Investigaciones Económicas. Volumen XVIII (3) - Septiembre 1994, pp. 403-419

BARGAINING THEORY WITHOUT TEARS

Ken BINMORE
University College London

The paper is an attempt to provide a compact but fairly comprehensive guide to the use of Rubinstein's bargaining model for applied workers. Particular attention is given to modeling the manner in which the negotiations may break down. An application to markets in which the price is determined by search and bargaining is given.

1. Introduction

The purpose of this article is two-fold. The primary aim is to provide a simple proof of a version of Rubinstein's (1982) bargaining theorem in a setting that is sufficiently general to cover the situations that typically arise in applications. In particular, the feasible set is not assumed to be convex and a reasonably general view is taken of the manner in which disagreement may arise.

The secondary aim pursues some points made in Binmore, Rubinstein and Wolinsky (1986). A detailed analysis of subgame-perfect equilibria in a complicated non-cooperative bargaining model is unnecessary for most applications. Much heavy computation can be short-circuited by applying certain simple principles directly rather than deriving them anew each time they are required. The methodology is illustrated in Section 8 for a model of decentralized price formation.

2. The alternating offers model

Rather than setting the story directly in utility space, it will be told in terms of the classic problem of «dividing the dollar». A philanthropist donates a dollar to Adam and Eve on condition that they are able to agree on how to share it. Disagreement may arise in various ways. A player may abandon the negotiations in favor of his or her best outside option leaving the other to do the same. Or the philanthropist may lose patience if agreement is delayed and withdraw his offer. If either of these eventualities occurs, the negotiations will be said to have broken down. Even without a breakdown, agreement may not be reached since it is open to the players to sit at the negotiation table for ever.

The result of bargaining under precisely specified rules will be studied. All action takes place at times $n\tau$ $(n = 0, 1, 2...)$, where $\tau > 0$. Adam is active when n is even. Eve is active when n is odd. If the game has not already ended at time $n\tau > 0$, the philanthropist begins by withdrawing his dollar with probability $\pi = \lambda\tau < 1$. The game continues with probability $\bar{\pi} = 1 - \pi$. The active player then decides whether to opt out. If the active player opts in, then he or she continues by making a proposal on how to split the dollar. The passive player then accepts or refuses. Only after a refusal does the clock advance by τ. The passive player then becomes active and the above sequence of events is repeated. The game begins at time $n = 0$ but, in this first period, the steps in which the dollar may be withdrawn and in which Adam may opt out are skipped. The very first move therefore consists of Adam's making a proposal. The second move consists of Eve's response. If she refuses, the sequence of events described in the previous paragraph commences with $n = 1$. Figure 1 illustrates the order of moves[1].

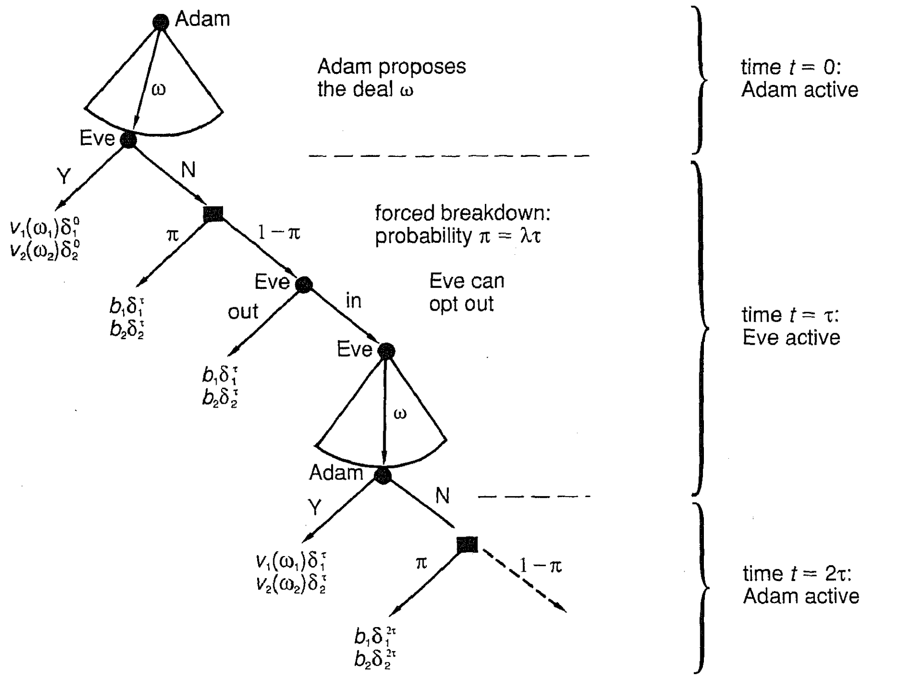

Figure 1
Rules of the bargaining game

[1] Most of the sequencing in this specification is unimportant to the results. That chosen is largely for mathematical convenience. However, it is important that the active player's opting out decision does not occur immediately *after* the passive player has refused an offer. Otherwise further equilibria appear (Shaked (1987)). But a model in which a player cannot leave without hearing one final offer from the opponent seems more realistic. The anomalous first period allows a more elegant statement of some results since then Adam will never actually opt out in equilibrium. It is, of course, trivial to extend the analysis using backwards induction to the case when the first period is not anomalous.

3. Preferences

Adam is taken to be player 1 and Eve to be player 2. The set of possible deals is identified with

$$\Omega = \{\omega : \omega_1 + \omega_2 \leq 1\}$$

Notice that it is assumed that money can be burned or borrowed and transferred freely between the players. Osborne and Rubinstein (1990), for example, assume that such transactions are impossible [2]. The point $\beta \in \Omega$ represents a pair of breakdown payments. These are the payments that each player will receive if the negotiations break down [3].

Player i's utility for the outcome $\omega \in \Omega$ at time t is taken to be

$$u_i(\omega, t) = v_i(\omega_i)\,\delta_i^t \qquad [1]$$

where the discount factor δ_i satisfies $0 < \delta_i \leq 1$. (The corresponding discount rate ρ_i is defined by $\delta_i = e^{-\rho_i}$). To economize on notation, we write $\Delta_i = \delta_i^\tau$. Recall that $\pi = \lambda\tau$. Thus, in what follows, both Δ_i and π are always functions of τ. The function v_i: $\mathbb{R} \to \mathbb{R}$ is assumed to be continuous and strictly increasing. Its range is an open interval R_1. We take $b_i = v_i(\beta_i)$. The breakdown point b then lies in $R_1 \times R_2$.

The Pareto frontier of the set X of utility pairs available at time 0 is the graph of the function f defined by

$$f(x_1) = v_2(1 - v_1^{-1}(x_1))$$

Both f: $R_1 \to R_2$ and its inverse f^{-1}: $R_2 \to R_1$ are continuous and strictly decreasing. Note that $b \in X$ so that $b_2 \leq f(b_1)$ and $b_1 \leq f^{-1}(b_2)$.

It remains to discuss how the players assess the consequences of the perpetual disagreement outcome D. This is assigned utility $u_i(D) = 0$ so as to be consistent with the result of allowing $t \to \infty$ in [1]. In addition, it is assumed that $0 \in \bar{R}_1 \times \bar{R}_2$ and $b \geq 0$.

A situation in which the players use strategies that rule out the possibility of agreement being reached but do not result in either player ever opting out will be called a deadlock. A deadlock leads to the outcome D with positive

[2] And hence are able to find equilibria in circumstances when they would not otherwise exist. The approach taken here seems more natural.

[3] The case when *different* payments are received depending on whether someone opts out or the dollar is withdrawn has been considered elsewhere (for example, Binmore, Osborne and Rubinstein (1992)). The only new difficulties are combinatorial, but the interest of this more general case does not seem sufficient to justify the extra algebra.

probability only when $\pi = 0$. Otherwise the expected utility to player i from a deadlock is

$$\pi b_i\Delta_i + \pi b_i\Delta_i(\bar{\pi}\Delta_i) + \pi b_i\Delta_i(\bar{\pi}\Delta_i)^2 + \ldots = \frac{\pi b_i\Delta_i}{1 - \bar{\pi}\Delta_i} \qquad [2]$$

If υ_1 and υ_2 are concave, then so is f. It follows that X is convex. If X is not convex, we specifically do *not* replace it by its convex hull as is normal practice. It is frequently unrealistic to suppose that agreement on a lottery is feasible. A union boss, for example, cannot report to his members that their wage settlement was decided by tossing a coin. Where lotteries are feasible, Ω should be replaced by the set of all lotteries with prizes in Ω. A simpler theory then results.

NOTES:

1) Fishburn and Rubinstein (1982) show that relatively mild assumptions on preference relations guarantee a utility representation of the form [1]. In particular, [1] substitutes in this paper for the conditions (A1) through (A5) of Osborne and Rubinstein (1990). Their condition (A6) is not used here.

2) Expressing the basic problem in terms of «dividing the dollar» clarifies the interpretation but may obscure the generality of the treatment. All that really matters is the shape of the set X of feasible utility pairs and the values of δ_1, δ_2, π and b. Geometric characterizations of the equilibrium outcomes in such a setting are easily obtained (Binmore (1987)). Algebraic characterizations, as studied here, require more labor.

3) If $\pi > 0$, the utility functions must be understood in the sense of Von Neumann and Morgenstern. If $\pi = 0$, expected utility calculations are not necessary and so any utility function representation will suffice. This observation permits the study of various cases not obviously included in the scope of the analysis. In particular, Rubinstein's (1982) case of «fixed costs of disagreement» is accessible. Player i's utility for the deal w at time t is then $U_i(w, t) = w_i - \gamma_i t$, where $\gamma_i > 0$. Rubinstein's version has to be supplemented here by requiring that, if someone opts out at time t, then player i gets $-\gamma_i t$. One takes $u_i(w, t) = \exp\, U_i(w, t)$ and $\delta_i = e^{-\gamma_i}$. Then $\upsilon_i(x) = e^x$ is far from concave. Notice that f: $(0, \infty) \to (0, \infty)$ is given by $f(x) = e/x$ and $b = (1, 1)$. The example is useful because it exhibits various pathologies.

4. Stationary subgame-perfect equilibria

In this section, η and ζ are Pareto-efficient deals in Ω. The utilities assigned to these deals at time 0 by the two players are given by $y_i = v_i(\eta_i)$ and $z_i = v_i(\zeta_i)$. Notice that $y_2 = f(y_1)$ and $z_2 = f(z_1)$.

Adam will be assumed to use a pure strategy s that requires him to:

a) Propose η whenever called upon for a proposal;

b) Accept ζ or better and to refuse anything worse whenever called upon for a response.

Eve will be assumed to use a pure strategy t which has the same properties except that the roles of η and ζ are reversed.

When do strategies with these properties constitute a subgame-perfect equilibrium? The current section explores this question in a series of lemmas. The simple proofs are relegated to an appendix.

Define: m_i: $R_i \to \mathbb{R}$ by

$$m_i(x) = \Delta_i(\pi b_i + \bar{\pi} \max \{b_i, x\}),$$

and restrict attention to those η and ζ for which

$$\left.\begin{aligned} y_2 &= m_2(z_2) \\ z_1 &= m_1(y_1) \end{aligned}\right\} \qquad [3]$$

These equations express the fact that the passive player will always be indifferent between accepting and refusing what is proposed.

LEMMA 1. If [3] holds, then $y_1 \geq b_1$ or $z_2 \geq b_2$. Also, $z_1 \leq y_1$ and $y_2 \leq z_2$.

PROOF. If $y_1 < b_1$, then $z_1 = \Delta_1 b_1 \leq b_1$. Since z is Pareto-efficient, it therefore cannot be that $z_2 < b_2$. If $y_1 \geq b_1$, then $z_1 = \Delta_1(\bar{\pi} y_1 + \pi b_1) \leq \Delta_1 y_1 \leq y_1$. Thus, $z_2 = f(z_1) \geq f(y_1) = y_2$. A similar argument applies if $z_2 \geq b_2$.

LEMMA 2. If y and z satisfy [3], then there exists a corresponding subgame-perfect equilibrium pair (s, t).

(The properties given for s and t do not specify whether or not a player should opt out when the opportunity arises. Adam should opt in if $y_1 \geq b_1$ and opt out if $y_1 < b_1$. Eve should opt in if $z_2 \geq b_2$ and opt out if $z_2 < b_2$).

PROOF. It needs to be confirmed that proposing decisions are optimal. They are optimal because the active player cannot demand more without being refused, and he or she prefers not to be refused. In checking this last point, Lemma 1 makes all the cases immediate except for that in which the active player opts out in equilibrium. Suppose, in particular, that $y_1 \geq b_1$ and $z_2 < b_2$. Then $y_2 = \Delta_2 b_2$. Hence,

$$z_2 \geq y_2 = \Delta_2 b_2 \geq \Delta_2(\bar{\pi}\Delta_2 b_2 + \pi b_2) = \Delta_2(\bar{\pi} y_2 + \pi b_2).$$

It follows that Eve prefers z to be accepted than to have her proposal refused. A similar argument applies when $y_1 < b_1$ and $z_2 \geq b_2$.

LEMMA 3. The pair (y, z) satisfies [3] and hence specifies a stationary subgame-perfect equilibrium if and only if y_1 is a zero of the function g: $R_1 \rightarrow \mathbb{R}$ defined by

$$g(x) = f(x) - (m_2 \circ f \circ m_1)(x)$$

and $y_2 = f(y_1)$, $z_1 = m_1(y_1)$, $z_2 = f(z_1)$.

It will be necessary to investigate the properties of the function g in some detail. Note to begin with that

$$g(x) = \begin{cases} f(x) - \Delta_2[\pi b_2 + \bar{\pi} f(\Delta_1 b_1)], & \text{if } x \leq b_1; \\ f(x) - \Delta_2[\pi b_2 + \bar{\pi} f(\Delta_1(\pi b_1 + \bar{\pi} x))] & \text{if } b_1 \leq x \leq c_1; \\ f(x) - \Delta_2 b_2, & \text{if } x \geq c_1; \end{cases}$$

where $\bar{\pi} c_1 = \Delta_1^{-1} f^{-1}(b_2) - \pi b_1$. The dependence of g on r is not made explicit since it turns out to be more convenient to study G: $R_1 \times (0, \infty) \rightarrow \mathbb{R}$, which is defined by

$$G(x, r) = \tau^{-1} g(x) \qquad [4]$$

when this dependence matters.

LEMMA 4. The function g: $R_1 \rightarrow \mathbb{R}$ has the following properties:

$$a) \; x < \Delta_1 b_1 \quad \Rightarrow g(x) > 0$$
$$b) \; x > f^{-1}(\Delta_2 b_2) \quad \Rightarrow g(x) < 0.$$

PROOF. *a)* If $x < \Delta_1 b_1$, then

$$\begin{aligned} g(x) &> f(\Delta_1 b_1) - \Delta_2[\pi b_2 + \bar{\pi} f(\Delta_1 b_1)] \\ &= (1 - \Delta_2 \bar{\pi}) f(\Delta_1 b_1) - \Delta_2 \pi b_2 \\ &\geq (1 - \Delta_2 \bar{\pi})\, b_2 - \Delta_2 \pi b_2 \\ &= b_2(1 - \Delta_2) \geq 0. \end{aligned}$$

b) If $x > f^{-1}(\Delta_2 b_2)$, then $f(x) < \Delta_2 b_2$. If it is also true that $g(x) \geq 0$, then

$$\Delta_2 b_2 - \Delta_2[\pi b_2 + \bar{\pi} f(\Delta_1(\pi b_1 + \bar{\pi} x))] > 0$$
$$b_2 > f(\Delta_1(\pi b_1 + \bar{\pi} x))$$
$$f^{-1}(b_2) < \Delta_1(\pi b_1 + \bar{\pi} x)$$
$$x > c_1.$$

Now suppose that $f^{-1}(\Delta_2 b_2) \leq c_1$. Then $g(x) < 0$ for $f^{-1}(\Delta_2 b_2) < x \leq c_1$. Hence, $g(x) < 0$ for $x > f^{-1}(\Delta_2 b_2)$, because g is continuous and decreases on (c_1, ∞). On the other hand, if $f^{-1}(\Delta_2 b_2) > c_1$, then $x > f^{-1}(\Delta_2 b_2)$ implies that $g(x) = f(x) - \Delta_2 b_2 < 0$.

LEMMA 5. The function g: $R_1 \rightarrow \mathbb{R}$ always has a zero in $[\Delta_1 b_1, f^{-1}(\Delta_2 b_2)]$ and hence a stationary subgame-perfect equilibrium exists.

PROOF. This follows from the previous lemma because g is continuous.

5. Non-stationary equilibria

This section proves a generalized version of a theorem of Rubinstein (1982). The proof follows Binmore (1987), Shaked and Sutton (1984) and Binmore, Shaked and Sutton (1989).

Let S be the set of all subgame-perfect equilibrium outcomes. The first result demonstrates that S is necessarily a large set when more than one stationary subgame-perfect equilibrium of the type considered in section 4 exists. In particular, Pareto-inefficient outcomes lie in S.

Multiple stationary subgame-perfect equilibria exist when g: $R_1 \rightarrow \mathbb{R}$ has more than one zero. Let $(\underline{s}, \underline{t})$ be the strategy pair that corresponds to the smallest zero $\underline{y}_1$ of g. Let $(\bar{s}, \bar{t})$ be the strategy pair corresponding to the largest zero $\bar{y}_1$.

Let T be the set of all feasible payoff pairs x that satisfy $x \geq (\underline{y}_1, \bar{y}_2)$.

LEMMA 6. If $x \in T$, then there exists a subgame-perfect equilibrium (s, t) in which Adam proposes a deal ξ at time 0 worth x and Eve accepts. Thus $T \subseteq S$.

PROOF. Three «states of mind», *UP*, *DOWN* and *MIDDLE* are distinguished. Players begin in the *MIDDLE* state. In this state, the subgame-perfect equilibrium (s, t) to be constructed requires Adam to propose ξ when called upon to make a proposal. Eve accepts ξ and anything at least as good as ξ. She refuses anything else.

In the *UP* state (s, t) requires that the players play according to $(\bar{s}, \bar{t})$ in the remainder of the game. In the *DOWN* state (s, t) requires playing according to $(\underline{s}, \underline{t})$ in the remainder of the game.

Once in the *UP* state, players remain there. The same applies to the *DOWN* state. Transitions from the *MIDDLE* state are made as follows. If Adam proposes ξ, then a refusal by Eve shifts both players to the *UP* state. If Eve refuses any other proposal, both players shift to the *DOWN* state.

Why is the schedule for proposal and response in the *MIDDLE* state optimal? Eve should accept ξ because $x_2 \geq \bar{y}_2 = m_2(z_2)$. Her response to other proposals is optimal because she gets $y_2 = m_2(z_2)$ from refusing. Adam should propose ξ because he gets at most $\underline{y}_1$ from deviating and $x_1 \geq \underline{y}_1$. (If Adam deviates to a proposal that is refused, either Eve opts out because $z_2 < b_2$ or she opts in because $z_2 \geq b_2$. In the latter case, Adam gets $\Delta_1 (\pi b_1 + \bar{\pi} z_1) \leq \Delta_1 z_1 \leq z_1 \leq \underline{y}_1 \leq x_1$. In the former case, he gets $\Delta_1 b_1 \leq b_1 \leq \underline{y}_2 \leq x_1$ by Lemma 1).

Next it will be shown that $S \subseteq T$. A preliminary lemma is needed.

LEMMA 7. The set Y of all subgame-perfect equilibrium payoffs to Adam is $[\underline{y}_1, \bar{y}_1]$.

PROOF. Let $a = \inf Y$ and $A = \sup Y$. Let Z be the set of all subgame-perfect equilibrium payoffs to Eve in the companion game in which it is Eve who makes the first proposal at time 0. Write $e = \inf Z$ and $E = \sup Z$.

1) It is open to Eve to refuse whatever Adam proposes at time 0. If equilibrium strategies are used in the continuation of the game, then Eve will get an expected payoff of at least $m_2(e)$, because the companion game will be played after a time delay of τ unless the dollar is withdrawn or Eve opts out in the interim. If equilibrium strategies are used, it follows that Eve gets at least $m_2(e)$ and so Adam gets at most $f^{-1}(m_2(e))$. Thus, $A \leq f^{-1}(m_2(e))$ and so

$$f(A) \geq m_2(e). \qquad [5]$$

On applying a similar argument in the companion game,

$$E \leq f(m_1(a)). \qquad [6]$$

2) It is optimal for Eve to accept any proposal from Adam that assigns her a payoff $w_2 > m_2(E)$, provided that equilibrium strategies are used after a refusal. Thus Adam must get at least $f^{-1}(w_2)$. Hence $f^{-1}(w_2)$ is a lower bound for S whenever $w_2 > m_2(E)$. Thus $a \geq f^{-1}(m_2(E))$ and so

$$f(a) \leq m_2(E). \qquad [7]$$

On applying a similar argument in the companion game,

$$e \geq f(m_1(A)). \qquad [8]$$

3) From [5] and [8],

$$f(A) \geq m_2(e) \geq (m_2 \circ f \circ m_1)(A)$$

and hence $g(A) \geq 0$. But $g(x) < 0$ for $x > \bar{y}_1$. Thus $A \leq \bar{y}_1$. But $\bar{y}_1 \in Y$ and so $A = \bar{y}_1$.

From [6] and [7],

$$f(a) \leq m_2(E) \leq (m_2 \circ f \circ m_1)(a)$$

and hence $g(a) \leq 0$. But $g(x) > 0$ for $x < \underline{y}_1$. Thus $a \geq \underline{y}_1$. But $\underline{y}_1 \in Y$ and so $a = \underline{y}_1$.

4) It remains to confirm that $\underline{y}_1 \leq y_1 \leq \bar{y}_1$ implies $y_1 \in Y$. This follows immediately from Lemma 6.

THEOREM 1. $S = T$.

PROOF. By the preceding Lemma, Adam's equilibrium payoffs lie in the set $[\underline{y}_1, \bar{y}_1]$. Similarly, Eve's equilibrium payoffs in the companion game lie in $[\bar{z}_2, \underline{z}_2]$. It follows that her equilibrium payoffs in the original game lie in $[\bar{y}_2, \underline{y}_2]$, since $\bar{y}_2 = m_2(\bar{z}_2)$ and $\underline{y}_2 = m_2(\underline{z}_2)$. Thus $S \subseteq T$. On the other hand, Lemma 6 shows that $T \subseteq S$.

NOTES:

1) When multiple equilibria exist, there may be subgame-perfect equilibria in which agreement is not reached immediately (section 3.10.1 of Osborne and Rubinstein (1990)).

The second and third notes concern two cases of special interest considered by Rubinstein (1982). The reason that the conclusions quoted differ from his is because the bargaining models considered are not identical.

2) In the notation of section 3, take $v_1(x) = v_2(x) = x$ and $\pi = 0$. Then $R_1 = R_2 = \mathbb{R}$ and $f(x) = 1 - x$. The zeros of the function g are the stationary subgame-perfect equilibrium outcomes for Adam. In this case, g is strictly decreasing. It follows that there is a unique equilibrium outcome. If $u = (1 - \Delta_2)/(1 - \Delta_1\Delta_2)$ satisfies $b_1 \leq u \leq \Delta_1^{-1}(1 - b_2)$, then Adam gets u. If $u < b_1$, then Adam will be planning to take his outside option should the opportunity arise. He gets $1 - \Delta_2(1 - \Delta_1 b_1)$ when proposing. If $u > \Delta_1^{-1}(1 - b_2)$, Eve is planning to take her outside option. Adam gets $1 - \Delta_2 b_2$.

3) The second case of interest is described in note 3 of section 3. When $\Delta_1 > \Delta_2$, g has a unique zero at Δ_2^{-1}. In terms of money payoffs, this means that, when $\gamma_1 < \gamma_2$, Adam gets $1 + \gamma_2\tau$ in equilibrium. (Eve plans to take her outside option). When $\Delta_1 < \Delta_2$, g has a unique zero at $\Delta_1\Delta_2^{-1}$. In terms of money payoffs this means that, when $\gamma_1 < \gamma_2$, Adam gets $(\gamma_2 - \gamma_1)\tau$ in equilibrium. (Adam plans to take his outside option). When $\Delta_1 = \Delta_2 = \Delta$, g is zero on $[1, e\Delta^{-1}]$. Thus, with the notation of section 4, $\underline{y}_1 = 1$ and, $\bar{y}_1 = e\Delta^{-1}$. Hence, $\bar{y}_2 = \Delta$. Any feasible pair $x \geq (1, \Delta)$ is therefore an equilibrium outcome. In terms of money payoffs this means that when $\gamma_1 = \gamma_2 = \gamma$, the set of equilibrium outcomes is $\{\omega : \omega_1 : + \omega_2 \leq 1, \omega_1 \geq 0, \omega_2 \geq -\gamma r\}$.

6. Generalized Nash bargaining solutions

Much of the interest of the bargaining model described in the previous sections lies in the fact that, in the limit as $\tau \to 0+$, the equilibrium outcomes can be characterized in terms of a suitably generalized version of Nash's bargaining solution. The case $\tau \to 0+$ deserves special emphasis for at least two reasons. The first reason is that there will often be nothing that constrains players to keep to the timetable specified in the model. After refusing a proposal, they will then wish to make a counterproposal at the earliest possible opportunity[4]. The second reason is that Adam's first-mover advantage disappears in the limit as $\tau \to 0+$. In this section, a generalized Nash bargaining solution will be described. Axiomatizations of the point-valued version can be found in Kalai (1977), Roth (1979) and elsewhere.

An abstract bargaining problem will be identified with a triple (X, b, d) in which X is interpreted as the set of feasible payoff pairs, b is a breakdown point whose coordinates are the players' outside options, and d is a deadlock point. As in section 2, R_1 and R_2 are open intervals, $f: R_1 \to R_2$ is a strictly decreasing surjection and $X = \{(x_1, x_2) \in R_1 \times R_2 : x_2 \leq f(x_1)\}$. Also $0 \in \bar{R}_1 \times R_2$ and $0 \leq d \leq b \in X$. For the remainder of the paper it will also be assumed that f is twice differentiable on R_1.

A generalized Nash product with bargaining powers $\alpha > 0$ and $\beta > 0$ is defined to be an expression of the form.

$$P(x_1, x_2) = (x_1 - d_1)^{\alpha} (x_2 - d_2)^{\beta} \qquad [9]$$

When X is convex and $b = d$, the regular Nash bargaining solution introduced by Nash [12] identifies the solution of (X, d, d) with the point n at which the Nash product P with $\alpha = \beta$ is maximized subject to the constraints $x \in X$ and $x \geq d$.

When X is not convex, a more elaborate definition is necessary. Let $p: R_1 \to \mathbb{R}$ be given by

$$p(x) = P(x, f(x)) \qquad [10]$$

Attention will be restricted to the case in which $p'(x)$ is zero neither at an endpoint of the interval $[b_1, f^{-1}(b_2)]$, nor at an interior point x where $p''(x) = 0$. Only pathological cases are excluded by this restriction. The function $H: R_1 \to \mathbb{R}$ is defined by

$$H(x) = \begin{cases} +\infty, & \text{if } x < b_1 \\ f'(x)\left(\dfrac{x-d_1}{\alpha}\right) + \left(\dfrac{f(x)-d_2}{\beta}\right), & \text{if } b_1 \leq x \leq f^{-1}(b_2) \\ -\infty, & \text{if } x > f^{-1}(b_2) \end{cases}$$

[4] The limiting factors will then be physical or physiological. Modeling these will involve a reinterpretation of δ_1 and δ_2.

This has the same sign as p' on the interval $[b_1, f^{-1}(b_2)]$. A 'zero' of H will be understood to be any $z \in R_1$ which has the property that all of its neighborhoods contain both positive and negative values of H. In view of the preceding restrictions on p, a 'zero' of H is either an interior point z of the interval $[b_1, f^{-1}(b_2)]$, at which $H(z) = 0$, or an endpoint of the interval.

A Nash bargaining point n can now be defined to be a Pareto-efficient point of X for which n_1 is a 'zero' of H. Of all Nash bargaining points, let $\bar{n}$ be that which assigns Adam the greatest payoff (and Eve the least). Let $\underline{n}$ be that which assigns Eve the greatest payoff (and Adam the least). The generalized Nash bargaining solution corresponding to the bargaining powers $\alpha > 0$ and $\beta > 0$ for the bargaining problem (X, b, d) is the set N of all feasible payoff pairs $x \geq (\underline{n}_1, \bar{n}_2)$. The definition is illustrated in Figure 2.

PROPOSITION 1. The point $\bar{n}$ is the local maximum of the Nash product [9] subject to $x \in X$ and $x \geq b$ that assigns Adam the greatest payoff. The point $\underline{n}$ is the local maximum that assigns Eve the greatest payoff.

PROPOSITION 2. A sufficient condition that N consist of a single point is that $(x - d_1) f'(x)$ be concave on $[b_1, f^{-1}(b_2)]$. In particular, it is sufficient if f is concave and so X is convex.

PROOF. The condition implies that H is strictly decreasing on $[b_1, f^{-1}(b_2)]$.

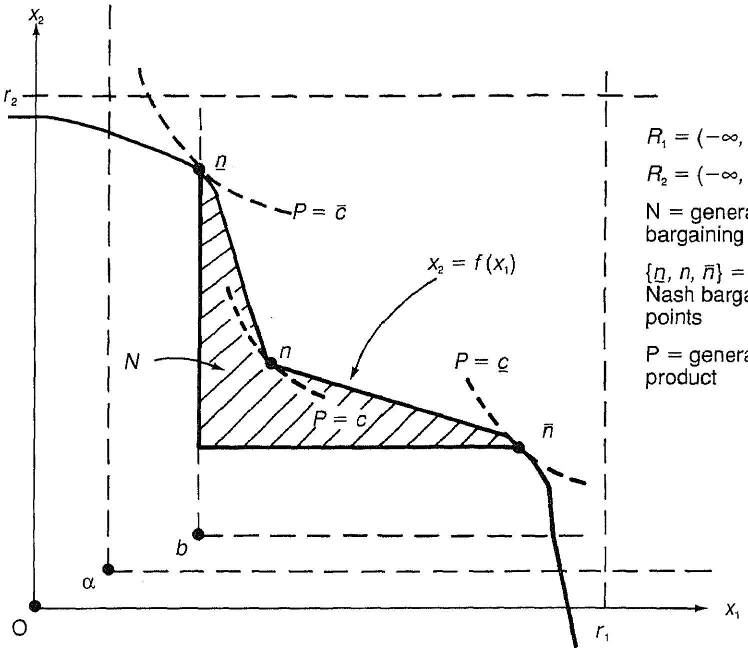

Figure 2
A Generalized Nash Bargaining Solution

7. The Nash program

The aim of the Nash program is to provide non-cooperative justifications, where possible, for the solution concepts of cooperative game theory. (See, for example, the introduction to Binmore and Dasgupta (1987). The alternating offers model of Sections 2-5 is important in this context since it provides a defense for the Nash bargaining solution.

THEOREM 2. As $\tau \to 0$, the set of subgame-perfect equilibrium outcomes in the alternating offers model converges to the generalized Nash bargaining solution $\mathcal{N}$ corresponding to the bargaining powers $\alpha = 1/(\lambda + \rho_1)$ and $\beta = 1/(\lambda + \rho_2)$ for the bargaining problem (X, b, d), in which the deadlock payoff $d_i = \lambda b_i/(\lambda + \rho_i)$ is the limiting value of [2] as $\tau \to 0+$.

PROOF. It will be shown that the set of values of x for which the function G defined by [4] is zero converges to the set of 'zeros' of the function H defined in the preceding section. The first step is the observation that, for each $x \in R_1$,

$$G(x, \tau) \to H(x) \text{ as } \tau \to 0+$$

On the interval $[b_1, f^{-1}(b_2)]$, this follows from L'Hôpital's rule because $c_1 > f^{-1}(b_2)$. Outside the interval, one may appeal to Lemma 4 to determine the sign of G.

1) Every neighborhood of a 'zero' of H contains a zero of G, provided τ is sufficiently small. With the restriction introduced in Section 6, each such neighborhood contains a point at which H is positive and a point at which H is negative. The same is therefore true of G if τ is sufficiently small. Since G is continuous, it follows that G has a zero in the neighborhood.

2) Every neighborhood that contains zeros of G for all sufficiently small τ also contains a 'zero' of H. The inverval $[b_1, f^{-1}(b_2)]$ is compact, and hence G converges uniformly to H on this interval. This observation takes care of neighborhoods centered at points of $(b_1, f^{-1}(b_2))$. A trivial argument extends the conclusion to neighborhoods centered at the endpoints.

In summary, where the defense of the use of the Nash bargaining theory is to be based on an alternating offers model[5], the «status quo» should correspond to the consequences of a deadlock (during which the players remain at the negotiation table but never reach an agreement). The outside options that they may obtain by abandoning the negotiations serve only as constraints on the range of validity of the Nash bargaining solution. Often, it is

[5] Rather than, for example, Nash's (1950) own model in which players simultaneously make take-it-or-leave-it offers.

convenient to apply these principles to payoff *flows*. In a wage negotiation, for example, the deadlock flows may be the income per period for the two sides during a strike.

The final result of this section is offered without a proof. It provides a criterion for the uniqueness of an equilibrium in the alternating offers model that does not depend on τ being small.

PROPOSITION 3. A necessary condition that the alternating offers model have a unique subgame-perfect equilibrium is that N consist of a single point.

Some special cases of Theorem 2 deserve mention. In each case, X is assumed to be convex so that N consists of a single point.

1) $\rho_1 = \rho_2 = 0$. In this case, the equilibrium outcome converges to the regular Nash bargaining solution for the problem (X, b, b). Here the breakdown and deadlock points are the same and there is no difficulty in deciding on an appropriate «status quo» in using Nash's theory. This case arises when it is not impatience that motivates an early agreement but fear that the opportunity to reach an agreement may disappear if an agreement is delayed.

2) $\rho_1 = \rho_2 = \rho > 0$. In this case, the equilibrium outcome converges to the regular Nash bargaining solution for the problem (X, b, d), where $d = \lambda b/(\lambda + \rho)$. Note the displacement of the «status quo» from b. In *symmetric* situations, this displacement leaves the location of the Nash bargaining solution unaltered. Models that mistakenly place the «status quo» at b will therefore nevertheless lead to the correct conclusions in symmetric situations.

3) $\lambda = 0$. In this case, the equilibrium outcome converges to an *asymmetric*, Nash bargaining solution with bargaining powers $\alpha = 1/\rho_1$ and $\beta = 1/\rho_2$ for the problem $(X, b, 0)$. Recall that the payoff pair 0 corresponds to the perpetual disagreement point D which therefore serves as the appropriate «status quo» under these circumstances. This case arises when the players are unconcerned about the risk of losing the opportunity to reach an agreement and are motivated simply by their impatience with delays.

8. Decentralized price formation

To illustrate the principles of the preceding section, a model will now be studied in which the price at which a good is traded is determined by bargaining between buyers and sellers rather than through some centralized auctioneering mechanism. Insofar as there is an innovation as compared with Rubinstein and Wolinsky (1985), Gale (1986) or Binmore and Herrero (1988), it lies in the more realistic modeling of the circumstances of a bargaining breakdown. Wolinsky (1988) and Bester (1988) consider other variants of the model.

Each seller owns a house. If he sells the house at time t for price p, his utility is $p\delta_1^t$. The buyer gets $(1 - p)\delta_2^t$. An agent who opts out of the market or who never succeeds in consummating a deal gets zero utility. After a sale, the buyer and seller leave the market, but are immediately replaced by a new buyer and a new seller so that things remain in a steady state. The market therefore always contains a pool of unmatched agents looking for a bargaining partner.

The price at which the house is exchanged is determined by bargaining between individual buyers and sellers who have succeeded in finding each other. The bargaining model is based on that of section 2, but various modifications are necessary. In particular, account needs to be taken of buyers and sellers who have yet to find a bargaining partner. Such unmatched agents are always deemed to be active.

At the beginning of each time period, *all* active agents are matched with a new partner with probability $\lambda_i\tau > 0$ $(i = 1, 2)$. A player who was passive in the preceding period and refused the proposal made by his or her partner may therefore have *two* partners in the current period. Such a player is in a powerful position because this creates an auctioning scenario[6]. The modeling of this scenario is discussed below. The next event is a decision by active players on whether or not to opt out. An unmatched player may opt out of the market altogether. A matched active player may do the same or abandon his or her current partner and so become unmatched. If a matched active player opts in, he or she makes a proposal that the passive player may accept or refuse. Before further events, the clock advances by τ. Any remaining passive players become active and the cycle of events is repeated.

This is a more complex problem than that discussed in preceding sections, but a full non-cooperative analysis will not be described. Instead, the result of such an analysis will be predicted using the principles outlined in section 7. These apply only in the limiting case when $\tau \to 0+$ (the case of 'no bargaining frictions'). The prediction is framed in terms of an appropriate Nash bargaining solution of an appropriate bargaining problem (X, b, d). Note first that the average probability that bilateral bargaining will break down during a period of length τ is $\lambda\tau = \frac{1}{2}(\lambda_1 + \lambda_2)\tau$. The appropriate bargaining powers are therefore $\alpha = 1/(\lambda + \rho_1)$ and $\beta = 1/(\lambda + \rho_2)$. The feasible set X is the unit simplex. The value of the generalized Nash bargaining solution is therefore a payoff pair of the form $(p, 1 - p)$, where p is the price at which the house is sold. Agreement on this price will be immediate when a buyer and seller get matched.

The seller's outside option b_1 is $\lambda_1\rho/(\lambda_1 + \rho_1)$. Similarly, $b_2 = \lambda_2(1 - \rho)/(\lambda_2 + \rho_2)$. Notice that, since $b_1 < \rho$ and $b_2 < 1 - \rho$, no player opts out in equilibrium. It

[6] Usually, this possibility is neglected by assuming that the rejection of a proposal or the discovery of a new partner dissolves the partnership.

remains to consider the deadlock point d. Matters are less simple than in section 7. We take

$$d_1 = \frac{1}{2(\rho_1 + \lambda)} \{\lambda_2 b_1 + \lambda_1(1 - b_2)\}$$

$$d_2 = \frac{1}{2(\rho_2 + \lambda)} \{\lambda_1 b_2 + \lambda_2(1 - b_1)\}.$$

The assumption is that, when *two* sellers are matched simultaneously with one buyer, the house is sold at a price equal to a seller's outside option b_1. The buyer then gets $1 - b_1$. Similarly, if *two* buyers are matched with one seller[7]. The deadlock payoffs are then calculated by considering the consequences of a matched buyer and seller continuing to negotiate without reaching agreement until one finds a second partner.

The equilibrium price is then found by solving the equation

$$p = \left(\frac{\alpha}{\alpha + \beta}\right)(1 - d_2) + \left(\frac{\beta}{\alpha + \beta}\right) d_1.$$

In the case when $\rho_1 = \rho_2 = \rho$, $p = (\lambda_1 + \rho)/2(\lambda + \rho)$, and so the unit of surplus gets divided in the ratio $(\lambda_1 + \rho) : (\lambda_2 + \rho)$. This is the conclusion reached in Rubinstein/Wolinsky [16][8].

Opting out plays no role in the preceding discussion. One may, however, follow Gale (1986), and enrich the model by replacing the assumption that agents get zero utility from leaving the market by something more realistic. To this end, continuous, strictly increasing functions $S : [0, 1] \to \mathbb{R}$ and $B : [0, 1] \to \mathbb{R}$ are introduced. In each period, it is assumed that $S(1)\tau$ and $B(1)\tau$ are the measures of sellers and buyers who appear in the market in one period[9]. The quantity $S(x_1)\tau$ is interpreted as the measure of these new sellers who can get a utility of at most x_1 outside the market. A similar interpretation applies to $B(x_2)\tau$.

For a steady-state equilibrium, the measures of new buyers and sellers who choose not to opt out by leaving the market must be equal. In the limiting case as $\tau \to 0+$, this reduces to the requirement that $S(b_1) = B(b_2)$. The measures S_τ^* and B_τ^* of sellers and buyers in the market at the beginning of a period will consist of $S(b_1)\ \tau$ and $B(b_2)\ \tau$ together with those sellers and buyers

[7] The auction envisaged can be modeled as a non-cooperative game as in Binmore [2] or Wilson [19].

[8] But note that the same conclusion would not be reached if $\rho_1 \neq \rho_2$ because the breakdown assumptions differ.

[9] They appear after the matching move but before the opting out move.

who were in the market in the previous period but did not get matched (A matched pair will agree immediately in equilibrium). The values S^* and B^* need to be related to the rates λ_1 and λ_2 at which agents are matched. One of many possible assumptions is that there is a fixed constant $k > 0$ for which

$$\lambda_1 = \frac{kB^*}{S^* + B^*} \; ; \; \lambda_2 = \frac{kS^*}{S^* + B^*} \, .$$

The measure of sellers who get matched in a period and hence leave the market after concluding a deal is then $\tau kS^*B^*/(S^* + B^*)$. This is equal to the measure of buyers who get matched in the same period. For a steady-state, it is therefore necessary that

$$S(b_1) = B(b_2) = kS^*B^*/(S^* + B^*). \qquad [11]$$

The principle that a bargainer's outside option acts only as a constraint on the range of validity of the bargaining solution is now applied. The conclusion is that the analysis that led to the equilibrium price ρ in the case when outside options were zero remains valid. The players do not even need to be informed of their partner's outside option.

In the case $\rho_1 = \rho_2 = \rho$, the value of the equilibrium price was given in terms of λ_1 and λ_2. This allows b_1 and b_2 to be calculated in terms of λ_1 and λ_2. However, λ_1 and λ_2 are functions of S^* and B^* and so the model can be solved.

When $\rho \rightarrow 0+$ (the case of 'no search frictions'), b_1 and b_2 reduce to ρ and $1 - \rho$ respectively. The equations $q = S(\rho)$ and $q = B(1 - \rho)$ can then be interpreted as defining supply and demand curves. The equilibrium price ρ is then simply the Walrasian price. Note, however, that the «law of one price» applies even when search frictions are not negligible.

It is interesting to explore the manner in which models like that discussed here relate to classical intuitions about price formation. However, my own view is that the value of such models lies more in their capacity to provide insight into situations which are not amenable to a classical approach because relevant frictions cannot be dismissed as negligible.

9. Conclusion

This paper has largely been an attempt to convince the reader that the material it covers is fairly straightforward as a piece of theory. However, one does not need to penetrate very deeply into the theory in order to be able to apply the principles to which it leads. In particular, a wide variety of matching-and-bargaining models is amenable to the analysis outlined in Section 8.

References

Bester, H. (1988): «Bargaining, Search Costs and Equilibrium Price Distributions», *Review of Economic Studies* 55, pp. 201-214.

Binmore, K. G. (1985): «Bargaining and Coalitions», in *Game-Theoretic Models of Bargaining*, Roth, A. E. (eds.), Cambridge University Press, Cambridge.

Binmore, K. G. (1987): «Perfect Equilibria in Bargaining Models», in *The Economics of Bargaining*, Binmore G. and Dasgupta, P. (ed.), Blackwell, Oxford.

Binmore, K. G. and Dasgupta, P. (1987): *The Economics of Bargaining*, Basil Blackwell, Oxford.

Binmore, K. G. and Herrero, M. J. (1988): «Matching and Bargaining in Dynamics Markets», *Review of Economic Studies* 55, pp. 17-31.

Binmore, K. G.; Osborne, M. and Rubinstein, A. (1992): «Noncooperative Models of Bargaining», in *Handbook of Game Theory* (Volumen I), Aumann, R. and Hart, S. (eds.) North Holland, Amsterdam.

Binmore, K. G.; Rubinstein, A. and Wolinsky, A. (1986): «The Nash Bargaining Solution in Economic Modeling», *Rand Journal of Economics* 17, pp. 176-188.

Binmore, K. G.; Shaked, A. and Sutton, J. (1989): «An Outside Option Experiment», *Quarterly Journal of Economics*, pp. 753-770.

Fishburn, P. C. and Rubinstein, A. (1982): «Time Preference», *International Economics Review* 23, pp. 677-694.

Gale, D. (1986): «A Simple Characterization of Bargaining Equilibriun in a Large Market without the Assumption of Dispersed Characteristics», Technical Report Working Paper 8605, Center for Analytic Research in Economics and Social Science, University of Pennsylvania.

Kalai, E. (1977): «Nonsymmetric Nash Solutions and Replications of 2-person Bargaining», *International Journal of Game Theory* 6, pp. 129-133.

Nash, J. (1950): «The Bargaining Problem», *Econometrica* 18, pp. 155-162.

Osborne, M. and Rubinstein, A. (1990): *Bargaining and Markets*, Academic Press, San Diego.

Roth, A. E. (1979): *Axiomatic Models of Bargaining*, Springer-Verlag, Berlin.

Rubinstein, A. (1982): «Perfect Equilibrium in a Bargaining Model», *Econometrica* 50, pp. 97-109.

Shaked, A. (1987): «Opting out: Bazaars versus High Tech' Markets». Technical Report Discussion Paper 87-159, Suntory Toyota International Centre for Economics and Related Disciplines.

Shaked, A. and Sutton, J. (1984): «Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model», *Econometrica* 52, pp. 1351-1364.

Wilson, R. (1984): «Notes on Market Games with Complete Information». Technical Report, Graduate School of Business, Standford University (Unpublished paper).

Wolinsky, A. (1988): «Dynamic Markets with Competitive Bidding», *Review of Economic Studies* 55, pp. 71-84.

Resumen

Este trabajo proporciona una guía compacta, aunque bastante completa, para la utilización del modelo de negociación de Rubinstein en distintas aplicaciones, prestando especial atención a la modelización de la manera en la que las negociaciones se pueden romper. El trabajo concluye con una aplicación a mercados en los que el precio se determina mediante la búsqueda y posterior negociación.

Recepción del original, marzo de 1994
Versión final, mayo de 1994