

Preferences Over Solutions to the Bargaining Problem

Author(s): Kim C. Border and Uzi Segal

Source: *Econometrica*, Vol. 65, No. 1 (Jan., 1997), pp. 1-18

Published by: The Econometric Society

Stable URL: <http://www.jstor.org/stable/2171811>

Accessed: 18-08-2017 21:26 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://about.jstor.org/terms>



The Econometric Society is collaborating with JSTOR to digitize, preserve and extend access to *Econometrica*

JSTOR

PREFERENCES OVER SOLUTIONS TO THE BARGAINING PROBLEM

BY KIM C. BORDER AND UZI SEGAL¹

There are several solutions to the Nash bargaining problem in the literature. Since various authors have expressed preferences for one solution over another, we find it useful to study preferences over solutions in their own right. We identify a set of appealing axioms on such preferences that lead to unanimity in the choice of solution, which turns out to be the solution of Nash.

KEYWORDS: Nash bargaining solution, quadratic utility functions.

1. INTRODUCTION

IN THE NASH (1950, 1953) APPROACH to bargaining, a bargaining game is described by a pair (S, d) where $S \subset \mathbb{R}^2$ is compact and convex and $d \in S$. Elements of S are interpreted as vectors of the von Neumann-Morgenstern utilities of the two players. The point d is the disagreement point, that is, a vector of utilities that either player can unilaterally enforce. A solution is a function f that assigns to each game (S, d) a point in S . Of course many solutions are possible, but Nash proposed internal consistency conditions on the values of a solution across different games. The only solution to satisfy his axioms is called the Nash bargaining solution, and is defined by $f(S, d) = \operatorname{argmax}_{x \in S, x \geq d} (x_1 - d_1)(x_2 - d_2)$.

Other notions of consistency can be imposed on a solution. Kalai and Smorodinsky (1975), for example, offered another set of axioms, with another notion of internal consistency, and got a different solution. It is therefore natural to ask, given that different notions of consistency exist, which is a better solution. We assume that there exist preferences over solutions satisfying certain sets of axioms and explore whether they admit a maximal element, and if so, what it is.

The preferences over solutions and the assumptions we make on them can be interpreted in more than one way. One interpretation is that the two bargainers hire an arbitrator to make choices for them. The arbitrator has a preference order over solutions that embodies his own notions of fairness. These preferences are similar to the social preferences discussed in the social choice literature (e.g., Harsanyi (1955)), but they differ in that they are over solution

¹We gratefully acknowledge stimulating suggestions from Larry Epstein, Peter Eso, Itzhak Gilboa, Ehud Kalai, Edi Karni, Bart Lipman, Joe Ostroy, Bob Pollak, Ariel Rubinstein, Zvi Safra, Lloyd Shapley, Joel Sobel, Dan Vincent, Bill Weiss, Bill Zame, the co-editor, two anonymous referees, and seminar participants at Caltech, Johns Hopkins, Northwestern, and Washington University. Uzi Segal also thanks the Social Sciences and Humanities Research Council of Canada for financial support.

concepts, rather than over final allocations. In other words, the arbitrator must come up with a decision rule that will apply to all possible games, and not to just one set of possible utility portfolios.

Of course, the bargainers have to agree to arbitration. We can therefore interpret our axioms on preferences as properties that both bargainers can accept in an arbitrator, subject to the constraint that both must agree to hire one. For instance, while my coauthor's preferences over solutions are simply to maximize his own payoff, and my preferences are to maximize mine, we both know that neither will agree to impose such a requirement on an arbitrator. Our aim is therefore to propose axioms that both bargainers can accept. Arguably, it will be easier for them to agree on a set of axioms as guidelines for the arbitrator than to agree on the allocation for each game.

One important feature of the bargaining game is the existence of a disagreement point, namely a point each player can impose by refusing to play the game. It is claimed in the social choice literature that lack of ex post fairness can be compensated by ex ante choice of procedures like randomization (see Diamond (1967) and Epstein and Segal (1992)). In the bargaining problem, restricting attention to ex ante fairness may not be enough. The arbitrator, and the bargainers instructing him, will have to make sure that the ex post result is not too unfair, at least to the extent that whatever game is to be played, each player must receive at least his disagreement level of utility. The need for ex post and ex ante fairness of procedures stands behind our two major axioms.

Our axioms imply that the Nash solution is the best, and we give a full description of all preference relations satisfying this set. These results are presented in Section 2. We discuss some aspects of the literature in Section 3 and conclude with an analysis of some possible objections in Section 4. The theorems are proved in the appendices.

2. THE NASH BARGAINING SOLUTION

For our purposes, a two-person bargaining game is represented by a compact and convex subset S of \mathbb{R}_+^2 such that S is disposable, that is, $[x \in S \text{ and } x \geq y] \Rightarrow y \in S$. (We use the following orders on vectors. $x \geq y$ means $x_i \geq y_i$ for all i , and $x \gg y$ means $x_i > y_i$ for all i). Each point $x = (x_1, x_2) \in S$ corresponds to an allocation of von Neumann-Morgenstern utilities for the two players in which player 1 receives x_1 and player 2 receives x_2 . In order to make our analysis more accessible, we assume throughout that the disagreement point is $(0, 0)$.²

Let \mathcal{S} denote the set of all games that are included in $[0, K] \times [0, K] \subset \mathbb{R}_+^2$ and contain the point (κ, κ) . All we require is that $\kappa > 0$, but it may be as close to zero as one wishes. A solution is a function $f: \mathcal{S} \rightarrow \mathbb{R}_+^2$ satisfying $f(S) \in S$ for every $S \in \mathcal{S}$. This definition and the assumption that the disagreement point is

²In the standard bargaining problem this point is not fixed, but our analysis can be extended to include this case (see the end of this section).

(0, 0) rule out the possibility that a solution will ever assign any of the players an outcome that is worse than his disagreement level. As mentioned in the introduction, we are looking for a mechanism that will capture both ex ante and ex post notions of fairness. Although one may argue for the ex ante fairness of a solution that benefits each player in some games at the cost of giving his opponent less than his disagreement level, such allocations are hardly fair ex post, and probably will be hard to enforce. We discuss this requirement further below.

The space \mathcal{S} of games is endowed with the Hausdorff metric, given by

$$\begin{aligned} \rho(S, T) &= \max \left\{ \sup_{x \in S} d(x, T), \sup_{x \in T} d(x, S) \right\} \\ &= \inf \{ \varepsilon > 0: S \subset N_\varepsilon(T) \text{ and } T \subset N_\varepsilon(S) \}, \end{aligned}$$

where $d(x, T) = \inf_{y \in T} \|x - y\|$ and $N_\varepsilon(T) = \{x \in \mathbb{R}_+^2: d(x, T) < \varepsilon\}$. Given this metric, the space \mathcal{S} is compact.³

Since games are convex and disposable, the outer boundary determines the distance between two games. If each point on the boundary of S is within ε neighborhood of a boundary point of T and vice-versa, then the Hausdorff distance between S and T is no more than ε , and conversely.

We assume throughout that solutions are continuous in the Hausdorff metric, that is, that each solution is a continuous function from \mathcal{S} to \mathbb{R}_+^2 . Both Nash's (1950) and Kalai-Smorodinsky's (1975) solutions satisfy this requirement. Let \mathcal{F} denote the set of continuous solutions. Then \mathcal{F} is a metric space under the metric

$$d(f, g) = \sup_{S \in \mathcal{S}} \|f(S) - g(S)\|.$$

This metric defines the topology of uniform convergence of solutions on \mathcal{S} , and \mathcal{F} is separable (see Lemma 3.72 in Aliprantis and Border (1994)).

Two solutions can be mixed as follows. For $f, g \in \mathcal{F}$ and $\alpha \in [0, 1]$, the solution $\alpha f + (1 - \alpha)g$ assigns game $S \in \mathcal{S}$ the outcome $\alpha f(S) + (1 - \alpha)g(S)$. All games S are assumed to be convex so this mixture is well defined. Since points in S are interpreted as vectors of von Neumann-Morgenstern utility levels, one can interpret the solution $\alpha f + (1 - \alpha)g$ as a lottery that yields the solution f with probability α and the solution g with probability $1 - \alpha$. This is because ex ante, each player's utility is the same under the mixed solution and under the lottery.

On \mathcal{F} we assume the existence of a quasiorder \succsim , that is, \succsim is reflexive, total, and transitive. Consider the following assumptions.

³This is because the space \mathcal{K} of compact convex subsets of $[0, K] \times [0, K]$ is a compact metric space under the Hausdorff metric (see Theorem 5.43 in Aliprantis and Border (1994)), and it is straightforward to show that \mathcal{S} is a closed subset of \mathcal{K} . Incidentally, this is why we need the assumption that all games in \mathcal{S} contain the point (κ, κ) . The set of games for which there only exists some strictly positive point is σ -compact, but not compact, which creates problems later.

ASSUMPTION C (Continuity): \succsim is a closed subset of $\mathcal{F} \times \mathcal{F}$.

ASSUMPTION M (Monotonicity): If $f(S) \geq g(S)$ for all $S \in \mathcal{G}$ and $f(T) \gg g(T)$ for some T , then $f \succ g$.

ASSUMPTION DI (Disagreement Indifference): If for all $S \in \mathcal{G}$, either $f_1(S) = 0$ or $f_2(S) = 0$ (or both), then $f \sim \mathbf{0}$ ($\mathbf{0}$ is the solution that always gives both players zero).

ASSUMPTION MS (Mixture Symmetry): $f \sim g$ implies that for each $\alpha \in [0, 1]$,

$$\alpha f + (1 - \alpha)g \sim \alpha g + (1 - \alpha)f.^4$$

The continuity assumption is on the preference relation \succsim , and is different from the assumption that each of the solutions is a continuous function. Regarding monotonicity, note that if $f(T) \gg g(T)$, then continuity of solutions guarantees that $f \gg g$ on an open set of games. Also note that condition DI does not require that it is always the same player who receives zero. This condition may be rationalized on the following grounds.

One important feature of the bargaining problem is that each player can enforce a minimum utility level for himself through the disagreement point. An arbitrator and the bargainers will have to take this into consideration. In social choice theory, a social planner can satisfy his concerns for fairness by ensuring ex ante equality, for example, by flipping a coin between members of society over who will go to war (see Diamond (1967) and Epstein and Segal (1992)). In bargaining, a solution must display some concern also for ex post fairness, otherwise it may be unacceptable by at least one of the players. So a good solution to the bargaining game must not be too inequitable. This requirement is behind the disagreement indifference assumption.

The arbitrator and the players do not know whether the latter are going to abide by the solution in any given game S . Clearly, the closer to zero a bargainer's utility is under a solution (when better outcomes are possible), the more likely he is to deviate from the solution. This probability becomes a virtual certainty when his outcome is zero. This is because in that case he has no incentive at all to play and may as well "punish" the other player. Thus any attempt to use a solution satisfying the hypotheses of condition DI is tantamount to using the zero solution.

There are of course situations where it must always be the case that ex post one player wins nothing (see Diamond (1967)). In such situations it is indeed impossible to assume that players have enough power to impose a bad outcome on everyone. It is precisely this extra strength each player has that will convince the arbitrator (or the bargainers) to avoid solutions where whatever the game is, someone will not mind blocking the outcome.

⁴Chew, Epstein, and Segal (1991) refer to this property as strong mixture symmetry. However, they show that it is equivalent to what they call mixture symmetry, so we adopt this definition.

But is it not true that letting each player win everything in some games is better than letting the same player win everything in all these games? Ex ante, the answer is of course yes. But if ex post both solutions will result in both players receiving zero, the answer is no.

If we interpret the axioms as characteristics of the arbitrator that both bargainers can accept, then player i 's true preferences over solutions need not satisfy disagreement indifference. Rather, the axiom represents his consent to the fact that he cannot impose his preferences on his opponent. This justification is somewhat similar in spirit to Binmore's (1991, Section 3.3) argument that players are constrained only by natural law, but not by ethical conventions. The reason a player agrees to disagreement indifference is not because of empathy with the other player, but because of his own self-interest, fearing a zero outcome for himself. Also note that delegating authority to an arbitrator requires both players to agree on some ground rules, which by their nature ought to treat both bargainers symmetrically. Disagreement indifference satisfies this requirement.

The rationale for the MS condition is the same as the one used in social choice theory (see Epstein and Segal (1992)).⁵ Recall that $\alpha f + (1 - \alpha)g$ is equivalent, ex ante, to the lottery that yields the solution f with probability α and the solution g with probability $1 - \alpha$. So we have to justify why if $f \sim g$, the lotteries over solutions $(f, \alpha; g, 1 - \alpha)$ and $(f, 1 - \alpha; g, \alpha)$ are equally attractive. The disagreement indifference and continuity assumptions imply that both players' claims must get some consideration. When such conflicting considerations balance each other, indifference between f and g will follow. A mathematically equivalent version of the mixture symmetry axiom (together with weak quasiconcavity) says that if $f \sim g$, then the $\frac{1}{2} - \frac{1}{2}$ lottery over the two solutions is not worse than any other possible lottery over them (see Lemma A1.2 (b) in Chew, Epstein, and Segal (1991)). If an arbitrator does not want to appear to be in favor of one of the two bargainers, such indifference will force the even chance lottery over the two solutions not to be inferior to any other possible lottery over them. Mixture symmetry thus follows.

If the axioms represent a common set of rules on which bargainers instructing an arbitrator can agree, then mixture symmetry becomes even more attractive. Suppose the arbitrator is going to be indifferent between two solutions because of conflicting claims by the bargainers. Then unless they want to work out such conflicts by themselves (which they probably wish not to do, since they hired an arbitrator), an even chance lottery over the two solutions seems the easiest compromise to reach.

⁵The mixture symmetry axiom is a weaker version of the independence axiom ($f \succcurlyeq g \Leftrightarrow \alpha f + (1 - \alpha)h \succcurlyeq \alpha g + (1 - \alpha)h \forall \alpha \in (0, 1]$ and $\forall h \in \mathcal{F}$), so any arguments in favor of the independence axiom in social choice (see Harsanyi (1955)) also apply here. Mixture symmetry has the advantage that it allows for preferences that some choices be made according to chance (e.g., draft lotteries). However, it is also consistent with the betweenness assumption, namely that $f \sim g \Rightarrow f \sim \alpha f + (1 - \alpha)g$ for all $\alpha \in [0, 1]$.

THEOREM 1: *If \succcurlyeq satisfies conditions C, M, DI, and MS, then there is a measure μ on \mathcal{G} with full support (that is, every open set in \mathcal{G} has a positive measure), such that the function*

$$V(f) = \int_{\mathcal{G}} f_1(S) f_2(S) d\mu(S)$$

represents \succcurlyeq on \mathcal{F} .

That is, solutions in \mathcal{F} are ranked on a weighted average of the “Nash social welfare function,” $w(x) = x_1 x_2$. This result immediately implies the next theorem.

THEOREM 2: *If \succcurlyeq satisfies conditions C, M, DI, and MS, then the Nash bargaining solution is the unique \succcurlyeq -best solution in \mathcal{F} .*

The formal proof of Theorem 1 is in Appendix A, but we shall explain the roles of the hypotheses here. We first pick one game, denoted T_1 . For each $x = (x_1, x_2) \in T_1$ we define a solution⁶ $f^x \in \mathcal{F}$ such that $f^x(T_1) = x$, and such that if $x_i = 0$, then for all S , $f_i^x(S) = 0$. The isomorphism $x \mapsto f^x$ is continuous and monotonic, in the sense that $x \succcurlyeq y$ implies for all S , $f^x(S) \succcurlyeq f^y(S)$, and satisfies $f^{\alpha x + (1-\alpha)y}(S) = \alpha f^x(S) + (1-\alpha)f^y(S)$. The order \succcurlyeq thus induces a continuous and monotonic order \succeq on T_1 , satisfying mixture symmetry. The following fact is implied by Theorem 6 and Appendix 2 in Chew, Epstein, and Segal (1991), and by Chew, Epstein, and Segal (1994).

FACT 1: *Let \succcurlyeq on a compact and convex subset $D \subset \mathbb{R}^n$ satisfy continuity, monotonicity, and mixture symmetry, and let V represent \succcurlyeq . Then D can be partitioned into three convex regions A , B , and C , so that V is quadratic and quasiconcave on A , quadratic and quasiconvex on C , and satisfies betweenness on B . Furthermore, $A \succcurlyeq B \succcurlyeq C$.*

Disagreement indifference rules out regions B and C , although this can be done by using much weaker assumptions. Next we extend the above construction to n games. Given $T_1, \dots, T_n \in \mathcal{G}$, we define for $x = (x^1, \dots, x^n) \in \prod T_i$ a solution f^x such that $f^x(T_i) = x^i$, and show that it satisfies the above requirements. Once again the induced order on $\prod T_i$ can be represented by a quadratic function. This quadratic function may however contain expressions of the form $x_k^i x_l^j$, $i, j \in \{1, \dots, n\}$, $k, l \in \{1, 2\}$, $i \neq j$, $k \neq l$. By taking a dense sequence of games $\{T_i\}$ in \mathcal{G} , we generate a unique limiting measure π on $\mathcal{G} \times \mathcal{G}$ such that \succcurlyeq can be represented by $\int_{\mathcal{G} \times \mathcal{G}} f_1(S) f_2(T) d\pi(S, T)$.

Disagreement indifference now implies that π must be supported by the diagonal $\Delta = \{(S, S) : S \in \mathcal{G}\}$ of $\mathcal{G} \times \mathcal{G}$. Otherwise, let $G_1 \times G_2 \subset \mathcal{G} \times \mathcal{G}$ be an

⁶In the Appendix, this solution is denoted $\Phi_{T_1, x}$.

open region such that $\Delta \cap (G_1 \times G_2) = \emptyset$ and $\pi(G_1 \times G_2) > 0$. Construct a solution f such that $f_i(S) = 0$ iff $S \notin G_i$. Then by disagreement indifference, $f \sim \mathbf{0}$, but $\int_{\mathcal{S} \times \mathcal{S}} f_1(S) f_2(T) d\pi(S, T) \neq 0$.

The full support of μ is guaranteed by the monotonicity condition. Otherwise, improving a solution on a zero measure set will not change the attractiveness of the solution, violating monotonicity. The compactness of the space \mathcal{S} is crucial to showing that a limiting measure exists, which is one reason we assume that (κ, κ) belongs to all games.⁷ Of course, once we have the utility representation, since the Nash bargaining solution maximizes the integrand at each point, it is the best solution.

The above analysis can be extended to the case where the disagreement point d is not fixed. Solutions are then defined over pairs (S, d) , where $x \in S \Rightarrow x \geq d$.⁸ These solutions are assumed to be continuous with respect to the second variable, while condition DI is modified as follows. If for all (S, d) , either $f_1(S) = d_1$ or $f_2(S) = d_2$, then $f \sim f_b$, where $f_b(S, d) = d$ for all (S, d) . The rationale for this axiom and for the requirement that all points in S yield both players at least as much utility as d is the same as the one offered for the original DI condition. Once they know what game S is being played, each player can block any outcome that does not yield him at least his disagreement utility level. Moreover, if he receives this utility level, he might as well force the minimal outcome on his opponent. The representation theorem now defines a measure over the set of pairs (S, d) with the integrand $(f_1(S) - d_1)(f_2(S) - d_2)$, and its proof is similar to the one outlined above. The major differences are that the isomorphism f^x will be defined taking d into consideration, and the sequence $\{T_i\}$ of games will be dense with respect to both S and d .

3. SOME REMARKS ON THE LITERATURE

Recently, Rubinstein, Safra, and Thomson (1992) analyzed the bargaining problem in light of the recent non-expected utility literature and presented an alternative model and axioms. They start with a certain set A of (physical) alternatives and a point D . Each of the two players has preferences over lotteries over elements of A . Given the utility functions u_1 and u_2 of the two players, points in A , and the point D , are transformed into points in \mathbb{R}^2 . Formally, $d(u_1, u_2; D) = (u_1(D), u_2(D))$ and $S(u_1, u_2; A) = \{(u_1(a), u_2(a)) : a \in A\}$. The authors point out that there are two possible interpretations of the bargaining model.

According to the first, the set A and the point D are fixed, say $A = A^*$ and $D = D^*$. The bargaining problem is to decide what point in $S(u_1, u_2; A^*)$ to choose for each pair of utilities (u_1, u_2) , given that the disagreement point is

⁷Otherwise, it might turn out that the utility of a solution f is determined by limits of values of $f(S)$ as S decreases to $\{(0, 0)\}$. There are of course other assumptions that could be used to rule out this implausible sort of preference order, but we feel our choice is as good as any other.

⁸Note that $(x_1 - d_1)(x_2 - d_2)$ is monotonic only where $x \geq d$.

$d(u_1, u_2; D^*)$. If f is a solution to that problem, then a physical outcome of the solution is a point a in A^* such that $(u_1(a), u_2(a)) = f(S(u_1, u_2; A^*))$. Rubinstein, Safra, and Thomson use this interpretation. Note that if u_1 and u_2 are von Neumann-Morgenstern utility functions, then the invariance with respect to utility transformations axiom is almost trivial—the outcome (in A^*) should depend on preferences, and not on the particular choice of a utility function.

In this paper we adopt an alternative approach which holds the players (with their preferences and utility functions) fixed, and lets the set A vary (we assumed a fixed point of disagreement). Formally, let u_1^* and u_2^* be two given von Neumann-Morgenstern utility functions. What rule should be used to determine how much utility each of the two players will get for different sets A ? This representation of the problem fits better into a social choice context. Note that Nash's axiom of invariance with respect to utility transformations does not fit into this interpretation, because it analyzes rescaling of the utility functions, which in this approach are assumed to be fixed. Such rescaling is possible, but leads to much weaker results. Suppose that instead of the utility function u_i , player i uses the utility function $\alpha_i u_i$. Then for each set A we get a new set of corresponding utility vectors. Formally, it defines an isomorphism $\mathcal{S} \mapsto \mathcal{S}_\alpha$, where for $S \in \mathcal{S}$, $S \mapsto S_\alpha = \{(\alpha_1 x_1, \alpha_2 x_2) : (x_1, x_2) \in S\}$. It is natural to require that possible solutions be transformed in the same way. Formally, $\mathcal{F} \mapsto \mathcal{F}_\alpha$, where $f_\alpha(S_\alpha) = (\alpha_1 f_1(S), \alpha_2 f_2(S))$. Unlike Nash's axiom, this is a transformation of solutions (which are functions), and not of the value of a solution for a given game. All it says is that the choice of points from the (different) sets A should depend on the players' preferences, and not on the chosen utility functions. A similar argument also implies that one should not assume symmetry. A symmetric set S depends on a specific choice of utility functions by the players and does not necessarily represent a true symmetric situation (see also Rubinstein, Safra, and Thomson (1992)).

A related analysis is offered by van Damme (1986). In his model players can choose a solution concept, and then bargain with their opponent, provided they adhere to their (perhaps different) solutions. Making some restrictions on the set of solutions players may adopt, he proves that the Nash solution constitutes the unique equilibrium of the game induced by this procedure. In his approach, the Nash solution is the outcome of strategic behavior, and not necessarily a preferred solution. We, on the other hand, are interested in solutions that are reached by agreement, and are preferred (by the players or by an arbitrator) to other solutions.

4. SOME POSSIBLE OBJECTIONS

One motivation for this paper is the fact that the solutions of both Nash and Kalai-Smorodinsky have a lot of appeal. Unfortunately, the two systems of axioms leading to these solutions are inconsistent. It is therefore natural to have

preferences over such systems, and therefore over solutions. One might argue, however, that we replaced systems of axioms about solutions by systems of axioms on preferences over solutions. And since other such systems are possible, we may once again face the same problem, namely, which do we prefer.

If we accept the interpretation that \succsim is the preference relation an arbitrator has over solutions, then this objection is faulty. Nash and Kalai-Smorodinsky axiomatize the notion of consistency of a solution. Clearly, both offer reasonable definitions of consistency, and one may like both. We, on the other hand, axiomatize individual (i.e., the arbitrator's) preferences over solutions. Standard models of consumer theory argue that each agent has one preference relation which is part of his characteristics, and is not an element of choice. A person may therefore prefer the axioms of Nash to those of Kalai-Smorodinsky, but it is meaningless for him to prefer the axioms of Section 2 to another set of axioms over his preferences.

This objection is however valid if one adopts the interpretation that our axioms are rules the players agree to impose on the arbitrator, where one may object to the merit of such an approach on the ground that we replace the requirement that players agree on what is a consistent solution (that is, axioms a-la-Nash and Kalai-Smorodinsky) with a requirement that the players agree on axioms ranking solutions. This is true, but even with this interpretation we believe our axioms to be better for two reasons. One is that (as mentioned above), in the context of fixed players and a variable set A , some of Nash's (and Kalai-Smorodinsky's) axioms lose their appeal. Another difference is that we allow players a lot more flexibility in deciding what parts of the problem they find relevant. Nash's IIA axiom holds for all games $S \subset T$, regardless of whether a player perceives some of them to be more likely to be played than others, or more important for him because of their form. In our approach, players do not have to agree on the relative importance of different games. As is reflected by Theorem 1, our axioms permit a lot of disagreement between the bargainers, even though enough agreement is built into the model to ensure that Nash's is the best solution.

Another possible objection argues that instead of agreeing on axioms over preference over solutions, the bargainers can try to solve a simpler problem. Even though they do not know now what game they will play, they can agree on (axioms about) a preference relation over possible allocations in the game they will play. Indeed, all our axioms can be easily applied to sets. For example, mixture symmetry will require that if two allocations $x, y \in S$ are indifferent to each other, then for all $\alpha \in [0, 1]$, $\alpha x + (1 - \alpha)y \sim (1 - \alpha)x + \alpha y$. Disagreement indifference will require that the set $\{x: x_1 x_2 = 0\}$ be an indifference set. The optimality of the Nash solution will then follow.

Although this approach is technically correct, we find it less convincing than ours. One reason a bargainer is willing to take his opponent into consideration is that he may ex ante compensate himself for an unfavorable outcome in one game by a good outcome in another. Once we agree that the induced prefer-

ences on a particular game depend on outcomes in other potential games, solutions, rather than outcomes, should be the correct object of preferences and choice.

In Section 2 we rationalized the mixture symmetry axiom by claiming that indifference between solution f , which is more favorable to one bargainer, and g , which is more favorable to the other, suggests the half-half point as the best mixture of the two. But there are situations where two solutions may be indifferent where no conflict between the bargainers is involved. Suppose $f_2(S) = g_2(S)$ for all $S \in \mathcal{S}$. So from the second player's selfish point of view it makes no difference whether f or g is chosen. Suppose further that player 1 too is indifferent between f and g , presumably because he deems f better for some games, and g for others. Why should the arbitrator now prefer $\frac{1}{2}f + \frac{1}{2}g$ to either f or g ? In fact, since player 1 is using a von Neumann-Morgenstern utility, and since no issue of fairness is involved, the arbitrator should be indifferent between f and $\alpha f + (1 - \alpha)g$, for all $\alpha \in [0, 1]$.

We agree. Indeed, this is exactly what is implied by the representation function of Theorem 1. Let $h(S) = f_2(S) = g_2(S)$. Then $f \sim g$ implies

$$\begin{aligned} \int_{\mathcal{S}} f_1(S)h(S) d\mu(S) &= \int_{\mathcal{S}} g_1(S)h(S) d\mu(S) \\ \Rightarrow \int_{\mathcal{S}} f_1(S)h(S) d\mu(S) & \\ &= \int_{\mathcal{S}} [\alpha f_1(S) + (1 - \alpha)g_1(S)]h(S) d\mu(S). \end{aligned}$$

There is one problem to which we do not have a simple answer. Our analysis depends strongly on the fact that there are only two players. The Nash solution for the n -person bargaining game involves an n -dimensional polynomial ($\max_{x \in S} \prod(x_i - d_i)$), while mixture symmetry is associated with quadratic functions. We offer another set of axioms that can deal with any number of players in Border and Segal (1995).

Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena CA 91125, U.S.A.

and

Department of Economics, University of Western Ontario, London N6A 5C2, Canada.

Manuscript received April, 1995; final revision received March, 1996.

APPENDIX A: PROOF OF THEOREM 1

The proof of Theorem 1 is divided into numerous lemmas. Given $T \in \mathcal{S}$, for $S \in \mathcal{S}$ define $\alpha_T(S) = \max\{\lambda: \lambda T \subset S\}$. That is, $\alpha_T(S)T$ is the largest multiple of T that fits in S . In particular, $\alpha_T(T) = 1$. Also note that if $S, T \in \mathcal{S}$, then $(\kappa/K) \leq \alpha_T(S) \leq (K/\kappa)$. The mapping $S \mapsto \alpha_T(S)$ is continuous for any T .

We now temporarily fix T and for each $x \in T$ we construct a solution $\Phi_{T,x}$ by

$$\Phi_{T,x}(S) = \alpha_T(S)x.$$

By definition, $\alpha_T(S)x \in S$, so $\Phi_{T,x}$ is truly a solution. Furthermore, observe that $\Phi_{T,x}(T) = x$.

LEMMA 1: *For every $S, T \in \mathcal{S}$ and for every $x \in T$, the solution $\Phi_{T,x}$ satisfies*

$$(1) \quad \|x - \Phi_{T,x}(S)\| \leq 4 \frac{K}{\kappa} \rho(S, T).$$

The proof of this lemma is a tedious exercise in elementary geometry, so we defer it to Appendix B.

The above construction induces an isomorphism between T and a subset of solutions.

LEMMA 2: *The isomorphism $x \mapsto \Phi_{T,x}$ from T into \mathcal{S} is continuous and one-to-one. It preserves mixtures in the sense that $\Phi_{T, \lambda x + (1-\lambda)y} = \lambda \Phi_{T,x} + (1-\lambda)\Phi_{T,y}$, and is monotonic (that is, if $x \gg y$, then $\Phi_{T,x}(S) \gg \Phi_{T,y}(S)$ for all S). Finally, if $x_i = 0$, then $\Phi_{T,x}$ gives player i the outcome 0 for all S .*

PROOF: Continuity follows from continuity of $\alpha_T(S)$. Mixture preservation follows from

$$\begin{aligned} (\lambda \Phi_{T,x} + (1-\lambda)\Phi_{T,y})(S) &= \lambda \Phi_{T,x}(S) + (1-\lambda)\Phi_{T,y}(S) \\ &= \lambda \alpha_T(S)x + (1-\lambda)\alpha_T(S)y \\ &= \alpha_T(S)(\lambda x + (1-\lambda)y) \\ &= \Phi_{T, \lambda x + (1-\lambda)y}(S). \end{aligned}$$

The other claims are obvious.

Q.E.D.

Our aim is to approximate arbitrary solutions by convex combinations of solutions of this form. In order to do this, we need a partition of unity with some special properties. So given n , let $Q_n = \{z \in \mathbb{R}_+^n : \text{at most one } z_i = 0\}$, and define $\lambda^n = (\lambda_1^n, \dots, \lambda_n^n) : Q_n \rightarrow \mathbb{R}_+^n$ by

$$\lambda_i^n(z_1, \dots, z_n) = \begin{cases} \frac{z_i^{-n}}{\sum_{j=1}^n z_j^{-n}} & \text{if each } z_j > 0, \\ 1 & \text{if } z_i = 0, \\ 0 & \text{if } z_j = 0 \text{ for some } j \neq i. \end{cases}$$

Observe that λ^n is continuous on Q_n and is a partition of unity, that is, $\sum_{i=1}^n \lambda_i^n(z) = 1$ for each $z \in Q_n$.

We now define the convex combinations. Given a vector T of distinct games $T = (T_1, \dots, T_n)$ and a vector x of points $x = (x_1, \dots, x_n)$ with $x_i \in T_i$ for each i , define the solution $\Phi_{T,x}$ by

$$\Phi_{T,x}(S) = \sum_{i=1}^n \lambda_i^n(\rho(S, T_1), \dots, \rho(S, T_n)) \Phi_{T_i, x_i}(S).$$

Observe that $\Phi_{T,x}$ is continuous, $\Phi_{T,x}(S) \in S$ for each S , and $\Phi_{T,x}(S) = x_i$ whenever $S = T_i$.

Our next claim is that solutions of this form are dense in \mathcal{S} . For this, fix a countable dense subset $\{T_1, T_2, \dots\}$ of \mathcal{S} . Let f be a fixed solution and define the solution

$$\Phi_n^f = \Phi_{T_1, \dots, T_n; f(T_1), \dots, f(T_n)}.$$

Observe that by construction, $\Phi_n^f(T_i) = f(T_i)$, $i = 1, \dots, n$.

LEMMA 3: For each $f \in \mathcal{F}$, $\Phi_n^f \rightarrow f$ uniformly on \mathcal{S} .

PROOF: Since f is continuous and \mathcal{S} is compact, f is uniformly continuous on \mathcal{S} . Therefore for each $\varepsilon > 0$, there is some $\delta(\varepsilon) > 0$ such that $\rho(S, T) < \delta(\varepsilon)$ implies $\|f(S) - f(T)\| < \varepsilon$.

To simplify notation, suppress the f and just write Φ_n . Now let $\varepsilon > 0$ be given. Set

$$\eta = \min \left\{ \frac{\kappa}{12K} \varepsilon, \delta \left(\frac{\varepsilon}{3} \right) \right\}.$$

Choose N_0 large enough so that for $n \geq N_0$, for every $T \in \mathcal{S}$, there is some $T_i \in \{T_1, \dots, T_n\}$ with $\rho(T, T_i) < (\eta/2)$. (This can be done since $\{T_1, T_2, \dots\}$ is dense.) Choose N_1 large enough so that for $n \geq N_1$, $(n/2^n)\sqrt{2}K < (\varepsilon/3)$, and set $N = \max\{N_0, N_1\}$. Choose an arbitrary $S \in \mathcal{S}$. Then recalling that $\sum_{i=1}^n \lambda_i^n = 1$, we get

$$\begin{aligned} & \|f(S) - \Phi_n(S)\| \\ &= \left\| f(S) - \sum_{i=1}^n \lambda_i^n (\rho(S, T_1), \dots, \rho(S, T_n)) \Phi_{T_i, f(T_i)}(S) \right\| \\ &= \left\| \sum_{i=1}^n \lambda_i^n (\rho(S, T_1), \dots, \rho(S, T_n)) (f(S) - \Phi_{T_i, f(T_i)}(S)) \right\| \\ (2) \quad & \leq \sum_{i=1}^n \lambda_i^n (\rho(S, T_1), \dots, \rho(S, T_n)) \|f(S) - \Phi_{T_i, f(T_i)}(S)\|. \end{aligned}$$

Now

$$(3) \quad \|f(S) - \Phi_{T_i, f(T_i)}(S)\| \leq \|f(S) - f(T_i)\| + \|f(T_i) - \Phi_{T_i, f(T_i)}(S)\|.$$

So break up the sum in (2) into two parts. Let $A = \{i : \rho(S, T_i) < \eta\}$ and $B = A^c$. Then for $i \in A$, inequality (3) implies

$$\|f(S) - \Phi_{T_i, f(T_i)}(S)\| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3}.$$

This is because the first term on the right-hand side of (3) is less than or equal to $\varepsilon/3$ by the definition of $\delta(\varepsilon)$, and the second term comes from Lemma 1. Also observe that there is some $i^* \in A$ with $\rho(S, T_{i^*}) < (\eta/2)$. Regarding the terms in B , since $S \in \mathcal{S}$ is bounded above by (K, K) , the distance $\|f(S) - \Phi_{T_i, f(T_i)}(S)\|$ is bounded above by $\sqrt{2}K$, which is independent of S . Then using the definition of λ^n in (2), we have

$$\begin{aligned} & \|f(S) - \Phi_n(S)\| \\ & \leq \sum_{i=1}^n \lambda_i^n (\rho(S, T_1), \dots, \rho(S, T_n)) \|f(S) - \Phi_{T_i, f(T_i)}(S)\| \\ & \leq \underbrace{\frac{\sum_{i \in A} \rho(S, T_i)^{-n}}{\sum_{i=1}^n \rho(S, T_i)^{-n}} \left(\frac{\varepsilon}{3} + \frac{\varepsilon}{3} \right)}_{\leq 1} + \frac{\sum_{i \in B} \rho(S, T_i)^{-n}}{\sum_{i=1}^n \rho(S, T_i)^{-n}} \sqrt{2}K \end{aligned}$$

$$\begin{aligned} &\leq \left(\frac{\varepsilon}{3} + \frac{\varepsilon}{3}\right) + \frac{n\eta^{-n}}{\sum_{i=1}^n \rho(S, T_i)^{-n}} \sqrt{2} K \\ &\leq \left(\frac{\varepsilon}{3} + \frac{\varepsilon}{3}\right) + \frac{n\eta^{-n}}{\left(\frac{\eta}{2}\right)^{-n} + \sum_{i \neq i^*} \rho(S, T_i)^{-n}} \sqrt{2} K \\ &\leq \left(\frac{\varepsilon}{3} + \frac{\varepsilon}{3}\right) + \frac{n\eta^{-n}}{\left(\frac{\eta}{2}\right)^{-n}} \sqrt{2} K \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{n}{2^n} \sqrt{2} K \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \end{aligned}$$

for $n \geq N$. Thus Φ_n converges uniformly to f .

Q.E.D.

We next define an ordering \succeq_n on $T_1 \times T_2 \cdots \times T_n$, where T_1, \dots, T_n are the first n elements of the fixed dense set $\{T_1, T_2, \dots\}$, by

$$x \succeq_n y \Leftrightarrow \Phi_{T_1, \dots, T_n; x} \geq \Phi_{T_1, \dots, T_n; y}.$$

It follows from Lemma 2 that the relation \succeq_n on $T_1 \times T_2 \cdots \times T_n$ is a monotonic convex continuous quasiorder, and satisfies mixture symmetry.

In the sequel we say that the function W of n variables is quadratic if it is of the form

$$W(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + \sum_{i=1}^n b_i x_i + c.$$

LEMMA 4: *The preference \succeq_n is represented by a utility function of the form*

$$V_n(x) = \sum_{i=1}^n \sum_{j=1}^n (x_{i1} x_{j2}) p_{ij},$$

where $p \geq 0$ and $\sum_{i,j=1}^n p_{ij} = 1$.

PROOF: We use Fact 1, and rule out areas B and C by showing that the coordinate axes are indifference sets, which is compatible only with a quasiconcave quadratic representation. Let $x^i \in T_1 \times T_2 \cdots \times T_n$ satisfy $x_{m2} = 0$ for all $m = 1, \dots, n$, $x_{m1} = 0$ for all $m \neq i$, and $x_{i1} = \max\{x_1 : x \in T_i\} \geq \kappa > 0$. It follows that $\Phi_{T, x^i}(S)_2 = 0$ for all S . So by condition DI, $\Phi_{T, x^i} \sim \Phi_{T, 0}$, which implies that the $i1$ -axis is a \succeq_n -indifference set. This shows that every axis is an indifference set, which rules out the regions B and C .

Thus \succeq_n has a representation of the form

$$(4) \quad V_n(x) = \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^2 \sum_{l=1}^2 (x_{ik} x_{jl}) p_{ikjl} + \sum_{i=1}^n \sum_{k=1}^2 x_{ik} b_{ik},$$

where p satisfies $p \geq 0$ (because of monotonicity) and $\sum_i \sum_j \sum_k \sum_l p_{ikjl} = 1$. (The sum cannot be zero, since V_n is not linear.) We now show that the b_{ik} terms in the representation in equation (4) are zero. Let x^i be as above. Then by condition DI, $\Phi_{T, x^i} \sim \Phi_{T, 0}$, so $V_n(x^i) = 0$, for all values of x_{i1} . This implies $p_{i1i1} = b_{i1} = 0$. Similar arguments show that each $p_{jkjk} = b_{jk} = 0$.

Next we show that if $i \neq j$ and $k = l$, then $p_{ikjl} = 0$. Let $x^{ij} \in T_1 \times T_2 \cdots \times T_n$ satisfy $x_{m2}^{ij} = 0$ for $m = 1, \dots, n$, $x_{m1}^{ij} = 0$ for $m \notin \{i, j\}$, $x_{m1}^{ij} = \max\{x_1 : x \in T_m\}$ for $m \in \{i, j\}$. Again $\Phi_{T, x^{ij}}(S)_2 = 0$ for all S , so as above $V_n(x^{ij}) = 0$, which implies $p_{i1j1} = 0$. Similarly each $p_{i2j2} = 0$. *Q.E.D.*

This construction defines a discrete probability measure π_n on $\mathcal{S} \times \mathcal{S}$ by $\pi_n(\{(T_i, T_j)\}) = p_{ij}$ for each (i, j) . This measure has the property that

$$V_n(x) = \int_{\mathcal{S} \times \mathcal{S}} f_1(S)f_2(T) d\pi_n(S, T) \quad \text{whenever } f_k(T_i) = x_{ik}$$

for $i = 1, \dots, n$ and $k = 1, 2$. So abuse notation slightly and define $V_n: \mathcal{F} \rightarrow \mathbb{R}$ by

$$V_n(f) = \int_{\mathcal{S} \times \mathcal{S}} f_1(S)f_2(T) d\pi_n(S, T)$$

and note that $V_n(f) = V_n(\Phi_n^f)$. This means that

$$V_n(f) \geq V_n(g) \Leftrightarrow (f(T_1), \dots, f(T_n)) \succeq_n (g(T_1), \dots, g(T_n)).$$

Since \mathcal{S} is compact, the set of probability measures on $\mathcal{S} \times \mathcal{S}$ is compact (in the topology of weak convergence of measures, see Theorem 12.10 in Aliprantis and Border (1994)), so there is a subsequence of π_n converging to a limit π , which is also a probability measure on $\mathcal{S} \times \mathcal{S}$. Define

$$V(f) = \int_{\mathcal{S} \times \mathcal{S}} f_1(S)f_2(T) d\pi(S, T).$$

Note that if $f_m \rightarrow f$ uniformly, then $\int f_{m1}f_{m2} d\pi_{n_m} \rightarrow \int f_1f_2 d\pi$ (see Corollary 12.6 in Aliprantis and Border (1994)). That is, $V_{n_m}(f_m) \rightarrow V(f)$. In particular, $V_{n_m}(\Phi_{n_m}^f) \rightarrow V(f)$.

LEMMA 5: *If $f \succcurlyeq g$, then $V(f) \geq V(g)$. In particular, $f \sim g$ implies $V(f) = V(g)$.*

PROOF: First suppose $f \succ g$. Since $\Phi_n^f \rightarrow f$ and $\Phi_n^g \rightarrow g$, for large enough n we have $\Phi_n^f \succ \Phi_n^g$, so $V_n(\Phi_n^f) > V_n(\Phi_n^g)$. Therefore $V(f) \geq V(g)$. By continuity of V and \succcurlyeq , we have $f \succcurlyeq g$ implies $V(f) \geq V(g)$. Q.E.D.

LEMMA 6: *The limit measure π is supported by the diagonal $\Delta = \{(S, S) : S \in \mathcal{S}\}$ of $\mathcal{S} \times \mathcal{S}$.*

PROOF: Let $G \times H \subset \mathcal{S} \times \mathcal{S}$ be an open region with $\Delta \cap (G \times H) = \emptyset$ and $\pi(G \times H) > 0$. Note that this implies that $G \cap H = \emptyset$. Pick $(S, T) \in G \times H$. Let $\lambda: \mathcal{S} \rightarrow \mathbb{R}$ be a continuous function satisfying $\lambda(\bar{S}) = 1, 1 > \lambda(U) > 0$ for all $U \in G$ with $U \neq \bar{S}$ and $\lambda(U) = 0$ for all $U \in G^c$. Similarly let $\alpha: \mathcal{S} \rightarrow \mathbb{R}$ be a continuous function satisfying $\alpha(\bar{T}) = 1, 1 > \alpha(U) > 0$ for all $U \in H$ with $U \neq \bar{T}$ and $\alpha(U) = 0$ for all $U \in H^c$. Let h be the solution that gives everything to 1, that is $h_1(U) = \max\{x_1 : x \in U\}$ and $h_2(S) = 0$ for all S . Similarly let g give everything to 2. Set $f(U) = \lambda(U)h(U) + \alpha(U)g(U)$. Then f is a continuous solution with $\{(S, T) : f_1(S)f_2(T) > 0\} = G \times H$. But by condition DI, $f \sim 0$, so by Lemma 5, $V(f) = V(0) = 0$. Thus $\pi(G \times H) = 0$. This implies $\pi(\Delta) = 1$. Q.E.D.

Set $\mu(G) = \pi(G \times G)$ for every Borel subset G of \mathcal{S} . Then clearly

$$V(f) = \int_{\mathcal{S}} f_1(S)f_2(S) d\mu(S).$$

LEMMA 7: *The measure μ has full support.*

PROOF: Let G be an open subset of \mathcal{S} . Let N denote the Nash bargaining solution (any positive solution will do), and let $\lambda: G \rightarrow \mathbb{R}$ be a continuous function satisfying $1 \geq \lambda(S) > 0$ for $S \in G$ and $\lambda(S) = 0$ for $S \in G^c$. Let g be the solution defined by $g(S) = \lambda(S)N(S)$. Then by monotonicity, $g > 0$. Thus for some $\varepsilon > 0$ small enough the solution f with $f_k(S) = \varepsilon \leq \kappa, k = 1, 2$, is a solution on \mathcal{S} satisfying $g \succ f$. Therefore by Lemma 5, $V(g) \geq V(f)$. Now $V_n(f) \rightarrow V(f)$, and $V_n(f) = \varepsilon^2$ for all n , so $V(g) \geq \varepsilon^2 > 0$, which implies $\mu(G) > 0$. Thus μ has full support. Q.E.D.

We can now finally prove Theorem 1.

PROOF OF THEOREM 1: We already know that $f \succsim g$ implies $V(f) \geq V(g)$. It only remains to show that $f \succ g$ implies $V(f) > V(g)$. Now if $f \succ g$, it follows by condition DI that $f_1(S)f_2(S) > 0$ for some S . (Otherwise $f \sim 0$.) Also, for some $1 > \alpha \geq 0$, we have $\alpha f \sim g$. But since μ has full support, we get $V(f) > V(\alpha f) \geq V(g)$, so $V(f) > V(g)$. *Q.E.D.*

APPENDIX B: PROOF OF LEMMA 1

We start with some simple facts.

FACT 2: If $\alpha_T(S) < 1$, then $\alpha_T(S \cap T) = \alpha_T(S)$.

FACT 3: $\rho(S \cap T, T) \leq \rho(S, T)$.

PROOF: We have to show that for $x \in T$, $d(x, S \cap T) \leq d(x, S)$. Let $x \in T \setminus S$ and $y \in S$ such that $d(x, S) = d(x, y)$. If $y \in S \cap T$, then $d(x, S \cap T) = d(x, S)$. Suppose $y \in S \setminus T$. Then since S is disposable, and by the triangle inequality, $d(x, S) < d(x, y)$, a contradiction. *Q.E.D.*

We now turn to the proof of the Lemma 1. We break the analysis into cases. For each of these cases we show that equation (1) is satisfied.

CASE 1: We consider first games S and T with $\alpha_T(S) < 1$. Let $x^* \in T$ be such that $\|x - \Phi_{T,x}(S)\| = (1 - \alpha_T(S))\|x\|$ is maximized over T at x^* . Note that $x^* \in \arg \max_{x \in T} \|x\|$ and that x^* is on the boundary of T . We shall replace S and T by sets S^* and T^* satisfying the following three requirements.

1. $x^* \in T^*$.
2. $\alpha_{T^*}(S^*) = \alpha_T(S)$.
3. $\rho(S^*, T^*) \leq \rho(S, T)$.

To simplify notation, replace $\alpha_T(S)$ by simply α . Also, for $x, y \in \mathbb{R}^2$, let $l(x, y)$ denote the line through these two points. The slope of a line H is denoted $\sigma(H)$.

By Facts 2 and 3, the three requirements are satisfied if $T^* = T$ and $S^* = S \cap T$. We can thus assume that $S \subset T$. By the definition of $\alpha_T(S)$, there is a point $y \in T$ such that αy is on the outer boundary of S . Of course, y is on the boundary of T . Assume without loss of generality that $y_1 \leq x_1^*$ and $y_2 \geq x_2^*$.

Let H be a supporting line to S at αy . Denote the area bounded by H and the two axes by S_1 , and let $S_2 = T \cap S_1$. Note that $\alpha_T(S_2) = \alpha$, but $\rho(S_2, T) \leq \rho(S, T)$.

Let T^* be the minimal convex disposable set containing y, x^* , and (κ, κ) , and let $S^* = S_2 \cap T^*$. Clearly, $\alpha_{T^*}(S^*) = \alpha$, but $\rho(S^*, T^*) \leq \rho(S, T)$. Observe that for all x on the boundary of T^* , $d(x, S^*) \leq d(y, S^*)$. This is because $\alpha x^* \in S^*$, hence $\sigma(H) \geq \sigma(l(y, x^*))$. On the other hand, the boundary of T^* to the left of y is either the horizontal segment $A = [(0, y_2), y]$ ($y_2 \geq \kappa$), or it is the two segments $B = [(0, \kappa), (\kappa, \kappa)]$ and $C = [(\kappa, \kappa), y]$. In the first case, no point in A can be further away from H than y . In the second case, since $(\kappa, \kappa) \in S^*$, it follows that $B \subset S^*$, and since $\sigma(H) \leq \sigma(l((\kappa, \kappa), \alpha y))$, it follows that $y = \arg \max_{x \in C} d(x, H)$. We thus obtain

$$(5) \quad \rho(S^*, T^*) = d(y, S^*).$$

There are now four subcases.

SUBCASE 1.A: Suppose first that $\sigma(H) < -1$ and $y_1 \geq \kappa$. Now $\sigma(l(y, x^*)) \leq \sigma(H) < -1$, so given that $x^* \in \arg \max_{x \in T^*} \|x\|$ it follows that y lies in the region bounded by the quarter circle of radius $\|x^*\|$ and the line of slope -1 through x^* . This line intersects this arc in exactly two places: x^* and its transpose (x_2^*, x_1^*) (which may coincide, in which case $x_1^* = x_2^*$ and $y = x^*$). In particular,

$x_1^* \geq x_2^*$ and since y lies to the right of the transpose (x_2^*, x_1^*) , we have $y_1 \geq x_2^*$. It also follows that $d(y, S^*) \geq (1 - \alpha)y_1$, so by equation (5) we have $(1 - \alpha)y_1 \leq \rho(S^*, T^*)$. Using $y_1 \geq \kappa$, and these other facts we get

$$\begin{aligned} \|x^* - \Phi_{T, x^*}(S)\| &= (1 - \alpha)\|x^*\| \\ &\leq (1 - \alpha)\|(K, x_2^*)\| \\ &\leq (1 - \alpha)\|(y_1, K)\| \\ &\leq \|(1, K/y_1)\| \rho(S^*, T^*) \\ &\leq \|(1, K/\kappa)\| \rho(S^*, T^*) \\ &\leq \sqrt{2} \frac{K}{\kappa} \rho(S, T). \end{aligned}$$

SUBCASE 1.B: Suppose now that $\sigma(H) < -1$ and $y_1 < \kappa$. Since y is on the boundary of T^* , $y_2 \geq \kappa$. Observe that $\|x^*\| \geq \|y\|$. Let $s^2 = x_1^{*2} + x_2^{*2} = \|x\|^2$, and obtain

$$\sigma(H) \geq \sigma(l(y, x^*)) \geq \sigma(l(s/\sqrt{2}, s/\sqrt{2}), (s, 0)) = -\frac{1}{\sqrt{2} - 1}.$$

Let H^* be the line through αy with slope $-1/(\sqrt{2} - 1)$. Then, since $\sigma(H) < -1$,

$$\begin{aligned} d(y, S^*) \geq d(y, H^*) &= \frac{(1 - \alpha)(y_1 + (\sqrt{2} - 1)y_2)}{\sqrt{4 - 2\sqrt{2}}} \\ &\geq \frac{(1 - \alpha)(\sqrt{2} - 1)\kappa}{\sqrt{4 - 2\sqrt{2}}}. \end{aligned}$$

In this calculation, we use the fact that the distance between the line $Ax_1 + Bx_2 + C = 0$ and the point (x_1^0, x_2^0) is

$$(6) \quad \frac{|(Ax_1^0 + Bx_2^0 + C)|}{\sqrt{A^2 + B^2}}.$$

Now

$$\begin{aligned} \|x^* - \Phi_{T, x^*}(S)\| &= (1 - \alpha)\|x^*\| \\ &\leq (1 - \alpha)\sqrt{2}K \\ &\leq \frac{\sqrt{2}K\sqrt{4 - 2\sqrt{2}}}{(\sqrt{2} - 1)\kappa} \rho(S, T) \\ &< 4 \frac{K}{\kappa} \rho(S, T). \end{aligned}$$

The last inequality follows from

$$\begin{aligned} \frac{\sqrt{2}\sqrt{4 - 2\sqrt{2}}}{(\sqrt{2} - 1)} &= \sqrt{2}(\sqrt{2} + 1)\sqrt{4 - 2\sqrt{2}} \\ &= \sqrt{8 + 4\sqrt{2}} < 4. \end{aligned}$$

SUBCASE 1.C: Suppose now that $\sigma(H) \geq -1$ and $y_2 \geq \kappa$. Since $\sigma(H) \geq -1$, we have $d(y, S^*) \geq (1 - \alpha)y_2$. Since $y_2 \geq \kappa$, we have

$$\begin{aligned} \|x^* - \Phi_{T, x^*}(S)\| &= (1 - \alpha)\|x^*\| \\ &\leq (1 - \alpha)\|(K, y_2)\| \\ &\leq \rho(S^*, T^*)\|(K/y_2, 1)\| \\ &\leq \rho(S^*, T^*)\|(K/\kappa, 1)\| \\ &\leq \sqrt{2} \frac{K}{\kappa} \rho(S, T). \end{aligned}$$

SUBCASE 1.D: The remaining case is $\sigma(H) \geq -1$ and $y_2 < \kappa$. Let $H^* = l((\kappa, \kappa), \alpha y)$. Since $(\kappa, \kappa) \in S^*$, $\sigma(H^*) \geq \sigma(H)$. Therefore, $\rho(S, T) \geq d(y, S^*) \geq d(y, H) \geq d(y, H^*)$. Note that H^* is given by $-(\alpha y_2 - \kappa)x_1 + (\alpha y_1 - \kappa)x_2 - \alpha\kappa(y_1 - y_2) = 0$. Therefore, by using equation (6), we obtain

$$\rho(S, T) \geq d(y, H^*) = \frac{(1 - \alpha)\kappa(y_1 - y_2)}{\sqrt{(\alpha y_1 - \kappa)^2 + (\alpha y_2 - \kappa)^2}}.$$

Hence

$$\begin{aligned} (7) \quad \|x^* - \Phi_{T, x^*}(S)\| &= (1 - \alpha)\|x^*\| \\ &\leq (1 - \alpha)\sqrt{2}K \\ &\leq \frac{\sqrt{(\alpha y_1 - \kappa)^2 + (\alpha y_2 - \kappa)^2}}{\kappa(y_1 - y_2)} \sqrt{2}K\rho(S, T). \end{aligned}$$

Since αy is on the boundary of S , it follows that $\alpha \geq \kappa/y_1$. Also, by assumption, $\alpha \leq 1$. Since $h(\alpha) := (\alpha y_1 - \kappa)^2 + (\alpha y_2 - \kappa)^2$ is a convex function of α , it follows that for $\alpha \in [\kappa/y_1, 1]$, $h(\alpha) \leq \max\{h(\kappa/y_1), h(1)\}$. For $\alpha = \kappa/y_1$, we obtain

$$\frac{\sqrt{\left(\frac{\kappa}{y_1}y_1 - \kappa\right)^2 + \left(\frac{\kappa}{y_1}y_2 - \kappa\right)^2}}{\kappa(y_1 - y_2)} = \frac{1}{y_1} \leq \frac{1}{\kappa}.$$

Now let $\alpha = 1$. Since $y_1 \geq \kappa$ and $y_2 \leq \alpha$, we obtain for $y \neq (\kappa, \kappa)$ that

$$\frac{\sqrt{(y_1 - \kappa)^2 + (y_2 - \kappa)^2}}{\kappa(y_1 - y_2)} \leq \frac{1}{\kappa}.$$

Thus inequality (7) becomes

$$\begin{aligned} \|x^* - \Phi_{T, x^*}(S)\| &\leq \frac{\sqrt{(\alpha y_1 - \kappa)^2 + (\alpha y_2 - \kappa)^2}}{\kappa(y_1 - y_2)} \sqrt{2}K\rho(S, T) \\ &\leq \sqrt{2} \frac{K}{\kappa} \rho(S, T). \end{aligned}$$

Finally, if $y = (\kappa, \kappa)$, then since αy is on the boundary of S , $\alpha = 1$, and the lemma is trivially true.

CASE 2: The analysis of the case $\alpha_T(S) > 1$ is similar, only we replace T with S^* , and S with T^* .
Q.E.D.

REFERENCES

- ALIPRANTIS, C. D., AND K. C. BORDER (1994): *Infinite Dimensional Analysis*. Berlin: Springer-Verlag.
- BINMORE, K. (1991): "Game Theory and the Social Contract," in *Game Equilibrium Models*, ed. by R. Selten. Berlin: Springer-Verlag, pp. 85–163.
- BORDER, K. C., AND U. SEGAL (1995): "Preferences over Solutions to the Bargaining Problem," Social Sciences WP #923, California Institute of Technology.
- CHEW, S. H., L. EPSTEIN, AND U. SEGAL (1991): "Mixture Symmetry and Quadratic Utility," *Econometrica*, 59, 139–163.
- (1994): "The Projective Independence Axiom," *Economic Theory*, 4, 189–215.
- DIAMOND, P. A. (1967): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparison of Utility: Comment," *Journal of Political Economy*, 75, 765–766.
- EPSTEIN, L. G., AND U. SEGAL (1992): "Quadratic Social Welfare Functions," *Journal of Political Economy*, 100, 691–712.
- HARSANYI, J. C. (1955): "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, 63, 309–321.
- KALAI, E., AND M. SMORODINSKY (1975): "Other Solutions to Nash's Bargaining Problem," *Econometrica*, 43, 513–518.
- NASH, J. F. (1950): "The Bargaining Problem," *Econometrica*, 18, 155–162.
- (1953): "Two Person Cooperative Games," *Econometrica*, 21, 128–140.
- RUBINSTEIN, A., Z. SAFRA, AND W. THOMSON (1992): "On the Interpretation of the Nash Bargaining Solution and its Extension to Non-Expected Utility Preferences," *Econometrica*, 60, 1171–1186.
- VAN DAMME, E. (1986): "The Nash Bargaining Solution is Optimal," *Journal of Economic Theory*, 38, 78–100.