A Characterization of the Nash Bargaining Solution

Published in *Social Choice and Welfare*, 19, 811-823, (2002) *

Nir Dagan † Oscar Volij ‡ Eyal Winter §

August 20, 2001

Abstract

We characterize the Nash bargaining solution replacing the axiom of Independence of Irrelevant Alternatives with three independent axioms: Independence of Non-Individually Rational Alternatives, Twisting, and Disagreement Point Convexity. We give a non-cooperative bargaining interpretation to this last axiom. JEL Classification, C72, C78.

*We thank Marco Mariotti, two anonymous referees and an associate editor for helpful comments.

†Nir Dagan, Academic Priority Ltd., www.nirdagan.com. Email: nir@nirdagan.com.

‡Department of Economics, Iowa State University, Ames, Iowa 50011, and Department of Economics, Hebrew University, Jerusalem 91905, Israel. Web site: http://volij.co.il. Email: oscar@volij.co.il.

§Department of Economics, Hebrew University, Jerusalem 91905, Israel. Web site: http://www.ma.huji.ac.il/~eyalw/. Email: mseyal@mscc.huji.ac.il.

1 Introduction

Since Nash (1950), a bargaining problem is usually defined as a pair (S, d) where S is a compact, convex subset of $\mathbb{R}^2$ containing both d and a point that strictly dominates d. Points in S are interpreted as feasible utility agreements and d represents the status-quo outcome. A bargaining solution is a rule that assigns a feasible agreement to each bargaining problem. Nash (1950) proposed four independent properties and showed that they are simultaneously satisfied only by the Nash bargaining solution.

While three of Nash's axioms are quite uncontroversial, the fourth one (known as *independence of irrelevant alternatives* (IIA)) raised some criticisms, which lead to two different lines of research. Some authors looked for characterizations of alternative solutions which do not use the controversial axiom (see for instance, Kalai and Smorodinsky (1975), and Perles and Maschler (1981)) while other papers provided alternative characterizations of the Nash solution without appealing to the IIA axiom. Examples of this second line of research are Peters (1986b), Chun and Thomson (1990), Peters and van Damme (1991), Mariotti (1999), Mariotti (2000), and Lensberg (1988). The first three papers replace IIA by several axioms in conjunction with some type of continuity. The next two papers replace IIA and other axioms by one axiom. Lastly, Lensberg (1988) replaces IIA with *consistency*, and consequently a domain with a variable number of agents is needed.

In this paper, we provide an alternative characterization of the Nash bargaining solution in which the axiom of independence of irrelevant alternatives is replaced by three different axioms. While all three of these axioms are known in the literature, they have never been used in combination. One of the axioms is *independence of non-individually rational alternatives*, which requires a solution to be insensitive to changes in the feasible set that involve only non-individually rational outcomes. This axiom neither implies nor is implied by IIA, but is weaker than IIA and Individual Rationality together.[1] The second axiom is *twisting*, which is a weak monotonicity requirement that is implied by IIA. The third axiom is *disagreement point convexity* which requires that the solution be insensitive to movements of the disagreement point towards the proposed compromise. This last axiom does not imply nor is implied by IIA. Further, the three axioms together do not imply IIA.

All of the axioms used in this paper have a straightforward interpretation except, perhaps, for disagreement point convexity. This axiom, however, has an interpretation that is closely related to non-cooperative models of bargaining. Assume that the solution recommends $f(S, d)$ when

[1] A solution is individually rational if it assigns each player a utility level that is not lower than its disagreement level. See next section.

the bargaining problem is (S, d). The players may postpone the resolution of the bargaining for t periods getting $f(S, d)$ only after t periods of disagreement. From today's point of view, knowing that one has the alternative of reaching agreement t periods later is as if the new disagreement point was $f(S, d)$ paid t periods later. Disagreement point convexity requires that the solution be insensitive to this kind of manipulation.

Our result, though not its proof, is closely related to Peters and van Damme (1991). The main difference is that we replace their *disagreement point continuity* axiom by the twisting axiom. In this way, we get rid of a mainly technical axiom and replace it by a more intuitive and reasonable one. Needless to say, disagreement point continuity and twisting, are not equivalent. Further, neither of them implies the other.

The paper is organized as follows: In Section 2, we present the preliminary definitions and the axioms used in the characterization. Section 3 gives the main result. Section 4 shows that the axioms are independent. Finally, Section 5 discusses the related literature.

2 Basic definitions

In this section, we present some basic definitions. Since most of them are standard, we do not provide their interpretation.

A *bargaining problem* is a pair (S, d) where $S \subseteq \mathbb{R}^2$ is a compact, convex set, $d \in S$ and there is $s \in S$ with $s \gg d$.[2] We denote by $\mathcal{B}$ the set of all bargaining problems. A *bargaining solution* is a set-valued function $f : \mathcal{B} \to 2^{\mathbb{R}^2} \setminus \emptyset$ such that for every bargaining problem $B = (S, d)$, $f(B) \subseteq S$. We allow for set-valued solutions to highlight the role of some of the axioms in the present characterization. Let (S, d) be a bargaining problem. We say that $s \in S$ is *individually rational* if $s \geq d$. We say that $s \in S$ is *weakly efficient* if there is no $s' \in S$ such that $s' \gg s$ and that s is *efficient* if there is no $s' \in S$, $s' \neq s$, such that $s' \geq s$. We denote by $\mathcal{IR}(S, d)$ the set of individually rational points in (S, d).

The *Nash bargaining solution* is the solution $n : \mathcal{B} \to 2^{\mathbb{R}^2} \setminus \emptyset$ that for each bargaining problem (S, d) selects the singleton $\{(s_1^*, s_2^*)\} \subseteq S$ that contains the only point in $\mathcal{IR}(S, d)$ which satisfies $(s_1^* - d_1)(s_2^* - d_2) \geq (s_1 - d_1)(s_2 - d_2)$, for all $(s_1, s_2) \in \mathcal{IR}(S, d)$.

We now turn to properties of bargaining solutions.

A bargaining problem (S, d) is *symmetric* if

- $d_1 = d_2$ and

[2] We adopt the following conventions for vector inequalities: $x \gg y \leftrightarrow x_i > y_i$ for all i, and $x \geq y \leftrightarrow x_i \geq y_i$ for all i.

- $(s_1, s_2) \in S$ implies $(s_2, s_1) \in S$.

We say that (S', d') is obtained from the bargaining problem (S, d) by the transformations $s_i \to \alpha_i s_i + \beta_i$, for $i = 1, 2$, if $d'_i = \alpha_i d_i + \beta_i$, for $i = 1, 2$ and

$$S' = \{(\alpha_1 s_1 + \beta_1, \alpha_2 s_2 + \beta_2) \in \mathbb{R}^2 : (s_1, s_2) \in S\}.$$

The following properties are standard:

Symmetry: A bargaining solution f satisfies *symmetry* if for all symmetric bargaining problems (S, d),

$$(s_1, s_2) \in f(S, d) \Leftrightarrow (s_2, s_1) \in f(S, d).$$

Weak Pareto optimality: A bargaining solution f satisfies *weak Pareto optimality* if for all bargaining problems (S, d), $f(S, d)$ is a subset of the weakly efficient points in S. It satisfies *Pareto optimality* if for all bargaining problems (S, d), $f(S, d)$ is a subset of the efficient points in S.

Invariance: A bargaining solution satisfies *invariance* if whenever (S', d') is obtained from the bargaining problem (S, d) by means of the transformations $s_i \to \alpha_i s_i + \beta_i$, for $i = 1, 2$, where $\alpha_i > 0$ and $\beta_i \in \mathbb{R}$, we have that $f_i(S', d') = \alpha_i f_i(S, d) + \beta_i$, for $i = 1, 2$.

IIA: A bargaining solution f satisfies *independence of irrelevant alternatives* if $f(S', d) = f(S, d) \cap S'$ whenever $S' \subseteq S$ and $f(S, d) \cap S' \neq \emptyset$.

Since we do not require solutions to be single-valued, the above properties are not enough to characterize the Nash bargaining solution. In order to establish what is essentially Nash's characterization we need the following property.

Single-valuedness in symmetric problems: A bargaining solution f satisfies *single-valuedness in symmetric problems* if for every symmetric problem $B \in \mathcal{B}$, $f(B)$ is a singleton.

As stated in the introduction, we shall replace the axiom of IIA by the following three independent properties:

Independence of non-individually rational alternatives: A bargaining solution satisfies *independence with respect to non-individually rational alternatives* if for every two problems (S, d) and (S', d) such that $\mathcal{IR}(S, d) = \mathcal{IR}(S', d)$ we have $f(S, d) = f(S', d)$.

Independence of non-individually rational alternatives requires that the solution be insensitive to changes in the feasible set that do not involve individually rational outcomes. It clearly implies that the solution always chooses a subset of the individually rational agreements. It can be checked that if a solution always chooses a subset of the individually rational agreements and also satisfies IIA then the solution satisfies independence of non-individually rational alternatives. This axiom was first discussed in Peters (1986a).

The following axiom says the following. Assume that the point $\hat{s} = (\hat{s}_1, \hat{s}_2)$ is chosen by the solution when the problem is (S, d). Assume further that the feasible set is modified so that all the subtracted points are preferred by one player to $\hat{s}$ while $\hat{s}$ is preferred by the same player to each of the added points. Then the axiom requires that $\hat{s}$ be weakly preferred by that same player to at least one point selected by the solution in the new problem (S', d).

Twisting: A bargaining solution f satisfies *twisting* if the following holds: Let (S, d) be a bargaining problem and let $(\hat{s}_1, \hat{s}_2) \in f(S, d)$. Let (S', d) be another bargaining problem such that for some agent $i = 1, 2$

$$\begin{aligned} S \setminus S' &\subseteq \{(s_1, s_2) : s_i > \hat{s}_i\} \\ S' \setminus S &\subseteq \{(s_1, s_2) : s_i < \hat{s}_i\}. \end{aligned}$$

Then, there is $(s'_1, s'_2) \in f(S', d)$ such that $s'_i \leq \hat{s}_i$.

Twisting is a mild monotonicity condition, which was introduced (in its single-valued version) by Thomson and Myerson (1980) who also showed that it is implied by IIA. Twisting is satisfied by most solutions discussed in the literature.

The next axiom was used in Peters and van Damme (1991). Thomson (1994), who calls it *star-shaped inverse* succinctly summarizes this axiom as saying "that the move of the disagreement point in the direction of the desired compromise does not call for a revision of this compromise".

Disagreement point convexity: A bargaining solution f satisfies *disagreement point convexity* if for every bargaining problem $B = (S, d)$, for all $s \in f(S, d)$ and for every $\lambda \in (0, 1)$ we have $s \in f(S, (1 - \lambda)d + \lambda s)$.

This axiom has a non-cooperative flavor and it is related to one of the properties of the Nash equilibrium concept for extensive form games, namely the property that one can "fold back the tree". Consider an extensive form game and fix a Nash equilibrium σ in it. For every node n in the tree, σ determines an outcome, $z(n, \sigma)$, which is the outcome that would result if σ was played in the subgame that starts at node n. In particular, σ determines a Nash equilibrium outcome $z(n_0, \sigma)$, where n_0 denotes the root of the tree. Now, $z(n_0, \sigma)$ remains a Nash equilibrium outcome if we replace any given node n by the outcome $z(n, \sigma)$. This "tree folding property" is also satisfied by the Subgame Perfect equilibrium concept. However, we want to stress that this property is so basic that it is even satisfied by the Nash equilibrium concept. The axiom of disagreement point convexity tries to capture the tree folding property when applied to the subgame perfect equilibrium of a specific class of bargaining games, which we turn to describe.

Many non-cooperative models of bargaining are represented by an infinite-horizon stationary extensive form game with common discount factor δ, Rubinstein's (1982) alternating offers model being the most prominent example. Further, the solution concept used is subgame perfect equilibrium. All these games have the following properties:

1. The disagreement outcome corresponds to the infinite history in which the current proposal is rejected at every period.

2. There is an agreement a^* such that the unique subgame perfect equilibrium of the game dictates that a^* is immediately agreed upon. Further, a^* is immediately agreed upon at every subgame that is equivalent to the original game.

To see an application of the tree folding property to one such game, consider a stationary extensive form bargaining game Γ with the properties 1 and 2 above[3] and fix a period t. Assume that at period t the proposer is the same one as in the first period so that all subgames that start at the beginning of period t are identical to Γ. Build a new game by replacing each subgame of Γ that starts at the beginning of period t by the subgame perfect equilibrium outcome of that subgame. (Note that an outcome will typically have the format of "disagreement until period t' and agreement a at t'". [4]). By property 2 above, this outcome is "disagreement until period t, and agreement a^* at t". The resulting game, $\Gamma(t)$, is a finite horizon extensive form game in which a history of constant rejections leads to a^* at period t. That is, in this new game disagreement leads to the subgame perfect equilibrium outcome a^*, but delayed by t periods during which there is disagreement. Still, the subgame perfect equilibrium outcome of this modified game $\Gamma(t)$ is an immediate agreement on a^*, which is what the tree folding property says.

Going back to the cooperative bargaining problem, let d be the present value of the utility stream of disagreement forever, and let s be the vector of utilities that correspond to the equilibrium outcome a^*. Then, the shifted disagreement point $(1-\lambda)d + \lambda s$ in the disagreement point convexity axiom corresponds precisely to the disagreement outcome of the amended game $\Gamma(t)$, λ being δ^t. To see this, note that the present value of a stream of t periods of disagreement and

[3] The reader may find it convenient to consider Rubinstein's (1982) game.

[4] We have in mind bargaining over a per-period payoff rather than over a stock. Both approaches are equivalent since every constant flow is equivalent to a stock and vice versa.

then agreement on a^* at t is $(1-\delta^t)d_i + \delta^t s_i$ for player i, for $i = 1, 2$.[5] Using this interpretation, disagreement point convexity simply says that if we amend the bargaining problem so that the consequence of no agreement is that players disagree for t periods, and receive $f(S, d)$ afterwards (yielding a payoff of $(1-\delta^t)d + \delta^t f(S, d)$), then they should agree on $f(S, d)$ to be paid from the outset. Note that for the disagreement point to move along the segment that connects d and s when we replace the subgame with its equilibrium outcome, it is essential to assume a common discount factor.

Disagreement point convexity seems to be an appropriate requirement, especially if one has in mind a stationary bargaining game. Dagan, Volij, and Winter (1999) exploit this axiom to give a characterization of the *time-preference Nash solution* in a setting with physical outcomes.[6]

3 The Main Result

We can now present the main result.

Theorem 1 A bargaining solution satisfies weak Pareto optimality, symmetry, invariance, single-valuedness in symmetric problems, independence with respect to non-individually rational allocations, twisting, and disagreement point convexity if and only if it is the Nash bargaining solution.

Proof: It is known that the Nash solution satisfies weak Pareto optimality, symmetry, invariance and single-valuedness in symmetric problems (see Nash (1950)). By its definition, the Nash solution also satisfies independence of non-individually rational alternatives. Also, the Nash solution satisfies twisting, since twisting is weaker than IIA (see Thomson and Myerson (1980) or the appendix for the set valued version used here), which is in turn satisfied by the Nash solution. Finally, Peters and van Damme (1991) showed that it also satisfies disagreement point convexity. This shows that the Nash solution satisfies all the axioms in the theorem. We now show that no other solution satisfies all of them together. Suppose that a solution f satisfies all the axioms.

First step. Consider first a triangular problem (S, d) where $S = \text{co}\{(d_1, d_2), (b_1, d_2), (d_1, b_2)\}$ with $b_i > d_i$ for $i = 1, 2$, and for any set $A \subseteq \mathbb{R}^2$, co$A$ is the convex hull of A. Since there are affine transformations by means of which (S, d) is obtained from $(\text{co}\{(0,0), (1,0), (0,1)\}, (0,0)) \equiv (I, (0,0))$ and since both f and n satisfy invariance, we have that $f(S, d) = n(S, d)$ if and only if

[5] If one considers a model without impatience but where after each rejected offer there is a probability $1-\delta$ of negotiations breakdown, resulting in d, then $(1-\delta^t)d + \delta^t s$ is the expected utility pair associated with a history of agreement on a^* after t rejections.

[6] See Binmore, Rubinstein, and Wolinsky (1986) for the difference between what they call the standard and the time-preference Nash solutions.

$f(I,(0,0)) = n(I,(0,0))$. But by single-valuedness in symmetric problems, weak Pareto optimality and symmetry of f we know that $f(I,(0,0)) = \{(1/2,1/2)\} = n(I,(0,0))$.

Second step. Consider a general bargaining problem (S,d) and let $\hat{s} \in f(S,d)$. Since both n and f satisfy independence of non-individually rational alternatives, we can assume without loss of generality that $\mathcal{IR}(S,d) = S$.

Case 1: $\hat{s} \gg d$**:** In this case, by invariance we can assume without loss of generality that $d = (0,0)$ and $\hat{s} = (1/2,1/2)$. It is enough to show that $\hat{s} \in n(S,d)$. Assume by contradiction that $\hat{s} \notin n(S,d)$ and consider the triangular problem $(\mathrm{co}\{(0,0),(1,0),(0,1)\},(0,0)) = (I,(0,0))$. We know that $n(I,(0,0)) = \{\hat{s}\}$. Since n satisfies IIA, we have that $S \not\subseteq I$. That is there exists $s^* = (s_1^*, s_2^*) \in S \setminus I$. By weak Pareto optimality of f, $\hat{s}$ is a weakly efficient point of S. Therefore it cannot be the case that $s^* \gg \hat{s}$. Also, we cannot have $s^* \leq \hat{s}$ because otherwise s^* would be in I. Therefore, either $s_1^* > \hat{s}_1$ or $s_2^* > \hat{s}_2$. Assume without loss of generality that $s_1^* > \hat{s}_1$ and $s_2^* < \hat{s}_2$ (if $s_1^* > \hat{s}_1$ and $s_2^* = \hat{s}_2$, then there must be another point $s^{**} = (s_1^{**}, s_2^{**}) \in S \setminus I$, close enough to s^* with $s_1^{**} > \hat{s}_1$ and $s_2^{**} < \hat{s}_2$). Also, since any convex combination of s^* and $\hat{s}$ is in $S \setminus I$, we can choose $s^* \gg d$. We now build two bargaining problems, both of which have (s_2^*, s_2^*) as disagreement point. The first problem is $(S', (s_2^*, s_2^*))$, where $S' = \mathcal{IR}(S, (s_2^*, s_2^*))$. The second problem is the individually rational region of the triangle whose hypothenuse is the line connecting s^* and $\hat{s}$ (see Figure 1) . Formally, the problem is $(\Delta, (s_2^*, s_2^*))$ where $\Delta = \mathrm{co}\{(s_2^*, s_2^*), (s_1^*, s_2^*), (s_2^*, s_2^* + \frac{\hat{s}_2 - s_2^*}{s_1^* - \hat{s}_1}(s_1^* - s_2^*)\}$.

Figure 1 here

By disagreement point convexity and independence of non-individually rational alternatives of f, we have

$$\hat{s} = (1/2, 1/2) \in f(S', (s_2^*, s_2^*)). \tag{1}$$

Further, we claim that

$$S' \setminus \Delta \subseteq \{(s_1, s_2) \in \mathbb{R}^2 : s_1 > \hat{s}_1\} \text{ and } \Delta \setminus S' \subseteq \{(s_1, s_2) \in \mathbb{R}^2 : s_1 < \hat{s}_1\}.$$

Indeed, if there was a point $(s_1, s_2) \in S' \setminus \Delta$ with $s_1 \leq \hat{s}_1 = 1/2$, then we would have that (s_1, s_2) is above the straight line that connects $\hat{s}$ and s^*. Therefore, the line segment that connects (s_1, s_2) with s^* is also above this line. But then, there would be a point in this segment which belongs to S and which dominates $\hat{s}$, which is impossible given that $\hat{s}$ is a weakly efficient point of S. Similarly, if there was a point $(s_1, s_2) \in \Delta \setminus S'$ with $s_1 \geq \hat{s}_1$, then (s_1, s_2) would be on or below the straight line that connects $\hat{s}$ and s^*. Therefore, it would be a convex combination of $\hat{s}$, s^* and (s_2^*, s_2^*). Since the three points are in S', so would (s_1, s_2), which contradicts the fact that $(s_1, s_2) \notin S'$.

Therefore, by twisting of f we have

$$\exists(\bar{s}_1, \bar{s}_2) \in f(\Delta, (s_2^*, s_2^*))$$

such that

$$\bar{s}_1 \leq \hat{s}_1 = 1/2. \tag{2}$$

On the other hand, since $(\Delta, (s_2^*, s_2^*))$ is a triangular problem, by the first step in the proof $f(\Delta, (s_2^*, s_2^*)) = n(\Delta, (s_2^*, s_2^*))$ which implies that

$$f(\Delta, (s_2^*, s_2^*)) = \{(\bar{s}_1, \bar{s}_2)\} = n(\Delta, (s_2^*, s_2^*)).$$

By construction of Δ, the Nash solution awards player 1 in $(\Delta, (s_2^*, s_2^*))$ more than 1/2, that is

$$\bar{s}_1 > 1/2$$

which contradicts (2).

Case 2: $\hat{s} \not\gg d$**:** Again, without loss of generality assume $d = (0, 0)$. In this case either $\hat{s} = (b_1, 0)$ or $\hat{s} = (0, b_2)$. Assume without loss of generality that $\hat{s} = (0, b_2)$ with $b_2 > 0$. Pick any $\lambda \in (0, 1)$ and let $S(\lambda) = \mathcal{IR}(S, \lambda\hat{s})$. Since $\lambda\hat{s}$ is an interior point of S in the space $\mathbb{R}_+^2$, we can find a triangular set $\Delta = \text{co}\{\lambda\hat{s}, \hat{s}, (c_1, \lambda\hat{s}_2)\}$ that is contained in $S(\lambda)$. Consider now the following two bargaining problems: $(S(\lambda), \lambda\hat{s})$ and $(\Delta, \lambda\hat{s})$ (see Figure 2).

Figure 2 here

By disagreement point convexity and independence of non-individually rational alternatives $f(S(\lambda), \lambda\hat{s}) = \hat{s} = (0, b_2)$. Since $(\Delta, \lambda\hat{s})$ is a triangular problem, by the first step in the proof we have

$$f(\Delta, \lambda\hat{s}) = n(\Delta, \lambda\hat{s}) = (s_1', s_2') \gg (0, 0). \tag{3}$$

By construction, we have

$$S(\lambda) \setminus \Delta \subseteq \{(s_1, s_2) : s_1 > \hat{s}_1\} \quad \text{and} \quad \Delta \setminus S(\lambda) \subseteq \{(s_1, s_2) : s_1 < \hat{s}_1\}.$$

Therefore, by twisting we must have $s'_1 \leq \hat{s}_1 = 0$ which contradicts equation 3.

$\square$

Remark. It should be clear that the statement of the theorem still holds if we restrict attention to the family of bargaining problems (S, d) that are comprehensive with respect to d. Namely, those bargaining problems (S, d) such that if $s \geq s' \geq d$ and $s \in S$, then $s' \in S$.

4 Independence of the axioms

The following examples show that the seven axioms used in the characterization are independent. Beside each axiom there is a solution that fails to satisfy that axiom but which satisfies the other six.

Weak Pareto optimality: The disagreement point solution: $f : (S, d) \to \{d\}$.

Symmetry: Any asymmetric Nash solution.

Invariance: The Lexicographic Egalitarian solution (see, Chun and Peters (1988)).

Single-valuedness in symmetric problems: The set of weakly efficient and individually rational points.

Independence of non-individually rational alternatives: The Kalai-Rosenthal solution: it selects the maximal point of S in the segment connecting d and $b(S, d)$, where $b_i(S, d) \equiv \max\{x_i : x \in S\}$ (see Kalai and Rosenthal (1978)).

Twisting: If B can be obtained by means of a pair of affine transformations from a bargaining problem $B' = (S', d')$, where $d = (0, 0)$ and $\mathcal{IR}(B') = \text{co}\{(0, 0), (1, 0), (1/3, 2/3)\}$, then $f(B)$ is the point that is obtained by means of these transformations from $(1/3, 2/3)$. Otherwise, f coincides with the Nash bargaining solution.

Disagreement point convexity: The Kalai-Smorodinsky bargaining solution: it selects the maximal point of S in the segment connecting d and $a(S, d)$, where $a_i(S, d) \equiv \max\{x_i : x \in \mathcal{IR}(S, d)\}$ (see, Kalai and Smorodinsky (1975)).

The reader may have noticed that we could have restricted solutions to be single valued instead of imposing single-valuedness in symmetric problems as an axiom. We chose this presentation to highlight the role of single-valuedness. There are many bargaining solutions that satisfy all the axioms except for single-valuedness. As mentioned above, the set of efficient and individually rational outcomes is one example but there are many more. For instance, if f^{α} is the asymmetric Nash solution that maximizes the asymmetric Nash product $s_1^{\alpha} s_2^{1-\alpha}$, for $\alpha \in (0,1)$, then the solution that selects for every (S,d), the set $f^{\alpha}(S,d) \cup f^{1-\alpha}(S,d)$ also satisfies all the axioms except for single-valuedness. Further, it can be easily checked that if $\{f_{\gamma}\}_{\gamma \in \Gamma}$ is a family of bargaining solutions that satisfy weak Pareto optimality, symmetry, invariance, independence of non-individually rational outcomes, twisting and disagreement point convexity, then the solution $\bigcup_{\gamma \in \Gamma} f_{\gamma}$ defined by $(\bigcup_{\gamma \in \Gamma} f_{\gamma})(S,d) = \bigcup_{\gamma \in \Gamma} f_{\gamma}(S,d)$ satisfies these axioms as well. Moreover, the set of efficient and individually rational points is the maximal (in the sense of set inclusion) bargaining solution that satisfies the above axioms. It is single-valuedness in symmetric problems what allows us to select the Nash bargaining solution out of the large family of solutions that satisfy the other axioms, including symmetry.

We also should note that the axioms of independence of non-individually rational alternatives, twisting and disagreement point convexity that we use to replace IIA, do not imply the independence of irrelevant alternatives axiom: the solution that selects the disagreement point if the feasible set is a line segment and the Nash outcome otherwise, satisfies all the three axioms (in fact, satisfies all the axioms except for weak Pareto optimality) but does not satisfy IIA.

5 Related Literature

This paper provides a characterization of the Nash bargaining solution on Nash's original domain of bargaining problems, and in which the independence axiom is replaced by three other axioms. Our result is closely related to Peters and van Damme (1991) and our contribution can be seen as eliminating of continuity axioms from the characterization. Continuity has been replaced by twisting, a mild axiom that, to our knowledge, is satisfied by most solution concepts discussed in the literature (the Perles-Maschler solution is one exception). Other characterizations of the Nash solution that use similar axioms, but still need continuity, are Peters (1986b) and Chun and Thomson (1990). Mariotti (1999) also provides a characterization of the Nash solution without appealing to IIA, but, as opposed to the other mentioned papers, he reduces the number of axioms. In fact, there are only two characterizing axioms: invariance and Suppes-Sen proofness. The same can be said about Mariotti (2000) who replaces IIA and symmetry by strong individual rationality and the axiom of Maximal Symmetry.

Chun and Thomson (1990) characterize the Nash bargaining solution using, along with Pareto optimality, symmetry, scale-invariance, independence of non-individually rational outcomes, and a continuity axiom. The two axioms, which capture features of bargaining with uncertain disagreement points can be stated as follows:[7]

R.D.LIN.: A single-valued bargaining solution f satisfies *restricted disagreement point linearity* if for every two problems (S, d) and (S, d'), and for all $\alpha \in [0, 1]$, if $\alpha f(S, d) + (1-\alpha)f(S, d')$ is efficient and S is smooth both at $f(S, d)$ and $f(S, d')$, then $f(S, \alpha d + (1 - \alpha)d') = \alpha f(S, d) + (1 - \alpha)f(S, d')$.

D.Q-CAV.: A single-valued bargaining solution f satisfies *disagreement point quasi-concavity* if for every two problems (S, d) and (S, d'), and for all $\alpha \in [0, 1]$, $f_i(S, \alpha d + (1 - \alpha)d') \geq \min\{f_i(S, d), f_i(S, d')\}$ for $i = 1, 2$.

We now investigate the relation between these two axioms and disagreement point convexity.

Claim 1 If a single-valued bargaining solution, f, satisfies Pareto optimality, independence of non-individually rational alternatives and D.Q-CAV., then it also satisfies disagreement point convexity.

Proof : Let (S, d) be a bargaining problem and let $s = f(S, d)$. Let $\lambda \in (0, 1)$ and assume that $f(S, (1-\lambda)d+\lambda s) \neq s$. Since f satisfies Pareto optimality, $f_i(S, d) > f_i(S, (1-\lambda)d+\lambda s)$ for some $i = 1, 2$, which, without loss of generality, can be taken to be agent 1. Therefore, we can find an $\alpha \in (0, 1)$ close enough to 1 such that the point $d' = (1-\alpha)d+\alpha s$ satisfies $d'_1 > f_1(S, (1-\lambda)d+\lambda s)$. Since f satisfies individual rationality, $f_1(S, d') > f_1(S, (1 - \lambda)d + \lambda s)$. This inequality, together with $f_1(S, d) > f_1(S, (1 - \lambda)d + \lambda s)$ imply $\min\{f_1(S, d), f_1(S, d')\} > f_1(S, (1 - \lambda)d + \lambda s)$. By the way d' was chosen, we know that $(1 - \lambda)d + \lambda s$ is a convex combination of d and d' and consequently the above inequality implies that f does not satisfy D.Q-CAV. □

As a corollary, we have that we could replace weak Pareto optimality and disagreement point convexity in our characterization by Pareto optimality and D.Q-CAV.

The relationship between disagreement point convexity and R.D.LIN. is not so clear, at least within the domain of problems considered in this paper. However, Pareto optimality, independence of non-individually rational alternatives and R.D.LIN. imply disagreement point convexity within the domain of bargaining problems with smooth Pareto frontiers provided we enlarge the

[7] Chun and Thomson (1990) define bargaining solutions as single-valued functions that select points from the set of feasible utilities. To facilitate comparison in what remains of this section, we use the single-valued versions of the axioms, including disagreement point convexity.

definition of bargaining problems to include those pairs (S,d) with efficient disagreement points.[8] To see this, consider a bargaining problem (S,d) in this domain and let f be a bargaining solution that satisfies Pareto optimality, independence of non-individually rational alternatives and R.D.LIN. By Pareto optimality, we have that $f(S,d)$ is efficient. By independence of non-individually rational alternatives, we have that $f(S,f(S,d)) = f(S,d)$. Since the efficient frontier is smooth, we can apply R.D.LIN. to conclude that $f(S,(1-\lambda)d + \lambda f(S,d)) = f(S,d)$ for all $\lambda \in (0,1)$. This means that f satisfies disagreement point convexity.

The Nash solution is not defined for the above domain. However, one can extend it, as Peters and van Damme (1991) do, so as to select the only efficient and individually rational point when the disagreement point is weakly efficient. In this case, our characterization goes through and the axioms of weak Pareto optimality and disagreement point convexity can, as a corollary of the observation of the previous paragraph, be replaced by Pareto optimality and R.D.LIN.

Our characterization is on Nash's original domain. In particular, we restrict attention to two-person bargaining problems. It is not clear whether the same axioms are sufficient to fully characterize the Nash bargaining solution for general n-person bargaining problems. The Nash bargaining solution does satisfy all the axioms. However, our proof makes use of the 2-dimensionality of the problem. In particular, when there are more than 3 players, it is not clear how to build the auxiliary set Δ with the critical properties used in step 2 of our proof.

Appendix

In this Appendix we show that the set valued version of the independence of irrelevant alternatives axiom that we use implies twisting. Formally:

Claim 2 If a bargaining solution satisfies independence of irrelevant alternatives, then it also satisfies twisting.

Proof : Let (S,d) be a bargaining problem and let $\hat{s} \in f(S,d)$. Let (S',d) be another bargaining problem such that for some agent $i = 1,2$

$$S \setminus S' \subseteq \{(s_1,s_2) : s_i > \hat{s}_i\} \tag{4}$$

$$S' \setminus S \subseteq \{(s_1,s_2) : s_i < \hat{s}_i\}. \tag{5}$$

[8] Peters and van Damme (1991) consider a domain of problems that contains pairs (S,d) where d is an efficient point of S.

We need to show that there is $(s'_1, s'_2) \in f(S', d)$ such that $s'_i \leq \hat{s}_i$. Assume now by contradiction that

$$f(S', d) \subseteq \{(s_1, s_2) : s_i > \hat{s}_i\} \tag{6}$$

and let $\hat{S} = S \cap S'$. Since $\hat{s} \in f(S, d) \cap \hat{S}$, we have by IIA that

$$\hat{s} \in f(\hat{S}, d). \tag{7}$$

Further, $f(S', d) \cap S \neq \emptyset$, for if $f(S', d) \subseteq S' \setminus S$, then by (5), $f(S', d) \subseteq \{(s_1, s_2) : s_i < \hat{s}_i\}$ which was assumed in (6) not to be true. Therefore, $\emptyset \neq f(S', d) \cap S \subseteq S' \cap S = \hat{S}$. This implies that $f(S', d) \cap \hat{S} \neq \emptyset$ and $\hat{S} \subseteq S'$. Then, by IIA $f(\hat{S}, d) = f(S', d) \cap \hat{S}$. But then, since by (7), $\hat{s} \in f(\hat{S}, d)$, we have that $\hat{s} \in f(S', d)$, which by (6) implies that $\hat{s}_i > \hat{s}_i$, which is absurd. $\square$

References

Binmore, K. G., A. Rubinstein, and A. Wolinsky (1986). The Nash bargaining solution in economic modeling. *Rand Journal of Economics 17*, 176–188.

Chun, Y. and H. Peters (1988). The lexicographic egalitarian solution. *Cahiers du CERO 30*, 149–156.

Chun, Y. and W. Thomson (1990). Nash solution and uncertain disagreement points. *Games and Economic Behavior 2*, 213–223.

Dagan, N., O. Volij, and E. Winter (1999). The temporal Nash solution. Unpublished manuscript: Department of Economics, Hebrew University.

Kalai, E. and R. W. Rosenthal (1978). Arbitration of two-party disputes under ignorance. *International Journal of Game Theory 7*, 65–72.

Kalai, E. and M. Smorodinsky (1975). Other solutions to Nash's bargaining problem. *Econometrica 43*, 513–518.

Lensberg, T. (1988). Stability and the Nash solution. *Journal of Economic Theory 45*, 330–341.

Mariotti, M. (1999). Fair bargains: Distributive justice and Nash bargaining theory. *Review of Economic Studies 66*, 733–741.

Mariotti, M. (2000). Maximal symmetry and the Nash solution. *Social Choice and Welfare 17*, 45–53.

Nash, J. F. (1950). The bargaining problem. *Econometrica 28*, 155–162.

Perles, M. and M. Maschler (1981). A superadditive solution to Nash bargaining games. *International Journal of Game Theory 10*, 163–193.

Peters, H. (1986a). Characterizations of bargaining solutions by properties of their status quo. Research Memorandum 86–012, University of Limburg.

Peters, H. (1986b). Simultaneity of issues and additivity in bargaining. *Econometrica 54*, 153–169.

Peters, H. and E. van Damme (1991). Characterizing the Nash and Raiffa bargaining solutions by disagreement point axioms. *Mathematics of Operations Research 16*(3), 447–461.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica 50*, 97–110.

Thomson, W. (1994). Cooperative models of bargaining. In R. J. Aumann and S. Hart (Eds.), *Handbook of Game Theory*, Volume 2, pp. 1237 – 1284. North–Holland.

Thomson, W. and R. B. Myerson (1980). Monotonicity and independence axioms. *International Journal of Game Theory 9*, 37–49.