

## REVOCABLE COMMITMENT AND SEQUENTIAL BARGAINING\*

*Abhinay Muthoo*

In the theory of strategic bargaining the two central and by far the most influential models are the Nash demand game (Nash, 1953) and the Rubinstein bargaining game (Rubinstein, 1982). These two models are often thought to be incompatible. In the Nash model the bargainers make simultaneous demands, while in the Rubinstein model the bargainers make alternating offers until an agreement is reached. Furthermore, the Nash model has a multiplicity of Nash equilibria; in particular, any individually rational, Pareto efficient outcome can be supported by an equilibrium. In contrast, the Rubinstein model possesses a unique subgame perfect equilibrium.

A key feature of the Nash demand game is that the bargainers make *irrevocable* commitments to their respective demands. At the other extreme and in contrast, in the Rubinstein bargaining game the bargainers cannot make any commitment to their respective offers. Clearly, if commitments can be revoked at zero cost they would not be credible; it would be *as if* no commitments can be made (cf. the Rubinstein model). On the other hand, if the cost of revoking a commitment is infinite, then the commitment would never be revoked; it would be *as if* the commitment is *irrevocable* (cf. the Nash model). We argue that it is hard to believe that the cost of revoking a commitment can literally be either zero or infinite. Indeed, the strategic bargaining models of Nash and Rubinstein embody the two extreme and polar commitment structures. Instead, it seems compelling to model the cost of revoking a commitment as being a strictly positive and finite real number. Naturally, what happens in the two limits, as the costs of revoking become arbitrarily small and arbitrarily large, would be of some considerable interest, in view of the importance of the Nash and Rubinstein models.

In this paper we therefore present a non-cooperative bargaining model in which both bargainers can make revocable commitments and where the costs to the bargainers of revoking their respective commitments are parameters of our model. Indeed, our framework unifies the Nash demand game and the Rubinstein bargaining game. We shall, of course, explore the robustness of the equilibrium set of these two bargaining games to a perturbation in their respective commitment structures. One main result of this analysis is an equilibrium selection procedure for the Nash demand game. We shall also address and investigate the relationship between the equilibrium commitment levels and the costs to the bargainers of revoking their respective commitments.

\* I thank Giovanni De Fraja, David Ulph and Spyros Vassilakis for many helpful comments and conversations. The comments of Norman Ireland and two anonymous referees has improved the exposition at several places.

The model to be presented in section I is a two-stage game in extensive form. In the first stage both bargainers, simultaneously, can make a commitment not to accept a share of the unit size cake which is strictly less than some (strategically) chosen share. The commitment can be revoked in the second stage at some cost to the bargainer. In the second stage the bargainers negotiate over the partition of the cake according to Ariel Rubinstein's alternating-offers procedure (see Rubinstein, 1982). Due to the simultaneity of the moves in the first stage this is an imperfect information game. However, it will be assumed that there is complete information.

It is shown in section II that, if the cost to each bargainer of revoking a commitment is finite, then our two-stage bargaining game has a *unique* subgame perfect equilibrium outcome, which is Pareto efficient. Moreover, the equilibrium shares of the unit size cake (received by the bargainers) are shown to depend on the relative (and not on the absolute) magnitudes of the costs of revoking a commitment. Furthermore, we show that a bargainer's equilibrium share of the cake is strictly increasing in his cost of revoking a commitment. If, on the other hand, both bargainers can make *irrevocable* commitments (i.e., the cost to each bargainer of revoking a commitment is infinite), then in section III it is shown that the two-stage bargaining game has a *continuum* of subgame perfect equilibrium outcomes. In particular, any individually rational, Pareto efficient outcome can be supported by a subgame perfect equilibrium.

In section IV we show that the limiting (unique equilibrium) outcome, as the cost to each bargainer of revoking a commitment diverges to plus infinity, depends on the *ratio* of these costs. This ratio is not determined by our model. However, the result suggests an interpretation of the indeterminacy of the bargaining outcome when both bargainers can make *irrevocable* commitments. Basically, the indeterminacy of the outcome is due to the indeterminacy of the value of this ratio.

We argue in section V that, if both bargainers can make *irrevocable* commitments, then our two-stage bargaining game is equivalent to the Nash demand game. It thus follows from our results that an asymmetric Nash bargaining solution can be justified as the *only* equilibrium outcome of Nash's demand game that is robust to a perturbation in the commitment structure of the Nash demand game. This result provides a new justification for the Nash bargaining solution *based* on the Nash demand game. As is well known, the Nash demand game has a multiplicity of equilibria. John Nash, however, argued that the Nash bargaining solution is the *only* equilibrium outcome of his demand game that is robust to a perturbation in the *informational* structure of this game (see Binmore, 1987*a*; *b* for a much clearer analysis of this point). Finally, we conclude with some comments on the role of delegation in bilateral bargaining. Before proceeding, it should be mentioned that Vincent Crawford has studied a model also based on the notion that bargainers can make revocable commitments, where revoking a commitment is costly (see Crawford, 1982). His model and objectives are different from ours.

## I. THE MODEL

Two players, denoted by  $A$  and  $B$ , bargain over the partition of a cake of unit size. The rules of the negotiation process are modelled as a two-stage game in extensive form. In the first stage both players, simultaneously, choose a number from the closed interval  $[0, 1]$ . Let  $z_A$  and  $z_B$  denote the numbers chosen by player  $A$  and player  $B$ , respectively. We shall interpret this stage of the game as: player  $i$  ( $i = A, B$ ) commits not to accept a share of the cake which is strictly less than  $z_i$ . It will be assumed that these commitments can be revoked at the second stage of our game, but only at some cost to the player.

Given the chosen pair of numbers,  $z_A$  and  $z_B$ , the players enter the second stage of the game where they bargain according to Rubinstein's alternating-offers procedure. We assume that no time elapses between the first stage and the beginning of the second stage. The second stage is as follows. At time 0 player  $A$  makes an offer to player  $B$ . An offer is a number  $x \in [0, 1]$ ;  $x$  (resp.  $1 - x$ ) denotes the share of the unit size cake to player  $A$  (resp. player  $B$ ). Player  $B$  can either accept or reject the offer. If the offer is accepted, then the game ends with the implementation of the accepted offer. Upon rejection, player  $B$  makes a counter-offer at time  $\Delta > 0$ . If player  $A$  accepts this offer, then agreement has been struck. Otherwise, player  $A$  makes a counter-offer at time  $2\Delta$ , and so on *ad infinitum*. This completes the description of the move-structure of our two-stage bargaining game.

Player  $i$ 's ( $i = A, B$ ) payoff function will depend on the number  $z_i$  that he chose in the first stage, the share of the cake  $x_i$  that he receives (given that an agreement is reached) and the time  $t$  at which agreement is struck. A key feature that this payoff function should incorporate is that, if  $x_i < z_i$  (i.e., if player  $i$  revokes his commitment made in the first stage), then player  $i$  should incur some cost. In this paper we shall adopt the following specific payoff function, which incorporates this feature in a particular manner. Two interpretations follow the description of this function. If player  $i$  ( $i = A, B$ ) chooses  $z_i \in [0, 1]$  in the first stage and if the offer  $x \in [0, 1]$  is accepted at time 0, then the payoff to player  $i$ , denoted by  $U_i(x, z_i)$ , is:

$$U_i(x, z_i) = \begin{cases} x_i & \text{if } 1 \geq x_i \geq z_i \\ x_i - c_i(z_i - x_i) & \text{if } z_i > x_i \geq 0, \end{cases}$$

where  $x_i = x$  if  $i = A$  and  $x_i = 1 - x$  if  $i = B$ , and  $c_i$  is a parameter. It is assumed that  $c_i$  is an element of the closed interval  $[0, +\infty]$ . Note that if  $c_i = +\infty$ , then  $U_i(x) = -\infty$  for all  $z_i > x_i \geq 0$ .

*Interpretations:* (1) The cost to player  $i$  of revoking his commitment (i.e., accepting a cake share  $x_i < z_i$ ) is (directly) proportional to the amount by which he has deviated from his commitment. (2) Player  $i$  (who, let us say, is the union leader if  $i = A$  and the management leader if  $i = B$ ) perceives that if he revokes his commitment, then with probability  $(z_i - x_i)$  he will be removed as leader, in which case his payoff is  $(x_i - c_i)$ , and with probability  $[1 - (z_i - x_i)]$  his payoff is  $x_i$ . Hence, if he revokes his commitment, his perceived expected payoff is:  $(z_i - x_i)(x_i - c_i) + [1 - (z_i - x_i)]x_i$  which equals  $x_i - c_i(z_i - x_i)$ .

The parameter  $c_i$  can be interpreted as representing the degree (or extent) to which player  $i$  can make a commitment stick. If  $c_i = +\infty$ , then it is as if player  $i$  can make an *irrevocable* commitment. At the other extreme, if  $c_i = 0$ , then it is as if player  $i$  cannot make any commitment at all. An intermediate value of  $c_i$  (i.e.,  $0 < c_i < +\infty$ ) captures that player  $i$  can make a commitment stick, but only to a certain degree.

We assume that the payoff to player  $i$  of reaching agreement  $x \in [0, 1]$  at time  $t \in (0, +\infty)$ , given that he chose  $z_i \in [0, 1]$  in the first stage, is  $U_i(x, z_i) e^{-rt}$ , where  $r > 0$  denotes the player's (common) rate of time preference. If the players do not reach agreement (i.e., they perpetually disagree) then each player receives a payoff of zero.

Denote by  $H(z_A, z_B)$  the (stage 2) subgame which is preceded by the players having chosen the numbers  $z_A$  and  $z_B$ . Moreover, denote the two-stage game by  $G$ . We shall assume that both the rules of the negotiation process and the payoffs, as described above, are common knowledge amongst the two players. Reinhard Selten's subgame perfect equilibrium solution concept (SPE, for short) will be employed to analyse our two-stage bargaining game  $G$ .

The concepts of an *irrevocable* commitment and *almost irrevocable* commitments will play an important role in our subsequent analysis. Thus:

**Definition 1:** Player  $i$  ( $i = A, B$ ) can make an *irrevocable* commitment if and only if  $c_i = +\infty$ .

**Definition 2:** A limiting SPE outcome of the game  $G$  as both  $c_A$  and  $c_B$  diverge to plus infinity is defined to be a SPE outcome of the game  $G$  when both players,  $A$  and  $B$ , can make *almost irrevocable* commitments.

The parameters of our two-stage game  $G$  are:  $c_A$ ,  $c_B$ ,  $r$  and  $\Delta$ , where  $0 \leq c_A, c_B \leq +\infty$  and  $0 < r, \Delta < +\infty$ . In the next section we compute the SPE of our game  $G$  for any *finite* values of  $c_A$  and  $c_B$ . Lemma 1, below, reveals that the method of proof used in that section cannot be used to compute the SPE of the game  $G$  when  $c_A = c_B = +\infty$ . However, both players,  $A$  and  $B$ , can make *irrevocable* commitments if and only if  $c_A = c_B = +\infty$  (cf. Definition 1). This case is important to our subsequent analysis. We therefore compute the SPE of our game  $G$  when  $c_A = c_B = +\infty$  separately in section III. We shall not examine the equilibria of our game  $G$  when  $c_i = +\infty$  and  $c_j < +\infty$  where  $i \neq j$  and  $i, j = A, B$ .

## II. FINITE COSTS OF REVOKING

### II.1 *Equilibrium in Stage 2*

The set of feasible payoff pairs at time 0 in the subgame  $H(z_A, z_B)$  denotes, in utility terms, the feasible set of agreements at time 0. Denote it by  $X_0$ . Formally,  $X_0 = \{(u_A, u_B) : \exists x \in [0, 1] \text{ s.t. } U_i(x, z_i) = u_i \text{ where } i = A, B\}$ . The following lemma (whose proof can be obtained from the author upon request) enables us to use the results contained in Binmore (1987a; c) on the SPE of Rubinstein-type bargaining games.

**LEMMA 1.** *Let  $(z_A, z_B)$  be an arbitrary element of the set  $[0, 1] \times [0, 1]$ . Then, for any finite values of  $c_A$  and  $c_B$  (i.e.,  $0 \leq c_A, c_B < +\infty$ ) the set  $X_0$  is the graph of a bijective,*

strictly decreasing and concave function mapping the closed interval  $[-c_A z_A, 1]$  into the closed interval  $[-c_B z_B, 1]$ .

Lemma 2, below, states that for any  $(z_A, z_B) \in [0, 1] \times [0, 1]$  the subgame  $H(z_A, z_B)$  has a unique SPE outcome, for any finite values of  $c_A$  and  $c_B$  and for any strictly positive values of  $r$  and  $\Delta$ . The proof of this lemma is based on the following observations. First of all, in the subgame  $H(z_A, z_B)$ , the set of feasible agreements (in utility terms) at time  $t$ , denoted by  $X_t$  (where  $X_t \subset \mathbb{R}^2$ ), is such that for  $t = \Delta, 2\Delta, \dots$ ,  $X_t = \{(u_A e^{-rt}, u_B e^{-rt}) : (u_A, u_B) \in X_0\}$ . Secondly, from lemma 1, for any finite values of  $c_A$  and  $c_B$  the set  $X_0$  is the graph of a bijective, strictly decreasing and concave function mapping the closed interval  $[-c_A z_A, 1]$  to the closed interval  $[-c_B z_B, 1]$ . It then follows from arguments contained in Binmore (1987a; c) that the subgame  $H(z_A, z_B)$  has a unique SPE outcome.

**LEMMA 2.** *Let  $(z_A, z_B)$  be an arbitrary element of the set  $[0, 1] \times [0, 1]$ . Then, for any finite values of  $c_A$  and  $c_B$  (i.e.,  $0 \leq c_A, c_B < +\infty$ ) and strictly positive values of  $r$  and  $\Delta$  (i.e.,  $0 < r, \Delta < +\infty$ ) the subgame  $H(z_A, z_B)$  has a unique SPE outcome.*

In fact, if  $X_0 \cap \mathbb{R}_+^2 \neq \emptyset$ , then, in any SPE, agreement is reached with no delay at time 0. If, on the other hand,  $X_0 \cap \mathbb{R}_+^2 = \emptyset$ , then, in any SPE, the players never reach agreement (i.e., there is perpetual disagreement).

In principle, characterising this unique SPE outcome of the subgame  $H(z_A, z_B)$  is not a problem. Rubinstein (1982) and Binmore (1987a; c) have provided the necessary methods of characterisation. It turns out, however, that these methods when applied to the subgame  $H(z_A, z_B)$  generate a large number of tedious and lengthy computations. There is an alternative, less costly, method of providing the characterisation of the limiting (unique SPE) outcome as  $\Delta \rightarrow 0+$ . Since, in this paper, we wish to minimise the impact of time on the bargaining outcome and instead focus on the role of commitment (as captured by the parameters  $c_A$  and  $c_B$ ) our prime interest is precisely on the limiting (unique SPE) outcome as  $\Delta \rightarrow 0+$ . It follows from arguments contained in Binmore (1987a; c) that the unique SPE payoffs of the subgame  $H(z_A, z_B)$  will converge as  $\Delta \rightarrow 0+$  to the symmetric Nash bargaining solution, computed with the feasible set  $X_0$  and *status quo* point  $(0, 0)$ . The solution to the constrained maximisation problem (which can be obtained from the author upon request) is not terribly important in itself. It is basically an (important) step in the characterisation of the set of SPE paths of our two-stage game  $G$ ; this particular set is our main concern.

## II.2 *Equilibrium in Stage 1*

Proposition 1 (whose proof can be obtained from the author upon request) describes the Nash equilibria of the stage 1 game. In accordance with the SPE solution concept, it is assumed that for any pair of numbers  $z_A$  and  $z_B$  (chosen in the first stage) the associated payoffs to the players are the (limiting) unique SPE payoffs of the subgame  $H(z_A, z_B)$ .

**PROPOSITION 1.** (a) If  $c_A \in (0, +\infty)$  and  $c_B \in (0, +\infty)$ , then  $z_A = (1 + c_A)/(2 + c_A + c_B)$  and  $z_B = (1 + c_B)/(2 + c_A + c_B)$  constitute the unique Nash equilibrium of the stage 1 game. The equilibrium payoff to player  $i$  ( $i = A, B$ ) is  $(1 + c_i)/(2 + c_i + c_j)$  where  $j \neq i$  and  $j = A, B$ .

(b) If  $c_i = 0$  and  $c_j \in (0, +\infty)$  where  $i \neq j$  and  $i, j = A, B$ , then  $z_j = (1 + c_j)/(2 + c_j)$  and any  $z_i \in [0, 1]$  constitutes a Nash equilibrium of the stage 1 game. All the Nash equilibria give player  $i$  a payoff of  $1/(1 + c_j)$  and player  $j$  a payoff of  $(1 + c_j)/(2 + c_j)$ .

(c) If  $c_A = c_B = 0$ , then any pair of numbers  $(z_A, z_B) \in [0, 1] \times [0, 1]$  constitutes a Nash equilibrium of the stage 1 game. All Nash equilibria give each player the same payoff, namely  $\frac{1}{2}$ .

**Conclusion 1.** For any finite values of  $c_A$  and  $c_B$  (i.e.,  $0 \leq c_A, c_B < +\infty$ ) the two-stage game  $G$  possesses a unique SPE outcome. The unique SPE payoffs to the players are: to player  $i$ ,  $(1 + c_i)/(2 + c_i + c_j)$  where  $i \neq j$  and  $i, j = A, B$ . In any SPE (that supports the unique SPE outcome), along the equilibrium path, the players will reach agreement on the partition of the unit size cake immediately at time 0 in stage 2. The equilibrium partition of the cake is such that player  $i$ 's share is  $(1 + c_i)/(2 + c_i + c_j)$  where  $i \neq j$  and  $i, j = A, B$ . Moreover, along the equilibrium path: (a) if both  $c_A$  and  $c_B$  are strictly positive, then neither player  $A$  nor player  $B$  will revoke their respective commitments (made in the first stage) and (b) if for some player  $i$ ,  $c_i = 0$ , then that player may revoke his commitment. Finally, note that the unique SPE outcome is Pareto efficient.

We now make three remarks. (1) For any finite values of  $c_A$  and  $c_B$  the equilibrium shares of the cake received by the players depend on the ratio  $(1 + c_A)/(1 + c_B)$ . Thus, the equilibrium shares depend on the relative (and not on the absolute) degrees to which the players can make their commitments stick. (2) Player  $i$ 's ( $i = A, B$ ) equilibrium share is strictly increasing in  $c_i$ . If weakness is interpreted as having a high cost of backing down, then this result gives formal content to Thomas Schelling's intuition that, '...in bargaining, weakness is often strength...' (Schelling, 1956, p. 282). (3) It is easy to verify that player  $A$ 's equilibrium payoff is the solution to the following constrained maximisation problem:  $\max [x^\alpha(1-x)^\beta]$  s.t.  $x \in [0, 1]$ , where  $\alpha = (1 + c_A)$  and  $\beta = (1 + c_B)$ . Hence, our two-stage game  $G$  (approximately) implements an asymmetric Nash bargaining solution, computed with feasible set  $\{(x, 1-x) : x \in [0, 1]\}$ , status quo point  $(0, 0)$  and bargaining powers  $\alpha$ , to player  $A$ , and  $\beta$ , to player  $B$ . Thus, it is the bargaining powers that are affected by the parameters  $c_A$  and  $c_B$ .

### III. IRREVOCABLE COMMITMENTS

We now compute the SPE of the game  $G$  when both players,  $A$  and  $B$ , can make irrevocable commitments (cf. Definition 1).

#### III.1 Equilibrium in Stage 2

**LEMMA 3.** Let  $c_A = c_B = +\infty$ . If  $(z_A, z_B) \in [0, 1] \times [0, 1]$  is such that  $z_A + z_B > 1$ , then in any Nash equilibrium (and hence, in any SPE) of the subgame  $H(z_A, z_B)$  the players never reach an agreement. The unique Nash equilibrium (and SPE) payoff to the player  $i$  ( $i = A, B$ ) is zero (i.e., the payoff associated with perpetual disagreement).

The proof of this lemma follows from the fact that if player  $i$  ( $i = A, B$ ) accepts a share of the cake  $x_i < z_i$ , then player  $i$  receives a payoff of  $-\infty$  (since  $c_i = +\infty$ ). Hence, there does not exist a partition of the unit size cake that can be acceptable to both players. In other words,  $X_0 \cap \mathbb{R}_+^2 = \emptyset$ .

**LEMMA 4.** *Let  $c_A = c_B = +\infty$ . Then, if  $(z_A, z_B) \in [0, 1] \times [0, 1]$  is such that  $z_A + z_B = 1$ , then the unique SPE payoff to the player  $i$  ( $i = A, B$ ) in the subgame  $H(z_A, z_B)$  is  $z_i$ .*

Again, the proof of this lemma is based on the fact that player  $i$  ( $i = A, B$ ) will not accept a share of the cake strictly less than  $z_i$  (since  $c_i = +\infty$ ). Hence, the only partition of the unit size cake that is acceptable to both players is such that for  $i = A, B$ ,  $x_i = z_i$ . In other words,  $X_0 \cap \mathbb{R}_+^2 = \{(z_A, z_B)\}$ . In any SPE this will be implemented at time 0.

### III.2 Equilibrium in Stage 1

**PROPOSITION 2.** *Let  $c_A = c_B = +\infty$ . Then any pair of numbers  $(z_A, z_B) \in [0, 1] \times [0, 1]$  such that either  $z_A + z_B = 1$  or  $z_A = z_B = 1$  constitutes a Nash equilibrium of the stage 1 game. The equilibrium payoffs are as follows. If  $z_A + z_B = 1$ , then player  $i$ 's ( $i = A, B$ ) payoff is  $z_i$ . And, if  $z_A = z_B = 1$ , then player  $i$ 's ( $i = A, B$ ) payoff is zero.*

*Proof.* Let  $c_A = c_B = +\infty$ . Any  $(z_A, z_B) \in [0, 1] \times [0, 1]$  such that  $2 > z_A + z_B > 1$  does not constitute a Nash equilibrium of the stage 1 game, since there exists a profitable deviation for at least one of the players. For such a pair of numbers both players receive a payoff of zero (cf. lemma 3). Since  $2 > z_A + z_B > 1$  it implies that for some  $i$  ( $i = A$  or  $B$ )  $z_i < 1$ . Thus, if player  $j$  ( $j \neq i$ ) deviates from  $z_j$  and instead chooses  $1 - z_i$ , then player  $j$  gets a payoff of  $1 - z_i$  (cf. lemma 4) which is strictly positive.

By a similar argument, any  $(z_A, z_B) \in [0, 1] \times [0, 1]$  such that  $z_A + z_B < 1$  does not constitute a Nash equilibrium of the stage 1 game. Let such a pair of numbers be given. Then, in any SPE of the subgame  $H(z_A, z_B)$ , each player  $i$  will receive at least a payoff equal to  $z_i$  (since  $c_i = +\infty$ ) and at most a payoff equal to  $1 - z_j$  (since  $c_j = +\infty$ ). If player  $i$ 's payoff is strictly less than  $1 - z_j$ , then he can profitably deviate from  $z_i$ , by instead choosing  $1 - z_j$ , because, from lemma 4, it follows that his payoff would now be equal to  $1 - z_j$ . If player  $i$ 's payoff equals  $1 - z_j$ , then it is player  $j$  who can profitably deviate.

$z_A = z_B = 1$  constitutes a Nash equilibrium, because neither player  $A$  nor player  $B$  can make a profitable deviation (cf. lemmas 3 and 4). Similarly, any  $(z_A, z_B) \in [0, 1] \times [0, 1]$  such that  $z_A + z_B = 1$  constitutes a Nash equilibrium. Again, there does not exist a profitable deviation for either player  $A$  or player  $B$  (cf. lemmas 3 and 4). Q.E.D.

**Conclusion 2.** If both players,  $A$  and  $B$ , can make *irrevocable* commitments, then our two-stage bargaining game  $G$  possesses a continuum of SPE outcomes. In particular, any individually rational, Pareto efficient outcome can be supported by a SPE.

## IV. ALMOST IRREVOCABLE COMMITMENTS

Corollary 1, below (which follows from Proposition 1 (a)) describes the set of SPE outcomes of the game  $G$  when both players,  $A$  and  $B$ , can make *almost irrevocable* commitments; Definition 2, in section 2, defines the concept of *almost irrevocable* commitments.

COROLLARY 1. Let the ratio  $c_A/c_B \rightarrow \lambda$  as  $c_A \rightarrow +\infty$  and  $c_B \rightarrow +\infty$ , where  $\lambda$  is an arbitrary number from the closed interval  $[0, +\infty]$ . Then, as  $c_A \rightarrow +\infty$ ,  $c_B \rightarrow +\infty$  and  $c_A/c_B \rightarrow \lambda$ , the unique SPE payoff pair of the game  $G$  converges to  $(u_A^*, u_B^*)$ , where:

$$\begin{aligned} u_A^* &= \lambda/(1+\lambda) \text{ and } u_B^* = 1/(1+\lambda) & \text{if } \lambda < +\infty, \text{ and} \\ u_A^* &= 1 \text{ and } u_B^* = 0 & \text{if } \lambda = +\infty. \end{aligned}$$

*Remark 1.* If the players can make *almost irrevocable* commitments, then the equilibrium shares (of the unit size cake) received by the players depend on the value of the parameter  $\lambda$ , i.e., they depend on the limiting value of the ratio  $c_A/c_B$  (cf. Corollary 1). Our model does not determine the value of  $\lambda$ ; it can take any value in the closed interval  $[0, +\infty]$ .

*Remark 2.* It follows from Corollary 1 and Proposition 2 that the SPE payoff pair correspondence, at the point  $(c_A, c_B) = (+\infty, +\infty)$ , is upper semi-continuous but not lower semi-continuous. In particular, note that the equilibrium payoff pair  $(0, 0)$  (cf. Proposition 2) cannot be approximated by a limiting equilibrium outcome (cf. Corollary 1).

*Remark 3.* Take any individually rational, Pareto efficient outcome of our game  $G$ . If the players can make *irrevocable* commitments, then there exists a SPE outcome of  $G$  which coincides with this outcome (cf. Proposition 2). Furthermore, if the players can make *almost irrevocable* commitments, then the unique SPE outcome of  $G$  also coincides with this outcome *provided* we choose the appropriate value for  $\lambda$  (cf. Corollary 1). We therefore suggest the following interpretation of the indeterminacy of the bargaining outcome when the players can make *irrevocable* commitments. The game  $G$  possesses a continuum of SPE outcomes because the value of the parameter  $\lambda$  has not been *pinned down*; if  $\lambda$  is given a particular value, then  $G$  would indeed have unique SPE outcome (as described by Corollary 1). Basically, the indeterminacy of the bargaining outcome is due to the indeterminacy of the value of  $\lambda$ .

## V. THE NASH DEMAND GAME

We now argue that our results indicate that an asymmetric Nash bargaining solution can be justified as the *only* equilibrium outcome (of Nash's demand game) that is robust to a perturbation in the *commitment* structure of the Nash demand game. Our argument is based on the following observation. If both players,  $A$  and  $B$ , can make *irrevocable* commitments (i.e., if  $c_A = c_B = +\infty$ ), then our two-stage bargaining game  $G$  is equivalent to the Nash demand game. Indeed, if the players' demands are incompatible (i.e., if  $z_A + z_B > 1$ ), then (as in the Nash demand game) the outcome is the *disagreement* outcome (cf. lemma 3). And, if  $z_A + z_B = 1$ , then again, as in the Nash demand game)



the players receive their respective demands, namely  $z_A$  and  $z_B$  (cf. lemma 4). The difference between our game  $G$  and Nash's demand game lies when  $z_A + z_B < 1$ . In Nash's game the players again receive only their respective demands, while in our game this need not be the case; it may be that at least one player receives more than his demand. However, this difference is not significant; the Nash equilibria of Nash's demand game exactly coincide with the SPE of our game  $G$  when  $c_A = c_B = +\infty$  (cf. Proposition 2).

We interpret this as follows. In the Nash demand game both players can make *irrevocable* commitments. Consequently, if both players can make *almost irrevocable* commitments, then our game  $G$  is equivalent to a perturbed Nash demand game; the perturbation is in the *commitment* structure of the Nash demand game. Hence, it follows from our results (cf. Corollary 1) that, the perturbed Nash demand game possesses a *unique* equilibrium. The outcome supported by this equilibrium depends on the value of the parameter  $\lambda$ . Furthermore, the equilibrium outcome can be described by an asymmetric Nash bargaining solution, computed with feasible set  $\{(x, 1-x) : x \in [0, 1]\}$ , *status quo* point  $(0, 0)$  and *bargaining powers*  $\lambda/(1+\lambda)$ , to player  $A$ , and  $1/(1+\lambda)$ , to player  $B$ .

## VI. DELEGATION

We conclude this paper with a comment on the role of delegation in bilateral bargaining. Quite often it is observed that bargainers do not negotiate directly with each other, but instead hire professional negotiators on their behalf. Why is this the case? A simple extension of our two-stage game (followed by a reinterpretation) provides some formal insight.

### VI.1 A Three-Stage Game

We introduce an additional stage, which precedes the two-stage game  $G$ . In this first stage both players, simultaneously, choose a number from the interval  $[0, +\infty)$ . The number chosen by player  $i$  ( $i = A, B$ ) denotes the value of  $c_i$ . Thus,  $c_i$  (a parameter of the game  $G$ ) is now determined endogenously. It is assumed that a player incurs a cost of  $c^2$  in choosing a number  $c \in [0, +\infty)$ ; the cost is incurred in the first stage. Assume that no time elapses between the first and the next stage.

*Interpretation.* In the first stage player  $i$  ( $i = A, B$ ) decides either to hire a professional negotiator, with characteristic  $c_i$  (where  $c_i > 0$ ), or not to hire one (i.e., put  $c_i = 0$ ). The underlying observation is that a player can make a commitment stick if and only if he hires a professional negotiator; moreover, the characteristic  $c$  (where  $c > 0$ ) represents the degree (or extent) to which a professional negotiator can make a commitment stick. A negotiator's interests is assumed to coincide with those of his employee (the player who has hired him). The fee of a professional negotiator with characteristic  $c \in (0, +\infty)$  is  $c^2$ ; the fee is paid in the first stage. If a professional negotiator is hired by a player, then it is he (rather than the player) who plays the next two stages.

## VI.2 Equilibria

For any pair of chosen numbers  $(c_A, c_B) \in [0, +\infty) \times [0, +\infty)$  the unique SPE payoffs to players  $A$  and  $B$  in the game  $G$  is: to player  $A$ ,  $(1 + c_A)/(2 + c_A + c_B)$  and to player  $B$ ,  $(1 + c_B)/(2 + c_A + c_B)$  (see Proposition 1 and Conclusion 1). These are also their respective equilibrium shares of the unit size cake.

Thus, in accordance with the SPE solution concept, the payoff to player  $i$  ( $i = A, B$ ) if he chooses  $c_i$  and his opponent (player  $j$  where  $j \neq i$  and  $j = A, B$ ) chooses  $c_j$  is  $f_i(c_i, c_j)$  where:  $f_i(c_i, c_j) = [(1 + c_i)/(2 + c_i + c_j) - c_i^2]$ .

**PROPOSITION 3.** *In the unique SPE outcome of the three-stage game both, player  $A$  and player  $B$ , hire professional negotiators with the same characteristic  $\hat{c}$ , where  $\hat{c} > 0$ .*

*Proof.* By differentiating the function  $f_i$  with respect to  $c_i$  ( $i = A, B$ ) we obtain first-order conditions, namely  $c_A/(1 + c_B) = 1/[2(2 + c_A + c_B)^2]$  and  $c_B/(1 + c_A) = 1/[2(2 + c_A + c_B)^2]$ . Solving these equations for  $c_A$  and  $c_B$  we obtain that  $c_A = c_B = \hat{c}$ , where  $\hat{c} = (\sqrt{6} - 2)/4$  which is *strictly* positive. The second-order conditions are satisfied. Q.E.D.

*Conclusion 3.* It follows from Proposition 3 that each player ( $A$  and  $B$ ) will receive (in equilibrium) one half of the unit size cake. Moreover, each player incurs the cost of hiring the professional negotiator with characteristic  $\hat{c}$  (where  $\hat{c} > 0$ ). Hence, the equilibrium payoff to each player in this three-stage game is  $1/2 - \hat{c}^2$  (which is strictly less than  $\frac{1}{2}$ ). Therefore, the outcome of this game is Pareto inefficient. Indeed, both players would prefer that both players do not hire professional negotiators (i.e., put  $c_A = c_B = 0$ ) and instead bargain directly. However, it is the case that for each player  $i$  the choice of hiring a professional negotiator (i.e., putting  $c_i > 0$ ) strictly dominates the choice of not hiring one (i.e., putting  $c_i = 0$ ). Thus, we conclude that a player hires a professional negotiator *not* because it gives the player a higher payoff, but because it is part of the noncooperative *equilibrium* of the game.

*University of Bristol*

*Date of receipt of final transcript: August 1991*

## REFERENCES

- Binmore, K. G. (1987a), 'Nash bargaining theory I & II', in Binmore and Dasgupta (1987).
- (1987b), 'Nash bargaining and incomplete information', in Binmore and Dasgupta (1987).
- (1987c), 'Perfect equilibria in bargaining models', in Binmore and Dasgupta (1987).
- and Dasgupta, P. S. (1987), *The Economics of Bargaining*, Oxford: Basil Blackwell.
- Crawford, V. (1982), 'A theory of disagreement in bargaining', *Econometrica*, vol. 50, pp. 607–38.
- Nash, J. (1953), 'Two person cooperative games', *Econometrica*, vol. 21, pp. 128–40.
- Rubinstein, A. (1982), 'Perfect equilibrium in a bargaining model', *Econometrica*, vol. 50, pp. 97–109.
- Schelling, T. C. (1956), 'An essay on bargaining', *American Economic Review*, vol. 46, pp. 281–306.