Econometrica, Vol. 49, No. 3 (May, 1981)

DISTORTION OF UTILITIES AND THE BARGAINING PROBLEM

BY JOEL SOBEL[1]

Given two agents with von Neumann-Morgenstern utilities who wish to divide n commodities, consider the two-person noncooperative game with strategies consisting of concave, increasing von Neumann-Morgenstern utility functions as well as rules to break ties and whose outcomes are some solution to the bargaining game determined by the strategies used. It is shown that, for a class of bargaining solutions which includes those of Nash and Raiffa, Kalai and Smorodinsky, any constrained equal-income competitive equilibrium allocation for the true utilities is a Nash equilibrium outcome for the noncooperative game.

1. INTRODUCTION

IT IS OFTEN THE CASE in economic or game theoretical models that predictions are based on information that is not observable. For example, Nash's [**15**] theory of bargaining determines an outcome that depends on the bargainers' von Neumann-Morgenstern utility functions. Kurz [**11, 12**] has recently introduced a technique for analyzing such models that yields predictions about the outcome of a game without relying on unobserved information. The technique of Kurz has been adopted by Crawford and Varian [**3**] to analyze the outcomes of the Nash bargaining solution over the division of a single commodity. The purpose of this paper is to extend the results of Crawford and Varian to bargaining over several commodities.

The approach used by Kurz [**11, 12**] and Crawford-Varian [**3**] is to embed the original game into a noncooperative *distortion game* in which the players' strategies consist of utility functions that may be distorted from their true utilities for strategic purposes. The outcomes are given by the solution to the underlying game determined by the reported utilities. If the Nash equilibria of the distortion game share common properties, then a description of the original game situation has been made without relying on information about the unobserved utility functions.

Kurz's [**11, 12**] papers are related to the work of Aumann and Kurz [**1, 2**] on the determination of taxes in an exchange economy. Aumann and Kurz postulate a particular solution concept and then characterize the income tax schedules and allocations that result from it. In [**11, 12**] Kurz observes that the Aumann and Kurz solution depends on the agents' von Neumann-Morgenstern utility functions. Since these functions are not directly observable, agents cannot be prevented from misrepresenting them if it is to their advantage to do so. Kurz therefore studies the game that results if each agent can report any utility function in an admissible class $\mathcal{U}$. $\mathcal{U}$ is intended to include all functions that are

[1] I am very grateful to Vincent Crawford, who suggested the problem considered in this paper and made several crucial suggestions that helped to make its solution possible. A conversation with Theodore Groves was also of value.

credible utilities for an agent. Kurz takes $\mathcal{U}$ to be the set of all von Neumann-Morgenstern utility functions that are increasing, concave, and continuously differentiable. In [11] he shows that for the one-commodity Aumann-Kurz model [1], reporting any linear function in $\mathcal{U}$ is a dominant strategy for each agent. The marginal tax rate implied by the use of linear strategies is 50 per cent. Kurz generalizes this result to the n-commodity case in [12]. Once again, players have dominant strategies that lead to a marginal tax rate of 50 per cent. In this case, however, the dominant strategy reported utility functions need not be linear. The significance of Kurz's results is that, regardless of the true preferences of the agents, the distortion game has a dominant strategy equilibrium that yields a Pareto-efficient outcome.

Crawford and Varian [3] use the methods of Kurz to analyze the effect that distortion of utilities has on the solutions to bargaining games. Assuming that agents may report any concave, increasing utility function, they find that in Nash [15] or Raiffa [17]-Kalai-Smorodinsky [8] bargaining over the division of a single good reporting linear utility functions constitutes a dominant strategy equilibrium. The allocation implied by the equilibrium reports is equal division. The purpose of this paper is to generalize this result to include bargaining over more than one commodity.

The main link between the one-commodity bargaining game and its multi-commodity generalization has to do with the effect a player's attitude towards risk has on the utility he receives at the solution. A utility function U is said to be *more risk averse* than V if there is an increasing concave function k with $U = k(V)$; an agent is more risk averse than another agent if his utility function is more risk averse than the other agent's. Kihlstrom, Roth, and Schmeidler [10][2] show that for a class of bargaining solutions that include the Nash [15] and the Raiffa [17]-Kalai-Smorodinsky [8] solution, a player's utility increases as his opponent becomes more risk averse. This result, which is related to a theorem of Kannai [9], makes it possible to deduce that players will report linear utilities in the one-commodity distortion game. This follows because all monotonic preferences defined over one commodity are (ordinally) equivalent. The Kihlstrom-Roth-Schmeidler results thus imply that the players will select the least risk averse representation of these preferences. In the one-commodity case this will be a linear function. As long as the solution for the bargaining problem satisfies the axioms of Pareto optimality, symmetry, and invariance with respect to affine transformations of utility, the linear strategies give rise to equal division.

The situation is made more complicated in the n-commodity case because there are many possible ordinal rankings of the outcomes. While the Kihlstrom, Roth, and Schmeidler result restricts the possible strategies that the players will find advantageous to report, a broad class of possible distortions (including any increasing linear function) cannot be excluded on the basis of their theorem. Consequently, it is not surprising that the characterization of equilibria for the distortion game is less satisfactory for multi-commodity bargaining than for

[2] This result is presented in Roth [18, pp. 38–48, 104–105].

one-commodity bargaining. Furthermore, although the solution to the bargaining problem specifies unique utility levels for both players with respect to their reported utilities, in general, there will be more than one outcome that gives rise to these utility levels. Since there is no reason to expect that these outcomes are utility equivalent for the true utilities, the strategy space must be augmented by tie-breaking rules that provide for a selection from the solution correspondence.

In spite of the greater complexity, the Nash equilibria for the n-commodity distortion game have some attractive properties. For a class of bargaining solutions that include those of Nash and Raiffa-Kalai-Smorodinsky, I show that any constrained equal-income competitive equilibrium (EICE) allocation (a constrained equilibrium allocation reached when agents have equal initial endowments) for the true utilities is a Nash equilibrium allocation for the distortion game. The equilibrium strategy for both players is to report linear utilities with indifference surfaces parallel to the hyperplane that supports the EICE. These strategies will result in a set of allocations that solve the bargaining problem but the players will be able to agree on a most preferred outcome. Since agents are assumed to have concave utility functions, this guarantees the existence of Pareto-efficient Nash equilibria. Moreover, the EICE allocations are the only Nash outcomes provided both players are required to report linear utilities. On the other hand, there is no reason to expect the competitive allocations to be unique. Thus, dominant strategy equilibria are impossible. Also, in general, the distortion game has inefficient Nash equilibria. However, all of the equilibria are "good" in a certain sense. Specifically, at any Nash equilibrium outcome, each agent prefers his allocation to that of the other agent.[3]

The class of bargaining game solutions for which these results are valid is described in Section 3. In addition to the axioms of Pareto optimality, symmetry, and independence of positive affine transformations of utility, axioms common to the Raiffa-Kalai-Smorodinsky and Nash theories, the solution is required to satisfy another property. In a bargaining game in which the players' utility functions are normalized so that each player receives utility 0 at the disagreement point and utility 1 at his most preferred outcome, a solution satisfies the axiom of *symmetric monotonicity* if each player receives utility of at least $\frac{1}{2}$ at the solution. In Section 3, it is shown that both the Nash and the Raiffa-Kalai-Smorodinsky solutions have this property. Furthermore, it is shown that symmetric monotonicity is guaranteed if the solution is symmetric, Pareto optimal, and *risk sensitive* as defined by Kihlstrom, Roth, and Schmeidler [**10**]. A solution is risk sensitive if a player prefers to bargain against the more risk averse of two players.

The problem considered in this paper may be viewed as an arbitration problem under ignorance. An arbitrator is assigned the task of determining a "good" outcome to the bargaining game. A possible technique for the arbitrator would be to ask the players to report their utility functions and then determine an outcome to the resulting bargaining game according to a fixed solution

[3] That no agent prefers another agent's allocation to his own is the definition of equity first used by Foley [**4**].

concept. If the arbitrator has no knowledge about the players' true utilities except that they are in the set $\mathcal{U}$, then the agents will be playing the distortion game described here. Notice that if the arbitrator can restrict the reported preferences to be linear, EICE allocations are assured. This restriction may be attractive to the arbitrator because it reduces the information which needs to be reported to an $n - 1$ dimensional vector (the constant marginal rates of substitution of each player, rather than entire utility functions) and the strategies used to resolve ties. The restriction may be acceptable to the agents because it guarantees EICE allocations.

Another model of arbitration under ignorance is discussed by Kalai and Rosenthal [7]. A cooperative two-player bimatrix game is transformed into a noncooperative game by an arbitrator. Players are asked to report a mixed strategy (threat) and two payoff matrices. The arbitrator then determines an outcome using a procedure that generalizes Nash's [16] extended bargaining solution if the players report the same payoff matrices. If the players report different payoff matrices then they receive the threat outcome. Assuming that the players know the underlying cooperative game and that the arbitrator knows only the dimensions of its payoff matrices, Kalai and Rosenthal show that reporting the true payoff matrices and appropriate mixed strategies forms a Nash equilibrium for the arbitration game. Moreover, the equilibrium outcome is Pareto efficient and individually rational.

Both Kalai and Rosenthal [7] and I assume that agents have perfect information about the game situation they are facing. The results would be more compelling if they remained valid under uncertainty. Suppose that the distortion game is being played with only linear strategies admissible, and the true utility functions of the agents are such that the EICE is unique. If player 1 is only slightly uncertain about his opponent's utility function (meaning that he knows with certainty that his opponent's utility function is "close" to a specific function), then player 1 knows—under certain regularity assumptions—that the true distortion game has a unique Nash equilibrium outcome, and he knows approximately where it is. Thus, there is a possibility that an adjustment process could be designed that converges to the EICE. In the Kalai-Rosenthal model, there appears to be no restriction that can be made that would make the Nash equilibrium unique or even locally unique. Therefore, while under certainty the players in the Kalai-Rosenthal game are likely to report truthfully and reach a "good" outcome, the introduction of a slight amount of uncertainty makes the argument for truthful reports lose much of its force: if the players' beliefs about the true game situation differ, then it is quite possible that reporting what are believed to be the true payoff matrices will lead to an equilibrium outcome inferior to the "good" outcome. Thus, in the Kalai-Rosenthal model, unless the players know the underlying game with certainty, the existence of multiple equilibria makes it unlikely that an adjustment process converging to the "good" Nash equilibria could be designed.

The requirement that players know each other's characteristics with certainty is a strong one. However, it appears that Pareto-efficient outcomes cannot be guaranteed in models in which the players are uncertain about their opponent's characteristics. In a model of arbitration under uncertainty, Myerson [**14**] observes that the set of allocations that arise from incentive-compatible mechanisms (allocations Myerson calls *incentive feasible*) is strictly contained in the set of all feasible allocations. Myerson argues that the arbitrator should be satisfied with selecting a "good" outcome which is undominated by any other incentive feasible allocation (but which may be Pareto inefficient), and proves that this can be done.

Closely related to my results are those of Thomson [**19**, **20**]. Thomson studies the Nash equilibria for the distortion game derived from a class of performance correspondences that yield individually rational and Pareto-efficient outcomes. Thomson [**19**] finds that if the reported utility functions are restricted to be twice continuously differentiable, concave, and have the transferable utility (t.u.) property, 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. This result is generalized to a broader class of performance correspondences in [**20**]. As the Nash bargaining solution and the Shapley value, with appropriate disagreement outcomes, coincide under transferable utility, these results are quite similar to mine. The main differences fall into two classes: the nature of allowable utilities and the range of generality.

When strategy spaces consisting of utility functions with transferable utility are used, the class of admissible utility functions is then broad enough to eliminate the need for explicit tie-breaking rules. If a tie occurs, a player typically has another admissible strategy that will allow him to break the tie so as to receive his most preferred outcome. In equilibrium, ties will occur unless the original endowments are Pareto-efficient for the true preferences. However, these ties can be broken because both players are able to agree on a most preferred outcome. On the other hand, when reports must be t.u. utility functions, Nash equilibrium outcomes other than the constrained competitive equilibria for the economy do not occur.[4] Inefficient Nash equilibria for the distortion game derived from the Nash bargaining solution may occur if any smooth report is allowed; an example is given in Section 5.

It seems unreasonable to require transferable utility reports, a priori. However, for a certain class of games this restriction may be justified. When tie-breaking rules are used, I can show that a player always has a linear best response for the Nash and Raiffa-Kalai-Smorodinsky distortion games. Thus, nonlinear strategies are dominated by linear strategies. It is reasonable to assume that only linear strategies will be used in this situation. A similar analysis may make it possible to delete all nontransferable utility preferences in Thomson's model.

Thomson's results apply to a different range of solutions than do mine. They are more general in one direction: his results are valid for any number of players.

[4]These results require that the reported preferences be twice continuously differentiable. If reports are not smooth, there may be other equilibria.

The fact that the Nash equilibrium outcomes include the competitive allocations remains valid in my model for any number of bargainers; however, other Nash equilibria will exist in general. This difference is probably the result of the different strategy spaces. Also, Thomson's results apply to any initial endowments. However, the constrained competitive equilibria with respect to any initial endowments can be obtained in my framework by varying the disagreement outcomes.

Besides requiring Pareto efficiency with respect to the reported preferences and individual rationality with respect to the given initial endowments, the class of solutions for which Thomson's results are valid have the property that equilibrium strategies have a "flatness" quality. Specifically, it is necessary that initial endowments be Pareto efficient with respect to equilibrium strategies. This will be true, for example, if all players report linear utilities with indifference surfaces parallel to the hyperplane that supports the competitive allocation. The underlying solutions that I deal with yield Pareto-efficient outcomes. The flatness property is satisfied when only linear reports are allowed, but not in general. Symmetric monotonicity can be viewed as an individual rationality requirement: provided that reports are linear, this assumption guarantees that outcomes are at least as good as equal division with respect to the reported utilities. The solutions I consider need not guarantee outcomes that are at least as preferred as equal division when nonlinear strategies are allowed.

It should be emphasized that my results depend on strategy spaces that are different from the set of admissible utility functions. The results of Hurwicz **[5]** guarantee that individually rational and Pareto-efficient allocations cannot coincide with the Nash equilibrium outcomes of a mechanism that has only preferences as strategies. Thus, my results depend in an essential way on the fact that a player's strategy includes tie-breaking rules as well as a utility function.

The distortion game is defined formally in Section 2. In Section 3, the class of bargaining solutions to be used is described. The main results are presented in Section 4. Finally, Section 5 further characterizes the equilibria of the Nash and the Raiffa-Kalai-Smorodinsky distortion game.

2. DEFINITIONS AND NOTATION

Consider two agents with von Neumann-Morgenstern utility functions who are to divide a bundle of n commodities. Units are chosen so that there is exactly one unit of each commodity. Letting $\mathbf{a} \equiv (a, \ldots, a)$, an outcome will be an element of the set

$$T = \{x \in \mathbb{R}^n : \mathbf{0} \leqslant x \leqslant \mathbf{1}\},$$

where agent 1 receives x and agent 2 receives $\mathbf{1} - x$. The true utility function of player 1 is denoted by u; of player 2, v. These functions are assumed to be concave and strictly increasing in T. Thus if $x, x' \in T$ and $x > x'$, then $u(x) > u(x')$ and $v(1 - x') > v(1 - x)$. The players report utilities that are restricted to lie in the class $\mathcal{U}$, where $\mathcal{U}$ consists of those functions: $U: T \to [0, 1]$ such that (i) U is continuous, strictly increasing, and concave in T; (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 their true preferences. Thus, 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. Since the solutions to the bargaining problem to be discussed are independent of affine transformations, condition (ii) is inessential.

The distortion game for the bargaining problem is played by each agent revealing a utility function in $\mathcal{U}$ and an element in a set $\mathfrak{M}$ that will be used to resolve ties. Typically, U will denote the function revealed by player 1; V that of player 2. Given these reports, a set of outcomes $B(U, V)$ is selected. $B(U, V)$ is the set of allocations that give rise to a bargaining solution determined by U and V. The properties of solution concepts used to define B will be discussed in Section 3. However, in order to define the distortion game, it is only necessary that $B(U, V)$ be a non-empty subset of T for all U and V in $\mathcal{U}$.

In order to completely characterize the strategies for the distortion game, the way in which a single element of B is selected must be described. In addition to a utility function, each player will report an element from a set $\mathfrak{M}$. An outcome will then be selected by a function $\bar{B}$. Thus, $\bar{B}: \mathcal{U} \times \mathfrak{M} \times \mathcal{U} \times \mathfrak{M} \to T$ with $\bar{B}(U, f; V, g) \in B(U, V)$ for all U and $V \in \mathcal{U}$ and f and $g \in \mathfrak{M}$. $(\mathfrak{M}, \bar{B})$ will be called the *tie-breaking pair* associated with B.

DEFINITION: The strategies $(U^*, f^*; V^*, g^*)$ constitute a *Nash equilibrium for the distortion game* determined by B with tie-breaking pair $(\mathfrak{M}, \bar{B})$ if and only if (i) $(U^*, f^*; V^*, g^*) \in \mathcal{U} \times \mathfrak{M} \times \mathcal{U} \times \mathfrak{M}$; (ii) $\bar{B}(U^*, f^*; V^*, g^*)$ solves: $\max u(x)$ subject to $x \in \{\bar{B}(U, f; V^*, g^*) : (U, f) \in \mathcal{U} \times \mathfrak{M}\}$; (iii) $\mathbf{1} - \bar{B}(U^*, f^*; V^*, g^*)$ solves: $\max v(y)$ subject to $\mathbf{1} - y \in \{\bar{B}(U^*, f^*; V, g) : (V, g) \in \mathcal{U} \times \mathfrak{M}\}$.

An appropriate choice of tie-breaking pair can allow the Nash equilibria for the distortion game to be characterized in terms of the reported utility functions and the correspondence B.

DEFINITION: A tie-breaking pair $(\mathfrak{M}, \bar{B})$ for the solution B is *unrestricted* if and only if, for all U and $V \in \mathcal{U}$ and f and $g \in \mathfrak{M}$,

$$B(U, V) = \left\{\bar{B}(U, f; V, h) : h \in \mathfrak{M}\right\} = \left\{\bar{B}(U, h; V, g) : h \in \mathfrak{M}\right\}.$$

Suppose player two reports a utility function $V \in \mathcal{U}$ and a tie-breaking strategy $g \in \mathfrak{M}$. If $(\mathfrak{M}, \bar{B})$ is unrestricted then, for any $U \in \mathcal{U}$, player one has a tie-breaking strategy that can cause any outcome in $B(U, V)$ to be selected. This makes the following description of Nash equilibria possible.

PROPOSITION 1: *If $(\mathfrak{M}, \bar{B})$ is an unrestricted pair for a solution B then, for any U^* and $V^* \in \mathcal{U}$, there exist f^* and $g^* \in \mathfrak{M}$ such that $(U^*, f^*; V^*, g^*)$ is a Nash* *equilibrium for the distortion game determined by B with tie-breaking pair $(\mathcal{M}, \bar{B})$ if and only if there exists $x^* \in B(U^*, V^*)$ such that x^* solves*:

$$\max u(x) \text{ subject to } x \in \{B(U, V^*) : U \in \mathcal{U}\}$$

and $\mathbf{1} - x^*$ *solves*:

$$\max v(y) \text{ subject to } \mathbf{1} - y \in \{B(U^*, V) : V \in \mathcal{U}\}.$$

PROOF: If $(U^*, f^*; V^*, g^*)$ is a Nash equilibrium then the conditions are satisfied when $x^* = \bar{B}(U^*, f^*; V^*, g^*)$. This follows since $(\mathcal{M}, \bar{B})$ is unrestricted and so

$$\{\bar{B}(U, f; V^*, g^*) : (U, f) \in \mathcal{U} \times \mathcal{M}\} = \{B(U, V^*) : U \in \mathcal{U}\} \qquad \text{and}$$

$$\{\bar{B}(U^*, f^*; V, g) : (V, g) \in \mathcal{U} \times \mathcal{M}\} = \{B(U^*, V) : V \in \mathcal{U}\}.$$

Conversely, if an x^* exists as described in the proposition then $(U^*, f^*; V^*, g^*)$ is a Nash equilibrium provided that f^* and g^* are selected so that $x^* = \bar{B}(U^*, f^*; V^*, g^*)$. This is possible because $(\mathcal{M}, \bar{B})$ is unrestricted. *Q.E.D.*

Thus, if an unrestricted pair can be found to break ties, Nash equilibria can be characterized in terms of reported utility functions and the solution correspondence.

The next result constructs unrestricted tie-breaking rules.

PROPOSITION 2: *Given any solution correspondence B, there exists an unrestricted pair $(\mathcal{M}, \bar{B})$ associated with B.*

PROOF: Let $\mathcal{M} = \mathbb{R}^n$ and define $\bar{B}$ by

$$\bar{B}(U, x; V, y) = \begin{cases} \frac{1}{2}(x+y) & \text{if } \frac{1}{2}(x+y) \in B(U, V), \\ \text{any element of } B(U, V) & \text{if } \frac{1}{2}(x+y) \notin B(U, V). \end{cases}$$

Clearly player one can obtain any $z \in B(U, V)$ given that player two is using strategy (V, y) by using a tie-breaking strategy x with $x = 2z - y$ and reporting the utility function U. Similarly, if player one reports the function U, player two can obtain any element of $B(U, V)$ by responding with the appropriate tie-breaking rule. *Q.E.D.*

In this paper all ties will be resolved using unrestricted tie-breaking rules. Proposition 2 says that this can be done, while Proposition 1 provides a characterization of Nash equilibria when unrestricted tie-breaking rules are used. Notice that it is not necessary for the players to actually report complicated tie-breaking strategies, provided that they know that some unrestricted procedure to break ties exists. Proposition 1 guarantees that if $B(U^*, V^*)$ consists of more than one point for equilibrium strategies U^* and V^*, then both players can agree on a most preferred outcome. Thus there will be no difficulty in making a selection from $B(U^*, V^*)$. The existence of unrestricted pairs rules out the possibility that, at an equilibrium, players cannot agree on a selection from the solution correspondence.

The characterization of Nash equilibria given by Proposition 1 will be used throughout this paper. Thus, a Nash equilibrium will be described by a triple (U^*, V^*, x^*) satisfying the cond tions of Proposition 1. The vector x^* will be called the Nash equilibrium outcome or allocation.

In what follows, Nash equilibrium allocations will be related to certain competitive outcomes.

DEFINITION: A constrained *equal-income competitive equilibrium* (EICE) is a pair, $(p^*; x^*)$ where (i) $p^* \in \mathbb{R}^n$, $p^* \geqslant \mathbf{0}$, $p^* \neq \mathbf{0}$; (ii) $x^* \in T$; (iii) x^* solves:

$$\max u(x) \text{ subject to } p^* \cdot x \leqslant \tfrac{1}{2} p^* \cdot \mathbf{1} \quad \text{and} \quad x \in T;$$

(iv) $\mathbf{1} - x^*$ solves:

$$\max v(y) \text{ subject to } p^* \cdot y \leqslant \tfrac{1}{2} p^* \cdot \mathbf{1} \quad \text{and} \quad y \in T.$$

In an equal-income competitive equilibrium, both agents make demands subject to a budget constraint only. A constrained equal-income competitive equilibrium requires that these demands do not exceed the total resources available.[5] It is well known that any equal-income competitive equilibrium is a constrained equal-income competitive equilibrium and that, provided preferences are convex, any interior constrained equal-income competitive equilibrium is an equal-income competitive equilibrium.

The vector x^* will be referred to as the competitive allocation.

On occasion, a vector $p = (p_1, \ldots, p_n)$ will be used to refer to the linear function from T to $\mathbb{R}$, where

$$p(x) \equiv p \cdot x \equiv \sum p_i x_i.$$

No confusion should arise.

3. SOLUTIONS TO BARGAINING GAMES

The underlying bargaining problem can be formulated as follows. A *bargaining game* is characterized by a pair (S, d), where: (i) $d = (d_1, d_2) \in \mathbb{R}^2$; (ii) $S \subset \mathbb{R}^2$ is compact, convex, and contains d as well as some point $x > d$.

The set S is interpreted as the set of feasible utility payoffs to the players. A point $x = (x_1, x_2)$ can be achieved if both players agree to it. In that case, player

[5] The concept of constrained competitive equilibria was introduced by Hurwicz, Maskin, and Postlewaite [6]. It is the smallest extension of the competitive correspondence that can be implemented in Nash strategies.

1 receives x_1 and player 2 receives x_2. If the players are unable to agree, then the outcome d, called the disagreement outcome, is the result.

In what follows, the set S will depend on the reported utilities U and V, and will be defined as the set of feasible utility payoffs. That is,

$$S = S(U,V) = \{(x_1,x_2) : 0 \leqslant x_1 \leqslant U(t), 0 \leqslant x_2 \leqslant V(\mathbf{1}-t) \text{ for some } t \in T\}.^6$$

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

Notice that when the functions U and V are in $\mathcal{U}$, the set S is compact, convex, and contains a point $x > (0,0)$. In fact, $(1,0) \in S$, $(0,1) \in S$, and $S \subset \{x_1, x_2) : 0 \leqslant x_1, x_2 \leqslant 1\}$. Such a game will be called 0-1 *normalized*.[7]

Nash [**15**] introduced the concept of a *solution* to a bargaining game. A solution is a function f, defined on the class of all bargaining games with $f(S,d) = (f_1(S,d), f_2(S,d)) \in S$ for all pairs (S,d). Nash characterized a particular solution in terms of the following axioms.

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

Axiom 2 (Symmetry): *If (S,d) is a symmetric game (that is, $(x_1,x_2) \in S$ if and only if $(x_2,x_1) \in S$, and $d_1 = d_2$) then $f_1(S,d) = f_2(S,d)$.*

Axiom 3 (Independence of Equivalent Utility Representations): *If (S,d) and (S',d') are bargaining games such that*

$$S' = \{(a_1x_1 + b_1, a_2x_2 + b_2) : (x_1,x_2) \in S\} \quad \textit{and}$$
$$d' = (a_1d_1 + b_1, a_2d_2 + b_2) \quad \textit{where} \quad a_1 \textit{ and } a_2 > 0, \quad \textit{then}$$
$$f(S',d') = (a_1f_1(S,d) + b_1, a_2f_2(S,d) + b_2).$$

Axiom 4 (Independence of Irrelevant Alternatives): *If (S,d) and (S',d) are bargaining games such that $S \subset S'$ and $f(S',d) \in S$, then $f(S,d) = f(S',d)$.*

Nash's result in [**15**] was that Axioms 1–4 characterize a solution, η. In terms of utilities, $U, V \in \mathcal{U}$,

$$N(U,V) = \{x \in T : x \in \arg\max\{U(y)V(\mathbf{1}-y) : y \in T\}\},$$

[6]The definition of $S(U,V)$ includes a free disposal assumption. Another (equivalent) definition would allow outcomes that do not distribute all of the commodities. That is,

$$S(U,V) = \{(a,b) : a = U(x), b = V(y), x, y \in T \quad \text{and} \quad x + y \leqslant 1\}.$$

[7]The normalization of $S(U,V)$ anticipates the axiom of independence of equivalent utility representations (Axiom 3). With that axiom, any game can be taken to be 0-1 normalized without loss of generality.

is the set of allocations that give rise to the Nash solution to the bargaining game with disagreement outcome $(0,0)$. That is, for all $x \in N(U, V)$,

$$\eta(S(U,V),(0,0)) = (U(x), V(1-x)).^{8}$$

Axiom 4 has been criticized for a variety of reasons (see, for example, Luce and Raiffa [13]). Another solution to the bargaining game has been presented by Raiffa [14] and axiomatized by Kalai and Smorodinsky [8]. Kalai and Smorodinsky replace Nash's Axiom 4 with an axiom of monotonicity. To state the axiom formally, it is necessary to define, for all bargaining games (S,d),

$$b_1(S) = \sup\{x_1 \in \mathbb{R} : \text{for some } x_2 \in \mathbb{R}\ (x_1,x_2) \in S\}$$

and

$$b_2(S) = \sup\{x_2 \in \mathbb{R} : \text{for some } x_1 \in \mathbb{R}\ (x_1,x_2) \in S\}.$$

Also, let g_S be a function defined for $x_1 \leqslant b_1(S)$ as follows:

$$\begin{aligned} g_S(x_1) &= x_2 \quad \text{if } (x_1,x_2) \in S \quad \text{and} \quad (x_1,x) \in S \quad \text{implies } x_2 \geqslant x, \\ &= b_2(S) \quad \text{if no such } x_2 \text{ exists.} \end{aligned}$$

Then $g_S(x)$ is the maximum player 2 can get if player 1 gets at least x. Because S is compact, $b_1(S)$ and $b_2(S)$ are finite and attained by points in S. I can now state the following axiom:

AXIOM 5 (Monotonicity): *If (S,d) and (S',d) are bargaining pairs such that*

$$b_1(S) = b_1(S') \quad \textit{and} \quad g_S \leqslant g_{S'}, \quad \textit{then} \quad f_2(S,d) \leqslant f_2(S',d).$$

This axiom says that if the maximum feasible utility level that player 2 can obtain is increased for every utility level that player 1 may demand, then the utility level assigned to player 2 according to the solution should also be increased.

Kalai and Smorodinsky [8] show that Axioms 1, 2, 3, and 5 characterize a solution, ξ, to the bargaining problem. In terms of utilities, $U, V \in \mathfrak{U}$,

$$K(U,V) = \{x \in T : x \in \{\arg\max U(y) : U(y) = V(1-y), y \in T\}\}$$

is the set of allocations that give rise to the Raiffa-Kalai-Smorodinsky (R-K-S) solution to the bargaining game with disagreement outcome $(0,0)$. That is, for all $x \in K(U,V)$,

$$\xi(S(U,V),(0,0)) = (U(x), V(1-x)).$$

Many of the results to be presented on the equilibria to distortion games are valid for a class of solutions that include both the Nash bargaining solution and

[8] The Nash solution gives rise to a single level of utility for each player. However, unless the utility functions are strictly concave, there may be several allocations that give rise to these utility levels.

the R-K-S solution. The crucial property, in addition to Axioms 1–3, seems to be the following.

AXIOM 6 (Symmetric Monotonicity): *If* $(S,(0,0))$ *is a* 0-1 *normalized bargaining game, then* $f(S,(0,0)) \geqslant (\frac{1}{2},\frac{1}{2})$.

Axiom 6 can be stated in a more general fashion, but the above formulation is sufficient for my purposes.

Any solution that satisfies Axioms 1, 2, 3, and 6 will be called *admissible*. It turns out that the equilibria for the distortion game can be characterized provided that the underlying bargaining solution is admissible.

Axiom 6 is a weaker assumption than Axiom 5. One consequence of Axiom 5 in this context is that if $d \in S \subset S'$, then $f(S',d) \geqslant f(S,d)$. For any 0-1 normalized game S', the convex hull of $(0,1)$, $(0,0)$, and $(1,0)$, S, is contained in S'. Thus, whenever Axioms 1, 2, 3, and 5 are satisfied by a solution f,

$$f(S',(0,0)) \geqslant f(S,(0,0)) = (\tfrac{1}{2},\tfrac{1}{2}).$$

It follows that the R-K-S solution is admissible. The Nash bargaining solution is also admissible. To see this, it is convenient to present another axiom, which was introduced by Kihlstrom, Roth, and Schmeidler **[10]** in order to study the effect that a player's attitude towards risk has on solutions to the bargaining game.

AXIOM 7 (Risk Sensitivity): *Suppose the bargaining game* (S,d) *is transformed into a game* (S',d') *by replacing player* 2, *say, with a more risk averse player* (*that is, if* $S = S(U,V)$ *then* $S' = S'(U,k(V))$ *where* k *is increasing and concave*); *then* $f_1(S',d') \geqslant f_1(S,d)$.

Any solution f that satisfies Axiom 7 describes a bargaining process in which it is advantageous to have a highly risk averse opponent. The axioms that have been presented are related by the following result.

LEMMA 1: *If* f *is a solution that satisfies Axioms* 1, 2, *and* 7, *then* f *satisfies Axioms* 3 *and* 6.

PROOF: Kihlstrom, Roth, and Schmeidler **[10]** show that Axioms 1 and 7 imply Axiom 3. It therefore suffices to show that Axioms 1, 2, 3, and 7 imply Axiom 6. Let $S = \{(a,b) \in \mathbb{R}^2 : a,b \geqslant 0,\ a+b \leqslant 1\}$; then $S = S(U,V)$ when $U(x) \equiv V(x) \equiv \frac{1}{n}\mathbf{1} \cdot x$. By Axioms 1 and 2, $f(S,(0,0)) = (\frac{1}{2},\frac{1}{2})$. I will show that if S' is 0-1 normalized, then it can be obtained from S by replacing player 2 by a more risk averse player. Then Axiom 7 will imply $f_1(S',(0,0)) \geqslant \frac{1}{2}$. The lemma will follow by symmetry. Let ϕ be a parametrization of the Pareto-efficient set of S'. That is, suppose the northeast boundary of S' can be written as

$$P' = \{(a,\phi(a)) : 0 \leqslant a \leqslant 1\}.$$

Clearly ϕ is decreasing and concave. Further, because of the normalization, $\phi(0)=1$ and $\phi(1)=0$. Now let k be the function defined by $k(a)=\phi(1-a)$. Then k is increasing, concave and satisfies $k(0)=0$, $k(1)=1$. Moreover,

$$P' = \{(a, k(1-a)) : 0 \leqslant a \leqslant 1\}.$$

Thus k takes the Pareto set of S onto the Pareto set of S'. It follows that S' can be derived from player 1 using the strategy $L(x) \equiv \frac{1}{n}\mathbf{1}\cdot x$ and player 2 using the strategy $k(L)$. As noted earlier this is sufficient to prove the lemma.

Q.E.D.

Kihlstrom, Roth, and Schmeidler **[10]** show that the R-K-S and the Nash solutions satisfy Axiom 7. Therefore, these solutions are admissible.

4. MAIN RESULTS

In this section, it will be assumed that the distortion game is determined by an admissible solution to the bargaining game, f. Such a game will be referred to as an *admissible distortion game*. Associated with a solution f and functions U and $V \in \mathcal{U}$ there is a set $B(U, V)$ defined by

$$B(U,V) = \{x \in T : f(S(U,V),(0,0)) = (U(x), V(\mathbf{1}-x))\}.$$

$B(U, V)$ is the set of outcomes giving rise to the utilities specified by the solution f.

The main theorem can now be stated.

THEOREM 1: *If $(p^*; x^*)$ is an EICE for the true preferences, then (p^*, p^*, x^*) is a Nash equilibrium for any admissible distortion game.*

PROOF: Since B is admissible, it follows that for all U and $V \in \mathcal{U}$,

(1) $\quad x \in B(U, p^*) \quad$ implies $\quad p^* \cdot (\mathbf{1}-x) \geqslant \frac{1}{2}$

and

(1′) $\quad x \in B(p^*, V) \quad$ implies $\quad p^* \cdot x \geqslant \frac{1}{2}$.

Moreover, by symmetry,

$$B(p^*, p^*) = \{x \in T : p^* \cdot x = \tfrac{1}{2}\}.$$

On the other hand, since $(p^*; x^*)$ is an EICE, x^* solves:

(A) $\quad \max u(x) \quad$ subject to $\quad p^* \cdot x \leqslant \frac{1}{2} \quad$ and $\quad x \in T$

and $\mathbf{1} - x^*$ solves:

(A′) $\quad \max v(y) \quad$ subject to $\quad p^* \cdot y \leqslant \frac{1}{2} \quad$ and $\quad y \in T$.

Since u and v are increasing, $p^* \cdot x^* = \frac{1}{2} = p^* \cdot (\mathbf{1} - x^*)$. It follows that $x^* \in B(p^*, p^*)$. Furthermore, since x^* solves (A), (1) implies that x^* solves:

$$\max u(x) \quad \text{subject to} \quad x \in B(U, p^*).$$

Similarly, combining (A′) and (1′) shows that $\mathbf{1} - x^*$ solves:

$$\max v(y) \quad \text{subject to} \quad \mathbf{1} - y \in B(p^*, V).$$

This establishes the theorem. *Q.E.D.*

The EICE is attained as follows: each player reports linear preferences with indifference surfaces parallel to the supporting prices. The set of solutions to the bargaining problem then consists of an entire hyperplane. However, since the hyperplane supports the EICE, both agents can agree on a most preferred point (with respect to their true preferences). This point is a competitive allocation.

Theorem 1 has a partial converse.

THEOREM 2: *If (p, q, x^*) is a Nash equilibrium for an admissible distortion game, and if p and q are linear, then x^* is an EICE allocation.*

PROOF: Since $x^* \in B(p, q)$,

$$p \cdot x^* \geqslant \tfrac{1}{2} \quad \text{and} \quad q \cdot (\mathbf{1} - x^*) \geqslant \tfrac{1}{2}.$$

Also, since $B(q, q) = \{x \in T : q \cdot x = \frac{1}{2}\}$ and u is increasing, it follows that x^* solves:

$$\max u(x) \quad \text{subject to} \quad q \cdot x \leqslant \tfrac{1}{2} \quad \text{and} \quad x \in T.$$

Similarly, $\mathbf{1} - x^*$ solves:

$$\max v(y) \quad \text{subject to} \quad p \cdot y \leqslant \tfrac{1}{2} \quad \text{and} \quad y \in T$$

and

$$(2) \qquad q \cdot x^* = \tfrac{1}{2} = p \cdot x^*.$$

To prove the theorem, it suffices to show that $p = q$. But this follows because x^* is a Pareto-efficient allocation with respect to the utilities p and q. Hence (2) implies that equal division must be Pareto efficient with respect to the utilities p and q. Since p and q are normalized, $p = q$. *Q.E.D.*

Informally Theorem 2 can be explained as follows. The use of a linear strategy by player two restricts player one to outcomes x that satisfy $q \cdot x \leqslant \frac{1}{2}$. Thus, the way to guarantee the most preferred outcome in this set is to use the strategy q; in this way $B(q, q) = \{x : q \cdot x = \frac{1}{2}\}$. Similarly, the best response player two can make to player one is to use the same strategy. It follows that two linear strategies p and q can comprise an equilibrium only if $p = q$ and the players are able to agree on the most preferred outcome in $\{x : p \cdot x = \frac{1}{2}\}$.

Taken together, Theorems 1 and 2 characterize the Nash equilibria for the distortion game if agents are restricted to linear strategies. In general, the equilibrium outcome of the distortion game cannot be guaranteed to be an EICE allocation. The next result shows that all equilibrium outcomes are envy-free. That is, each agent weakly prefers his allocation to the allocation of the other player.

THEOREM 3: *If (U, V, x^*) is a Nash equilibrium for an admissible distortion game and the true utilities are concave, then*

$$u(x^*) \geqslant u(\tfrac{1}{2}) \geqslant u(\mathbf{1} - x^*) \quad \textit{and}$$

$$v(\mathbf{1} - x^*) \geqslant v(\tfrac{1}{2}) \geqslant v(x^*).$$

PROOF: $u(x^*) \geqslant u(y)$ where y solves:

$$\max u(x) \quad \text{subject to} \quad x \in B(V, V).$$

Since $\frac{1}{2} \in B(V, V)$, $u(x^*) \geqslant u(\frac{1}{2})$. The concavity of u and the fact that $u(x^*) \geqslant u(\frac{1}{2})$ guarantee $u(\frac{1}{2}) \geqslant u(\mathbf{1} - x^*)$. Identical arguments establish the statements about v. *Q.E.D.*

By using his opponent's strategy, a player can guarantee himself the outcome $\frac{1}{2}$. Thus, any Nash equilibrium must yield each an outcome at least as preferred as equal division.

Theorem 3 is true even if the underlying distortion game is not admissible. Since both players have the same strategy set, a player is able to use his opponent's strategy. Also, any solution B that is Pareto efficient and symmetric satisfies $\frac{1}{2} \in B(V, V)$ for all V. It follows that a player is able to guarantee an outcome that is at least as preferred as equal division by using his opponent's strategy.

The following consequence of Theorem 3 is immediate.

COROLLARY: *If equal division is efficient, then all Nash equilibria to the distortion game give player* 1 *utility $u(\frac{1}{2})$, and player* 2 *utility $v(\frac{1}{2})$.*

In particular, if the agents have identical preferences, equal division—or a utility equivalent allocation— will be the unique outcome of the distortion game.

5. NASH AND RAIFFA-KALAI-SMORODINSKY DISTORTION GAMES

The previous section proved existence of Nash equilibria for admissible distortion games. The characterization can be made more explicit if the nature of the solution to the bargaining problem is restricted. In this section, the Nash and the R-K-S solutions will be considered specifically.

For the results of this section, I shall assume that the true utilities are differentiable on the interior of T, and that reported utilities are twice continuously differentiable on the interior of T. I shall denote the partial derivative of a function U with respect to its ith argument by U_i. ∇U denotes the gradient vector $(U_1, \ldots, U_n)$, and second partial derivatives are denoted by U_{ij}.

The function, V, reported by agent 2 constrains the possible equilibria of the distortion game. In order for x to be a Nash equilibrium of the distortion game determined by the Nash bargaining solution, there must exist a $U \in \mathcal{U}$ such that x solves:

$$\text{(B)} \qquad \max U(y)V(\mathbf{1}-y) \quad \text{subject to} \quad y \in T.$$

Since U and V are concave, the first order conditions associated with this maximization problem are necessary and sufficient. That is, x solves (B) if and only if, for all i,

$$\text{(3)} \qquad U_i(x)V(\mathbf{1}-x) - U(x)V_i(\mathbf{1}-x) \leqslant 0 \quad \text{if} \quad x_i < 1$$

and

$$\text{(3')} \qquad U_i(x)V(\mathbf{1}-x) - U(x)V_i(\mathbf{1}-x) \geqslant 0 \quad \text{if} \quad x_i > 0.$$

Hence,

$$x_i U_i(x)V(\mathbf{1}-x) \geqslant U(x)V_i(\mathbf{1}-x)x_i \quad \text{for all } i.$$

Summing and using the fact that, for all $y \in T$,

$$U(y) \geqslant \nabla U(y) \cdot y \quad \text{whenever} \quad U \in \mathcal{U},$$

it follows that

$$\text{(4)} \qquad V(\mathbf{1}-x) \geqslant \nabla V(\mathbf{1}-x) \cdot x$$

for any potential Nash equilibrium allocation x.

Thus, given the report V, the best possible Nash outcome for player 1 is the solution, x^*, to:

$$\max u(x) \quad \text{subject to} \quad V(\mathbf{1}-x) \geqslant \nabla V(\mathbf{1}-x) \cdot x.$$

An identical argument can be used to deduce the restrictions a reported strategy U has on the possible outcomes for player 2 and these results can be used to characterize the equilibria of the distortion game.

LEMMA 2: *Let U^*, $V^* \in \mathcal{U}$. Then (U^*, V^*, x^*) is a Nash equilibrium for the Nash distortion game if and only if*

$$x^* \in N(U^*, V^*),$$

x^* solves:

$$\text{(C)} \qquad \max u(x) \quad \textit{subject to} \quad V^*(\mathbf{1}-x) \geqslant \nabla V^*(\mathbf{1}-x)\cdot x \quad \textit{and} \quad x \in T,$$

and $\mathbf{1} - x^*$ solves:

$$\text{(C}'\text{)} \qquad \max v(y) \quad \textit{subject to} \quad U^*(\mathbf{1}-y) \geqslant \nabla U^*(\mathbf{1}-y)\cdot y \quad \textit{and} \quad y \in T.$$

PROOF: If x^* solves problem (C), then x^* is the best outcome player 1 can obtain given that player 2 reports V^*; and if $\mathbf{1} - x^*$ solves (C′), then $\mathbf{1} - x^*$ is the best that player 2 can obtain given that player 1 reports U^*. Thus, (U^*, V^*, x^*) is a Nash equilibrium provided $x^* \in N(U^*, V^*)$.

To prove that the conditions are necessary, it suffices to show that, given V^*, player 1 can always report a utility function, U, so that the element in $P = \{z \in T: V^*(\mathbf{1}-z) \geqslant \nabla V^*(\mathbf{1}-z)\cdot z\}$ that he most prefers is contained in $N(U, V^*)$. Since $x^* \in P$ by (4), (U^*, V^*, x^*) will be a Nash equilibrium only if x^* solves (C). Suppose z^* solves:

$$\max u(z) \quad \text{subject to} \quad z \in P.$$

Then, since u is strictly increasing,

$$V^*(\mathbf{1}-z^*) = \nabla V^*(\mathbf{1}-z^*)\cdot z^*.$$

Let

$$U(x) = \nabla V^*(\mathbf{1}-z^*)\cdot x / \nabla V^*(\mathbf{1}-z^*)\cdot \mathbf{1}.$$

It is easy to check that $U \in \mathcal{U}$ and $z^* \in N(U, V^*)$. Therefore, the earlier comments guarantee that $u(x^*) = u(z^*)$ if (U^*, V^*, x^*) is a Nash equilibrium. A similar argument shows $\mathbf{1} - x^*$ solves (C′) and proves the lemma. *Q.E.D.*

Notice that a player can select a best response which is linear. In this sense, nonlinear strategies are dominated. It is unlikely that dominated strategies will be used.

A similar result is true for the R-K-S distortion game.

LEMMA 3: *A triple (U^*, V^*, x^*) is a Nash equilibrium for the R-K-S distortion game if and only if*

$$x^* \in K(U^*, V^*),$$

x^* solves

$$\text{(D)} \qquad \max\ u(x) \quad \textit{subject to} \quad (\nabla V^*(\mathbf{1}-x)\cdot(\mathbf{1}-x))V^*(\mathbf{1}-x)$$
$$\geqslant (1 - V^*(\mathbf{1}-x))\nabla V^*(\mathbf{1}-x)\cdot x,$$

and $\mathbf{1} - x^*$ *solves*:

$$\text{(D}'\text{)} \qquad \max v(y) \quad \textit{subject to} \quad (\nabla U^*(\mathbf{1}-y)\cdot(\mathbf{1}-y))U^*(\mathbf{1}-y) \geqslant (1 - U^*(\mathbf{1}-y))\nabla U^*(\mathbf{1}-y)\cdot y.$$

The proof of Lemma 3 is analogous to that of Lemma 2, and is omitted.

In order to identify other possible Nash equilibria, implications of the necessary conditions given in the previous lemmas must be examined.

LEMMA 4: *Suppose* (U, V, x) *is a Nash equilibrium for the distortion game determined by the Nash or R-K-S bargaining solution. Then*

$$(5) \qquad x\cdot\nabla U(x) = (\mathbf{1}-x)\cdot\nabla U(x) = U(x),$$

$$(5') \qquad (\mathbf{1}-x)\cdot\nabla V(\mathbf{1}-x) = x\cdot\nabla V(\mathbf{1}-x) = V(\mathbf{1}-x),$$

$$(6) \qquad U(\lambda x) = \lambda U(x) \quad \textit{for} \quad 0\leqslant\lambda\leqslant 1,$$

$$(6') \qquad V(\lambda(\mathbf{1}-x)) = \lambda V(\mathbf{1}-x) \quad \textit{for} \quad 0\leqslant\lambda\leqslant 1,$$

$$(7) \qquad \nabla U(\lambda x) = \nabla U(x) \quad \textit{for} \quad 0\leqslant\lambda\leqslant 1,$$

$$(7') \qquad \nabla V(\lambda(\mathbf{1}-x)) = \nabla V(\mathbf{1}-x) \quad \textit{for} \quad 0\leqslant\lambda\leqslant 1,$$

$$(8) \qquad \sum_i U_{ij}(x)x_i = 0 \quad \textit{for all } j, \quad \textit{and}$$

$$(8') \qquad \sum_i V_{ij}(\mathbf{1}-x)(1-x_i) = 0 \quad \textit{for all } j.$$

The proof of Lemma 4 is given in the Appendix. Properties (5) and (5′) are derived directly from Lemma 2 or 3, the remaining properties follow for any elements of $\mathcal{U}$ satisfying (5) or (5′). The restrictions placed on the reported strategies at a Nash equilibrium can be interpreted in the context of the Nash bargaining solution. The Nash solution to the bargaining problem depends on the local properties of the reported preferences and of their derivatives. Properties (6), (6′), (7), and (7′) say that utility, as measured by the reported preferences, increases linearly along the segment connecting $\mathbf{0}$ to the outcome, the direction and magnitude of increase along the ray being constant.

The characterization of equilibrium strategies given in Lemma 4 is suggested by the results of Kihlstrom, Roth, and Schmeidler [**10**] and Kannai [**9**]. Their results show that a player's utility at the Nash and R-K-S bargaining solutions increases as his opponent becomes more risk averse. Thus, one would expect equilibrium strategies to be "least concave"[9] representations of some ordinal

[9] A least concave utility function is a minimal element in the set of continuous, concave functions on T ordered by $\succsim$, where $U \succsim V$ if U is more risk averse than V.

preferences. The fact that equilibrium strategies must be linearly homogeneous along a segment is consistent with this expectation.

Notice that, at least when there are only two commodities, (7) and (7′) imply Thomson's [**19**] results. U is required to be of the form $U(x_1, x_2) = x_1 + W(x_2)$ with W concave, and (7) implies W is linear. Similarly (7′) requires that V must be linear. Therefore, the results of Section 3 imply that the equilibria of the Nash and R-K-S distortion games coincide with the EICE's when reported strategies are smooth and there is transferable utility.

The next results are true for the distortion games determined by both the R-K-S and the Nash bargaining solutions. Proofs are given only in the Nash case.

THEOREM 4: *If (U, V, x^*) is a Nash equilibrium for the distortion game, then $(\nabla U(x^*); x^*)$ is an EICE for the reported utilities.*

PROOF: Lemma 4(6) guarantees that

$$\nabla U(x^*) \cdot x^* = \nabla U(x^*) \cdot \tfrac{1}{2} = \nabla U(x^*) \cdot (\mathbf{1} - x^*)$$

and clearly x^* solves:

$$\max U(x) \quad \text{subject to} \quad \nabla U(x^*) \cdot x \leqslant \nabla U(x^*) \cdot \tfrac{1}{2} \quad \text{and} \quad x \in T.$$

To show that $\mathbf{1} - x^*$ solves:

$$\max V(y) \quad \text{subject to } \nabla U(x^*) \cdot y \leqslant \nabla U(x^*) \cdot \tfrac{1}{2} \quad \text{and} \quad y \in T,$$

it must be verified that $\mathbf{1} - x^*$ satisfies the first order conditions

$$V_i(\mathbf{1} - x^*) - \lambda U_i(x^*) \geqslant 0 \quad \text{if} \quad x_i^* < 1 \quad \text{and}$$

$$V_i(\mathbf{1} - x^*) - \lambda U_i(x^*) \leqslant 0 \quad \text{if} \quad x_i^* > 0.$$

Since $x^* \in N(U, V)$, these conditions follow from (3) and (3′) with $\lambda = V(\mathbf{1} - x^*)/U(x^*)$. *Q.E.D.*

The next result is used to characterize the efficient Nash equilibria.

LEMMA 5: *If (U, V, x^*) is a Nash equilibrium then*

$$\nabla u(x^*) \cdot x^* \geqslant \nabla u(x^*) \cdot \tfrac{1}{2} \quad \textit{and}$$

$$\nabla v(\mathbf{1} - x^*) \cdot (\mathbf{1} - x^*) \geqslant \nabla v(\mathbf{1} - x^*) \cdot \tfrac{1}{2}.$$

PROOF: By Lemma 2, x^* must satisfy the first order conditions for the maximization problem (C). Thus, there is a $\mu > 0$, such that for all i,

$$u_i(x^*) - \mu\left[2V_i(\mathbf{1} - x^*) - \sum_j V_{ij}(\mathbf{1} - x^*)x_j^*\right] \geqslant 0 \quad \text{if} \quad x_i^* > 0,$$

and

$$u_i(x^*) - \mu\left[2V_i(1 - x^*) - \sum_j V_{ij}(1 - x^*)x_j^*\right] \leqslant 0 \quad \text{if} \quad x_i^* < 1.$$

By (8′) and the concavity of V it follows that

$$\nabla u(x^*) \cdot x^* \geqslant 2\mu \nabla V(1 - x^*) \cdot x^*$$

and

$$\nabla u(x^*) \cdot (1 - x^*) \leqslant 2\mu \nabla V(1 - x^*) \cdot (1 - x^*).$$

Thus,

$$\nabla u(x^*) \cdot x^* \geqslant \nabla u(x^*) \cdot \tfrac{1}{2},$$

since

$$\nabla V(1 - x^*) \cdot x^* = \nabla V(1 - x^*) \cdot (1 - x^*)$$

by (5′). Similar arguments establish the inequality involving v. *Q.E.D.*

If x^* were a competitive allocation, $\nabla u(x^*)$ could be taken to be the supporting prices. Lemma 5 then guarantees that the value of the allocation x^* at these competitive prices is at least as great as the value of the equal division allocation, $\frac{1}{2}$. The next consequence is therefore evident.

THEOREM 5: *If x^* is an efficient Nash equilibrium, then x^* is an EICE.*

PROOF: I assume, for convenience, that x^* is in the interior of T. In this case,

$$\nabla u(x^*) = \lambda \nabla v(1 - x^*) \quad \text{for some} \quad \lambda > 0.$$

Thus, Lemma 5 implies

$$\nabla u(x^*) \cdot x^* \geqslant \nabla u(x^*) \cdot \tfrac{1}{2} \tag{9}$$

and

$$\nabla u(x^*) \cdot (1 - x^*) \geqslant \nabla u(x^*) \cdot \tfrac{1}{2}. \tag{9′}$$

It follows that (9) and (9′) must hold as equalities. Hence, $(\nabla u(x^*); x^*)$ is an EICE. When $x^* \in$ Boundary (T) there may be several equilibrium prices associated with it. The argument given above can be modified to show that one of these supporting prices makes the allocation x^* an EICE. *Q.E.D.*

Inefficient Nash equilibria exist in nonpathological settings, if nonlinear strategies are allowed. The following example shows that an EICE allocation may not be Pareto superior to other Nash equilibria of the distortion game.

EXAMPLE: Let $u(x, y) = x^{5/6}y^{1/6}, v(x, y) = x^{1/2}y^{1/2}$, $U(x, y) = (5x + 3y)/8$, and $V(x, y) = x^{1/2}y^{1/2}$. Then u, v, U, and $V \in \mathfrak{U}$, and routine verification, using Lemma 2, shows that (U, V, x^*) is a Nash equilibrium for the distortion game when $x^* = (3/5, 1/3)$. In this example, there is a unique EICE for the true preferences. It is $(p^*; y^*) = ((2/3, 1/3); (5/8, 1/4))$. A computation shows that $v(\mathbf{1} - y^*) > v(\mathbf{1} - x^*)$ and that $u(x^*) > u(y^*)$. Thus, the first player prefers the inefficient outcome x^* to the EICE outcome y^*. Also, the second player is worse off at the Nash equilibrium (u, v, x^*) even though he is reporting his true utility function.

The example also shows that the set of equilibria to the Nash distortion game is not equal to the set of equilibria of the R-K-S distortion game. It is easy to check $x^* \notin K(U, V)$ so (U, V, x^*) is not a Nash equilibrium for the R-K-S distortion game.

University of California-San Diego.

Manuscript received September, 1979; revision received January, 1980.

APPENDIX

This appendix will prove Lemma 4 for the Nash distortion game. The proof for R-K-S bargaining is similar. A preliminary result must be established first.

FACT: Let $A = (a_{ij})$, $1 \leqslant i, j \leqslant n$ by a (symmetric) negative semi-definite matrix, and suppose $\sum_{i,j} a_{ij}x_i x_j = 0$ for some $x = (x_1, \ldots, x_n)$. Then, for all i, $\sum_j a_{ij}x_j = 0$.

PROOF: Let P be an orthogonal matrix that diagonalizes A, and let $D = P'AP$ be a diagonal matrix with diagonal entries λ_i (P' denotes the transpose of P). By assumption, $\lambda_i \leqslant 0$ for all i. Finally, let $y = P'x$. It follows that $0 = x'Ax = y'Dy = \sum_i \lambda_i y_i^2$. Since $\lambda_i \leqslant 0$ for all i, $\lambda_i y_i = 0$ for all i. Therefore,

$$Ax = Dy = 0.$$

This proves the Fact. *Q.E.D.*

LEMMA: 4: *Suppose (U, V, x) is a Nash equilibrium for the distortion game. Then*

$$x \cdot \nabla U(x) = (\mathbf{1} - x) \cdot \nabla U(x) = U(x), \tag{5}$$

$$(\mathbf{1} - x) \cdot \nabla V(\mathbf{1} - x) = x \cdot \nabla V(\mathbf{1} - x) = V(\mathbf{1} - x), \tag{5'}$$

$$U(\lambda x) = \lambda U(x) \quad \text{for} \quad 0 \leqslant \lambda \leqslant 1, \tag{6}$$

$$V(\lambda(\mathbf{1} - x)) = \lambda V(\mathbf{1} - x) \quad \text{for} \quad 0 \leqslant \lambda \leqslant 1, \tag{6'}$$

$$\nabla U(\lambda x) = \nabla U(x) \quad \text{for} \quad 0 \leqslant \lambda \leqslant 1, \tag{7}$$

$$\nabla V(\lambda(\mathbf{1} - x)) = \nabla V(\mathbf{1} - x) \quad \text{for} \quad 0 \leqslant \lambda \leqslant 1, \tag{7'}$$

$$\sum_i U_{ij}(x)x_i = 0 \quad \text{for all} \quad j, \quad \text{and} \tag{8}$$

$$\sum_i V_{ij}(\mathbf{1} - x)(1 - x_i) = 0 \quad \text{for all } j. \tag{8'}$$

PROOF: It follows from Lemma 2 and the fact that the true utilities are strictly increasing, that

(A1) $\quad (\mathbf{1} - x) \cdot \nabla U(x) = U(x)$ and

(A1′) $\quad x \cdot \nabla V(\mathbf{1} - x) = V(\mathbf{1} - x).$

Also, since $x \in N(U, V)$, (3) and (3′) hold. Therefore, for all i,

$$(1 - x_i)U_i(x)V(\mathbf{1} - x) - U(x)V_i(\mathbf{1} - x)(1 - x_i) \leqslant 0$$

and

$$x_i U_i(x)V(\mathbf{1} - x) - U(x)V_i(\mathbf{1} - x)x_i \geqslant 0.$$

Summing, applying (A1), (A1′), and the concavity of U and V show that

$$U(x) = x \cdot \nabla U(x) \quad \text{and} \quad V(\mathbf{1} - x) = (\mathbf{1} - x) \cdot \nabla V(\mathbf{1} - x).$$

This establishes (5) and (5′).

To show (6) observe that because of concavity and the fact that $U(x) = x \cdot \nabla U(x)$,

$$\begin{aligned} U(x) - U(\lambda x) &\geqslant \nabla U(x) \cdot (x - \lambda x) \\ &= U(x) - \lambda \nabla U(x) \cdot x \\ &\geqslant U(x) - \lambda \nabla U(\lambda x) \cdot x \\ &\geqslant U(x) - U(\lambda x) \end{aligned}$$

whenever $0 \leqslant \lambda \leqslant 1$. It follows that all of the inequalities above hold as equalities. Hence, for $0 \leqslant \lambda \leqslant 1$,

$$U(\lambda x) = \lambda \nabla U(x) \cdot x = \lambda \nabla U(\lambda x) \cdot x = \lambda U(x).$$

Thus, for $0 \leqslant \lambda \leqslant 1$,

(A2) $\quad \nabla U(\lambda x) \cdot x = U(x).$

Identity (A2) can be differentiated with respect to λ. This yields

$$\sum_{i,j} U_{ij}(\lambda x)x_i x_j = 0 \quad \text{for} \quad 0 \leqslant \lambda \leqslant 1.$$

The preliminary Fact now implies that for all j and $0 \leqslant \lambda \leqslant 1$,

$$\sum_i U_{ij}(\lambda x)x_i = 0.$$

This establishes (8). Equation (7) follows since

$$\frac{d}{d\lambda}(U_j(\lambda x)) = \sum_i U_{ij}(\lambda x)x_i = 0.$$

Similar arguments establish (6′), (7′), and (8′). *Q.E.D.*

REFERENCES

[1] AUMANN, ROBERT J., AND MORDECAI KURZ: "Power and Taxes," *Econometrica*, 45 (1977), 1137–1161.

[2] ———: "Power and Taxes in a Multi-Commodity Economy," *Israel Journal of Mathematics*, 27 (1977), 185–234.

[3] CRAWFORD, VINCENT P., AND HAL R. VARIAN: "Distortion of Preferences and the Nash Theory of Bargaining," *Economics Letters*, 3 (1979), 203–206.

[4] FOLEY, DUNCAN K.: "Resource Allocation and the Public Sector," *Yale Economic Essays*, 7 (1967), 45–98.

[5] HURWICZ, LEONID: "On Informationally Decentralized Systems," in *Decision and Organization: A Volume in Honor of Jacob Marshak*, ed. by C. B. McGuire and R. Radner. New York: American Elsevier Publishing Company, 1972.

[6] HURWICZ, LEONID, ERIC MASKIN, AND ANDREW POSTLEWAITE: "Constrained Walrasian Equilibria," mimeo, University of Minnesota, 1978.

[7] KALAI, EHUD, AND ROBERT W. ROSENTHAL: "Arbitration of Two-Party Disputes Under Ignorance," *International Journal of Game Theory*, 7 (1978), 65–72.

[8] KALAI, EHUD, AND MEIR SMORODINSKY: "Other Solutions to Nash's Bargaining Problem," *Econometrica*, 43 (1975), 513–518.

[9] KANNAI, YAKAR: "Concavifiability and Constructions of Concave Utility Functions," *Journal of Mathematical Economics*, 4 (1977), 1–56.

[10] KIHLSTROM, RICHARD E., ALVIN E. ROTH, AND DAVID SCHMEIDLER: "Risk Aversion and Solutions to Nash's Bargaining Problem," mimeo., University of Illinois, March, 1979.

[11] KURZ, MORDECAI: "Distortion of Preferences, Income Distribution, and the Case for a Linear Income Tax," *Journal of Economic Theory*, 14 (1977), 291–298.

[12] ———: "Income Distribution and Distortion of Preferences: The l-Commodity Case," *Journal of Economic Theory*, 22(1980), 99–106.

[13] LUCE, R. DUNCAN, AND HOWARD RAIFFA: *Games and Decisions: Introduction and Critical Survey*, New York: John Wiley and Sons, Inc., 1957.

[14] MYERSON, ROGER: "Incentive Compatibility and the Bargaining Problem," *Econometrica*, 47 (1979), 61–73.

[15] NASH, JOHN: "The Bargaining Problem," *Econometrica*, 18 (1950), 155–162.

[16] ———: "Two-Person Cooperative Games," *Econometrica*, 21 (1953), 128–140.

[17] RAIFFA, HOWARD: "Arbitration Schemes for Generalized Two-Person Games," in *Contributions to the Theory of Games II*, ed. by H. W. Kuhn and A. W. Tucker. Princeton: Princeton University Press, 1953.

[18] ROTH, ALVIN: *Axiomatic Models of Bargaining*. Berlin and Heidelberg: Springer-Verlag, 1979.

[19] THOMSON, WILLIAM: "The Manipulability of the Shapley-Value," University of Minnesota Discussion Paper No. 79-115, September, 1979.

[20] ———: "On the Manipulability of Resource Allocation Mechanisms Designed to Achieve Individually-Rational and Pareto-Optimal Outcomes," University of Minnesota Discussion Paper No. 79-116, September, 1979.