The Effect of Disagreement on Noncooperative Bargaining

Paola Manzini

Department of Economics, Queen Mary, University of London and IZA, London, United Kingdom

and

Marco Mariotti

Department of Economics, University of Exeter, United Kingdom

m.mariotti@exeter.ac.uk

Received October 1, 2000; final version received October 3, 2001

A seemingly mild assumption of the standard alternating offers bargaining model under risk is that the breakdown event is not strictly worse than the worst agreement. When this assumption is relaxed the structure of the equilibrium set of agreements changes in an interesting way. We analyse the effect of disagreement on equilibrium, and relate our result to a class of outside option models. Journal of Economic Literature Classification Number: C78. © 2002 Elsevier Science (USA)

Key Words: disagreement; bargaining.

1. INTRODUCTION

In many bargaining situations, breakdown of negotiations may result in an outcome that one or both contenders find less attractive than the most disadvantageous of the feasible agreements. As an extreme case, think of when the breakdown outcome is nuclear war. Less dramatically, several

1 We thank Jess Benhabib, an associate editor, a referee, and József Sákovics for helpful comments. Any errors are our own.

2 Department of Economics, Queen Mary and Westfield College, Mile End Road, London E1 4NS. Fax: 020-8983 3580; URL: www.qmw.ac.uk/~ugte172).

3 Department of Economics, University of Exeter, Streatham Court, Rennes Drive, Exeter EX4 4PU, UK. Fax: 01392- 263242. URL: www.ex.ac.uk/sobe/Staff/MMariotti/MMariotti.html.
institutional constraints such as a minimum wage, minimum profitability, and price ceilings and floors may operate.

While this does not require any special consideration in the axiomatic approach to bargaining, some analysis is needed in an alternating offers strategic setting. Indeed, alternating offers models of bargaining in the presence of risk are usually based on an assumption implying that for any lottery \( l = px \oplus (1 - p)b \) between the breakdown event \( b \) and a feasible agreement \( x \), there exists another feasible agreement \( y \) such that \( y \sim l \) (see, e.g., [1, 2, 5–7]). In his original model with time preferences [9] considers an example of fixed costs of delay for the two bargainers, which exhibits an analogous feature in that context (i.e., lack of a present value for some outcome–time pairs). In that example, although existence of an equilibrium is maintained, uniqueness may fail.

In this paper we show that when the breakdown event is strictly worse than the worst agreement for at least one player, there exists a unique subgame perfect equilibrium. Depending on the magnitude of the breakdown utilities and on the probability of the breakdown event, the unique equilibrium is either alternating dictatorship (i.e., the first mover in each subgame gets all the surplus), one player’s dictatorship, or the standard outcome. Our characterisation highlights two interesting points:

- Dictatorship equilibria can be obtained even in the limit when the probability of breakdown (hence the expected loss) goes to zero.
- There is a \textit{formal} analogy with the outside option model, although the forces driving the equilibria in our model are different.

2. THE MODEL

Two players must reach an agreement on an outcome in a set \( A \). A breakdown of negotiations is a distinct outcome denoted by \( b \). To simplify notation, we will work directly in utility space and assume that the set \( A \) maps onto a compact and convex feasible set of utility pairs, \( S \subseteq \mathbb{R}^2 \). Utilities are viewed as von Neumann–Morgenstern representations of players’ risk preferences. Our only and crucial departure from the standard model is to let at least one of the players strictly prefer any agreement in \( A \) to \( b \). From now on, we call \( S \) the set of alternatives. It is convenient to assume that \( S \) can be described in the following way. Denote by \( d_i \) player \( i \)’s utility.

\footnote{A recent model that stresses the importance of disagreement in bargaining is [4]. In their model, however, the assumption that disagreement is strictly better than the worst agreement is made.}

\footnote{The case of fixed costs of delay crucially also fails to satisfy strict monotonicity of loss to delay.
if the breakdown event occurs. Denote by $u_i \geq d_i$ the utility of the worst outcome in $A$ for player $i$, with strict inequality for at least one $i$. Denote by $\bar{u}_i$, with $\bar{u}_i > u_i$, the utility of the best outcome in $A$ for player $i$, $i = 1, 2$. Let the continuous strictly decreasing concave function $f_j : [u_j, \bar{u}_j] \rightarrow [u_j, \bar{u}_j]$ denote the maximum utility that player $i$ can get for any given feasible level of player $j$’s utility. That is, $f_j(x) = \{s_j| (s_j, s_i) \in S$ and $s_j = x\}$. So $S = \{s| u_j \leq s_i \leq \bar{u}_i$ and $u_j \leq s_j \leq f_j(s_j)\} = \{s| u_j \leq s_i \leq \bar{u}_i$ and $u_j \leq s_j \leq f_j(s_j)\}$. For future reference note that obviously we have the identity $f_i(f_j(x)) = x$, $x \in [u_j, \bar{u}_j]$.

There is an unbounded number of potential rounds, $r = 0, 1, ...$, and at any round $r$ there is an exogenous probability $p \in (0, 1)$ that, following disagreement, negotiations continue to round $r+1$ and a probability $1-p$ that they break down irretrievably. Players alternate in proposing an alternative (in utility space), which can be either accepted, ending the game, or rejected. If it is the latter, with probability $p$ the responding player follows with a counter-offer in round $r+1$, and so on.

The standard alternatives $\rho^1, \rho^2 \in S$ are defined as the (unique) solution pair to the standard system $f_i(\rho^j) = pp^j + (1-p) d^j, i, j = 1, 2$; thus:

$$\rho^i = f_i(p f_j(p +(1-p) d_j) + (1-p) d_j).$$ (1)

Note that $\rho^j = pp^j + (1-p) d_j, i, j = 1, 2$ and that $\rho'$ and $\rho'$ lie on the same rectangular hyperbola with origin in $d \equiv (d_1, d_2)$. Also, unlike in the standard game, in our model such a pair may not be defined. We discuss this further below.

3. RESULTS

To simplify notation assume w.l.o.g. that $\bar{u}_i = \bar{u}_j = 1$ and $u_i = u_j = 0$. Then the equilibria are characterised as follows.

THEOREM 1. Consider the following exhaustive\(^6\) parameter configurations:

- **Alternating dictators:** $-d_i > \frac{p}{1-p}$ for all $i$.
- **Player $i$’s dictatorship:** $-d_i \leq \frac{p}{1-p}$ and $-d_j > \frac{p}{1-p}$, or

$$\frac{p}{1-p} f_i(p +(1-p) d_i) < -d_j \leq -d_j \leq \frac{p}{1-p} f_i(p +(1-p) d_j) < \frac{p}{1-p}$$

- **Standard outcome:** $-d_i < \frac{p}{1-p} f_i(p +(1-p) d_j)$ for all $i, j$.

\(^6\) See Lemma 4 below.
TABLE I

<table>
<thead>
<tr>
<th>Alternating</th>
<th>Player 1’s</th>
<th>Player 2’s</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>dictators</td>
<td>dictatorship</td>
<td>dictatorship</td>
<td>outcome</td>
</tr>
<tr>
<td>(Case 1)</td>
<td>(Case 2)</td>
<td>(Case 3)</td>
<td>(Case 4)</td>
</tr>
</tbody>
</table>

player 1 proposes $(1, 0)$ $(1, 0)$ $(f_1(p+(1-p) d_1), p+(1-p) d_1)$ $\rho^1$
accepts $x \geq 0$ $x \geq p+(1-p) d_1$ $x \geq 0$ $x \geq \rho^1$

player 2 proposes $(0, 1)$ $(p+(1-p) d_1, f_2(p+(1-p) d_1))$ $(0, 1)$ $\rho^2$
accepts $x \geq 0$ $x \geq 0$ $x \geq p+(1-p) d_1$ $x \geq \rho^2$

Then for each case there exists a unique s.p.e. with strategies as described in Table I.7 (At the boundaries of the four parameter configurations the relevant equilibria coincide.)

Proof. Checking the optimality of the strategies described above is straightforward and omitted. We then turn to uniqueness. To avoid repetitions, throughout we use the convention that all rectangular hyperbolas are measured with respect to the $(d_1, d_2)$ origin. The following lemmas are useful:

**Lemma 2.** Assume the standard system has a solution. Then:

(a) if $x > \rho^1$, then $-d_2(1-p) < pf_2(px+(1-p) d_1) - f_2(x)$;
(b) if $x < \rho^1$, then $-d_2(1-p) > pf_2(px+(1-p) d_1) - f_2(x)$.

Proof. We use a geometric argument. Consider Fig. 1, which depicts the bargaining set. Take $x > \rho^1$. Then $px+(1-p) d_1 > pp^1+(1-p) d_1$. It is clear that the point $B \equiv (px+(1-p) d_1, f_2(px+(1-p)))$ lies on a higher rectangular hyperbola than the point $A \equiv (x, f_2(x))$. Consequently

$$
(x - d_1)(f_2(x) - d_2) < (px+(1-p) d_1 - d_1)(f_2(px+(1-p) d_1) - d_2) \iff -d_2(1-p) < pf_2(px+(1-p) d_1) - f_2(x). \quad (2)
$$

The argument for case (b) is symmetric.8

7 If $-d_1 > 0$ and $-d_2 = 0$ the result is modified as follows. If $-d_1 \geq \frac{1}{p} \frac{1}{d_2}$ or $\frac{1}{d_2} f_1(p) < -d_1$, then the unique s.p.e. yields immediate agreement on $(f_1(p), p)$. If $-d_1 < \frac{1}{d_2} f_1(p)$ then the unique s.p.e. yields immediate agreement on the standard alternative $\rho^1$.

8 Incidentally, this also constitutes a geometric proof of the uniqueness of the solution of the standard system.
**Lemma 3.** Suppose that $-d_j > \frac{p}{1-p} f_j(p + (1-p) d_i)$ for at least one $j$ and $\frac{p}{1-p} \geq -d_j$ for all $j$ (so that $f_j(p + (1-p) d_i)$ is defined). Then the standard system has no solution.

**Proof.** For definiteness, let $i = 1$ and $j = 2$. Suppose by contradiction that the standard system has a solution. Since $\rho'_1 < 1$, by Lemma 2 it must be that

$$-d_2(1-p) < pf_2(p + (1-p) d_1) - f_2(1) = pf_2(p + (1-p) d_1),$$

with strict inequality if $d_i \neq d_j$.

$$-d_2 < \frac{p}{1-p} f_2(p + (1-p) d_1),$$

(3)

a contradiction. Analogously for the other case. \[\Box\]

**Lemma 4.** Suppose $-d_i \leq \frac{p}{1-p} f_j(p + (1-p) d_i)$ for all $i, j$ (so that $f_j(p + (1-p) d_i)$ is defined). Let $-d_i = \min \{-d_i, -d_j\}$. Then $-d_i \leq \frac{p}{1-p} f_j(p + (1-p) d_j)$, with strict inequality if $d_i \neq d_j$.

**Proof.** Let $i = 1$. Then point $(0, 1)$ lies on a lower hyperbola than all Pareto optimal points, in particular than the point $(f_1(p + (1-p) d_2), p + (1-p) d_2)$, that is.
\[(0 - d_i)(1 - d_j) \leq (p + (1 - p) d_z - d_z)(f_z(p + (1 - p) d_z) - d_i) \iff -d_i \leq \frac{p}{1 - p} f_z(p + (1 - p) d_i) \] (4)

with strict inequality if \( d_i \neq d_j \).

To show that no other equilibrium payoff exists, first consider the case where

\[-d_i > \frac{p}{1 - p} \text{ for at least one } i. \] (5)

If the inequality holds for both \( i \), then for each player \( 0 > p + (1 - p) d_i > px + (1 - p) d_i \), so that each player when responding prefers the worst alternative, 0, to any continuation payoff. Therefore the unique optimal proposal by \( i \) is \((1, 0)\) (where the first entry is \( i \)'s payoff). Similarly, if

\[-d_i > \frac{p}{1 - p} \text{ and } -d_j \leq \frac{p}{1 - p} \] (6)

player \( i \) when responding prefers the worst alternative to any continuation payoff, so that \( j \)'s unique optimal proposal is \((0, 1)\). This implies that \( i \)'s unique optimal proposal is \((f_i(p + (1 - p) d_i), p + (1 - p) d_i)\). This proposal is feasible as long as \( p + (1 - p) d_i \geq 0 \); that is, \(-d_i \leq \frac{p}{1 - p}\). Now consider the last possibility, namely that

\[-d_i \leq \frac{p}{1 - p} \text{ for both } i. \] (7)

Let \( M_i \) and \( m_i \) denote the supremum and infimum s.p.e. payoff for player \( i \), respectively, in subgames where he or she is the proposer. Since there may be alternatives \( x \in S \) for which \( f_i(px + (1 - p) d_i) \) is not defined, the usual \([10]\) system of inequalities must be amended as follows, for \( i, j = 1, 2 \):

\[m_j \geq f_j(\max\{0, pM_j + (1 - p) d_j\})\]

\[M_j \leq f_j(\max\{0, pm_j + (1 - p) d_j\}).\] (8)

This generates a number of possible cases, depending on whether \(-d_i \leq \frac{p}{1 - p} M_i \) and \(-d_j \leq \frac{p}{1 - p} m_i \) for both players, one of them, or none.

Consider first

\[-d_i \leq \frac{p}{1 - p} M_i \text{ for all } i. \] (9)
Then system (8) becomes
\[ m_i \geq f_i(pM_j + (1-p)d_j) \]
\[ M_j \leq f_j(\max\{0, pm_i + (1-p)d_i\}). \]  
(10)

In the subcase where \(-d_i \leq \frac{p}{1-p} m_i\) for both agents the standard system of inequalities is obtained. Then, when \(-d_i > \frac{p}{1-p} f_i(p + (1-p)d_i)\) for at least one \(i\), by Lemma 3 the system admits no solution with the property required by this subcase. When \(-d_i \leq \frac{p}{1-p} f_i(p + (1-p)d_i)\) for both players, a unique solution exists, and \(m_i = M_i = p'\) (note that \(\frac{p}{1-p} f_i(p + (1-p)d_i) < \frac{p}{1-p} m_i\)). In the other subcase where \(-d_i > \frac{p}{1-p} m_i\) for some \(i\) we have
\[ m_i \geq f_i(pM_j + (1-p)d_j) \]
\[ M_j \leq f_j(0) = 1. \]  
(11)

It follows that \(m_i \geq f_i(pM_j + (1-p)d_j) \geq f_i(p + (1-p)d_j),\) so that
\[ -d_i > \frac{p}{1-p} m_i \geq \frac{p}{1-p} f_i(p + (1-p)d_j). \]  
(12)

Since by Lemma 4 this inequality cannot be true if \(-d_i \leq -d_j\), it must be \(-d_i \leq -d_j\), with \(-d_j \leq \frac{p}{1-p} m_j\) and \(-d_j \leq \frac{p}{1-p} f_j(p + (1-p)d_j)\). Then we have
\[ m_j \geq f_j(p + (1-p)d_j) \Leftrightarrow \]
\[ f_j(m_j) - d_j \leq p + (1-p)d_j - d_j \Leftrightarrow \]
\[ (m_i - d_i)(f_j(m_j) - d_j) \leq p(1-d_j)(m_i - d_i) < (0-d_j)(1-d_j), \]  
(13)

where the last inequality follows from \(-d_i > \frac{p}{1-p} m_i\), the definition of this subcase. In summary this yields
\[ (m_i - d_i)(f_j(m_j) - d_j) < (0-d_j)(1-d_j). \]  
(14)

This means that point \((m_i, f_j(m_j))\) lies on a lower hyperbola than point \((0, 1)\). Consider now Fig. 2. From the above inequality and \(M_i \geq m_i\) it must be that point \(C \equiv (M_i, f_j(M_i))\) is also on a lower hyperbola than point \((0, 1)\), that is, \((M_i - d_i)(f_j(M_i) - d_j) < (0-d_j)(1-d_j)\). Therefore, point \(D \equiv (pM_i + (1-p)d_i, f_j(pM_i + (1-p)d_i))\) must lie on a higher hyperbola than \(C\), or
\[ (M_i - d_i)(f_j(M_i) - d_j) < (pM_i + (1-p)d_i - d_i)(f_j(pM_i + (1-p)d_i) - d_i) \Leftrightarrow \]
\[ (f_j(M_i) - d_j) < p(f_j(pM_i + (1-p)d_i) - d_i). \]  
(15)
FIGURE 2

But this contradicts the system inequalities, which yield

\[ f_j(M_i) \geq p f_j(p M_i + (1 - p) d_i) + (1 - p) d_j \iff \]

\[ (f_j(M_i) - d_i) \geq p(f_j(p M_i + (1 - p) d_i) - d_j). \]  \hspace{1cm} (16)

The other cases (namely \(-d_i \leq \frac{p}{1-p} M_i\) and \(-d_j > \frac{p}{1-p} M_j\); and \(-d_i > \frac{p}{1-p} M_i\) for all \(i\) can be dealt with in a similar way\(^9\) to show that for each parameter configuration s.p.e. payoffs are unique and correspond to those given in the statement. Having shown uniqueness of s.p.e. payoffs, completing the proof to show uniqueness of the s.p.e. itself is routine.  

4. DISCUSSION

Our result is intuitive in the sense that, loosely speaking, the dictatorship equilibria are supported by the fact that if a player’s loss in case of breakdown is sufficiently large he or she may be induced to accept a low offer to avoid breakdown. What is more striking, however, is that the corner equilibria may survive even as the probability of breakdown goes to zero, and

\(^9\) Details are available from the authors.
therefore the expected loss when rejecting an offer vanishes (recall that the penalty for disagreement is always finite). The size of the expected penalty, in itself, cannot determine the bargaining outcome. Furthermore, even small expected penalties may drive the outcome away from the standard one.

More precisely, given the parameters other than \( p \), configurations with \( -d_i > \frac{1}{1-p} \) obviously become impossible as \( p \) grows sufficiently large. However, single player dictatorship type of equilibria do not necessarily disappear as \( p \to 1 \), since for some feasible sets the inequality \( -d_i > \frac{1}{1-p} f_i(p + (1-p)d_j) \) holds for one player \( i \) for all \( p \). In this case the Nash bargaining solution is in a corner of the feasible set. Therefore, even when the standard alternative is not defined, in the limit as \( p \to 1 \) the equilibrium payoffs approach the Nash bargaining solution.

The floor principle (FP) entailed by dictatorship equilibria is somehow reminiscent of the outside option principle (OOP), in the following sense. If the standard equilibrium is defined, then it yields both players a larger utility than the minimum utility, and this is the unique equilibrium. If instead the standard equilibrium is not defined, the minimum utility becomes a binding constraint. To clarify the relationship with outside option models, consider the following alternative normalisation. Let \( d_1 = d_2 = 0 \), \( \bar{u}_1 = \bar{u}_2 = 1 \). System (8) would be written as

\[
\begin{align*}
  m_i & \geq f_i(\max\{u_i, pM_j\}) \\
  M_j & \leq f_j(\max\{u_i, pm_i\}).
\end{align*}
\]  

(17)

Its solution would characterise the equilibrium payoffs of a standard bargaining game with outside option (of the bazaar type) \( u_i \) for player \( i \), and in which the Pareto frontier of the feasible set has been extended so as to make it d-comprehensive\(^{10}\) (preserving convexity).\(^{11}\) This analytical correspondence is remarkable for two reasons.

First, the OOP is sustained by introducing an additional action in the model (i.e., walk out). On the contrary, for the FP to drive the equilibrium outcome no such action is required from the player who benefits from it. The two models are structurally different and the similar equilibrium outcomes are supported by different strategic forces.

Second, in the model of this paper, the standard points are not defined when the FP is active. This is why the parameter inequalities discriminating the four types of equilibria of Theorem 1 are so different from the equivalent inequalities in the game with outside options. For example, in the case

\(^{10}\) That is, if \( s \in S \) and \( d \leq s' \leq s \), then \( s' \in S \).

\(^{11}\) Without the extension, the Pareto frontier would not be defined for alternatives where one of the players obtains less than his or her outside option.
when player 2 only has an outside option, this is binding only if $u_2 > p\rho_2^2$ (with $u_2$ interpreted as the outside option). In our model the corresponding inequality obviously cannot involve the undefined number $\rho_2^2$ and it is expressed only in terms of the primitives: $u_2 > p\bar{u}$ or $p\bar{f}(p\bar{u}) < u_2 < p\bar{u}$.

The result above is useful to assess the validity of some disagreement point axioms. Consider for example the characterisation by [8], who use an axiom called star-shaped inverse (SSI). This axiom\textsuperscript{12} states that shifting the disagreement point in the direction of the solution point should not change the solution point itself. However, in our framework this is not necessarily the case.

On the other hand, another common disagreement point axiom, disagreement point monotonicity, is confirmed in our model. It is easy to verify that whenever the disagreement point moves favouring only player $i$, his or her equilibrium payoff (weakly) increases.

REFERENCES

\textsuperscript{12} The same axiom has been used more recently in [3], under the more descriptive name of disagreement point convexity.