$ and $\mathcal{B}$, or the opportunity set F, are strictly convex. Let $X_{\mathcal{A}}$ and $X_{\mathcal{B}}$ be the points in $F \setminus \{(0,0)\}$ chosen when preferences are respectively $\mathcal{A}$ and $\mathcal{B}$. Then*

(1) *$\mathcal{B}$ **MAT** $\mathcal{A}$ implies $d_{X_{\mathcal{B}}} \geq d_{X_{\mathcal{A}}}$ for homothetic preferences.*
(2) *If $X \in \partial_E F$ lies on a ray with slope between $d_{X_{\mathcal{A}}}$ and $d_{X_{\mathcal{B}}}$, then there are preferences $\mathcal{P}$ with $\mathcal{B}$ **MAT** $\mathcal{P}$ **MAT** $\mathcal{A}$ such that X is the $\mathcal{P}$-chosen point in F.*
(3) *There are admissible preferences for which the chosen point is arbitrarily close to S_F, and other admissible preferences for which the chosen point is arbitrarily close to N_F.*

Propositions 1 and 2 deal with a fixed opportunity set. Often we need predictions of how an agent with given preferences will choose in a new opportunity set. Textbook preference theory offers such predictions in the case of standard budget sets and convex monotone preferences. We will get weaker predictions because we deal with more general opportunity sets and with preferences that are convex but not necessarily monotone in other's income y. The following example illustrates this.

Example 5.1. Figure 9 shows standard budget sets F with $I = 1, p = 1$ (solid line) and G with $I = 2, p = 4$ (dashed line). Suppose that a player with preferences $\mathcal{P}$ picks X from F. What can we predict about his choice W from G? If it happens that X is in G then textbook preference theory tells us that W is not in F; it must be on the segment of the G budget line that lies outside F. Using homotheticity, we can strengthen the prediction: W lies on the sub-segment between $Y = tX$ and the South point S_G, which is $(2, 0)$, as indicated in the figure.

The result in Example 5.1 generalizes to nonlinear convex opportunity sets, as shown by the following proposition and illustrated in Figure 10. The proposition constructs a point Z which solves $\mathbf{NTP}_{\partial F}(X) = \mathbf{NTP}_{\partial G}(Z)$. The Appendix shows how to extend the definition so that Z is well defined even with corners and kinks at which **NTP** is not single valued.

Proposition 3. *Let an agent with strictly convex and homothetic preferences $\mathcal{A}$ choose X and W respectively from some opportunity sets F and G. Let $Y = tX$ be the most distant point from the origin on the ray through X in $\partial_E G$, and let $Z \in \partial_E G$ solve $\mathbf{NTP}_{\partial F}(X) = \mathbf{NTP}_{\partial G}(Z)$. Then*

(1) *if $X \in G$ then $W \in F^c$ or $W = X$,*
(2) *$d_W \leq \min\{d_X, d_{N_G}\}$ iff $\mathbf{NTP}_{\partial F}(X) \leq \mathbf{NTP}_{\partial G}(Y)$, and*
(3) *$d_W \geq d_Z$ iff $d_Z \leq d_X$.*

Figure 9 makes it transparent that statement (1) in the proposition is equivalent to the Weak Axiom of Revealed Preference for standard budget sets defined by price vectors p_G and p_F: if $p_G \cdot W \geq p_G \cdot X$ (i.e. $X \in G$) then $p_F \cdot W > p_F \cdot X$ (i.e. $W \in F^C$) or $W = X$. Statements (1) - (3) extend these traditional ideas to convex opportunity sets with nonlinear boundaries in the case of homothetic preferences, as shown in Figure 10.[5]

Propositions 1-3 do not invoke Axioms **R** and **S**. These axioms sometimes sharpen and sometimes weaken the predictions of standard preference theory, as illustrated in the rest of this section.

Example 5.2. What happens in example 5.1 if preferences $\mathcal{A}$ are altered by the choice of G over F? Were G less generous than F, then reciprocity assumption R would imply that the choice W would shift southward, towards the corner $(2, 0)$ of the budget set, i.e., the earlier prediction would hold *a fortiori.* However, G **MGT** F for reasons explained in the second to last paragraph of the previous section. Consequently, Axiom **R** implies that W will shift northward. The prediction reduces to saying that W is north of the South corner. But this tells us nothing; no choice along the Eastern portion of ∂G is ruled out. The problem here is that the reciprocity effect doesn't reinforce the usual substitution effect in preference theory, but rather counteracts it and we have no indication which effect is stronger.

Sharper results often arise from closer examination of specific games. We illustrate by deriving testable predictions for the Investment game of Example 3.5.

[5]For convex opportunity sets, Matzkin (1991) derives revealed preference inequalities in terms of the supporting hyperplanes. For more general opportunity sets, Forges and Minelli (2006) derive revealed preferences that are not necessarily convex.

Proposition 4. *Let the FM in the Investment game choose F_s as the SM's opportunity set, and let $r(s)$ be the SM's response. Also let the same SM be given the same opportunity set F_s in a dictator game, and let $r^o(s)$ be his response there. Assume that* $\mathbf{WTP} \leq 1$ *and* $\partial \mathbf{WTP}/\partial \mathbf{m} \geq 0$. *Then:*

(1) *convexity implies that $r^o(s)$ is increasing in s;*
(2) *Axiom **R** implies that $r(s)$ is increasing in s;*
(3) *Axiom **S** implies that $r(s) \geq r^o(s)$ for all feasible s.*

The assumption $\mathbf{WTP} \leq 1$ says that the Follower would not favor an inefficient adverse transfer; at the margin he loves his neighbor no more than himself. The assumption $\partial \mathbf{WTP}/\partial \mathbf{m} \geq 0$ simply says that the Follower becomes no less generous when her income increases, other things equal.

The standard Stackelberg game in Example 3.9 is especially useful for our purposes. Figure 8 suggests (and Appendix A.7 verifies) that smaller output choices by the Stackelberg Leader create **MGT** opportunity sets for the Follower. By Axiom **R** we expect this to induce **MAT** preferences in the Follower. It seems that this preference shift should induce the Follower to choose smaller output. But of course we must also take into account preference convexity, and also the changing curvature of the opportunity set. The next proposition sorts out these effects and expresses them in terms of the Follower's deviation from selfish best reply.

Proposition 5. *In the standard Stackelberg game of Example 3.9 let $Q_D(q_L) = q_F - q_F^o$ be the deviation of the Follower's output choice q_F from the selfish best reply $q_F^o = 12 - \frac{1}{2}q_L$ when the Leader chooses output q_L. One has*

$$\frac{dQ_D}{dq_L} = -\frac{1}{2}w - \frac{dw}{dq_L}q_L$$

where $w = \mathbf{WTP}(Mq_F, Mq_L)$ is willingness to pay at the chosen point. Furthermore,

(1) *If Follower's preferences $\mathcal{A}$ are fixed and linear, then w is constant with respect to q_L and $\frac{dQ_D}{dq_L}$ is positive if and only if preferences at the chosen point are malevolent.*
(2) *If Follower's preferences $\mathcal{A}$ are fixed and convex, then w is decreasing in q_L and $\frac{dQ_D}{dq_L}$ contains an additional term that is positive provided that $q_L \leq 12$, $w \leq 1$, $w_m \geq 0$ and $w_m + w_y \geq 0$.*
(3) *If Follower's preferences satisfy Axiom **R** strictly, then w is decreasing in q_L and $\frac{dQ_D^r}{dq_L}$ contains an additional positive term.*

(4) *If Follower's preferences satisfy Axiom **S** strictly, then* **WTP** *is decreasing in* q_L *and* $\frac{dQ_D^s}{dq_L}$ *has an additional positive (negative) term if the status quo is smaller (larger) than* q_L.

Proposition 5 shows that an increase in q_L has three different effects:

- A reciprocity effect, items (3) - (4) in the Proposition. If Axiom **R** holds strictly, then the less generous opportunity set decreases the Follower's **WTP**, increasing q_F and q_D. Axiom **S** moderates or intensifies this effect, depending on the status quo.

- A preference convexity (or substitution) effect, item (2) in the Proposition. The choice point is pushed northwest, where (subject to some technical qualifications) **WTP** is less, again increasing q_F and q_D.

- An opportunity set shape effect (in some ways analagous to an income effect), item (1). The curvature of the parabola decreases. Holding $w = \mathbf{WTP}$ constant, q_D increases when the Follower is malevolent ($w < 0$, hence $q_D > 0$), and decreases when the Follower is benevolent ($w > 0$, hence $q_D < 0$).

The technical qualifications for the preference convexity effect are not especially restrictive. In a sense, Leader choices q_L exceeding the monopoly (and standard Stackelberg) level 12 are dominated: they produce choice sets for the Follower that are strict subsets of those produced by $q_L \leq 12$. The mild conditions $w \leq 1$ and $w_m \geq 0$ were used in the previous Proposition. Finally, the condition $w_m + w_y \geq 0$ says that equal increases in income do not push preferences towards malevolence.

A parametric example may clarify the logic. For given $q_L \in [0, 24]$, the Follower's choice set is the parabola $\{(m, y) : m = Mq_F, y = Mq_L, M = 24 - q_L - q_F, q_F \in [0, 24 - q_L]\}$, with $\mathbf{NTP} = -\frac{dm/dq_F}{dy/dq_F} = \frac{24 - q_L - 2q_F}{q_L}$. Suppose that the Follower has fixed Cobb-Douglas preferences represented by $u(m, y) = my^\theta$, so **WTP** is $\theta m/y = \theta q_F/q_L$. Solving $\mathbf{NTP} = \mathbf{WTP}$, one obtains $q_F = Q(q_L|\theta) = (24 - q_L)/(2 + \theta)$. Noting that the selfish best reply is $q_F^o = Q(q_L|0)$, one obtains a closed form expression for the deviation, $q_D = -\frac{\theta}{4+2\theta}(24 - q_L)$. For fixed θ positive (benevolent preferences) or smaller than -2 (pathologically malevolent preferences), the deviation is negative but increasing in the Leader's output; the opposite is true when θ is negative but larger than -2 (moderately malevolent). This is the combined impact of the convexity (or substitution) and shape (or income) effects noted above. Of course, reciprocity effects will decrease θ and hence increase q_D.

A direct way to test for reciprocity is to manipulate the first mover's choice collection $\mathcal{C}$ in the laboratory so that a fixed opportunity set is more or less generous. For example, suppose restrictions on the Stackelberg Leader's choice set make a given output choice q^* the most generous (smallest) possible in one situation, and the same output q^* the least generous (largest) possible in another situation. If a given Follower reacts differently in the two situations, it must be due to reciprocity effects, since by holding q^* constant we have eliminated the convexity and shape effects. This is the idea behind the Stackelberg mini-game introduced in the last empirical section.

Corollary 1. *Stackelberg Mini-Game. For* $0 < x < s < z < 24$, *suppose the Stackelberg Leader has restricted output choices* $q_L \in \{x, s\}$ *in situation (a) and* $q_L \in \{z, s\}$ *in another situation (b). Let the Leader choose* s *in both situations. If Follower's preferences satisfy Axioms* $\mathbf{R}$ *and* $\mathbf{S}$ *then* $Q_D^a(s) \geq Q_D^b(s)$, *and at each possible Follower choice* q_F, $\mathbf{WTP}^a(Mq_F, Ms) \leq \mathbf{WTP}^b(Mq_F, Ms)$.

6. Investment Game Data

We begin illustration of empirical applications with the Investment game of Example 3.5. Using a double-blind protocol, Cox (2004) gathers data from a one-shot investment game (Treatment A) with 32 pairs of FMs and SMs. Cox also reports parallel data (Treatment C) with another 32 pairs in which SMs are "dictators" with exactly the same opportunity sets given to them by the experimenter. In both treatments, the choices s and r are restricted to integer values but the conclusions of Proposition 4 still hold. Axiom $\mathbf{S}$ immediately implies that a SM with any particular F_s would have more altruistic preferences and hence would return more in Treatment A than in Treatment C, as noted in part (3) of Proposition 4. The first two parts of the proposition predict that the SM will return more to the FM when s is larger in both treatments.

To test these predictions, construct the dummy variable $D = 1$ for Treatment C data, so $D = 0$ for Treatment A data. Regress the SM choice r on the amount sent s and its interaction with D, using the 2-sided Tobit procedure to account for the limited range of SM choices in the 54 relevant observations ($r \in [0, 3s]$).[6] The estimated coefficient for s is 0.58 ($\pm$ standard error of 0.22) with one-sided p-value of 0.006, consistent with reciprocity and part (2) of Proposition 4. The

[6]The five observations for each treatment in which $s = 0$ are not used in the estimation for two reasons: (a) since the SM opportunity set is a singleton, there is nothing for a theory about SM choices to explain; and (b) since the left ($r = 0$) and right ($r = 3s$) censors in the Tobit estimation are equal, the estimation algorithm would not be well defined.

estimated coefficient for $D \times s$ is -0.69 (± 0.32, $p = 0.018$), consistent with Axiom **S** and part (3) of Proposition 4. Since the coefficient sum is statistically indistinguishable from 0, the convexity prediction in part (1) of Proposition 4 is neither supported nor contradicted.

We confirm the Axiom **S** result by direct hypothesis tests on the most relevant subset of data, where $s = 5$ (with 7 observations in each treatment) and $s = 10$ (with 13 observations in each treatment). The Mann-Whitney and t-test both reject the null hypothesis of no difference between the amounts returned in favor of the strict Axiom **S** alternative hypothesis that returns are larger in Treatment A. The one-sided p-values for the t-test (respectively the Mann-Whitney test) are 0.027(0.058) for the $s = 5$ data and are 0.04(0.10) for the $s = 10$ data.

7. Stackelberg Duopoly Data

The investment game data are consistent with the theory but they do not permit crisp tests of reciprocity because, among other limitations, (a) the opportunity sets are linear and hence can't reveal much about **WTP**, and (b) only one choice is observed per subject, precluding direct observation of changed preferences. Limitations (a) and (b) are overcome in the Stackelberg duopoly data of Huck, Müller, and Normann (2001, henceforth HMN).

The HMN data consist of 220 output choices (q_L, q_F) by 22 FMs (or Leaders) choosing $q_L \in \{3, 4, 5, \ldots, 15\}$ randomly rematched for 10 periods each with 22 SMs (or Followers). As noted in Example 3.9 and elsewhere, the SM's choice $q_F \in \{3, 4, 5, \ldots, 15\}$ determines payoffs (m, y) within an opportunity set of discrete points on a parabolic arc. Specifically, payoffs are $m = M\, q_F$ and $y = M\, q_L$, where $M = 24 - q_L - q_F$ is the profit margin. The **WTP** can be inferred at a chosen point (q_L, q_F) by the **NTP** at that point, $(24 - 2q_F - q_L)/q_L$.

Recall that Proposition 5 predicts that the SM's output choice reveals a constant **WTP** if her preferences are linear and unaffected by the FM's output choice q_L. The corresponding deviation Q_D from her selfish best reply output is linearly decreasing in q_L if her preferences are benevolent. Convexity and Axiom **R** effects produce a revealed **WTP** that is decreasing (and Q_D that is increasing) in q_L.

Table 1 reports tests of these predictions on the HMN data, omitting the 26 data points where the Proposition's hypothesis $q_L \leq 12$ fails.[7] To check for asymmetric responses to large and small FM choices (relative to the Cournot choice $q_L = 8$), we define the dummy variable $DP = 1$ if $q_L \leq 8$. All columns

[7]The results are substantially unchanged when the data points for $q_L > 12$ are included.

in the Table report panel regressions with individual subject fixed effects. The first column, with dependent variable $\mathbf{WTP} \times 100$, firmly rejects the hypothesis of benevolent linear and fixed preferences: the coefficient for q_L is significantly negative, not positive. The second column, with dependent variable Q_D, confirms this result. We infer that Q_D is an increasing function of FM output q_L, consistent with convexity and reciprocity. The last column reports that there is a stronger response to "greedy" FM choices in excess of the Cournot output 8 than to "generous" FM choices below or equal to output 8. Some supplementary regressions are noted in the Appendix, also consistent with Proposition 5.

$Dep.Variable$	$\mathbf{WTP} \times 100$	Q_D	Q_D
q_L	$-5.43 \pm 1.11^{0.000}$	$0.35 \pm 0.06^{0.000}$	$0.24 \pm 0.08^{0.002}$
$DP \times q_L$			$-0.10 \pm 0.05^{0.023}$
$constant$	$28.87 \pm 11.01^{0.005}$	$-2.13 \pm 0.58^{0.000}$	$-0.81 \pm 0.87^{0.176}$

TABLE 1. Panel Regressions with fixed effects. Data consist of 194 choices by 22 Followers in HMN experiment when $q_L < 13$. The coefficient point estimates are shown $\pm$ the standard error, with one-sided p-values in superscripts.

8. Stackelberg Mini-Game Data

The HMN data still do not permit tests of some of our most distinctive predictions. All FMs (Leaders) have the same choice set, eliminating variability that could help separate the convexity effect from the reciprocity effect. Also, due in part to differing experiences, SMs may have different views on the generosity of a given output choice q_L. In order to overcome these limitations while preserving the nice quadratic shape of the SM choice sets, we created a new version of the Stackelberg game that restricts FMs to binary choices.

In our Stackelberg mini-game, each subject in the FM role twice chooses $q_L \in \{6, 9\}$ and twice chooses $q_L \in \{9, 12\}$ without feedback. Each subject in the SM role is then paired simultaneously with four different FMs and chooses an integer value of $q_F \in \{5, 6, ..., 11\}$ with no feedback. The corresponding payoffs (m, y) are clearly displayed. The final payoff is given by one of the four choices, selected randomly at the end of the session. The "double blind" procedures are detailed in the instructions to subjects, reproduced in Appendix B.

Figure 11 summarizes the data. More generous (smaller) choices by the FM seem to be associated with more altruistic or less malevolent (smaller) choices by the SM, but it is hard to tell from the figure whether the effect is significant. For example, there are only five observations at $q_L = 6$. Most importantly, the scatterplot doesn't show which of the 24 subjects made which choices.

To infer how individual subjects respond to reciprocity concerns, we turn again to panel regressions with individual subject fixed effects. The second column in Table 2 reports that, consistent with Corollary 1, SMs' average **WTP** decreased by almost 8 cents per dollar when $q_L = 9$ was the less generous choice (indicated by $D9 = 1$). The second column reports the same data in a different way: SM output choice increased by 0.34 on average, significant at the $p = 0.016$ level (one-sided). Since the opportunity set F_9 is constant in these 72 data points, the result cannot be due to convexity or shape effects; it must be pure reciprocity. The last column of Table 2 reports regressions for Q_D for the entire data set, using the additional dummy variable $D12$, which takes value 1 if $q_L = 12$, and 0 otherwise.[8] The signs of all coefficient estimates are consistent with Axiom **R** and convexity.

	$WTP{\times}100$ $(q_L{=}\ 9)$	Q_D $(q_L{=}\ 9)$	Q_D
$nobs(gr)$	72(24)	72(24)	91(24)
$D9$	$-7.65 \pm 3.05^{0.008}$	$0.34 \pm 0.14^{0.008}$	$0.32 \pm 0.14^{0.013}$
$D12$			$0.37 \pm 0.19^{0.028}$
$constant$	$-5.93 \pm 2.31^{0.007}$	$0.27 \pm 0.10^{0.007}$	$0.19 \pm 0.11^{0.046}$

TABLE 2. Panel Regressions with fixed effects for Stackelberg mini-game data. Entries are coefficient estimates ± standard errors with one-sided p-values in superscripts.

9. DISCUSSION

Classic choice theory (e.g., Hicks, 1939; Samuelson, 1947) clarified and unified earlier work on how preferences and opportunities affect outcomes. The present paper applies those classic ideas to social preferences. We focus on willingness to pay (**WTP**), the reciprocal of the marginal rate of substitution between own income and others' income. Increasing **WTP** along indifference curves is simply convexity, and convex social preferences provide a unified account of several

[8]We omit here a dummy variable that takes value 1 for $q_L = 6$ because there are only five such observations. When the dummy is included, the coefficient estimate has the predicted sign but of course is insignificant statistically, while the other coefficient estimates change only slightly.

social motives previously considered separately, such as efficiency, maximin, and inequality aversion.

The same classic ideas also permit a unified definition of reciprocity. We say that one set of preferences is more altruistic than (**MAT**) another if it has a larger **WTP** at every point. We formalize reciprocity as a **MAT**-shift in preferences following more generous behavior by others. The definitions apply to malevolent (**WTP** < 0) as well as benevolent (**WTP** > 0) preferences, and automatically combine positive and negative reciprocity.

Convexity and reciprocity are quite different formally and conceptually, but we show that empirical work has a natural tendency to confound the two notions. The problem is simply that more generous behavior by a first mover tends to push the second mover's opportunities "southeast," towards larger income for the first mover and smaller income for the second mover. Convexity typically implies greater **WTP** as one pushes southeast, even when there is no **MAT**-shift in preferences due to reciprocity.

Several theoretical propositions develop the observable consequences of convexity and reciprocity. We show that more northerly choices on the Eastern boundary of an opportunity set reveal more altruistic (or less malevolent) preferences. For fixed preferences, choices in an opportunity set reveal bounds on preferences that we translate into bounds on choices in new opportunity sets. With standard budget sets, these bounds are equivalent to the Weak Axiom of Revealed Preference, and sharper versions are derived for homothetic preferences. In the context of the well-known Investment and Stackelberg duopoly games, the last two theoretical propositions relate the testable implications of traditional convex preferences to our formalization of reciprocity and status.

Finally, to illustrate the empirical content of the theory, we examine two existing data sets and one new data set. Existing investment game data are consistent with convexity and reciprocity, and confirm that people respond more strongly to acts of commission than to default choices. Existing Stackelberg data confirm reciprocity/convexity effects and suggest a stronger negative response to greedy behavior than the positive response to generous behavior. The new Stackelberg mini-game data allow us to separate convexity from reciprocity effects, and confirm that reciprocity has a significant impact.

Theoretical clarification sets the stage for further empirical work. One can now refine earlier empirical studies that examine altruism and reciprocity. Such work should shed light on the extent to which typical human preferences depart from selfishness, and to what extent they are altered by experiencing generous or selfish behavior.

Further theoretical work is also in order. Other definitions of the "more generous than" relation should be considered and examined empirically. Also, Axiom **S** invokes the status quo to distinguish between acts of commission and omission, and between generous and greedy acts. But what does it take for a choice to become generally recognized as the status quo? What if an act has beneficial short run impact but is harmful in the long run? Answers to these and other questions await further theoretical development.

Appendix A. Mathematical proofs and derivations

A.1. **Relative income sensitivity and homothetic preferences.**

Lemma A.1. *Preferences are homothetic on* $\Re^2_+$ *iff* **WTP** *is constant along every ray* $R_d = \{(t, td) : t > 0\} \subset \Re^2_+$.

Proof. By definition, preferences are homothetic iff they can be represented by a utility function $u(m, y)$ whose ratio of partial derivatives u_m/u_y depends only on the ratio m/y, not on m and y separately (see Simon and Blume, p. 503). But the ratio $d = m/y$ is constant on the ray R_d by construction. □

A.2. **Curvature Formula.** For a curve in the plane, curvature at a point has absolute value $|\mathbf{K}\ | = 1/r$, where r is the radius of the circle that is second-order tangent to the curve at the given point. Intuitively, curvature of a curve is the rate at which the curve turns and formally it is given by the derivative of the curve's tangential angle, i.e. $\mathbf{K}\ = d\varphi/ds$ where $\tan\varphi = -dy/dm$ and $ds = \sqrt{dm^2 + dy^2}$. Substituting for ds, and inserting $-dy/dm = MRS$ and $d\varphi = d(MRS)/(1 + MRS^2)$, one has

$$\mathbf{K} = \frac{d(MRS)}{(1 + MRS^2)\sqrt{dm^2 + dy^2}} = \frac{1}{\sqrt{1 + MRS^2}^3}\frac{d(MRS)}{dm} \tag{A.1}$$

See e.g., Gray, 1997, p. 14-17. If the indifference curve is given implicitly by $u(m, y) = 0$ then $MRS = u_m/u_y$, hence

$$\begin{aligned}
\frac{d(MRS)}{dm} &= \frac{(u_{mm} + u_{my}dy/dm)u_y - (u_{ym} + u_{yy}dy/dm)u_m}{u_y^2} \\
&= \frac{(u_{mm} - u_{my}u_m/u_y)u_y - (u_{ym} - u_{yy}u_m/u_y)u_m}{u_y^2} \\
&= \frac{u_{mm}u_y - 2u_{my}u_m u_y + u_{yy}u_m^2}{u_y^3}
\end{aligned}$$

and therefore

$$\mathbf{K} = \frac{u_{mm}u_y^2 - 2u_{my}u_m u_y + u_{yy}u_m^2}{(u_m^2 + u_y^2)^{3/2}}.$$

A.3. **MAT is a partial ordering.**

Proof. Properties of reflexivity and transitivity are straightforward whereas the antisymmetry property follows from Hicks Lemma (Hicks, 1939, Appendix): if preferences have the same **MRS** everywhere in $\mathfrak{D}$ then they are the same. □

A.4. **Relative curvature and homothetic preferences.**

Lemma A.2. *If preferences are homothetic then $m\mathbf{K}(m, dm)$ is constant along the ray R_d.*

Proof. It is well known that homothetic preferences can be represented by a utility function $u(m, y)$ that is homogenous of degree 1, and that first (second) partial derivatives of homogeneous functions of degree 1 are homogeneous of degree 0 (-1) (e.g. Varian, 1992, p. 482). It then follows directly from equation (2.1) that $m\mathbf{K}\ (m, md) = \mathbf{K}\ (1, d)$, $\forall m > 0$. □

A.5. **Stackelberg Follower's opportunity set. NTP and Curvature.**
The Follower's opportunity set $F(q_L)$ has Eastern boundary $S(q_L) = \{(m, y) : m = Mq_F, y = Mq_L, q_F \in [0, 24 - q_L]\}$ where $M = 24 - q_L - q_F$. Along this boundary **NTP** and curvature **K** (as in A.1) are given by

$$\begin{aligned}
\mathbf{NTP} &= -\frac{dm/dq_F}{dy/dq_F} = \frac{24 - q_L - 2q_F}{q_L}, \\
\mathbf{K} &= \frac{1}{\sqrt{1 + MRT^2}^3}\frac{d(MRT)}{dm} \\
&= \frac{1}{\sqrt{1 + q_L^2/(24 - q_L - 2q_F)^2}^3}\frac{d(q_L/(24 - q_L - 2q_F))}{dq_F}\frac{dq_F}{dm} \\
&= \frac{2q_L}{\left(\sqrt{(24 - q_L - 2q_F)^2 + q_L^2}\right)^3}.
\end{aligned}$$

Note that **NTP** varies smoothly from positive to negative values as increasing q_F passes through $q_F^o = 12 - q_L/2$, the selfish best response. At corners $(0, 0)$ and $(0, (24 - q_L)q_L)$ curvature is not defined.

A.6. **MGT is a partial ordering.**

Proof. First note that

(1) $\rho(m) - m$ is a (weakly) positive increasing function of m. This follows from $\rho(m) = \rho(0) + \rho'(z)m \geq m$, for some $z \in [0, m]$, and $(\rho(m) - m)' = \rho'(m) - 1 \geq 0$.
(2) $\vartheta(z) \leq \rho(z)$ for all z, by definition.

Now we verify the three properties of a partial order.

: reflexivity, F **MGT** F. Simply take $\rho(z) = \vartheta(z) = z$.

: transitivity, G **MGT** F and H **MGT** G implies H **MGT** F. By definiton of **MGT** for $C = H, G$ there exist $\mathbb{T}^C : (m, y) \longmapsto (\rho_c(m), \vartheta_c(y))$ such that $\rho_c(0) \geq 0$, $\rho_c'(m) \geq 1$, $\vartheta_c(z) \leq \rho_c(z)$ for all z, and $G = \mathbb{T}^G F$, $H = \mathbb{T}^H G$. Take $\mathbb{T} = \mathbb{T}^H \circ \mathbb{T}^G : \Re_+^2 \longmapsto \Re_+^2$ such that $(m, y) \longmapsto (\rho_H \circ \rho_G(m), \vartheta_H \circ \vartheta_G(y))$. Verify that $\mathbb{T}$ has properties required in Definition 2: $H = \mathbb{T}^H G = \mathbb{T}^H \mathbb{T}^G F$ and $\rho'(m) \geq 1$ are straightforward. Now $\vartheta(z) = \vartheta_H \circ \vartheta_G(z) \leq \rho_H \circ \vartheta_G(z) \leq \rho_H \circ \rho_G(z) = \rho(z)$ follows from (2) and monotonicity of ρ_H. Finally $\rho(0) = \rho_H \circ \rho_G(0) \geq \rho_G(0) \geq 0$, by (1) and $\rho_H(0) \geq 0$.

: antisymmetry, G **MGT** F and F **MGT** G implies $F = G$. By definition of **MGT** there exist transformations $\mathbb{T} = (\rho, \vartheta)$ and $\mathbb{T}^* = (\rho^*, \vartheta^*)$ such that $G = \mathbb{T}F$ and $F = \mathbb{T}^* G$. Let $m^C = \sup\{m : \exists y \ s.t. \ (m, y) \in C\}$ be the largest feasible own income in $C = F, G$. Note that G **MGT** F and the definition of m^G imply $m^F \leq \rho(m^F) \leq m^G$, while F **MGT** G and the definition of m^F imply $m^G \leq \rho^*(m^G) \leq m^F$. Hence $m^F = m^G$. Using the definition of m^C we conclude that $\rho(m^C) - m^C = 0$ for $C = F, G$. Statement (1) now tells us that $\rho(m) = m = \rho^*(m)$, for all $m \in [0, m^C]$. Statement (2) shows that in fact $\mathbb{T}$ maps an arbitrary point $(m, y) \in F$ into the point $(m, \vartheta(y)) \in G$ with $\vartheta(y) \leq y$ whereas $\mathbb{T}^*$ maps some point $(m, z) \in G$ into the point (m, y) with $y = \vartheta^*(z) \leq z$. Convexity of G implies $(m, y) \in G$, and therefore $F \subset G$. Likewise for $G \subset F$.

□

A.7. **Examples of MGT-ordered Opportunity Sets.**
Opportunity sets in the Investment game are MGT ordered by s. We show that $F(s^*)$ **MGT** $F(s)$ for all $s^* > s \geq 0$.

Proof. Recall that $\partial_E F_s = \{(m, y) : m = I + ks - r, y = I - s + r, r \in [0, ks]\}$. Let $\rho, \vartheta : \Re_+ \longrightarrow \Re_+, \rho(m) = (m - I)s^*/s + I$, if $m \geq I$; and $\rho(m) = m$, if $m < I$; and $\vartheta(y) = (y - I)s^*/s + I$. It can be easily verified that $\rho'(m) \geq 1$, $\rho(0) = 0$ and

$\rho(z) \geq \vartheta(z)$, for all z. So $\mathbb{T} : (m, y) \longmapsto (\rho(m), \vartheta(y))$ has the definitive properties in Definition 2.[9]

To verify that $F_{s^*} = \mathbb{T}F_s$, take an arbitrary point $P = (P_1, P_2) \in \mathbb{T}\partial_E F_s$. Then $P_1 = \rho(I + ks - r) = (ks - r)s^*/s + I$; $P_2 = \vartheta(I - s + r) = (-s + r)s^*/s + I$ for some $r \in [0, ks]$. Note that $r^* = \frac{s^*}{s} r \in [0, ks^*]$ iff $r^* \in [0, ks^*]$ and that $P_1 + P_2 = (k-1)s^* + 2I$. Hence, $\mathbb{T}$ maps $\partial_E F_s$ to $\partial_E F_{s^*}$. Recall that all other points in F_s are of the form (aP_1, P_2), for some $a \in [0, 1]$. Since $\rho(0) = 0$ and ρ is continuous and increasing, we see that $\mathbb{T}$ is well-behaved everywhere, indeed mapping F_s onto F_{s^*}. □

Opportunity sets in the Stackelberg duopoly game are MGT ordered by q_L. We show that $F(q_L^*)$ **MGT** $F(q_L)$ for all $q_L > q_L^* \geq 0$.

Proof. Given $q_L > q_L^*$ define $\rho, \vartheta : \Re_+ \longrightarrow \Re_+$ by

$$\rho(t) = \left(\frac{24 - q_L^*}{24 - q_L}\right)^2 t; \ \vartheta(t) = \frac{\left(24 - q_L^*\right) q_L^*}{\left(24 - q_L\right) q_L} t$$

and verify that $\rho(0) = 0$, $\rho'(t) = [\left(24 - q_L^*\right) / (24 - q_L)]^2 > 1$ and $\vartheta(z) \leq \rho(z)$, for all z. Transformation $\mathbb{T} : \Re_+^2 \to \Re_+^2$ s. t. $(m, y) \longmapsto (\rho(m), \vartheta(y))$ has all properties required for $\mathbb{T}$ in Definition 2. Hence to complete the proof all we need to show is that $\mathbb{T}F(q_L) = F(q_L^*)$. As in the previous proof, it sufficies to show that $\mathbb{T}\partial F(q_L) = \partial F(q_L^*)$. Let P be an arbitrary point from $\mathbb{T}\partial F(q_L)$. This implies $\exists q_F \in [0, 24 - q_L]$ s.t.

$$P_1 = \rho(Mq_F) = \left(\frac{24 - q_L^*}{24 - q_L}\right)^2 Mq_F; P_2 = \vartheta(Mq_L) = \frac{\left(24 - q_L^*\right) q_L^*}{\left(24 - q_L\right) q_L} Mq_L$$

Write P_1 and P_2 as

$$P_1 = \left(24 - q_L^* - \frac{24 - q_L^*}{24 - q_L} q_F\right) \frac{24 - q_L^*}{24 - q_L} q_F = M^* q_F^*,$$

$$P_2 = \left(24 - q_L^* - \frac{24 - q_L^*}{24 - q_L} q_F\right) q_L^* = M^* q_L^*$$

where $q_F^* = q_F(24 - q_L^*)/(24 - q_L)$ and $M^* = 24 - q_L^* - q_F^*$. It can be verified that q_F^* as just defined is from $[0, 24 - q_L^*]$ iff $q_F \in [0, 24 - q_L]$, hence $P \in \partial F(q_L^*)$. Similarly for $\mathbb{T}\partial F(q_L) \supset \partial F(q_L^*)$. □

[9]Smoothness fails along the line $(m, y) = (I, y)$. This can be patched either by relaxing the definition slightly to require that the transformation be smooth almost everywhere and continuous, or alternatively by smoothing the transformation used here using a standard partition of unity construction, as in Rudin (1966, p. 40).

A.8. **Proposition 1.** *Suppose that either preferences A or the opportunity set F are strictly convex, and let X be the $A-$chosen point in F. Then $X \in \partial_E F$.*

Proof. Suppose that $X = (m, y) \notin \partial_E F$. Then by definition of $\partial_E F$ there exists $z > m$ such that $M = (z, y) \in F$. Positive monotonicity in own payoff implies that M is strictly preferred to X, contradicting the hypothesis that X is the $\mathcal{A}$-preferred point in F. □

A.9. **Proposition 2. Theoretical predictions for fixed opportunity set.** *Suppose that either preferences A and B, or the opportunity set F, are strictly convex. Let $X_{\mathcal{A}}$ and $X_{\mathcal{B}}$ be the points in F chosen when preferences are respectively $\mathcal{A}$ and $\mathcal{B}$. Then*

(1) *$\mathcal{B}$ **MAT** $\mathcal{A}$ implies $d_{X_{\mathcal{B}}} \geq d_{X_{\mathcal{A}}}$ for homothetic preferences.*
(2) *If $X \in \partial_E F$ lies on a ray with slope between $d_{X_{\mathcal{A}}}$ and $d_{X_{\mathcal{B}}}$, then there are preferences P with $\mathcal{B}$ **MAT** $\mathcal{P}$ **MAT** $\mathcal{A}$ such that X is the $\mathcal{P}$-chosen point in F.*
(3) There are admissible preferences for which the chosen point is arbitrarily close to S_F, and other admissible preferences for which the chosen point is arbitrarily close to N_F.

Proof. First, recall that for homothetic preferences (i.e. **WTP** is constant on a ray) strict convexity is equivalent with **WTP** decreasing as d increases. Formally, for preferences $\mathcal{J}$, $\forall Y, Z \in \partial_E F$, $d_Y < d_Z$ iff

$$\mathbf{WTP}_{\mathcal{J}}(Y) > \mathbf{WTP}_{\mathcal{J}}(Z) \tag{A.2}$$

Recall as well that along $\partial_E F$, **NTP** increases as d increases (by convexity of F.)

Part 1. Suppose that $\mathcal{B}$ **MAT** $\mathcal{A}$. This and optimality of $X_{\mathcal{A}}$ imply

$$\mathbf{NTP}(X_{\mathcal{A}}) = \mathbf{WTP}_{\mathcal{A}}(X_{\mathcal{A}}) \leq \mathbf{WTP}_{\mathcal{B}}(X_{\mathcal{A}})$$

Since all points from F in lower rays than $X_{\mathcal{A}}$ have larger **WTP** (see (A.2)) and not smaller **NTP** (by convexity of F) than $X_{\mathcal{A}}$ none of them is among the $\mathcal{B}-$preferred point. Hence, $d_{X_{\mathcal{B}}} \geq d_{X_{\mathcal{A}}}$.

Part 2. *Let* X be given such that $d_{X_a} < d_X < d_{X_b}$. Let w_a and w_b denote continuous **WTP** functions for $\mathcal{A}$ and $\mathcal{B}$ preferences. If $w_b(X) = w_a(X)$ then X is the chosen point for both $\mathcal{A}$ and $\mathcal{B}$ preferences. If $w_b(X) > w_a(X)$ then consider some preferences $\mathcal{P}$ such that, for all Y, $w_p(Y) = kw_b(Y) + (1-k)w_a(Y)$

where

$$k = \frac{\mathbf{NTP}(X) - w_a(X)}{w_b(X) - w_a(X)}.$$

There exists a utility function with $w_P(Y)$ since the latter is continuous (Hurewicz, 1958, p. 7-10; see also Hurwicz and Uzawa, 1971). Let $\mathcal{P}$ denote preferences being represented by this utility function. $\mathcal{B}$ **MAT** $\mathcal{P}$ **MAT** $\mathcal{A}$ follows from k between 0 and 1 (implied by $d_{X_a} < d_X < d_{X_b}$). X is $\mathcal{P}$-chosen since straightforwardly, $w_P(X) = \mathbf{NTP}(X)$.

Part 3. Linear preferences with w going to $-\infty$ $(+\infty)$ have the chosen point arbitrarily close to S_F (N_F). □

A.10. **Proposition 3. Theoretical predictions for different opportunity set.** *Let an agent with strictly convex and homothetic preferences A choose X and W from some opportunity sets F and G, respectively. Let $Y = tX \in \partial G$ be the most distant point from the origin on the ray through X in the opportunity set G, and let $Z \in \partial_E G$ solve $\mathbf{NTP}_{\partial F}(X) = \mathbf{NTP}_{\partial G}(Z)$. Then, for preferences A,*

1. *if $X \in G$ then $W \in F^c$ or $W = X$,*
2. *$d_W \leq d_X$ iff $\mathbf{NTP}_{\partial F}(X) \leq \mathbf{NTP}_{\partial G}(Y)$, and*
3. *$d_W \geq d_Z$ iff $d_Z \leq d_X$.*

Proof. Suppose that X is a regular point from $\partial_E F$. Then $x = \mathbf{NTP}(X)$ is unique. Let the **NTP** of points from $\partial_E G$ take values between $[\gamma_*, \gamma^*]$. Z is: N_F, if $\mathbf{NTP}(X) > \gamma^*$, S_F, if $\mathbf{NTP}(X) < \gamma_*$; otherwise Z is the point of $\partial_E G$ with $x \in \mathbf{NTP}(Z)$. Such a point exists by the Intermediate Value Theorem and is unique because G is convex. If X is not a regular point then $\mathbf{NTP}(X)$ takes values from some $[\delta_*, \delta^*]$. Make the arbitary convention that $x = \delta^*$ and proceed as with a regular point.

Part 1. Follows from standard revealed preference theory e.g., Varian (1992) p. 131-133.

Part 2. $d_W \geq d_X$ is equivalent with *(a)* $\mathbf{NTP}_{\partial G}(Y) \leq \mathbf{NTP}_{\partial G}(W)$ by convexity of G and construction of Y; and *(b)* $\mathbf{WTP}(W) \leq \mathbf{WTP}(Y)$ from (A.2). On the other hand, by construction of Y and homotheticity *(c)* $\mathbf{WTP}(\mathbf{Y}) = \mathbf{WTP}(\mathbf{X})$, and *(d)* $\mathbf{WTP}(\mathbf{X}) = \mathbf{NTP}_{\partial F}(\mathbf{X})$, $\mathbf{NTP}_{\partial G}(W) = \mathbf{WTP}(W)$ since X, W are the most preferred points in respectively, F, G. $\mathbf{NTP}_{\partial G}(Y) \leq \mathbf{NTP}_{\partial F}(\mathbf{X})$ follows from (a)-(d) and transitivity.

Part 3. $d_W > d_Z$ iff $d_Z < d_X$. Referring to (A.2), $\mathbf{WTP}(Y) < \mathbf{WTP}(Z)$ iff $(d_X =)d_y > d_Z$ and

$$\begin{aligned}\mathbf{WTP}(Y) = \mathbf{WTP}(X) = \mathbf{NTP}_{\partial F}(X) = \mathbf{NTP}_{\partial G}(Z) \leq \mathbf{NTP}_{\partial G}(W) \\ = \mathbf{WTP}(W) < \mathbf{WTP}(Z)\end{aligned}$$

where the first and the third equalities are true by construction of Y and Z, the second and the fourth equalities follow from the optimality of X and W, whereas the first and the last inequalities are, for convex opportunity sets and preferences, equivalent with $d_W > d_Z$. □

A.11. **Proposition 4. Investment Game.** *Let the FM in the Investment game choose F_s as the SM's opportunity set, and let $r(s)$ be the SM's response. Also let the same SM be given the same opportunity set F_s in a dictator game, and let $r^o(s)$ be his response there. Assume that* $\mathbf{WTP} \leq 1$ *and* $\mathbf{WTP}_m \geq 0$. *Then:*

(1) *convexity implies that $r^o(s)$ is increasing in s;*
(2) Axiom **R** *implies that $r(s)$ is increasing in s;*
(3) *Axiom* **S** *implies that $r(s) \geq r^o(s)$ for $s = 0, 1, 2, ..., 10$.*

Proof. **Part 1.** To streamline notation, let $w = \mathbf{WTP}(m, y)$, where $m = 10 + 3s - r(s)$ and $y = 10 - s + r(s)$. By hypothesis, $w \leq 1$ and (a) $w_m \geq 0$. By Lemma A.3 below, strict convexity implies that $w_m w - w_y > 0$. It follows that (b) $w_m - w_y > 0$. Since $\mathbf{NTP} = 1$, a constant, along the Eastern boundary of any opportunity set determined by s, the first order condition for optimality requires w also to remain constant. Therefore

$$\begin{aligned}0 = \frac{dw}{ds} &= w_m \frac{dm}{ds} + w_y \frac{dy}{ds} \\ &= [(3 - dr/ds)w_m + (-1 + dr/ds)w_y] \\ &= 2[w_m] + (1 - dr/ds)\,[w_m - w_y]\,.\end{aligned}$$

The bracketed expressions are positive by (a) and (b) above, so we must have $dr/ds > 1$ for choices not at the corner. For corner choices and weak convexity, the argument allows only to conclude that $r(s)$ is nondecreasing.

Part 2. Applying Axiom **R** in the argument above, we see that it increases dr/ds, so the preceding argument holds *a fortiori*.

Part 3. Axiom **S** has the indicated impact since, as shown in the previous subsection, F_s is **MGT** ordered by s. □

Lemma A.3. *Strict convexity implies that $w_m w - w_y > 0$.*

Proof. Recall that $\mathbf{K} < 0$. We claim that

$$\mathbf{K} = \frac{(w_y - w_m w)}{(1+w^2)^{3/2}},$$

from which the lemma follows immediately. To verify the claim, simply substitute

$$w = \frac{u_y}{u_m}, w_m = \frac{u_{ym}u_m - u_{mm}u_y}{u_m^2}, w_y = \frac{u_{yy}u_m - u_{my}u_y}{u_m^2}$$

into the above expression for $\mathbf{K}$ and recover equation (2.1). □

A.12. **Proposition 5. Stackelberg Duopoly Game.** *In the standard Stackelberg game of Example 3.9 let $Q_D(q_L) = q_F - q_F^o$ be the deviation of the Follower's output choice q_F from the selfish best reply $q_F^o = 12 - \frac{1}{2}q_L$ when the Leader chooses output q_L. One has*

$$\frac{dQ_D}{dq_L} = -\frac{1}{2}w - \frac{dw}{dq_L}q_L$$

where $w = WTP(Mq_F, Mq_L)$ is willingness to pay at the chosen point. Furthermore,

(1) *If Follower's preferences A are fixed and linear, then w is constant with respect to q_L and $\frac{dQ_D}{dq_L}$ is positive if and only if preferences at the chosen point are malevolent.*
(2) *If Follower's preferences A are fixed and convex, then w is decreasing in q_L and $\frac{dQ_D}{dq_L}$ contains an additional term that is positive provided that $q_L \leq 12$, $w \leq 1$, $w_m \geq 0$ and $w_m + w_y \geq 0$.*
(3) *If Follower's preferences satisfy Axiom **R** strictly, then w is decreasing in q_L and $\frac{dQ_D^r}{dq_L}$ contains an additional positive term.*
(4) *If Follower's preferences satisfy Axiom **S** strictly, then WTP is decreasing in q_L and $\frac{dQ_D^s}{dq_L}$ has an additional positive (negative) term if the status quo is smaller (larger) than q_L.*

Proof. The FOC can be written as $w(q_F, q_L) \equiv \mathbf{WTP}(Mq_F, Mq_L) = \mathbf{NTP} = \frac{24-2q_F}{q_L} - 1$, which can be rewritten as

$$q_F = 12 - \frac{w(q_F, q_L) + 1}{2}q_L. \tag{A.3}$$

Inserting the definition of Q_D from the statement of the proposition, we obtain

$$Q_D = -\frac{w(q_F, q_L)}{2}q_L. \tag{A.4}$$

Part 1. *Linear preferences* If Follower's preferences are fixed and linear with **WTP** $= w$ then differentiation of (A.4) with respect to q_L gives

$$\frac{dQ_D}{dq_L} = -\frac{w}{2}.$$

Part 2. *Convex Preferences* If Follower's preferences are fixed and convex then

$$\frac{dQ_D}{dq_L} = -\frac{w(q_F, q_L)}{2} - \frac{q_L}{2}\frac{dw(q_F, q_L)}{dq_L}.$$

The additional (second) term above is positive because, as we now will verify, $\frac{dw(q_F,q_L)}{dq_L}$ is negative. Indeed,

$$\begin{aligned}\frac{dw(q_F, q_L)}{dq_L} &= w_m\frac{dm}{dq_L} + w_y\frac{dy}{dq_L}\\ &= w_m((-1-\frac{dq_F}{dq_L})q_F + M\frac{dq_F}{dq_L}) + w_y((-1-\frac{dq_F}{dq_L})q_L + M)\end{aligned}$$

which after substituting $M = 24 - q_L - q_F$, $q_F = 12 - \frac{w(q_F,q_L)+1}{2}q_L$ and $\frac{dq_F}{dq_L} = -\frac{w(q_F,q_L)+1}{2} - \frac{dw(q_F,q_L)}{dq_L}q_L$ and solving for $\frac{dw(q_F,q_L)}{dq_L}$ we get

$$\frac{dw(q_F, q_L)}{dq_L} = \frac{B}{A}$$

where

$$A = 2 + [w_m w - w_y]\, q_L^2 > 0$$

by Lemma (A.3), and

$$B = 24(w_y - w_m) + q_L(1-w)(w_m - w_y + ww_m - w_y).$$

To sign B, recall that as in the proof of the previous proposition and Lemma (A.3), $w_y - w_m \leq 0$ due to convexity, $w \leq 1$ and $w_m \geq 0$. Suppose $q_L \leq 12$. Then $2q_L \leq 24$ and $w_y - w_m \leq 0$ imply

$$B < -\left(w^2 w_m - 2ww_y + w_m\right) q_L.$$

If $w \in [0,1]$ then the expression in brackets is non-negative (write it as $w(ww_m - w_y + w_m/w - w_y)$ which is larger than $2w(ww_m - w_y) \geq 0$). If however $w < 0$ then the term in brackets is positive since $w_y^2 - w_m^2 \leq 0$. To see this, recall that $w_y + w_m \geq 0$ and $w_y - w_m \leq 0$. Hence B is negative and the additional term is indeed positive.

Part 3. Axiom **R** *Effect.* Let $w^r(q_F, q_L)$ denote **WTP** for changed preferences as in Axiom **R**. Then

$$Q_D^r = -\frac{w_r(q_F, q_L)}{2}q_L$$

for all q_L, and

$$\frac{dQ_D^r}{dq_L} = -\frac{w^r(q_F,q_L)}{2} - \frac{q_L}{2}\frac{dw^r(q_F,q_L)}{dq_L}$$
$$= -\frac{w(q_F,q_L)}{2} - \frac{w^r(q_F,q_L)-w(q_F,q_L)}{2} - \frac{q_L}{2}\frac{dw^r(q_F,q_L)}{dq_L}.$$

From Axiom **R** the second term is positive and similarly as in part 2 the third term is positive if induced preferences are benevolent ($w^r \geq 0$) or malevolent with $w_y^r + w_m^r \geq 0$.

Part 4. *Axiom **S** Effect.* Let $w^s(q_F,q_L)$ denote **WTP** for changed preferences as in Axiom **S.** Then

$$Q_D^c = -\frac{w^s(q_F,q_L)}{2} q_L$$

is smaller (larger) than Q_D^r if the status quo is smaller (larger) than q_L, and

$$\frac{dQ_D^s}{dq_L} = -\frac{w^s(q_F,q_L)}{2} - \frac{q_L}{2}\frac{dw^s(q_F,q_L)}{dq_L}$$

has an additional positive (negative) term if the status quo is smaller (larger) than q_L. □

A.13. **Corollary 1. Stackelberg Mini-Game.** *For* $0 < x < s < z < 24$, *suppose the Stackelberg Leader has restricted output choices* $q_L \in \{x, s\}$ *in situation (a) and* $q_L \in \{z, s\}$ *in another situation (b). Let the Leader choose* s *in both situations. If Follower's preferences satisfy Axioms **R** and **S** then* $Q_D^a(s) \geq Q_D^b(s)$, and *at each possible Follower choice* q_F, $\mathbf{WTP}^a(Mq_F, Ms) \leq \mathbf{WTP}^b(Mq_F, Ms)$.

Proof. In situation (a) induced preferences $\mathcal{A}_{F_s}^a$ are $\mathcal{A}_{F_s^c}^a$ or $\mathcal{A}_{F_s^*}^a$ depending on whether output x is considered as status quo by the Follower. Axiom **S** implies $\mathcal{A}_{F_s^o}^a$ **MAT** $\mathcal{A}_{F_s^c}^a$ and $\mathcal{A}_{F_s^o}^a$ **MAT** $\mathcal{A}_{F_s^*}^a$. Similarly, in situation (b) Axiom **S** implies $\mathcal{A}_{F_s^c}^b$ **MAT** $\mathcal{A}_{F_s^o}$ and $\mathcal{A}_{F_s^*}^b$ **MAT** $\mathcal{A}_{F_s^o}$. By transitivity $\mathcal{A}_{F_s}^b$ **MAT** $\mathcal{A}_{F_s}^a$. Then the the last inequality is straightforward by definition of **MAT** whereas for the first one recall that: (i) q_F^o stays constant (it depends only on s), and **NTP** along ∂F_s decreases as q_F increases. □

A.14. **Alternative Regressions.** The last proof suggests alternative specifications for the HMN regressions. Let $a = w_m w - w_y$, $b = w_y - w_m$ and $c = (1-w)(w_m - w_y + w w_m - w_y)$. Recall from the proof that $a \geq 0, b \leq 0$

and $c \geq 0$ for all q_L , and

$$\begin{aligned} \frac{dw(q_F, q_L)}{dq_L} &= \frac{24b}{2 + aq_L^2} + \frac{c}{2 + aq_L^2} q_L, \\ \frac{dQ_D}{dq_L} &= -\frac{w(q_F, q_L)}{2} - \frac{12b}{2 + aq_L^2} q_L - \frac{c}{2(2 + aq_L^2)} q_L^2. \end{aligned}$$

The first order Taylor expansion $Q_D(q_L) \approx C + \frac{dQ_D}{dq_L} q_L$ then suggests fitting Q_D to a cubic expression in q_L,

$$Q_D = \beta_0 + \beta_1 q_L + \beta_2 q_L^2 + \beta_3 q_L^3 + u_i + \varepsilon_i.$$

The predicted coefficient signs then are $\beta_2 > 0, \beta_3 < 0$, consistent with the results reported in Table 3 below. Likewise, $\frac{dw}{dq_L} \approx \gamma_1 + \gamma_2 q_L$ suggests the quadratic specification $\mathbf{WTP} = \beta_0 + \beta_1 q_L + \beta_2 q_L^2 + u_i + \varepsilon_i$. The predictions $\beta_2 > 0, \beta_1 < 0$ are consistent with the results but are not significant in this specification. Allowing asymmetric responses to Leader choices more or less generous than Cournot produces more significant estimates, as reported in the last column.

$Dep.Variable$	Q_D	$\mathbf{WTP} \times 100$	$\mathbf{WTP} \times 100$
q_L	$-2.13 \pm 0.93^{0.013}$	$-6.07 \pm 5.24^{0.124}$	$-3.98 \pm 5.19^{0.222}$
$DP \times q_L$			$2.62 \pm 0.92^{0.003}$
q_L^2	$0.29 \pm 0.11^{0.005}$	$0.08 \pm 0.27^{0.385}$	$0.09 \pm 0.26^{0.365}$
q_L^3	$-0.01 \pm 0.004^{0.005}$		
$constant$	$4.41 \pm 2.63^{0.048}$	$28.26 \pm 24.87^{0.129}$	$0.39 \pm 26.29^{0.494}$

Table 3. Panel Regressions with fixed effects. Data consist of 220 choices by 22 Followers in HMN experiment. One-sided p-values are reported and $\pm$ refers to standard error.

Appendix B. Instructions

Welcome

This is an experiment about decision-making. You will be paid a $5 participation fee plus an additional positive or zero amount of money determined by the decisions that you and the other participants make, as explained below. Payment is in cash at the end of the experiment. A research foundation has provided the funds for this experiment.

No Talking Allowed

Now that the experiment has begun, we ask that you do not talk. If you have a question, please raise your hand and an experimenter will approach you and answer your question in private.

A Monitor and Two Groups

A monitor will be selected randomly from among those of you who came here today. The rest of you have been divided randomly into two groups, called the First Mover Group and the Second Mover Group.

Complete Privacy

The experiment is structured so that no one — not even the experimenters, the monitor, and the other subjects — will ever know your personal decision in the experiment. You collect your cash payment from a staff person in the Economics Department office who has no other role in the experiment. Your payment is in a sealed envelope with a code letter (A, B, C, etc). Your privacy is guaranteed because neither your name nor your student ID number will appear on any decision records. The only identifying mark on the decision forms will be a code letter known only to you. You will show your code letter to the staff person and nobody else will see it. The experimenters will not be in the department office when you collect you cash payment. This procedure is used to protect your privacy.

The Idea of the Game

The game involves two players, called the First Mover (FM) and the Second Mover (SM), in the roles of producers of an identical good. Each decides how much to produce. The profit for each player is the number of units he decides to produce times price, net of cost. The price of the good decreases as total production increases. If you and the other player produce too much, you will drive down the price and your profits. Of course, if you don't produce much you won't have many units to sell.

To simplify your task, the profits will be calculated for you and shown in an easy-to-read table. Your cash payment will include the profit you earn in one round of the game. The round will be selected randomly at the end of the experiment.

Game Details

Each round the FM chooses between two possible amounts to produce, as shown in a table with two rows. The SM sees the choice of the FM, and then decides among seven possible amounts to produce, as shown in seven columns of the same table. The table shows the profits for both players. The FM's profit is shown in italics in the lower left corner of each box, and the SM's profit is shown in bold in the upper right corner. For example, in Table B.1 below, if FM chooses

Output=6 and SM then chooses Output=4, then FM's profit is 84 and SM's profit is 56.

Table B.1

	SM's Choice of Output Quantity:							
	4	**5**	**6**	**7**	**8**	**9**	**10**	**11**
FM's Output=6	*84* **56**	*78* **65**	*72* **72**	*66* **77**	*60* **80**	*54* **81**	*48* **80**	*42* **77**
FM's Output=9	*99* **44**	*90* **50**	*81* **54**	*72* **56**	*63* **56**	*54* **54**	*45* **50**	*36* **44**

Different Subject Pairs in Every Decision

Each First Mover and each Second Mover will make four decisions. But the pairing of First Movers with Second Movers will be different in every decision. This means that you will interact with a DIFFERENT person in the other group in every decision that you make.

Experiment Procedures and the Monitor

At the beginning of the experiment, the monitor will walk through the room carrying a box containing unmarked, large manila envelopes. Each subject in the First Mover Group will take one of these envelopes from the box. This envelope will contain the experiment decision forms and a code letter.

After the First Movers have made their decisions, they return the experiment decision forms to their large manila envelopes and then walk to the front of the room and deposit the envelopes in the box on the table. It is very important that the First Movers do NOT return their code letters to the large manila envelopes, because they will need them to collect their payoffs.

After all First Movers have deposited their envelopes in the box, the Monitor will take the box to another room in which the experimenters will sort the decision forms and place them in the correct large manila envelopes for the Second Movers. The experimenters will also put code letters in the envelopes for the Second Movers.

Next, the Monitor will walk through the room carrying a box containing unmarked, large manila envelopes. Each subject in the Second Mover Group will take one of these envelopes from the box. This envelope will contain the experiment decision forms and a code letter.

After the Second Movers have made their decisions, they return the experiment decision forms to their large manila envelopes and then walk to the front of the room and deposit the envelopes in the box on the table. It is very important that the Second Movers do NOT return their code letters in the large manila envelopes because they will need them to collect their payoffs.

After all Second Movers have deposited their envelopes in the box, the Monitor will take the box to another room in which the experimenters will record the profits and cash payments determined by the subjects' decisions.

A Roll of a Die Determines Which Decision Pays Money

Although you will make four decisions, only one will pay cash. Which of these decisions will pay cash will be determined by rolling a six-sided die. The experimenter will roll the die in front of you and the monitor will announce which of the numbered sides has ended up on top. The first number from 1 to 4 that ends up on top will determine the page number of the decision that pays cash.

The monitor's cash payment will be the average of all First Movers and Second Movers payments.

Be Careful

Be careful in recording your decisions. If a First Mover forgets to circle one of the rows in the table, or circles both rows on the same decision page, then it will be impossible to ascertain what decision the First Mover made. In that case, the First Mover will get paid 0 and the Second Mover will get paid 60 if that decision page is selected for payoff by the roll of the die. If a Second Mover doesn't circle a column, then it will be impossible to ascertain what decision the Second Mover made. In that case, the Second Mover will get paid 0 and the First Mover will get paid 60 if that decision page is selected for payoff by the roll of the die.

Pay Rates

For each point of profit you earn, the experimenter will put a fixed number of dollars in your envelope. This fixed number is called the pay rate and is written on the board at the front of the room. Today's pay rate is $0.25, which means that every participant earns 25 cents for each profit point shown in the table.

Frequently Asked Questions

Q1: Exactly how are profits calculated in the Tables?

A: Price is 30 minus the sum of FM output and SM output. Marginal cost is 6. Profit is output times (price minus marginal cost). But you don't have to worry about doing the calculation; the Tables do it for you.

Q2: Who will know what decisions I make?

A: Nobody else besides you; that is the point of the private envelopes etc. The experimenters are only interested in knowing the distribution of choices for FMs and SMs, not in the private decisions of individual participants.

Q3: Is this some psychology experiment with an agenda you haven't told us?

A: No. It is an economics experiment. If we do anything deceptive, or don't pay you cash as described, then you can complain to the campus Human Subjects

Committee and we will be in serious trouble. These instructions are on the level and our interest is in seeing the distribution of choices made in complete privacy.

Any More Questions?

If you have any questions, please raise your hand and an experimenter will approach you and answer your question in private. Make sure that you understand the instructions before beginning the experiment; otherwise you could, by mistake, mark a different decision than you intended.

Quiz

(1) In Table B.2 below, what are the two possible output choices for the FM?

__

(2) Does the SM see the FM's choice? (Y or N)

(3) In Table B.2, can the SM choose:

   (a) Output=5? __(Y or N)

   (b) Output =7?__(Y or N)

   (c) Output=12?__(Y or N)

(4) Suppose the FM chooses the top row (Output = 9) in Table B.2 and the SM chooses a middle column (Output = 8).

   (a) How many points will the FM get? __ points

   (b) How much money is that if this is the decision that pays money? $ __

   (c) How much will the SM get in this case? __ points, $ __

(5) In the previous question, if SM chose Output=9 instead of Output=8,

   (a) how many more or fewer points would the SM get? __ more/fewer points

   (b) how many more or fewer points would the FM get? __ more/fewer points

(6) If the FM chooses the top row, what is the maximum number of points that the SM can get? __ the minimum number?__

(7) If the FM chooses the bottom row, what is the maximum number of points that the SM can get? __ the minimum number?

(8) Will the SM ever be able to tell which person made any FM choice? (Y or N)

(9) Will the FM ever be able to tell which person made any SM choice? (Y or N)

(10) Will the experimenter ever be able to tell who made any choices? (Y or N)

Table B.2

	SM's Choice of Output Quantity:							
	4	**5**	**6**	**7**	**8**	**9**	**10**	**11**
FM's Output=9	*99* **44**	*90* **50**	*81* **54**	*72* **56**	*63* **56**	*54* **54**	*45* **50**	*36* **44**
FM's Output=12	*96* **32**	*84* **35**	*72* **36**	*60* **35**	*48* **32**	*36* **27**	*24* **20**	*12* **11**

References

[1] Abbink, K. , B. Irlenbusch, and E. Renner (2000): "The Moonlighting Game: An Empirical Study on Reciprocity and Retribution," *Journal of Economic Behavior and Organization* **42**, 265-77.

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

[3] Andreoni, J., M. Castillo and R. Petrie (2003): "What Do Bargainers' Preferences Look Like? Experiments with a Convex Ultimatum Game," *American Economic Review* **93**, 672-685.

[4] Berg, J., J. Dickhaut, and K. McCabe (1995): "Trust, Reciprocity, and Social History," *Games and Economic Behavior* **10**, 122-42.

[5] Bolton, G. E. and A. Ockenfels (2000): "ERC: A Theory of Equity, Reciprocity, and Competition," *American Economic Review* **90**, 166-193.

[6] Bosman, R. and F. van Winden (2002), "Emotional Hazard in a Power-to-Take Experiment," *Economic Journal* **112**, 147-69.

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

[8] Cox, J. C. (2004), "How to Identify Trust and Reciprocity," *Games and Economic Behavior* **46**, 260-281.

[9] Cox, J. C., D. Friedman and S. Gjerstad (forthcoming): "A Tractable Model of Reciprocity and Fairness," *Games and Economic Behavior*.

[10] Cox, J. C. and V. Sadiraj (2004): "Direct Tests of Models of Social Preferences and a New Model," University of Arizona Department of Economics, Discussion Paper.

[11] Dufwenberg, M. and G. Kirchsteiger (2004): "A Theory of Sequential Reciprocity," *Games and Economic Behavior* **47**, 268-98.

[12] Falk, A. and U. Fischbacher (2006), "A Theory of Reciprocity," *Games and Economic Behavior* **54**, 293-315.

[13] Fehr, E. and S. Gaechter (2000): "Fairness and Retaliation: The Economics of Reciprocity," *Journal of Economic Perspectives* **14**, 159-181.

[14] Fehr, E., G. Kirchsteiger, and A. Riedl (1993): "Does Fairness Prevent Market Clearing? An Experimental Investigation," *Quarterly Journal of Economics* **108**, 437-460.

[15] Fehr, E. and K. M. Schmidt (1999): "A Theory of Fairness, Competition, and Cooperation," *Quarterly Journal of Economics* **114**, 817-868.

[16] Forges, F., E. Minelli (2006): "Afriat's Theorem for General Budget Sets," *CESIfo Working Paper, No.* 1703.

[17] Gray, A. (1997): *Modern Differential Geometry of Curves and Surfaces with Mathematica*, 2nd ed. Boca Raton, FL: CRC Press, p. 14-17. see as well, http://mathworld.wolfram.com/Curvature.html.

[18] Guth, W., R. Schmittberger, and B. Schwarze (1982): "An Experimental Analysis of Ultimatum Barga:ining," *Journal of Economic Behavior and Organization* **4**, 367-88.

[19] Hicks, J. (1939). *Value and Capital*. Oxford: Clarendon Press.

[20] HUCK, S., W. MULLER AND H. NORMANN (2001): "Stackelberg beats Cournot: On collusion and efficiency in experimental markets," *Economic Journal* **111**,749-766.

[21] HUREWICZ, W., (1958): *Lectures on Ordinary Differential Equations*. New York: Wiley.

[22] HURWICZ, L., AND H. UZAWA (1971): "On the Integrability of Demand Functions," in J.S.Chipman et al. (eds.), *Studies in the Mathematical Foundations of Utility and Demand Theory*. New York: Harcourt, Brace & World.

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

[24] LIEBRAND, W.B.G. (1984): "The effect of social motives, communication and group sizes on behavior in an n-preson multi stage mixed motive game," *European Journal of Social Psychology* **14**, 239-264.

[25] MATZKIN, R.L. (1991): "Axioms of Revealed Preferences for Nonlinear Choice Sets," *Econometrica* **59**, 1779-1786.

[26] ROCKAFELLAR, R. T. (1970): *Convex Analysis*, Princeton University Press.

[27] RUDIN, W. (1966): *Real and Complex Analysis*, McGraw-Hill, New York, New York, USA.

[28] SAMUELSON, P. (1947): *Foundations of Economic Analysis*. Cambridge, Mass., Harvard University Press.

[29] SIMON, C.P., AND L. BLUME, (1994): *Mathematics for Economists*, W.W. Norton & Company, Inc.

[30] SMITH, A. <1759> (1976): *The Theory of Moral Sentiments*. Indianapolis, Liberty Classics.

[31] SOBEL, J. (2005): "Interdependent Preferences and Reciprocity," *Journal of Economic Literature* **43**, 392-436.

[32] SONNEMANS, J, F. VAN DIJK AND F. VAN WINDEN (2005): "On the Dynamics of Social Ties Structures in Groups," *Journal of Economic Psychology,* forthcoming.

[33] VARIAN, H. R. (1992): *Microeconomic Analysis*, third ed., Norton.

[34] VARIAN, H. R. (1994): "Sequential provision of public goods," *Journal of Public Economics* **53**, 165-186.

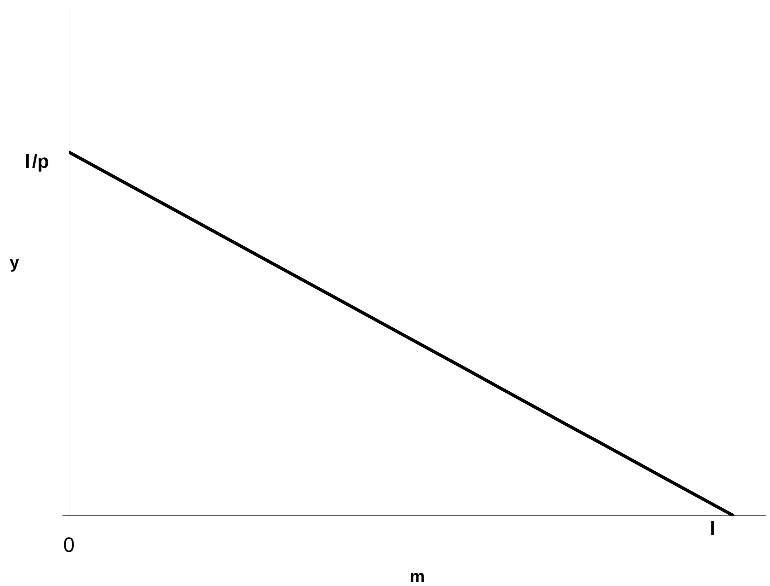

FIGURE 1. Standard Budget Set.

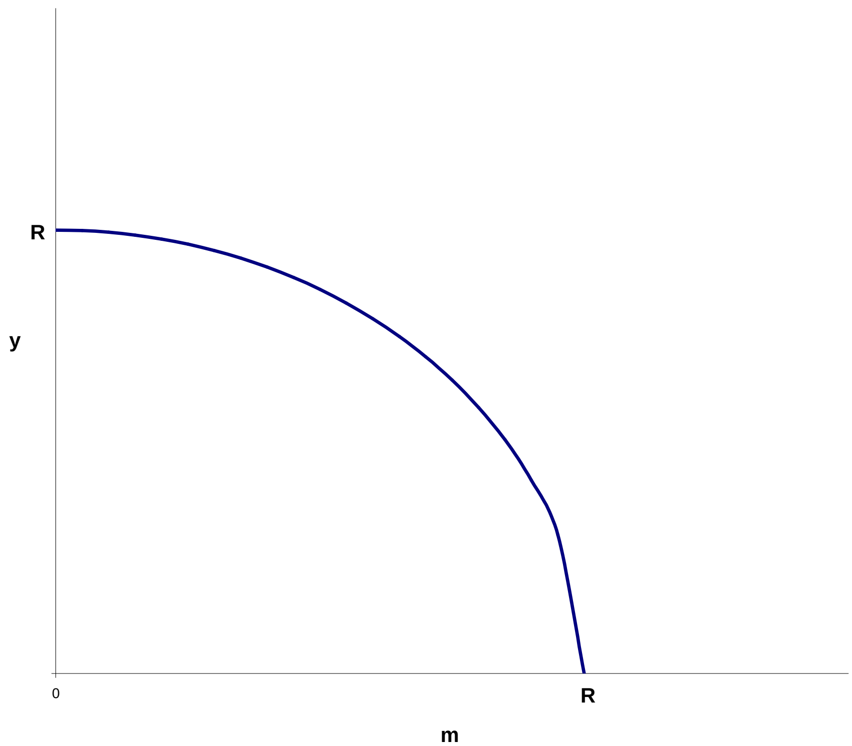

FIGURE 2. Ring test Budget Set.

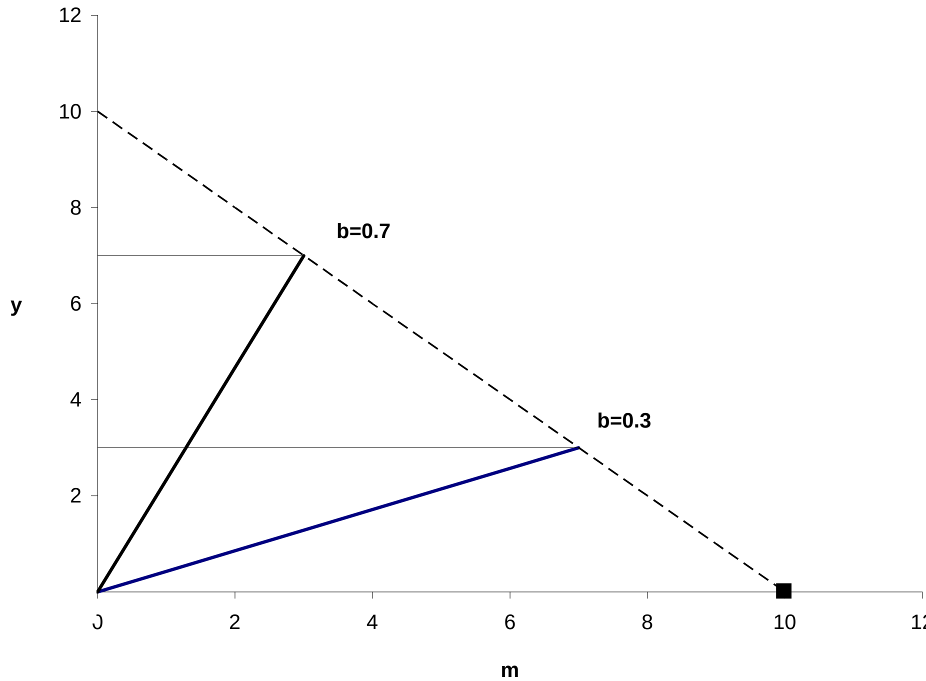

FIGURE 3. Power to Take Game, Second Mover's Opportunity Set.

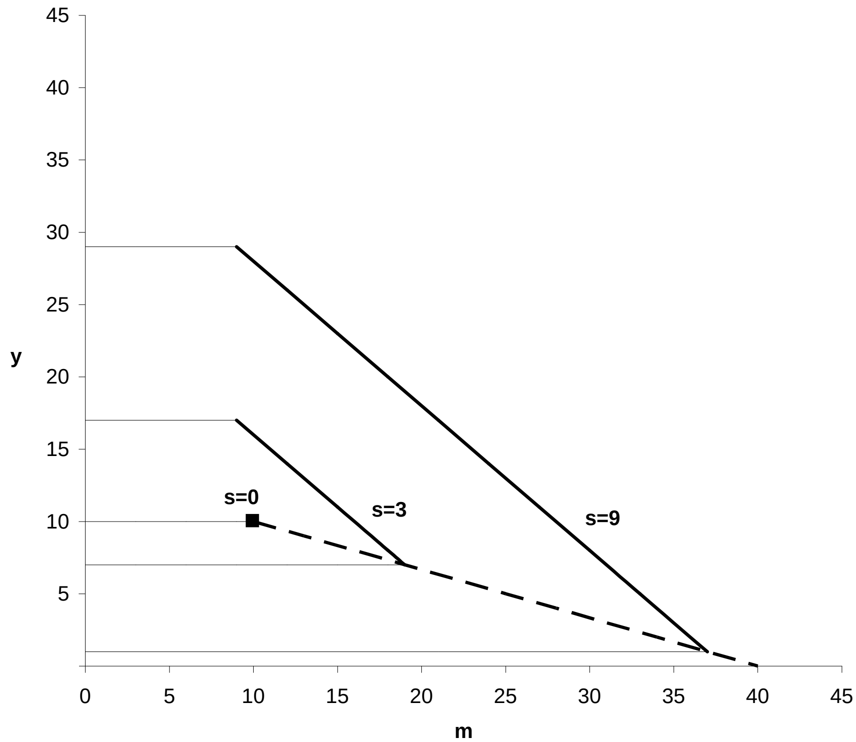

FIGURE 4. Investment Game, Second Mover's Opportunity Set.

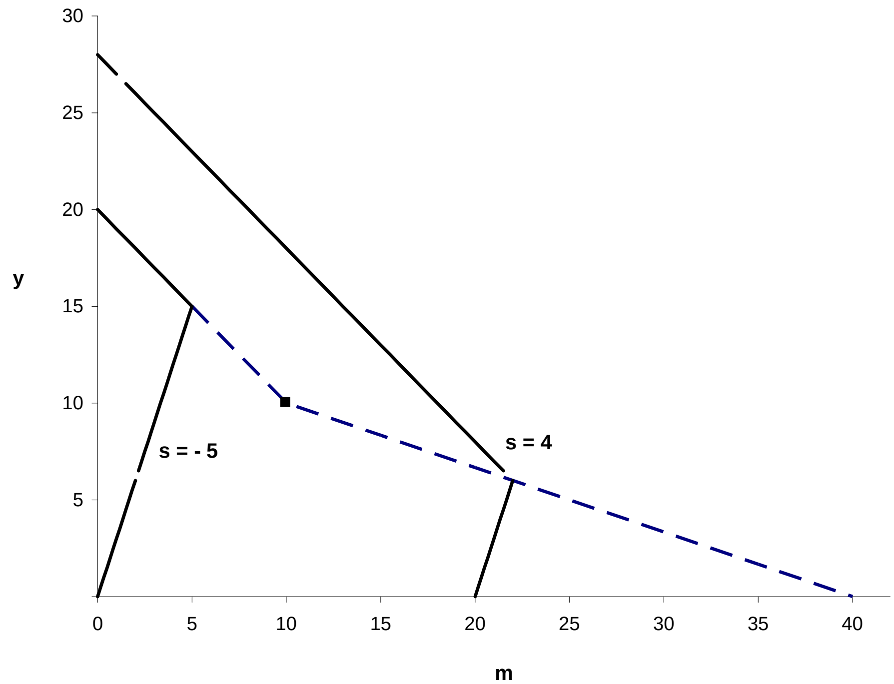

FIGURE 5. Moonlighting Game, Second Mover's Opportunity Set.

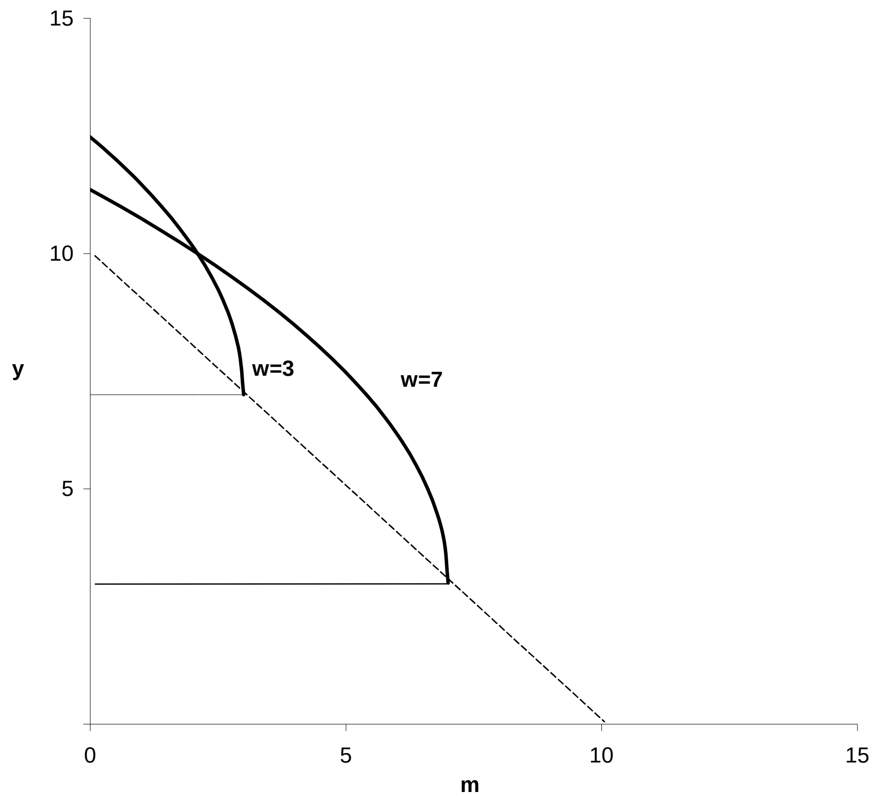

FIGURE 6. Gift Exchange Labor Markets, Second Mover's Opportunity Set.

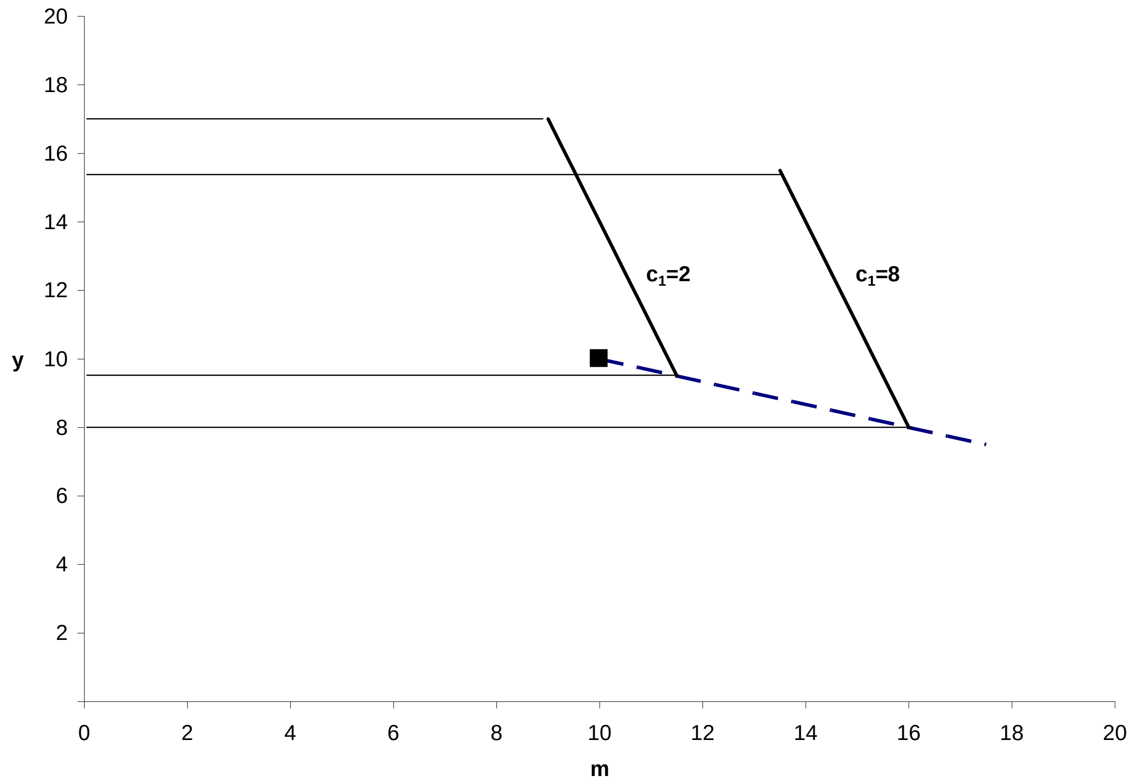

FIGURE 7. Sequential VCM Public Goods Game with two players (a=0.75).

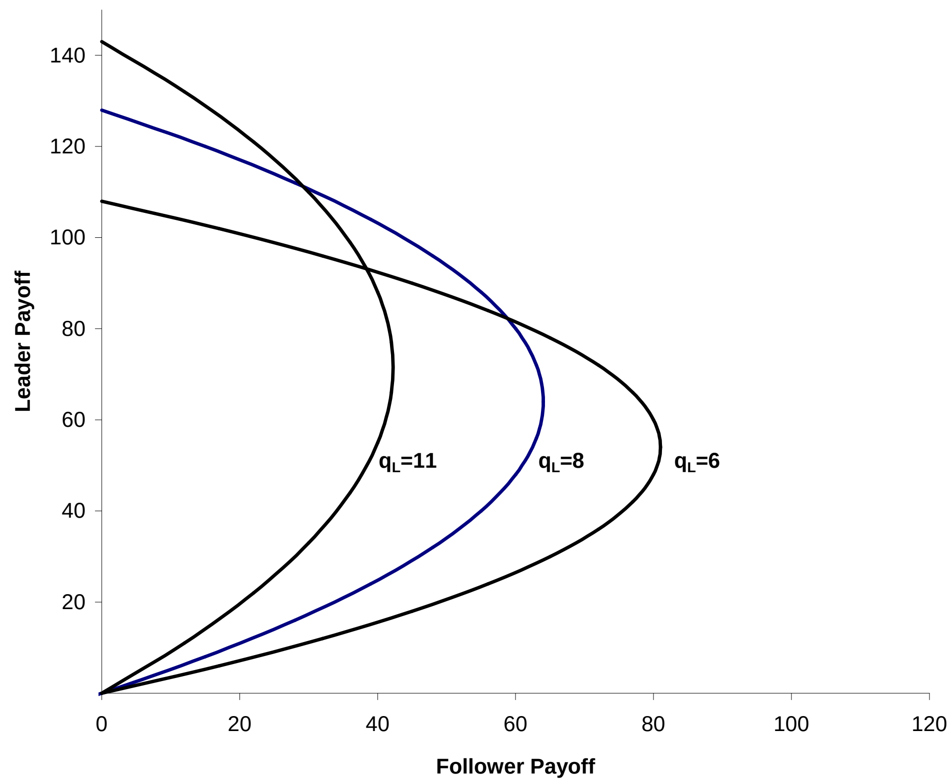

FIGURE 8. Stackelberg Duopoly Game, Follower's Opportunity Set.

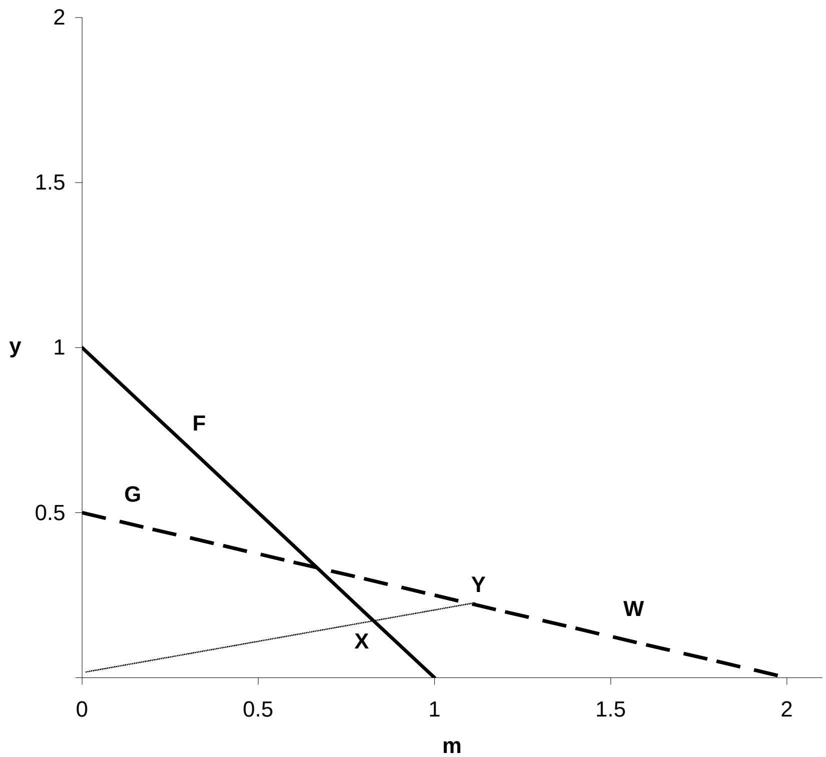

FIGURE 9. Illustration of Example 5.1.

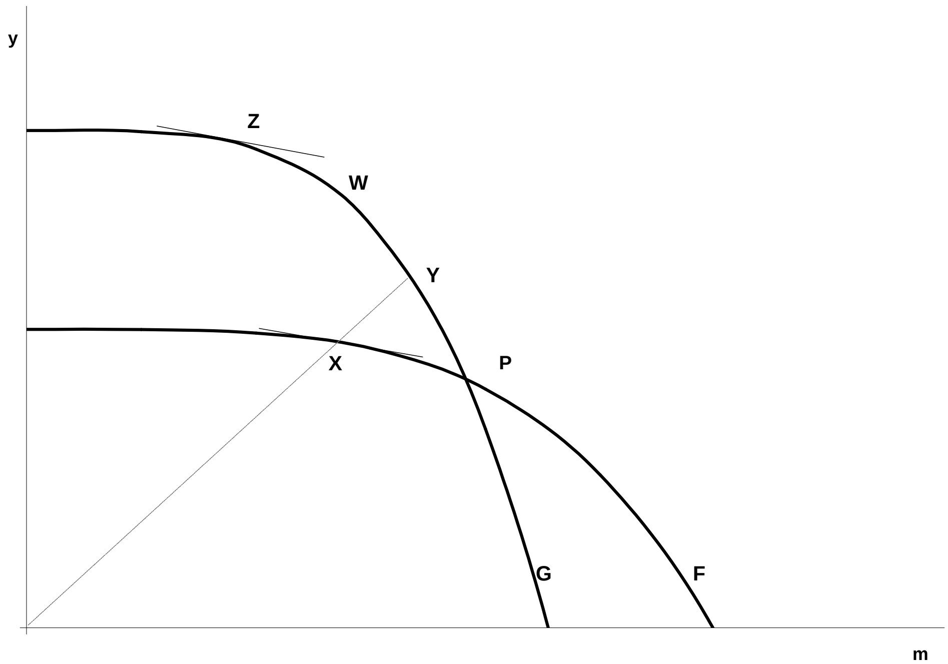

FIGURE 10. Proposition 3 predicts that, with unchanged homothetic preferences, the choice W will lie on the Eastern boundary of opportunity set G between points Y and Z. The prediction (Part (1) of the proposition) is that W is north of point P on $\partial_E G$.

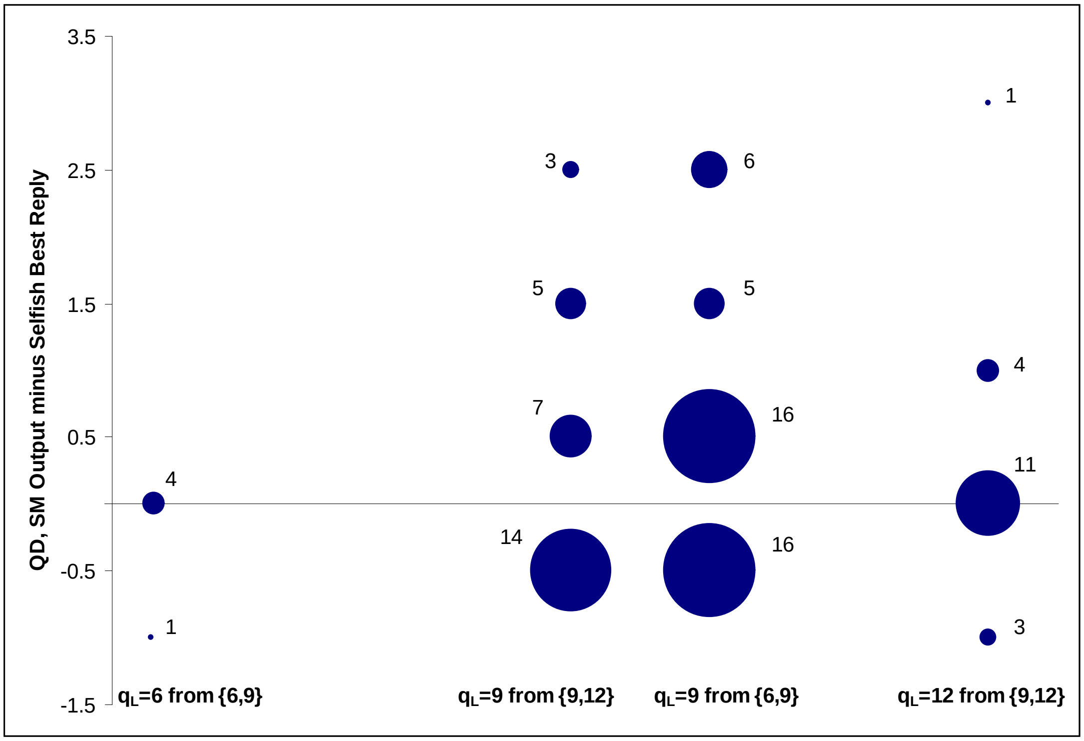

FIGURE 11. Data from Stackelberg mini-game. Deviations Q_D of SM actual output choice from the selfish best response are shown in four situations. From left to right, the situations are: FM chose $q_L = 6$ from $\{6, 9\}$, FM chose $q_L = 9$ from $\{9, 12\}$, FM chose $q_L = 9$ from $\{6, 9\}$, and FM chose $q_L = 12$ from $\{9, 12\}$. The size of the dot reflects the fraction of observations in each situation. The number of observations is shown next to each dot.