Nash’s bargaining solution when the disagreement point is random

Rann Smorodinsky*

Faculty of Industrial Engineering and Management Technion—Israel Institute of Technology, Haifa 32000, Israel

Received 1 January 2003; received in revised form 1 November 2004; accepted 1 February 2005
Available online 19 April 2005

Abstract

In his seminal work, Nash (1950) derives a solution for two-person bargaining problems, within a cooperative setup. Nash assumes that the result of disagreement is known to both players and is not stochastic. We study the same problem, where the last assumption is relaxed. We provide a set of axioms which characterizes a natural generalization of the Nash solution to bargaining problems with a random point of disagreement.

Keywords: Bargaining; Nash solution; Random disagreement point

JEL classification: C71; D74

1. Introduction

In a seminal work, Nash (1950) studies 2-person bargaining problems via an axiomatic model. A bargaining problem is composed of two elements: (1) the possible set of outcomes that the bargaining parties can potentially agree on, and (2) the result of disagreement (so-called disagreement point). Using a set of four axioms, Nash finds a unique solution to the problem. Nash’s work induced a large volume of papers which characterized various solutions adhering with different sets of axioms.

*Tel.: +972 4 8294422; fax: +972 4 8235194.
E-mail address: rann@ie.technion.ac.il.

0165-4896/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
Random disagreement points were later introduced by various authors (e.g., Livne, 1988; Chun and Thomson, 1990b; Peters and van Damme, 1991). Random disagreement points can be interpreted in two ways. One interpretation is that players must agree on some outcome or otherwise receive the realization of a random disagreement point. Assuming the Expected Utility paradigm, the modeler may replace the random disagreement point with its expectation. Thus, one can model a bargaining problem as Nash originally did, namely as a pair composed of a set in $\mathbb{R}^2$ and a disagreement point in $\mathbb{R}^2$. This is exactly the approach taken by Chun and Thomson (1990b) and Peters and van Damme (1991).1

However, there exists another interpretation, which is the one taken by Livne (1988) and is also the interpretation we pursue. Assume players may choose a point of agreement or otherwise be notified of the realized disagreement point, and then get a second chance of agreeing. If we assume that the set of feasible agreement points has a convex Pareto frontier then expected utility maximizing players have an incentive to decide at the first stage, rather than wait for the realized bargaining problem. With this interpretation it is not obvious that a random disagreement point should be replaced with its expectation. Rather, the natural modelling choice is to consider a richer class of bargaining problems, namely all pairs consisting of sets of alternatives in $\mathbb{R}^2$ and a random variable in $\mathbb{R}^2$.

In this paper we indeed begin with the richer class of bargaining pairs while maintaining all four axioms originally proposed by Nash. To these we add a continuity axiom as well as the following two axioms, which are in context of the more general framework of random disagreement points. The axioms we add are: (1) Betweeness—the solution for a random disagreement point cannot be better, for any player, than all solutions associated with the certain disagreement points in its support; and (2) Replaceability—it is possible to replace a random disagreement point with a certain disagreement point (not necessarily the expectation). We show that these seven axioms characterize a unique solution which generalizes the Nash solution in the most natural way.

Section 2 provides the model and axioms. Section 3 states and proves the main results and discusses the two key additional axioms we introduce (Betweeness and Replaceability). Finally, Section 4 discusses related literature.

2. The model

A two-person bargaining problem is formulated as a pair $(\tilde{a}, S)$, where $S$ is a subset of $\mathbb{R}^2$ and $\tilde{a} \in \tilde{D}$, the set of all random variables with values in $\mathbb{R}^2$. The pair $(\tilde{a}, S)$ has the following intuitive interpretation: for any $a = (a_1, a_2)$ in the support of $\tilde{a}$, denoted $\text{Supp}(\tilde{a})$, $a_i$ is the level of utility for player $i$ which is received if the players do not cooperate and if $a$ is the realization of $\tilde{a}$. Every $s = (s_1, s_2) \in S$ represents levels of utility for players 1 and 2, respectively, that can be reached by an outcome of the game which is feasible for both players when they do cooperate.

---

1The reader is referred to Safra and Zilcha (1993), among others, for an explicit treatment of the bargaining problem without the expected utility hypothesis, and to Rubinstein et al. (1992) for a discussion of bargaining theory based on ordinal preferences.
We begin with some traditional assumptions regarding the pair \((\tilde{a}, S)\):

**Assumption 1.** \(S\) is a nonempty compact and convex set.

**Assumption 2.** \(\tilde{a}\) has a finite support, contained in \(S\).

We call a pair \((\tilde{a}, S)\) satisfying the above assumptions a bargaining problem and we denote by \(U\) the class of all bargaining problems.\(^2\) A bargaining solution is a function \(\phi : U \to \mathbb{R}^2\) satisfying \(\phi(\tilde{a}, S) \subseteq S\).

For any one-to-one and onto function \(H : \mathbb{R}^2 \to \mathbb{R}^2\) (similarly for \(H : \mathbb{R}^2 \to \mathbb{R}\)) and any set \(A \subseteq \mathbb{R}^2\) we denote \(H(A) = \{H(x) | x \in A\}\). For any two dimensional random variable \(\tilde{a}\) we denote by \(H(\tilde{a})\) the random variable satisfying \(\text{Prob}(H(\tilde{a}) \leq (a, b)) = \text{Prob}(\tilde{a} \leq H^{-1}(a, b))\) for any \((a, b) \in \mathbb{R}^2\).

We consider the following axioms on \(\phi\). The first four axioms are straightforward extensions of the axioms used by Nash.

**Axiom 1 (Pareto optimality).** For every \((\tilde{a}, S)\) there is no \(y \in S\) such that \(y \geq \phi(\tilde{a}, S)\) (i.e., \(y_i \geq \phi_i(\tilde{a}, S)\) for \(i = 1, 2\)) and \(y \neq \phi(\tilde{a}, S)\).

**Axiom 2 (Symmetry).** Let \(T : \mathbb{R}^2 \to \mathbb{R}^2\) be defined by \(T(x, y) = (y, x)\). For every \((\tilde{a}, S) \in U\) \(\phi(T(\tilde{a}), T(S)) = T(\phi(\tilde{a}, S))\).\(^3\)

**Axiom 3 (Invariance with respect to Positive Affine Transformations).** For any affine transformation \(A : \mathbb{R}^2 \to \mathbb{R}^2\) (i.e., \(A = (A_1, A_2)\) and \(A_i : \mathbb{R} \to \mathbb{R}\) is defined by \(A_i(x) = c_i x + d_i\) with \(c_i, d_i \in \mathbb{R}, c_i > 0\), \(\phi A(\tilde{a}), A(S)) = A(\phi(\tilde{a}, S))\).\(^4\)

**Axiom 4 (Independence of irrelevant alternatives).** If \((\tilde{a}, S_1), (\tilde{a}, S_2) \in U\), \(S_1 \subset S_2\) and \(\phi(\tilde{a}, S_1) \subseteq S_1\) then \(\phi(\tilde{a}, S_1) = \phi(\tilde{a}, S_2)\).

In addition to Nash’s axioms we consider axioms which treat the randomness of the disagreement point. For these axioms we introduce the following notation: For \(x \in [0, 1]\) \(x\tilde{a} \oplus (1 - x)\tilde{b}\) denotes the compound two-stage lottery, where an \((x, 1 - x)\) coin is used at the first stage and \(\tilde{a}\) or \(\tilde{b}\) is then used accordingly.

**Axiom 5 (Continuity).** Let \(\{(\tilde{a}_n, S)\}_{n=0}^\infty \subseteq U\) be an infinite sequence of bargaining problems. If \(\tilde{a}_n\) converges to \(\tilde{a}\) in the weak topology, then \(\phi(\tilde{a}_n, S) \rightarrow \phi(\tilde{a}, S)\).\(^5\)

**Axiom 6 (Betweeness).** Let \(\phi_i\) denote player \(i\)'s outcome in the solution. If \(\phi_i(\tilde{a}, S) < \phi_i(\tilde{b}, S)\) then \(\forall \ x \in (0, 1) \ \phi_i(\tilde{a}, S) < x_i(\tilde{a} \oplus (1 - x)\tilde{b}) < \phi_i(\tilde{b}, S)\).\(^6\)

The betweenness axiom excludes a situation where one of the players would surely prefer to wait for the realization of the disagreement point before cooperating with his opponent.

---

\(^2\)The traditional approach furthermore assumes that the support of \(\tilde{a}\) is a singleton.

\(^3\) Note that if \((\tilde{a}, S) \in U\) then \((T(\tilde{a}), T(S)) \in U\).

\(^4\) Note that if \((\tilde{a}, S) \in U\) then \((A(\tilde{a}), A(S)) \in U\).

\(^5\) We say that \(\tilde{a}_n\) converges to \(\tilde{a}\) in the weak topology if for every bounded continuous function, \(\rho : \mathbb{R}^2 \to \mathbb{R}\), the sequence of expectations, \(E(\rho(\tilde{a}_n))\), converges to \(E(\rho(\tilde{a}))\).

\(^6\) We borrow terminology from the decision theoretic literature (e.g., Coombs and Huang, 1976; Dekel, 1986; Chew, 1989).
The following axiom states that the bargaining problem \((\mathbf{\bar{a}}, S)\) is, conceptually, resolved by the players in two stages. In the first stage, the players replace the random disagreement point with a (virtual) fixed point and in the second stage they resolve the new bargaining problem.

**Axiom 7** (Replaceability of random disagreement point). There exists a function \(\mathcal{R} : \hat{D} \rightarrow \mathbb{R}^2\) such that for all \((\mathbf{\bar{a}}, S) \in U, \phi(\mathbf{\bar{a}}, S) = \phi(\mathcal{R}(\mathbf{\bar{a}}), S)\) and \((\mathcal{R}(\mathbf{\bar{a}}), S) \in U\).

By applying Axiom 7 to the case where \(S\) is actually the convex hull of \(\mathbf{\bar{a}}\) we may conclude:

**Lemma 1.** Axiom 7 implies that \(\mathcal{R}(\mathbf{\bar{a}})\) must be a point in the convex hull of \(\mathbf{\bar{a}}\).

### 3. Results

The statement of the main result makes use of Nash’s (1950) original solution for the class of bargaining problems with a fixed disagreement point:

\[
N(a, S) = \arg \max_{x_1 \geq a_1, x_2 \geq a_2, (x_1, x_2) \in S} (x_1 - a_1)(x_2 - a_2).
\]

**Theorem A.** The unique solution \(\phi\) satisfying Axioms 1–7 is: \(\phi(\mathbf{\bar{a}}, S) = N(E(\mathbf{\bar{a}}), S), \forall(\mathbf{\bar{a}}, S) \in U\), where \(E(\cdot)\) denotes the expectation operator.

Note that if we restrict attention to bargaining problems with a fixed disagreement point then we are back to Nash’s original model. Therefore, for any fixed point, \(a \in \mathbb{R}^2\) we must have \(\phi(a, S) = N(a, S)\). The next lemma then follows:

**Lemma 2.** Axioms 1–4 imply that for any two fixed disagreement points \(a \neq b \in \mathbb{R}^2\) there exists some set \(S \subset \mathbb{R}^2\) such that \(a, b \in S, (a, S) \in U, (b, S) \in U\) and \(\phi(a, S) \neq \phi(b, S)\).

#### 3.1. The replaceable disagreement point

We begin the first step by studying the properties of the replaceable disagreement point, defined as a function \(\mathcal{R} : \hat{D} \rightarrow \mathbb{R}^2\) (see Axiom 7), as implied by Axioms 1–6.

**Lemma 3.** Axioms 1–4 and 7 imply that for any affine transformation \(A : \mathbb{R}^2 \rightarrow \mathbb{R}^2\), and any \(\mathbf{\bar{a}} \in \hat{D}, R(A(\mathbf{\bar{a}})) = A(\mathcal{R}(\mathbf{\bar{a}}))\).

**Proof.** Assume this is not true. Then there exists some affine transformation \(A\) and some \(\mathbf{\bar{a}} \in \hat{D}\) such that \(\mathcal{R}(A(\mathbf{\bar{a}})) \neq A(\mathcal{R}(\mathbf{\bar{a}}))\). By Lemma 2 we may choose a set, \(S\), satisfying \(\mathcal{R}(A(\mathbf{\bar{a}})) \in S, A(\mathcal{R}(\mathbf{\bar{a}})) \in S, (\mathcal{R}(A(\mathbf{\bar{a}})), S) \in U, (A(\mathcal{R}(\mathbf{\bar{a}})), S) \in U\) and \(\phi(A(\mathbf{\bar{a}}), S) \neq \phi(A(\mathcal{R}(\mathbf{\bar{a}})), S)\). Let \(S'\) be such that \(A(S') = S\) (such a set, \(S'\), exists as \(A\) is affine). We conclude that \(\phi(A(\mathbf{\bar{a}}), A(S')) = \phi(A(\mathcal{R}(\mathbf{\bar{a}})), A(S')) = A(\phi(\mathcal{R}(\mathbf{\bar{a}}), S')) = A(\phi(\mathbf{\bar{a}}, S'))\), which contradicts Axiom 3. \(\square\)

---

\(^7\)This kind of replaceability is implicitly assumed in Chun and Thompson (1990a,b) and Peters and van Damme (1991). In fact, they assume that the random disagreement point can be replaced by its expectation.
Lemma 4. Let \( \{\tilde{a}_n\}_{n=1}^{\infty} \) be a sequence of random variables in \( \hat{D} \) that converges to some \( \tilde{a} \in \hat{D} \), in the weak topology. Then, Axioms 1–5 and 7 imply that \( \lim_{n \to \infty} R(\tilde{a}_n) = R(\tilde{a}) \).

**Proof.** Assume this is not the case. Namely, \( \tilde{a}_n \to \tilde{a} \) and \( \lim_{n \to \infty} (R(\tilde{a}_n)) \neq R(\tilde{a}) \) (without loss of generality we assume \( \lim_{n \to \infty} R(\tilde{a}_n) \) exists. Otherwise, we use an arbitrary convergent subsequence). In the spirit of the proof of Lemma 3, let \( S \) be the set satisfying the argument of Lemma 2 for the fixed disagreement points \( \lim R(\tilde{a}_n) \) and \( R(\tilde{a}) \). Assume, without loss of generality that \( \lim \phi(\tilde{a}_n, S) \) exists (otherwise take a subsequence). Now \( \lim \phi(\tilde{a}_n, S) = \lim \phi(R(\tilde{a}_n), S) = \phi(\lim R(\tilde{a}_n), S) \neq \phi(R(\tilde{a}), S) = \phi(\tilde{a}, S) \), which contradicts Axiom 5. \( \square \)

For any \( \tilde{a} \in \hat{D} \), we denote by \( \tilde{a}_x \) and \( \tilde{a}_y \) the marginals on the \( X \)-axis and \( Y \)-axis, respectively. Also, for any real random variables \( \tilde{a}_1 \) and \( \tilde{a}_2 \) we denote by \( \langle \tilde{a}_1, \tilde{a}_2 \rangle \) the \( \mathbb{R}^2 \) valued random variable with independent marginals \( \tilde{a}_1 \) and \( \tilde{a}_2 \).

Similarly we denote by \( R_x \) and \( R_y \) the projection of \( R \) onto the \( X \)-axis and the \( Y \)-axis, correspondingly.

Corollary 1. Axioms 1–5 and 7 imply that \( R_x \) and \( R_y \) are invariant with respect to affine transformations and are continuous with respect to the weak topology.

**Proof.** Follows directly from Lemmas 3, 4 and the definitions of \( R_x \) and \( R_y \). \( \square \)

Lemma 5. Axioms 1–5 and 7 imply that the function \( R : \hat{D} \to \mathbb{R}^2 \) satisfies \( R_x(\tilde{a}) = R_x(\tilde{a}_x, 0) \) and similarly \( R_y(\tilde{a}) = R_y(0, \tilde{a}_y) \).

In words, the value of the replaceable disagreement point of \( \tilde{a} \) is determined by the marginals of \( \tilde{a} \) on the axis.

**Proof.** Let \( \tilde{a} \) be an arbitrary element in \( \hat{D} \) and set \( f_n(x, y) = (x, y/n) \). By Lemmas 3 and 4 \( \lim f_n(R(\tilde{a})) = \lim f_n(R(\tilde{a}_x, \tilde{a}_y)) = \lim R(f_n(\tilde{a}_x, \tilde{a}_y)) = \lim R(\tilde{a}_x, \tilde{a}_y/n) = \lim R(\tilde{a}_x, 0) = R(\tilde{a}_x, 0) \). Now take the first coordinate in the last equation and conclude that \( R_x(\tilde{a}) = R_x(\tilde{a}_x, 0) \). A similar proof holds for the \( Y \)-axis. \( \square \)

Note that we can now conclude that the support of \( R(\tilde{a}_x, 0) \) \( \left( R_y(0, \tilde{a}_y) \right) \) must be on the \( X \)-axis (\( Y \)-axis).

Lemma 6. Axioms 1–5 and 7 imply that \( R_x : \hat{D} \to \mathbb{R} \) (and similarly \( R_y : \hat{D} \to \mathbb{R} \)) satisfies \( R_x(-\tilde{a}) = -R_x(\tilde{a}) \).

**Proof.** Given the result of Lemma 5 it is sufficient to prove that \( -R_x(\tilde{a}_x, 0) = R_x(-\tilde{a}_x, 0) \). In fact, we assume, without loss of generality, that \( R_x(\tilde{a}_x, 0) = 0 \), and prove that \( R_x(-\tilde{a}_x, 0) = 0 \) as well.

By Assumptions 1, 2 and the Replaceability Axiom we deduce that \( R(\tilde{a}) \) is in the convex hull of the support of \( \tilde{a} \). Consider the random threat point, \( \tilde{a} \), of which the support is on the line \( X+Y=0 \) and its marginal on the \( X \)-axis is \( \tilde{a}_x \). Note that its marginal on the
Lemma 7. Axioms 1–7 imply that if $\mathcal{R}_x(\tilde{a}) > \mathcal{R}_x(\tilde{b})$ then $\mathcal{R}_x(\tilde{a}) > \mathcal{R}_x(z\tilde{a} \oplus (1 - z)\tilde{b}) > \mathcal{R}_x(\tilde{b})$ for any $0 < z < 1$.

**Proof.** Without loss of generality we will assume the support of the distributions $\tilde{a}$ and $\tilde{b}$ are on the $X$-axis. Let $\tilde{a}'$ and $\tilde{b}'$ denote two random disagreement points which marginal on the $X$-axis is $\tilde{a}$ and $\tilde{b}$, respectively, and of which the support is on the line $x + y = 0$. Let $S$ be the compact convex hull of the support of $\tilde{a}'$ and $\tilde{b}'$. Axiom 1 (Pareto optimality) implies that $\phi(\tilde{a}', S) = \mathcal{R}(\tilde{a}')$ and $\phi(\tilde{b}', S) = \mathcal{R}(\tilde{b}')$. Similarly $\phi(x\tilde{a}' \oplus (1 - z)\tilde{b}', S) = \mathcal{R}(x\tilde{a}' \oplus (1 - z)\tilde{b}')$. We can conclude that $\mathcal{R}_x(\tilde{a}') > \mathcal{R}_x(x\tilde{a}' \oplus (1 - z)\tilde{b}') > \mathcal{R}_x(\tilde{b}')$.

The desired conclusion follows directly from the last inequality and the fact that $\mathcal{R}_x(\tilde{a}) = \mathcal{R}_x(\tilde{a}')$, $\mathcal{R}_x(x\tilde{a}' \oplus (1 - z)\tilde{b}') = \mathcal{R}_x(x\tilde{a} \oplus (1 - z)\tilde{b})$ and $\mathcal{R}_x(\tilde{b}) = \mathcal{R}_x(\tilde{b}')$.

We can now use the main result of Smorodinsky (2000), which states:

**Lemma 8.** Let $D$ be the set of all bounded real valued random variables. A function $g : D \rightarrow \mathbb{R}$ satisfies the following 4 properties

1. (P1) If $\tilde{a}_n \rightarrow \tilde{a}$ (in the weak topology) then $g(\tilde{a}_n) \rightarrow g(\tilde{a})$.
2. (P2) $g(c \cdot \tilde{a} + b) = c \cdot g(\tilde{a}) + b$ for any $c, b \in \mathbb{R}$, where $c > 0$.
3. (P3) $g(-\tilde{a}) = -g(\tilde{a})$.
4. (P4) $g(\tilde{a}) > g(\tilde{b}) \Rightarrow g(\tilde{a}) > g(x \tilde{a} \oplus (1 - z)\tilde{b}) > g(\tilde{b})$ for any $0 < z < 1$.

If and only if $g(\tilde{a}) = \text{arg min}_t E|\tilde{a} - t|^c$ for some $c > 0$.

Note that for the case $c = 2$, $g(\tilde{a}) = \text{arg min}_t E|\tilde{a} - t|^2 = E(\tilde{a})$, namely the above equals the expectation.

**Lemma 9.** There exist $c_x, c_y \in (0, \infty)$ s.t.

$$\mathcal{R}_x(\tilde{a}_x, 0) = \text{arg min}_t E|\tilde{a}_x - t|^{c_x} \text{ and } \mathcal{R}_y(0, \tilde{a}_y) = \text{arg min}_t E|\tilde{a}_y - t|^{c_y}.$$

**Proof.** Follows directly from Corollary 1 and Lemmas 6, 7 and 8. □

The following 2 examples provide additional information on the value of $c_x$ and $c_y$:

**Example 1.** Let $S$ be the triangle with vertices $(0, 0)$, $(1, 0)$ and $(0, 1)$. Let $\tilde{a}$ assign probability $p$ to $(1/2, 1/2)$ and $1 - p$ to $(0, 0)$. As

$$\phi((0, 0), S) = \phi\left(\left(\frac{1}{2}, \frac{1}{2}\right), S\right) = N((0, 0), S), \mathcal{R}(\tilde{a})$$

must be on the line connecting $(0, 0)$ and $(1/2, 1/2)$ for any $p$. 


We now conclude that $c_x = c_y$.

**Example 2.** Let $S = \{(x, y) | x, y \geq -1/2, x + y \leq 1\}$ and let $\tilde{a}$ be the random disagreement point assigning probabilities 1/2 to (1/2, 1/2), 1/4 to (1/2, 0) and 1/4 to (−1/2, 0). Let $\tilde{b}$ assign probability 1 to (1/2, 1/2) and let $\tilde{c}$ assign probability 1/2 to (1/2, 0) and 1/2 to (−1/2, 0). Independently of the values of $c_x$ and $c_y$, $\varphi(\tilde{b}, S) = \varphi(\tilde{c}, S) = (1/2, 1/2)$. It now follows that $\mathcal{R}(\tilde{a})$ is of the form $(x, x)$. By the symmetry of the distribution of $\tilde{a}_y$ around 1/4 we know that for any value of $c_y$, $\mathcal{R}_y(0, \tilde{a}_y) = 1/4$. This in turn implies that $\mathcal{R}_x(\tilde{a}_x, 0) = 1/4$.

Note that the unique value of $c_x$ that satisfies $\mathcal{R}_x(\tilde{a}_x, 0) = 1/4$ is $c_x = 2$ and consequently $c_y = 2$ as well. Thus, we have just proved:

**Corollary 2.** $\mathcal{R}(\tilde{a}) = E(\tilde{a})$.

### 3.2. Proof of main result

We can now conclude and prove the main result:

**Proof of Theorem A.** It is easy to verify that $N(E(\tilde{a}), S)$ indeed satisfies Axioms 1–7. Suppose $\varphi$ is a solution to $U$ satisfying all these axioms. By Axioms 7 and Corollary 2 $\varphi(\tilde{a}, S) = \varphi(E(\tilde{a}), S)$ on $U$. As $E(\tilde{a})$ is a fixed disagreement point we conclude from the main result in Nash (1950) that $\varphi(E(\tilde{a}), S) = N(E(\tilde{a}), S)$ on $U$. 

### 3.3. Independence of axioms

It is quite obvious that Axioms 1–4 cannot be omitted. To see this, consider Nash’s original setting. For each of these axioms there exists a solution, different from Nash’s solution, that satisfies the other three axioms. By applying this to the expected disagreement point one easily generates a solution, in the current setting, that satisfies the other six axioms. The difficulty is, therefore, providing examples for Axioms 5–7. Below are alternative solutions when either of the two latter axioms is dropped. Whether Axioms 5 can be dropped is an open question we leave open.

**Example 3** (Omitting betweeness). Let $f : \hat{D} \rightarrow \mathbb{R}^2$ be an arbitrary operator, different than the expectation, such that (a) $f(\tilde{a})$ is in the convex hull of the support of $\tilde{a}$; (b) $f$ is continuous with respect to the weak topology; (c) $f$ is invariant with respect to affine transformations; and (d) $f$ is symmetric. Now, set $u^6(\tilde{a}, S) = \arg \max_{(x, y) \in \mathcal{P}(S)} \Sigma p_i \log(x - a_i)(y - b_i)$, where $(a_i, b_i)$ are the

$E.g., f(\tilde{a}) = \arg \max_{(a, b) \in \hat{D}} E(\tilde{a}_x - a)^2 + |\tilde{a}_y - b|$. Continuity and invariance of $f(\tilde{a})$ are straightforward. Proving that the image of $f(\tilde{a})$ is in the convex hull of the support $\tilde{a}$ follows from first order conditions. We leave the full proof to the reader. We thank Ron Holzman for this.
atoms of $\tilde{a}$ and $p_i$ are the respective probabilities. We leave it to the reader to check that $\varphi^7(\tilde{a}, S) \neq N(E(\tilde{a}), S)$ (note that for fixed disagreement points $\varphi^7(\tilde{a}, S) = N(a, S)$), yet it satisfies Axioms 1–6, which proves lack of uniqueness.

4. Related literature

Random disagreement points have been studied in the literature and various solution concepts have been characterized. The closest papers to ours are Livne (1988), Chun and Thomson (1990b) and Peters and van Damme (1991).

Livne (1988) looks at a model where players may get a signal relating to the random disagreement point. Players can consequently agree on a mapping from the set of possible signals to points in $S$. The setup we use is similar to that of Livne, namely, a bargaining pair is composed as a set in $\mathbb{R}^2$ coupled with a random variable in $\mathbb{R}^2$. Livne provides axioms that yield the following extension of Nash’s original solution—first compute the expected value of the Nash solution for each point in the support of the random disagreement point. Then use this point as a new disagreement point and compute the Nash solution for the new problem. Apparently, this is a different solution than what we suggest, namely applying the Nash solution to the expectation of the disagreement point. In particular, Livne’s solution does not satisfy the ‘Independence of Irrelevant Alternatives’ (IIA) axiom.

Of all four axioms proposed by Nash, the axiom relating to IIA is the most controversial (see Kalai and Smorodinsky, 1975). Much of the follow up literature concentrated on replacing the IIA axiom with other axioms (e.g., Kalai and Smorodinsky, 1975; Kalai, 1977; Roth, 1979; Perles and Maschler, 1981). Consequently these papers provide solutions that are different from Nash’s original solution. The papers of Chun and Thomson (1990b) and of Peters and van Damme (1991), however, use the random setup to provide a characterization of Nash’s original solution without the IIA axiom.

As mentioned in the introduction these papers stem from a different interpretation of the disagreement point and consequently assume a random disagreement point can be replaced with its expectation and so they work within Nash’s original framework of bargaining pairs consisting of sets in $\mathbb{R}^2$ and certain disagreement points.

Chun and Thomson’s (1990b) main axiom is called the Restricted Disagreement Point Linearity (R.D.Lin). It compares between a bargaining problem at stage one, where the disagreement point is a convex combination of two disagreement points, $d^1$ and $d^2$, and the two possible realizations at stage 2, where the disagreement point is either $d^1$ or $d^2$. The axiom states that if the expectation of the solution at stage 2 is Pareto optimal, then it must be the solution to the problem at stage one. This axiom, together with continuity of the solution with respect to minor changes in the set of possible agreements, is sufficient in replacing Nash’s IIA axiom. In fact, for the two player case Chun and Thomson show that the result can be obtained by replacing R.D.Lin with a weaker one, the Disagreement Point Quasi-Concavity (D.Q.- Cav) axiom. Apparently, D.Q.-Cav is equivalent to our Betweeness axiom.

Peters and van Damme (1991) take quite a similar approach. The convexity assumption on the set of disagreement points they introduce is called Disagreement Point convexity
(DVEX). It asserts that when the set of possible agreements is held fixed and a new disagreement point is the result of a lottery between a preliminary fixed disagreement point and the corresponding solution, then the solution to the new problem must be the same as the solution to the preliminary problem. This axiom, together with continuity of the solution, this time with respect to changes in disagreement point, is the driving force behind the characterization of the Nash solution.

We finish by reminding the reader that the one cannot take for granted the replacement of a random disagreement point by its expectation. In fact, we have shown that a weaker axiom, our Replaceability axiom, is needed for the results of this paper, and without it one may have several solutions (recall Example 4).

Acknowledgements

I would like to thank Ron Holzman, Zvi Safra, Eilon Solan, Leeat Yariv, the associate editor and an anonymous referee for valuable comments. Financial support by the Technion’s fund for the promotion of research, the Technion VPR Fund and the William Davidson Fund are gratefully acknowledged.

References