

# UC San Diego

## Recent Work

### Title

Manipulation of Preferences and Relative Utilitarianism

### Permalink

<https://escholarship.org/uc/item/1x32m32n>

### Author

Sobel, Joel

### Publication Date

1998-07-01

98-17

**UNIVERSITY OF CALIFORNIA, SAN DIEGO**

DEPARTMENT OF ECONOMICS

MANIPULATION OF PREFERENCES AND RELATIVE UTILITARIANISM

BY

JOEL SOBEL

**DISCUSSION PAPER 98-17  
JULY 1998**

# Manipulation of Preferences and Relative Utilitarianism

JOEL SOBEL\*

July 3, 1998

ABSTRACT. Given  $n$  agents with von Neumann-Morgenstern utility functions who wish to divide  $m$  commodities, consider the  $n$ -person noncooperative game with strategies consisting of concave, increasing von Neumann-Morgenstern utility functions, and whose outcomes are the relative utilitarian solution. It is shown that any constrained equal-income competitive equilibrium allocation for the true utilities is a Nash equilibrium outcome for the noncooperative game. Conditions are presented under which these are the only pure strategy equilibrium outcomes.

## 1. INTRODUCTION

When solution concepts make predictions on the basis of unobservable information, agents will have incentives to distort their private information. The paper adds to a literature that studies the robustness of axiomatic bargaining models to strategic misrepresentation of preference information.

The paper identifies what outcomes can be expected if a social planner selects an allocation that maximizes a weighted average of the agents' reported utility functions. It studies a particular distortion game. There is a fixed quantity of a finite number of commodities. Players' strategies consist of utility functions defined over these commodities. Given the reported utility functions, the social planner provides an allocation that maximizes the weighted sum of reported preferences according to the relative utilitarian solution (an agent's weight is equal to the inverse of his maximum utility gain). Motivation for this solution and axiomatic derivations for it, are provided in recent work by Dhillon and Mertens [1996] and Segal [1998]. If reports are limited to continuous, increasing, concave functions, then any constrained equal-income competitive allocation for the true preferences (that is, a constrained equilibrium allocation reached when agents have equal initial endowments) is an equilibrium allocation to the distortion game. The paper also provides conditions under which constrained equal-income competitive allocations are the only allocations that arise in pure-strategy equilibria of the distortion game. In equilibrium, agents report

---

\* I thank a UCLA seminar audience, Uzi Segal, and Joel Watson for comments and NSF for financial support.

that their marginal utility is constant and equal to the price that supports the competitive equilibrium for quantities within the equal-income budget set and that their marginal utility is zero otherwise.

Sobel [1981] studies the distortion game induced by a class of bargaining game solutions that satisfy axioms of Pareto optimality, symmetry, independence of equivalent utility representations, and midpoint domination. The first three axioms are familiar. The fourth axiom requires that the bargaining solution of an  $n$ -player game gives each player a utility that exceeds the disagreement utility by at least one- $n$ th of the difference between his largest feasible utility and his disagreement utility. Sobel shows that the Nash bargaining solution and the Kalai-Smorodinsky solution satisfy these four assumptions. The paper goes on to show that any constrained equal-income competitive allocation for the true utilities is an equilibrium payoff for the distortion game derived from one of these bargaining game solutions; and that if players are restricted to reporting linear strategies, then the constrained equal-income competitive allocations are the only pure-strategy Nash equilibria of the distortion game in two-player games. The relative utilitarian solution fails to satisfy the midpoint domination axiom. Nevertheless, it is still true that any constrained equal-income competitive allocation for the true utility functions is an equilibrium allocation for the distortion game. Furthermore, if players are restricted to truncated linear reports, the only pure-strategy Nash equilibrium allocations of a modification of the distortion game are the constrained equal-income competitive allocations in two-player games.

Several other paper study distortion games and establish related results. Crawford and Varian [1979] study the distortion game derived from the Nash bargaining solution when there is only one commodity. They show that it is a dominant strategy to report a linear utility function. The model predicts equal division of the commodity.

Thomson ([1979] and [1984]) characterizes Nash equilibria for the distortion game derived from performance correspondences that yield individually rational and Pareto-efficient outcomes. Thomson [1979] finds that if reported utility functions must be twice continuously differentiable, concave, and have transferable utility, then the Nash equilibria for the distortion game derived from the Shapley value with fixed initial endowments are exactly the constrained competitive allocations with respect to those endowments. Thomson [1984] generalizes this result in an environment with transferable utility to a broader class of performance correspondences. Thomson [1987] identifies a class of performance correspondences for which equilibrium outcomes of the distortion game are equal-income Walrasian equilibria for the reported preferences.

Kurz ([1977] and [1980]) studies the distortion game derived from his model of tax determination in an exchange economy. Regardless of the true preferences of the agents, the distortion game has a dominant strategy equilibrium that yields a Pareto-efficient outcome. Mas-Colell [1980] presents analogous results in his study

of the manipulability of the Shapley-value in an economy with an infinite number of agents falling into a finite number of types.

There are three lessons to learn from this paper. First, the distortion game can be viewed as another way to implement the constrained Walrasian correspondence. This point is minor, given the existence of many other procedures that do the same thing. Second, the paper generalizes the results of Sobel and Thomson. The generalization may be surprising because, in contrast to the earlier results, equilibrium strategies are necessarily non linear and dominant strategy equilibria do not arise even in the one-dimensional case. Third, the equilibria of the distortion game predict a particular qualitative behavior for the players. In the bargaining solutions considered in this paper and earlier work, weight is given to a player's valuation of the last unit he receives. It is to a player's advantage to claim that his marginal unit is as valuable as possible. Since reports are limited to concave functions, equilibrium reports tend to be linear in order to maximize the marginal utility of marginal units. The earlier papers emphasized the importance of linear strategies. For the case of the relative utilitarian solution, it is advantageous for agents to claim to be uninterested in units they will not receive. This creates a role for truncated linear reports in which agents report that anything that they receive beyond their equilibrium allocation is worthless to them.

Section 2 describes the distortion game formally. Section 3 introduces the relative utilitarian solution. Section 4 briefly describes the results for the special case of two agents and one commodity. Section 5 presents the main results. Section 6 attempts to reconcile the results with earlier findings. Section 6 introduces a one-parameter family of bargaining solutions, which includes the relative utilitarian solution, the Nash bargaining solution, and the Kalai-Smorodinsky solution. Each of these games has the property that constrained equal-income competitive allocations are equilibrium outcomes of the distortion game, but results differ as to whether equilibria are supported by linear or truncated linear strategies.

## 2. DEFINITIONS AND NOTATION

Consider  $n$  agents with von Neumann-Morgenstern utility functions who are to divide a bundle of  $m$  commodities. Let  $\Delta_m$  denote the  $m - 1$  dimensional simplex and let  $e_k \in \Delta_m$  be the unit vector in the  $k$ th direction. Let  $\mathbf{a} \equiv (a, \dots, a)$ . An outcome will be an element of the set

$$T = \left\{ x = (x_1, \dots, x_n) \in \mathcal{R}^{nm} : x_i \in \mathcal{R}^m \text{ for } i = 1, \dots, n, \text{ and } \mathbf{0} \leq \sum_{i=1}^n x_i \leq \mathbf{1} \right\}$$

where agent  $i$  receives  $x_i$ . The true utility function of player  $i$  is denoted by  $u_i$ . These functions are assumed to be concave and strictly increasing. The players report utilities that are restricted to lie in the class  $\mathcal{U}$ , where  $\mathcal{U}$  consists of those

functions:  $U : \mathcal{R}^m \rightarrow [0, 1]$  such that (i)  $U$  is continuous, increasing, and concave in  $\mathcal{R}^m$ ; (ii)  $U$  is normalized so that  $U(\mathbf{0}) = 0$  and  $U(\mathbf{1}) = 1$ . The class of admissible utilities should include those functions that are credible representations of the agents' true preferences. Condition (i) is a regularity assumption on the range of potential players. The concavity assumption means that the agents cannot pretend to be risk lovers. In order to guarantee existence of equilibrium, it is necessary to permit  $\mathcal{U}$  to include functions that are not strictly increasing. Since the solution to the bargaining problem studied in this paper is independent of affine transformations, condition (ii) is not essential.

The distortion game is played by each agent revealing a utility function in  $\mathcal{U}$ . Let  $U_i$  denote the utility function revealed by player  $i$ . Given the report  $U = (U_1, \dots, U_n)$ , a set of outcomes  $B(U)$  is selected.  $B(U)$  is the set of allocations that give rise to a bargaining solution determined by utility functions  $U$ . Section 3 discusses the particular bargaining solution that is the focus of this paper. However, in order to define the distortion game, it is only necessary that  $B(U)$  be a non-empty subset of  $T$  for any admissible reported utilities.

If  $B(U)$  is single-valued, then all of the elements of the definition of the distortion game would be in place. Players would be restricted to admissible strategies and, denoting the  $i$ th component of  $B(U)$  by  $B_i(U)$ , the payoff to player  $i$  would be  $u_i(B_i(U))$ . While it is typical to assume that the bargaining solution selects a unique utility for each player, there is no reason to believe that there will be a unique allocation associated with these utilities.  $B(U)$  will typically not be single valued, so it is necessary to specify how to make a selection from the bargaining solution. I assume that agents have the right to select their most preferred outcome in  $B(\cdot)$ . This assumption leads to the following definition.

**Definition 1.**  $(U^*; x^*) = ((U_1^*, \dots, U_n^*); (x_1^*, \dots, x_n^*))$  is an equilibrium for the distortion game determined by  $B$  if and only if  $x^* \in B(U^*)$  and for each  $i$ ,  $(U_i^*, x_i^*)$  solves  $\max u_i(x_i)$  subject to  $x_i \in \{B_i(U_i, U_{-i}^*) : U_i \in \mathcal{U}\}$ .

The definition implies that for equilibrium reports  $U^*$ , there exists an allocation  $x^* \in B(U^*)$  such that  $x_i^*$  solves  $\max u_i(x_i)$  subject to  $x_i \in B_i(U^*)$ . That is, players' true preferences can agree on a most preferred allocation in  $B(U^*)$ .

In Sobel [1981], strategy sets of the distortion game consist of both reported preferences and tie-breaking rules that make unique selections from the correspondence  $B$ . The paper introduces a class of tie-breaking rules that permit agents to select their most preferred outcome in  $B$ . Consequently, an equilibrium for the distortion game corresponds to a Nash equilibrium of an extended game in which each agent reports a tie-breaking rule in addition to a utility function. Thomson [1984] adopts the definition above.

Bargaining solutions uniquely determine the utilities of agents according to the agents' reported preferences. One might think that players would be indifferent between all of the allocations in  $B(U)$ . However, unless an agent reports his true utility function, there is no reason to expect that he would be indifferent over all elements in the bargaining solution according to his true preferences. In the equilibria described in Section 5,  $B(U)$  consists of a budget hyperplane, and agents will generally not be indifferent over the elements in  $B$ .

In what follows, I prove that the Nash equilibrium outcomes in the distortion game correspond to certain competitive outcomes.

**Definition 2.** *A constrained equal-income competitive equilibrium (EICE) is a pair,  $(p^*, x^*)$  where (i)  $p^* \in \Delta_m$ ; (ii)  $x^* \in T$ ; (iii)  $x_i^*$  solves:*

$$\max u_i(x_i) \text{ subject to } p^* \cdot x_i \leq \frac{1}{n} p^* \cdot \mathbf{1} \text{ and } \mathbf{0} \leq x_i \leq \mathbf{1}.$$

In an exchange economy, an equal-income competitive equilibrium is a competitive equilibrium to the economy in which agents have equal initial endowments. In a constrained equal-income competitive equilibrium, agents' demands are restricted to be both in their budget set and feasible relative to the total resources of the economy. The constraint that an agent  $i$ 's demand  $x_i$  satisfy  $\mathbf{0} \leq x_i \leq \mathbf{1}$  is not placed on (unconstrained) equal-income competitive equilibria.<sup>1</sup> Any equal-income competitive equilibrium is a constrained equal-income competitive equilibrium and, provided that preferences are convex, any interior constrained equal-income competitive equilibrium is an equal-income competitive equilibrium. I will call  $x^*$  in the definition of EICE a competitive allocation.

### 3. THE RELATIVE UTILITARIAN SOLUTION

This paper studies the distortion game induced by the relative utilitarian solution proposed by Cao [1981], Dhillon and Mertens [1996], Karni [1996], Karni and Safra [1996], and Segal [1998]. Cao's approach focuses on a bargaining problem. The other papers present the functional form described below for social-choice problems. This section presents additional detail on bargaining games. The discussion this section is a bit more general than needed for the main results in Section 5. I include the additional detail in order to contrast this paper with earlier work and analyze the more general games in Section 6.

A *bargaining game* is described by a pair  $(S, d)$ , where: (i)  $d = (d_1, \dots, d_n) \in \mathcal{R}^n$ ; (ii)  $S \subset \mathcal{R}^n$  is compact, convex, and contains  $d$  as well as some point  $s > d$ .

---

<sup>1</sup>Hurwicz, Maskin, and Postlewaite [1995] introduced the concept of constrained competitive equilibrium. It is the smallest extension of the competitive correspondence that can be implemented in Nash strategies.

The set  $S$  is normally interpreted as the set of feasible utility payoffs to the players. A point  $s \in S$  can be achieved if all players agree to it. In case of agreement, player  $i$  receives  $s_i$ . If the players are unable to agree, then the outcome  $d$ , called the disagreement outcome, is the result.

In this paper, the reported utilities  $U = (U_1, \dots, U_n)$  with each  $U_i \in \mathcal{U}$  determine the set  $S$ . That is,

$$S = S(U) = \{s = (s_1, \dots, s_n) : s_i = U_i(x_i) \text{ for some } x \in T\}.$$

The disagreement outcome will always be taken to be  $\mathbf{0} = U(\mathbf{0})$ .

When the functions  $U_i$  are admissible, the set  $S$  is compact, convex, and contains a point  $s > \mathbf{0}$ . In fact, the  $i$ th unit vector  $e_i \in S$  for  $i = 1, \dots, n$  and  $\Delta_n \in S$ . Such a game will be called 0 – 1 normalized.<sup>2</sup>

Beginning with Nash [1951], axiomatic bargaining theory has studied *solutions* to bargaining games. A solution is a function  $f$ , defined on the class of all bargaining games with  $f(S, d) = (f_1(S, d), \dots, f_n(S, d)) \in S$  for all pairs  $(S, d)$ . Nash characterized his bargaining solution in terms of four axioms, three of which are relevant here.

**Axiom 1.** (*Pareto Efficiency*): If  $f(S, d) = x$  and  $y \geq x$ , then either  $y = x$  or  $y \notin S$ .

**Axiom 2.** (*Independence of Equivalent Utility Representations*): If  $(S, d)$  and  $(S', d')$  are bargaining games such that

$$S' = \{s' = (s'_1, \dots, s'_n) : s'_i = a_i s_i + b_i \text{ for } s = (s_1, \dots, s_n) \in S\}$$

and

$$d' = (a_1 d_1 + b_1, \dots, a_n d_n + b_n)$$

where  $a_i > 0$  for all  $i$ , then

$$f(S', d') = (a_1 f_1(S, d) + b_1, \dots, a_n f_n(S, d) + b_n).$$

**Axiom 3.** (*Symmetry*): If  $(S, d)$  is a symmetric game (that is,  $s = (s_1, \dots, s_n) \in S$  if and only if  $(s_{\pi(1)}, \dots, s_{\pi(n)}) \in S$  for any permutation  $\pi$ ), then  $f_i(S, d)$  is independent of  $i$ .

---

<sup>2</sup>The normalization of  $S(U)$  anticipates the observation that the relative utilitarian solution is invariant with respect to equivalent utility representations. With that property, any game can be taken to be 0 – 1 normalized without loss of generality.



Given a bargaining game  $(S, d)$ , let  $\bar{s}_i \equiv \max \{s_i : s = (s_1, \dots, s_n) : s \in S\}$  be player  $i$ 's maximum utility. The relative utilitarian quasi-solution consists of

$$\left\{ s^* = (s_1^*, \dots, s_n^*) : s^* \text{ solves } \max \sum_{i=1}^n \left( \frac{s_i}{\bar{s}_i - d_i} \right) \text{ subject to } s = (s_1, \dots, s_n) \in S \right\}. \quad (1)$$

Note that if  $(S, d)$  is 0 – 1 normalized, then (1) simplifies to

$$\left\{ s^* = (s_1^*, \dots, s_n^*) : s^* \text{ solves } \max \sum_{i=1}^n s_i \text{ subject to } s = (s_1, \dots, s_n) \in S \right\}. \quad (2)$$

Since (1) may contain more than one point, it defines a quasi-solution rather than a solution. One forms a relative-utilitarian solution by making a selection from (1). It is convenient to work with symmetric relative-utilitarian solutions. In what follows, therefore, assume that the social planner makes a selection from (1) whenever (1) contains more than one point. The selection can be arbitrary, although when  $S$  is symmetric, it makes sense to assume that the selection is symmetric. Denote the selection by  $RI(S)$ .

Admissible reported utilities  $U = (U_1, \dots, U_n)$  determine a set  $S(U)$  of feasible utilities with respect to the reported preferences. The social planner will select an allocation according to the relative utilitarian solution. This selection determines the set  $B(U)$  resource allocations that give rise to the relative utilitarian solution:

$$B(U) = \{x = (x_1, \dots, x_n) \in T : (U_1(x_1), \dots, U_n(x_n)) = RI(S(U))\}.$$

It follows that if  $x^* \in B(U)$ , then  $x^*$  solves:

$$\max \sum_{i=1}^n U_i(x_i) \text{ subject to } x = (x_1, \dots, x_n) \in T \quad (3)$$

In the equilibria described in Sections 4 and 5 agents report utility functions that are not necessarily strictly increasing. This creates the possibility that there exist  $x$  and  $y \in B(U)$  such that  $\sum_{i=1}^n y_i < \sum_{i=1}^n x_i$ ; that is, both  $x$  and  $y$  solve the planner's optimization problem, but  $y$  does so using strictly less of each commodity. When this situation arises, the social planner might restrict attention to allocations that are *resource conserving*.

$$B^*(U) = \left\{ x \in B(U) : \text{if } \sum_{i=1}^n y_i \leq \sum_{i=1}^n x_i, \text{ then } y \notin B(U) \text{ or } y = x \right\}.$$

If all of the reported utilities are strictly increasing, then  $B^*(U) = B(U)$ . In general,  $B^*(U) \subseteq B(U)$ .  $B^*(U)$  consists of the allocations that give rise to utilities that

solve the social planner's problem without wasting resources. The distortion game determined by a relative utilitarian solution  $B(\cdot)$  is called the relative utilitarian distortion game. The corresponding game determined by  $B^*(U)$  is called the resource-conserving relative utilitarian distortion game.

It is apparent that the relative utilitarian quasi-solution is invariant with respect to equivalent utility representations and is Pareto efficient with respect to the reported utilities. Any symmetric relative-utilitarian solution also satisfies the symmetry axiom.

Sobel [1981] describes equilibria of distortion games for bargaining solutions that satisfy Axioms 1, 2, 3, as well as Axiom 4.

**Axiom 4.** (*Midpoint Domination*):  $f_i(S, d) \geq \frac{1}{n(\bar{s}_i - d_i)}$  for all  $i$ .

If  $(S, d)$  is 0 – 1 normalized, then Axiom 4 becomes:  $f_i(S, d) \geq \frac{1}{n}$  for all  $i$ . Sobel [1981] introduced Axiom 4 under the name symmetric monotonicity. Thomson [1994] refers to it as midpoint domination.

Sobel [1981] demonstrates that any bargaining solution for a two-person game that is risk dominant (see Roth [1979]) satisfies Axiom 4. It follows from Roth [1979] that the Nash and Kalai-Smorodinsky bargaining solutions satisfy Axiom 4. On the other hand, the relative utilitarian solution does not satisfy midpoint domination. For instance, let

$$S = \{s = (s_1, s_2) : s_1 \in [0, 1] \text{ and } s_2 \in [0, 1 - as_1]\}$$

When  $a \in (0, 1)$  the solution to  $\max s_1 + s_2$  subject to  $(s_1, s_2) \in S$  is  $(1, 1 - a)$ . For  $a > .5$ , player 2 receives utility less than .5.

The example demonstrates that the theorem of Sobel [1981] does not apply to the relative utilitarian solution. Nevertheless, as I now demonstrate, the qualitative conclusions of the earlier work do apply.

#### 4. THE ONE-COMMODITY CASE

To get an intuition for the results, consider the special case in which there are two players and one commodity. When there is only one commodity, all monotonic preferences are ordinally equivalent. Given the symmetry of the game, one would expect that equal division is the only equilibrium outcome of the distortion game. This conjecture is correct. Each agent can guarantee at least one half of the commodity by reporting the utility function  $U^*(x) = \min[2x, 1]$ . This utility function is concave, continuous, and increasing. Furthermore, a player who reports this utility must receive at least .5. To see this, note that under  $U^*(\cdot)$  the agent asserts that the marginal utility of each  $x < .5$  is two. If the solution were to give the other player more than one half, then the other player's reported marginal utility must be greater

than or equal to two for each of the first  $1 - x > .5$  units. This is not consistent with monotonicity and the restriction that  $U(1) = 1$ .

The one-commodity case provides several lessons. First, the outcome of the game is efficient with respect to the true preferences. (Indeed, it is the unique EICE.) Second, equilibrium strategies of the agents are not linear and the equilibrium is not a dominant strategy equilibrium. Notice that if one agent reported the linear utility function  $U(x) = x$ , then for any  $\epsilon > 0$  the opponent could respond with the function  $V(x) = \min[(1 + \epsilon)x, 1]$  and obtain almost the entire unit of the commodity (the share  $\frac{1}{1+\epsilon}$ ). Third, the equilibrium strategies are not strictly increasing. Equilibria would not exist if players were restricted to strictly increasing strategies. Agents could guarantee arbitrarily close to .5 by playing continuous approximations to  $U^*(\cdot)$ . Hence, in equilibrium, both players must receive .5. However, given any continuous strategy of one player, it is possible for the other to obtain strictly more than one half by choosing a better approximation to  $U^*(\cdot)$ . Fourth, while the equilibrium strategies are not linear, they are linear when restricted to  $x \leq .5$ , the quantities that the agent consumes in equilibrium. In order to make a claim for a marginal unit, it is in the interest of bargainers to claim that the marginal unit is as valuable as possible to them. They deny any interest in units that they are not going to receive and report nondecreasing marginal utility over the units that they do receive. These observations generalize to the case of more than one commodity.

While agents lack strictly dominant strategies in the one-commodity model, when players are restricted to truncated linear strategies, the equilibrium strategies are the only strategies that satisfy iterated deletion of weakly dominated strategies provided that the selection from the relative utilitarian correspondence is symmetric. Limit players to strategies in the set  $\mathcal{L} = \{U(x) = \min[ax, 1] \text{ for } a \geq 1\}$ . This restriction is plausible because given any  $U \in \mathcal{U}$ , a player always has a best response in  $\mathcal{L}$ . Let  $X_b = \{\min[ax, 1] \in \mathcal{L} : 2 \geq a \geq b\}$ . Suppose agent  $i$  uses strategies in  $X_b$ . Player  $j \neq i$  receives less than .5 when he uses  $\min[ax, 1]$  for  $a < b$  or  $a > 2$ . If  $a = b$ , then the player receives exactly .5 when the other player uses the same strategy (by symmetry) and strictly less against any other strategy in  $X_b$ . On the other hand, the strategy  $U^*$  guarantees that agent  $j$  receive  $x = .5$ . Consequently,  $U^*(\cdot)$  is the only strategy that survives iterative deletion of weakly dominated strategies. The same argument applies when there are  $n > 2$  players (each agent uses the function  $\min[nx, 1]$ ).

The one-commodity case also provides a clear indication of the importance of the assumption that agents' reports are concave functions. Assume instead that agents could report arbitrary monotonic functions and that agent one reports  $[x]$ , where  $[\cdot]$  is the greatest integer function.<sup>3</sup> A solution to the problem:

$$\max [x] + V(1 - x) \text{ subject to } x \in [0, 1] \quad (4)$$

---

<sup>3</sup>The argument does not depend on the discontinuity of  $[\cdot]$ .

is to give the first agent everything for all  $V \in \mathcal{U}$  (given agent one's report, the value of the objective function in (4) can be no greater than one; this upper bound is attained if player one receives everything). An equilibrium in which all players report "all-or-nothing" preferences exists whenever arbitrary monotonic reports are permitted.<sup>4</sup>

### 5. THE GENERAL CASE

When there is more than one commodity, agents may not have the same ordinal preferences. It need not be efficient to split each unit equally. This section demonstrates that the equilibrium outcomes of the distortion game are still efficient in the general setting.

For  $p \in \Delta_m$ , let  $L(z; p) \equiv \min[np \cdot z, 1]$ . Given  $V \in \mathcal{U}$ ,  $\lim_{h \rightarrow 0^+} \frac{V(z) - V(z - he_k)}{h}$  exists; let  $D_k^- V(z)$  denote the limit. When  $V(\cdot)$  is differentiable at  $x$ ,  $D_k^-(z)$  is a partial derivative; in general it is a one-sided partial derivative. Let  $\nabla V(z) = (D_1^-(z), \dots, D_m^-(z))$ . Since  $V(\mathbf{0}) = 0$  and  $V$  is concave,

$$V(z) \geq \nabla V(z) \cdot z \text{ for all } z \geq \mathbf{0}. \tag{5}$$

**Lemma 1.** *If player  $j$  reports the truncated utility  $L(\cdot; p)$ , then for all admissible  $U_{-j}$ , if  $x = (x_1, \dots, x_n) \in B(L(\cdot; p), U_{-j})$ , then  $np \cdot x_j \geq 1$ .*

Lemma 1 is a special case of Lemma 2, which I state and prove in Section 6. The lemma establishes that using a truncated linear report, a player can guarantee an allocation that is at least as good as equal division according to the reported preferences. If players are restricted to reported preferences of the form  $L(z; p)$  for  $p \in \Delta_m$ , then the result is particularly easy to understand. In this case, equal division yields the maximum value of the relative-utilitarian social welfare function. Hence any allocation in  $B(\cdot)$  for truncated linear reports must give each agent utility one with respect to his reported preference. This is the conclusion of the lemma. The actual proof is a bit more involved because agents are not limited to truncated linear reports.

**Theorem 1.** *If  $(p^*, x^*)$  is an EICE for the true preferences, then  $(L(\cdot; p^*), \dots, L(\cdot; p^*); x^*)$  is a Nash equilibrium for the relative utilitarian distortion game.*

**Proof.** Assume that all of the players except player  $i$  use the strategy  $L(\cdot; p^*)$ . Suppose that if player  $i$  selects the strategy  $V$ , then the allocation is  $x = (x_1, \dots, x_n)$ . It follows from Lemma 1 that for all  $i \neq j$ ,  $np^* \cdot x_j \geq 1$ . Since  $x \in T$ ,  $np^* \cdot x_i \leq 1$ . It

---

<sup>4</sup>When reports need not be concave, it is natural to permit the bargaining solution to randomize. Symmetry suggests that the solution to the bargaining problem when all agents report "all-or-nothing" preferences is to give everything to a randomly selected player.

follows that the maximum utility player  $i$  could receive in response to his opponents' strategies is

$$\max u_i(x_i) \text{ subject to } np^* \cdot x_i \leq 1 \text{ and } \mathbf{0} \leq x_i \leq \mathbf{1}. \quad (6)$$

Since  $(p^*, x^*)$  is an EICE,  $x_i^*$  solves (6). Since  $x^* \in B(L(\cdot; p^*), \dots, L(\cdot; p^*))$  it follows that it is a Nash equilibrium for player  $i$  to report  $L(\cdot, p^*)$  and obtain the EICE allocation.

The theorem demonstrates that any EICE allocation is a Nash equilibrium of the distortion game. The lemma guarantees that if each agent reports the same truncated linear strategy  $L(\cdot, p^*)$ , then the set of solutions to the planner's problem consists of the entire budget hyperplane with normal  $p^*$  through equal division. If  $p^*$  supports an EICE, then all players can agree upon a most preferred outcome in the set of allocations that maximize the relative utilitarian objective function. Consequently, an EICE allocation is an equilibrium of the distortion game.

There is a curious aspect of the equilibrium strategies that support an EICE allocation. The value of the social planner's optimization problem is equal to  $n$ , the maximum feasible value. Even though there is conflict over the goods, the equilibrium allocation provides each agent utility one (according to reported preferences). If the social planner were naive enough to believe that the reported preferences were honest representations of the agent's motives, then she would be delighted to know that every agent received his maximum utility.

In the one-commodity case, the equilibrium allocation was unique. Since there may be multiple EICE allocations, the uniqueness property will not carry over into the many commodity case. The best that one might hope for is that the only equilibrium allocations are the EICE allocations. This result fails to hold, as the next example demonstrates. The example is a modification of an example in Sobel [1981].

**Example 1.** *In a two-consumer, two-commodity model, let  $u_1(x_1, x_2) = x_1^{5/6} x_2^{1/6}$ ,  $u_2(x_1, x_2) = x_1^{1/2} x_2^{1/2}$ ,  $U_1(x_1, x_2) = \min[(5x_1 + 3x_2)/4, 1]$ , and  $U_2(x_1, x_2) = x_1^{1/2} x_2^{1/2}$ . For  $i = 1$  and  $2$ ,  $u_i$  and  $U_i \in \mathcal{U}$ . A routine computation shows that  $(U_1, U_2, x^*)$  is a Nash equilibrium for the distortion game. The allocation  $x^*$  gives agent 1  $(3/5, 1/3)$  and agent 2  $(2/5, 2/3)$ . In this example, there is a unique EICE for the true preferences. It is  $(p^*, y^*) = ((2/3, 1/3), (5/8, 1/4))$  [ $y^*$  is player 1's allocation in the equilibrium. Player 2 receives  $\mathbf{1} - y^* = (3/8, 3/4)$ ]. A computation shows that the first player prefers the inefficient outcome  $x^*$  to the EICE. Also, the second player is worse off at the Nash equilibrium even though he is reporting his true utility function.*

Sobel [1981] shows that in two-player games, the only equilibrium allocations of distortion games derived from bargaining solutions satisfying Axioms 1 through 4 are EICE allocations provided that agents are restricted to linear strategies. This restriction makes no sense in the present context, since linear strategies do not support

the equilibria described in Theorem 1. It does suggest that a partial converse to Theorem 1 holds if players are limited to truncated linear strategies. Even this result fails to hold, as the next example demonstrates.

**Example 2.** *In a two-consumer, two-commodity model, let  $u_1(x_{11}, x_{12}) = x_{11}^{2/3} x_{12}^{1/3}$ ,  $u_2(x_{21}, x_{22}) = x_{21}^{1/3} x_{22}^{2/3}$ ,  $U_1(x_{11}, x_{12}) = \min[2(x_{11} + 2x_{12})/3, 1]$ , and  $U_2(x_{21}, x_{22}) = \min[2(2x_{21} + x_{22})/3, 1]$ . For  $i = 1$  and  $2$ ,  $u_i \in \mathcal{U}$ . Let the allocation  $x^*$  give both agents  $(.5, .5)$ . A computation shows that  $(U_1, U_2, x^*)$  is a Nash equilibrium for the distortion game. Lemma 1 implies that an upper bound to the utility that agent 1 can attain is the value of  $\max u_1(x_{11}, x_{12})$  subject to  $2(2x_{11} + x_{12})/3 \leq 1$  (because the lemma guarantees that player 2's allocation satisfies  $2(2x_{21} + x_{22})/3 \geq 1$ ) and similarly that the upper bound for player 2 is the value of  $\max u_2(x_{21}, x_{22})$  subject to  $2(x_{21} + 2x_{22})/3 \leq 1$ . For the reported preferences  $(U_1, U_2)$ ,*

$$B(U_1, U_2) = \{(x_{11}, x_{12}, x_{21}, x_{22}) \in T : 2(x_{11} + 2x_{12})/3 \geq 1 \text{ and } 2(2x_{21} + x_{22})/3 \geq 1\}$$

*For the true preferences, both agents agree that  $x^*$  is the best allocation in  $B(U_1, U_2)$ . It is apparent that the allocation  $x^*$  is not efficient. The only EICE in this example is  $p = (.5, .5)$   $(x_{11}, x_{12}) = (\frac{2}{3}, \frac{1}{3})$  and  $(x_{21}, x_{22}) = (\frac{1}{3}, \frac{2}{3})$ .*

In the example,  $B(U)$  contains allocations that are not resource conserving. It is not necessary to use all of the endowment to attain the maximum in the social planner's utility function. This suggests the possibility that the inefficient equilibrium would disappear if the planner used a resource-conserving selection from the set of maximizing allocations.

Imposing the additional requirement that the solution be resource conserving shrinks the set of allocations. The planner allocates the minimum quantity of resources needed to maximize the relative utilitarian objective function. For the reported preferences

$$B^*(U_1, U_2) = \left\{ \left( 0, \frac{3}{4}, s, \frac{3}{2} - 2s \right), s \in \left( \frac{5}{8}, \frac{3}{4} \right) \right\} \cup \left\{ \left( \frac{3}{2} - 2t, t, \frac{3}{4}, 0 \right), t \in \left( \frac{5}{8}, \frac{3}{4} \right) \right\},$$

$x^* \notin B^*(U_1, U_2)$  and the agents have conflicting preferences over allocations in  $B^*(U_1, U_2)$  (Agent one prefers the allocation  $(x_{11}, x_{12}) = (\frac{1}{4}, \frac{5}{8})$  and  $(x_{21}, x_{22}) = (\frac{3}{4}, 0)$ , while agent two prefers the allocation  $(x_{11}, x_{12}) = (0, \frac{3}{4})$  and  $(x_{21}, x_{22}) = (\frac{5}{8}, \frac{1}{4})$ ).

The two examples suggest that a partial converse to Theorem 1 requires both a restriction on strategies and a selection from the set of maximizing allocations. The next theorems provide the results.

**Theorem 2.** *Let  $n = 2$ . If  $(L(\cdot; p_1), L(\cdot; p_2); x^*)$  is a Nash equilibrium for the resource conserving relative-utilitarian distortion game, then  $x^*$  is an EICE allocation.*

**Proof.** It follows from Lemma 1 that for all  $i$ ,  $2p_i \cdot x_i^* \geq 1$ . Let  $j \neq i$ . Since  $x^* \in T$ ,  $2p_j \cdot x_i^* \leq 1$ . It follows that the maximum utility player  $i$  could receive in response to his opponent's strategy is

$$\max u_i(x_i) \text{ subject to } 2p_j \cdot x_i \leq 1 \text{ and } \mathbf{0} \leq x_i \leq \mathbf{1}. \tag{7}$$

Player  $i$  can indeed achieve this payoff by using the strategy  $L(\cdot; p_j)$ . It follows from monotonicity of  $u_i(\cdot)$ , that  $2p_j \cdot x_i^* = 1$ . The resource conserving condition therefore implies that  $p_i = p_j$  for otherwise the planner would solve (3) with an allocation  $x^*$  such that  $2p_j \cdot x_i^* < 1$ . It follows that in equilibrium  $p_i = p_j$ . Consequently, (7) implies that  $x^*$  is an EICE allocation.

I conjecture that Theorem 2 holds when  $n > 2$ .

There are two justifications for the restriction to truncated linear strategies. From the point of view of implementation, the planner can make whatever restrictions on strategies needed in order to guarantee desirable outcomes. From the point of view of the strategic behavior of the agents, it is possible to show that even when there are no restrictions on admissible reports (other than continuity, monotonicity, and concavity), agents always have a truncated linear best response.

## 6. RECONCILIATION

Earlier results demonstrate that the Nash equilibria of distortion games derived from admissible bargaining games are EICE allocations and that in these equilibria, players report linear utility functions. This paper studies the distortion game derived from a bargaining game that fails to satisfy the assumption of midpoint domination. Still, the Nash equilibria of distortion games are EICE allocations. Equilibrium reports are non-linear, but they are piecewise linear. There is no logical contradiction here. The relative utilitarian solution does not satisfy the assumptions in the earlier paper.

It may be instructive to reflect more on the sense in which the results generalize the earlier ones. First, the restriction to linear strategies in the earlier work is not always necessary. For the distortion game derived from the Nash Bargaining Solution, one can support the same outcomes as equilibria by using piecewise linear reports of the kind introduced in this paper. Second, while the relative utilitarian solution fails midpoint domination, it does not fail midpoint domination if players are restricted to truncated linear strategies that are used to support equilibria. This is essentially the conclusion of Lemma 1.

To make these points formally, consider a family of bargaining solutions. For  $s = (s_1, \dots, s_n)$  and  $a \leq 1$ , let  $W(s, a) = \left\{ \sum_{i=1}^n \frac{1}{n} s_i^a \right\}^{\frac{1}{a}}$ .<sup>5</sup> Let  $(S, d)$  be 0–1 normalized. For  $a < 1$ ,  $W(\cdot, a)$  is strictly concave and generates a bargaining solution

$$f(S; a) = \{s^* = (s_1^*, \dots, s_n^*) : s^* \text{ solves } \max W(s, a) \text{ subject to } s \in S\}.$$

---

<sup>5</sup>When  $a = 0$ , define  $W(s, 0) = \lim_{a \rightarrow 0} W(s, a) = \left\{ \prod_{i=1}^n s_i \right\}^{\frac{1}{n}}$ .

Of course,  $W(s, 1)$  is the utilitarian objective function. Since  $W(s, 0) = \{\prod_{i=1}^n s_i\}^{\frac{1}{n}}$  and  $\lim_{a \rightarrow -\infty} W(s, a) = \min\{s_1, \dots, s_n\}$ ,  $f(S; 0)$  corresponds to the Nash Bargaining Solution and  $f(S; -\infty)$  corresponds to the Kalai-Smorodinsky solution. Increases in  $a$  tend to make  $W(\cdot)$  more like the utilitarian solution, while decreases in  $a$  make the solution more sensitive to the relative welfare of agent who receives the lowest (normalized) utility.

Since the family of bargaining solutions are defined as solutions to optimization problems, one might be tempted to conclude that they all satisfy Nash's Independence of Irrelevant Alternatives Axiom. In fact, only  $f(S; 0)$  (the Nash Bargaining Solution) satisfies this axiom. The possible confusion arises because I have defined the solution for  $0 - 1$  normalized bargaining games. In general, changing the maximum utility available to a bargainer will change that bargainer's weight in  $W(\cdot)$ , leading to the possibility of a violation of the Independence of Irrelevant Alternatives Axiom.

Imagine a family of distortion games that varies with the parameter  $a$ . For  $a \leq 0$ , one can show that the bargaining solutions satisfy the midpoint domination axiom. Hence results from Sobel [1981] can be used to characterize the solution to the distortion game. For  $a \geq 0$ , Lemma 1 of Section 5 continues to hold. Hence the results from the previous section apply. In all cases, EICE allocations are equilibria of the distortion game. In the case of solutions with non-negative parameter  $a$ , these equilibria are supported by truncated linear strategies. When  $a$  is nonpositive, the equilibrium is supported by linear strategies. Curiously, the Nash bargaining solution is the only one in which both kinds of strategy support the equilibrium.

**Lemma 2.** *If player  $j$  reports the truncated utility  $L(\cdot; p)$ , then for all admissible  $U_{-j}$ , if  $x = (x_1, \dots, x_n) \in B(L(\cdot; p), U_{-j}; a)$  and  $a > 0$ , then  $np \cdot x_j \geq 1$ .*

**Proof.** Assume that player  $j$  uses the strategy  $L(\cdot; p)$ . Suppose that if player  $i$  selects the strategy  $V$ , then the allocation would be  $x = (x_1, \dots, x_n)$ . In order to obtain a contradiction, assume that

$$np \cdot x_j < 1. \tag{8}$$

Since  $x$  must maximize (3) for the reported preferences,

$$[np \cdot (x_j + h)]^a + [V(x_i - h)]^a \leq [np \cdot x_j]^a + [V(x_i)]^a \tag{9}$$

for all  $0 \leq h \leq x_i$ . By (9), the partial derivative of  $[np \cdot x_j]^a + [V(x_i)]^a$  in the direction of  $e_k$  must be less than or equal to zero whenever  $x_{ik} > 0$ . It follows that  $np_k \leq D_k^- V(x_i)$  whenever  $x_{ik} > 0$ ; consequently

$$a [np \cdot x_j]^{a-1} np_k x_{ik} \leq a [V(x_i)]^{a-1} D_k^- V(x_i) x_{ik} \text{ for all } k.$$



Summing this inequality over  $k$  yields for all  $i \neq j$ ,

$$[np \cdot x_j]^{a-1} np \cdot x_i \leq [V(x_i)]^{a-1} \nabla V(x_i) \cdot x_i. \quad (10)$$

It follows from (5), (10), and  $1 \geq V(x)$ , that for all  $i \neq j$ ,

$$np \cdot x_i \leq [np \cdot x_j]^{1-a}. \quad (11)$$

Since  $a < 1$ , (8) and (11) imply

$$np \cdot x_i \leq 1. \quad (12)$$

Since  $\sum_{i=1}^n x_i = \mathbf{1}$  and  $p \in \Delta_m$ , summing (12) over all  $i \neq j$  contradicts (8) and establishes the lemma.

Lemma 2 demonstrates that when  $a > 0$ , the crucial lemma from Section 5 continues to hold. One can establish the same result for the Nash Bargaining solution ( $a = 0$ ) directly or through a limiting argument. Lemma 2 states that a player can obtain an allocation at least as good as a  $\frac{1}{n}$ th share of each commodity according to his reported preferences, provided that his reported preferences are in an appropriate class. One can state this condition directly on bargaining solutions. I have no interpretation for this property, however.

**Lemma 3.** *When  $a < 0$ ,  $f(S, a)$  satisfies Axiom 4 (midpoint domination).*

**Proof.** Let  $S \subset \mathcal{R}^n$  be a 0 – 1 normalized compact, convex set. When  $a < 0$ , maximizing  $W(\cdot, a)$  is equivalent to minimizing  $\sum_{i=1}^n s_i^a$ . Therefore, it suffices to show that if  $s^* = (s_1^*, \dots, s_n^*)$  solves:

$$\min \sum_{i=1}^n s_i^a \text{ subject to } s \in S \quad (13)$$

then  $s_i^* \geq \frac{1}{n}$  for all  $i$ . Without loss of generality, assume that  $s_1^* = \min \{s_1^*, \dots, s_n^*\}$ . Define the function  $\varphi(s_1) = \max \{\lambda : (s_1, \lambda s_{-1}) \in S\}$ . Since  $S$  is 0 – 1 normalized and convex,  $\varphi$  is a real-valued, decreasing, concave function satisfying:  $\varphi(s_1^*) = 1$ ,  $\varphi(1) \geq 0$ , and  $(s_1, \varphi(s_1)s_{-1}) \in S$  for all  $s_1 \in [0, 1]$ . Since  $s^*$  solves (13), it must be the case that  $s_1^*$  solves

$$\min \left\{ s_1^a + \sum_{i=2}^n [\varphi(s_1)s_i]^a \right\} \text{ subject to } s_1 \in [s_1^*, 1] \quad (14)$$

Differentiating the objective function in (14) and using  $a < 0$ , it follows that

$$(s_1^*)^{a-1} + \varphi'(s_1^*)\varphi(s_1^*)^{a-1} \sum_{i=2}^n (s_i^*)^a \leq 0, \quad (15)$$

where,  $\varphi'(s_1^*) = \lim_{h \rightarrow 0^+} \frac{\varphi(s_1^*+h) - \varphi(s_1^*)}{h}$ . Since  $\varphi(\cdot)$  is concave,

$$\varphi'(s_1) \geq \frac{\varphi(1) - \varphi(s_1)}{1 - s_1}. \quad (16)$$

Substituting (16) into (15) yields

$$(1 - s_1^*) (s_1^*)^{a-1} + [\varphi(1) - \varphi(s_1^*)] \varphi(s_1^*)^{a-1} \sum_{i=2}^n (s_1^*)^a \leq 0 \quad (17)$$

Since  $\varphi(s_1^*) = 1$  and  $\varphi(1) \geq 0$ , (17) implies that  $(s_1^*)^{a-1} \leq \sum_{i=1}^n (s_i^*)^a$  or

$$\frac{1}{s_1^*} \leq \sum_{i=1}^n \left[ \frac{s_i^*}{s_1^*} \right]^a \leq n \quad (18)$$

where the second inequality in (18) follows from  $a < 0$  and  $s_1^* = \min \{s_1^*, \dots, s_n^*\}$ .

As in the case of Lemma 2, Lemma 3 also holds for the Nash Bargaining solution ( $a = 0$ ).

Combining Lemmas 2 and 3, one obtains a generalization of Theorem 1.

**Theorem 3.** *If  $(p^*, x^*)$  is an EICE for the true preferences, then  $x^*$  is a Nash equilibrium allocation for the distortion game determined by the solution  $f(\cdot; a)$ . If  $a \in [0, 1]$ , then it is an equilibrium strategy for all players to report the truncated linear utility function  $L(\cdot; p^*)$ . If  $a \leq 0$ , then it is an equilibrium strategy for all players to report the linear utility function  $U(y) = p^* \cdot y$ .*

**Proof.** When  $a \in [0, 1]$ , result follows from Lemma 2 (and the same argument used to establish Theorem 1). When  $a \leq 0$ , the result follows from Lemma 3 and Sobel [1981].

The family of bargaining solutions shares the property that agents want to act as if the marginal unit they are allocated is valuable. The optimal manipulation always involves claiming that the first units allocated add the minimum possible utility (subject to the concavity constraint). Strategies differ as to how agents should represent their utility for what they do not receive in equilibrium. Utilitarian solutions ( $a > 0$ ) induce agents to claim that the units that they do not receive are worthless. Denying an interest in these units permits agents to exaggerate their interest for the units that they actually receive. Egalitarian solutions ( $a < 0$ ) favor an agent who appears to be getting a small utility relative to the maximum available utility for that agent. Consequently, it is in the interest of an agent faced with an egalitarian planner to report preferences in a way that exaggerate his maximum utility relative to the utility he receives. This force creates an incentive to report linear preferences.

## REFERENCES

- [1] Cao, X. [1981], "Preference Functions and Bargaining Solutions," IEEE 161-71.
- [2] Crawford, V. and H. Varian [1979], "Distortion of Preferences and the Nash Theory of Bargaining," *Economics Letters* 3, 203-6.
- [3] Dhillon, A. and J.-F. Mertens [1996], "Relative Utilitarianism: An Improved Axiomatisation," CORE.
- [4] Hurwicz, L., E. Maskin, and A. Postlewaite [1995], "Feasible Nash Implementation of Social Choice Rules When Designer Does not Know Endowments or Production Sets," in The Economics of Informational Decentralization, John O. Ledyard (ed.), Boston: Kluwer Academic Publishers.
- [5] Karni, E. [1996], "Impartiality and Interpersonal Comparisons of Ordinal Well-Being," Johns Hopkins.
- [6] Karni, E. and Z. Safra [1996], "Individual Sense of Justice," Johns Hopkins.
- [7] Kurz, M. [1977], "Distortion of Preferences, Income Distribution, and the Case for a Linear Income Tax," *Journal of Economic Theory* 14, 291-8.
- [8] Kurz, M. [1980], "Income Distribution and Distortion of Preferences: The *l*-commodity Case," *Journal of Economic Theory* 22, 99-106.
- [9] Mas-Colell, A. [1980], "Remarks on the Game-Theoretic Analysis of a Simple Distribution of Surplus Problem," *International Journal of Game Theory* 9, 125-40.
- [10] Nash, J. [1951], "The Bargaining Problem," *Econometrica* 28, 155-62.
- [11] Roth, A. [1979], Axiomatic Models of Bargaining. Berlin and Heidelberg: Springer Verlag.
- [12] Segal, U. [1998], "Let's Agree that All Dictatorships are Equally Bad," University of Western Ontario, March.
- [13] Sobel, J. [1981], "Distortion of Utilities and the Bargaining Problem," *Econometrica* 49, 597-619.
- [14] Thomson, W. [1979], "The Manipulation of the Shapley-Value."

- [15] Thomson, W. [1984], "The Manipulability of Resource Allocation Mechanisms," *Review of Economic Studies* 51, 447-60.
- [16] Thomson, W. [1987], "The Vulnerability to Manipulative Behavior of Resource Allocation Mechanisms Designed to Select Equitable and Efficient Outcomes," in Groves, T., R. Radner, and S. Reiter (eds.), Information, Incentives, and Economic Mechanisms. Minneapolis: University of Minnesota Press.
- [17] Thomson, W. [1994], "Cooperative Models of Bargaining," in Aumann, R. and S. Hart (eds.), Handbook of Game Theory, volume II. Amsterdam: North Holland.