Centre de Referència en Economia Analítica

Endogenous bargaining power

Joan Esteban and József Sákovics

July, 2003

Endogenous bargaining power[*]

Joan Esteban

Institut d'Anàlisi Econòmica, CSIC

and Universitat Pompeu Fabra, Barcelona

József Sákovics[§]

University of Edinburgh

Abstract

We present a novel approach to N-person bargaining, based on the idea that the agreement reached in a negotiation is determined by how the direct conflict resulting from disagreement would be resolved. Our basic building block is the *disagreement function*, which maps each set of feasible outcomes into a disagreement point. Using this function and a weak axiom based on individual rationality we reach a unique solution: the *agreement in the shadow of conflict, ASC*. This agreement may be construed as the limit of a sequence of partial agreements, each of which is reached as a function of the parties' relative *power* in the non-cooperative scenario.

Key words: Bargaining, conflict, disagreement.
JEL Numbers: C78, D74.

[§] Corresponding author: Edinburgh School of Economics, University of Edinburgh, 50 George Square, EDINBURGH, EH8 9JY. E-mail: jozsef.sakovics@ed.ac.uk

«The rich get the law passed by means of force and arms or get it accepted by fear to their might, aren't things this way?» Plato, *Republic*.

1. Introduction

Standard bargaining theory arrives at solutions in two steps. The first step consists in the reduction of a bargaining situation into the confines of a *bargaining problem*, defined by Nash (1950) as the set S of feasible utility allocations and the threat point d. The latter is meant to be the outcome of some (presumably) wasteful interaction that follows disagreement. Thus, different "disagreement games" – reflecting altered distributions of power among the players – are summarized as different threat points. The second step –which has concentrated the efforts of bargaining theory proper so far – consists in selecting a solution to this simplified problem. The solution to this second step is based on either a set of plausible axioms or on the outcome of a posited extensive-form game that is completely independent from the one which determines the disagreement point. This game might admit differential bargaining power by the different agents, as in the generalized Nash solution. But the essential point is that current bargaining theory establishes no link between the power of the players in the first step – determining d – and in the second step, when fixing the shares of the surplus over and above d. Our paper is an attempt at integrating the above two steps in a consistent manner.

Specifically, we consider that there is no other source of differential power than the one underlying the non-cooperative game supporting the disagreement point. That is, our interpretation of *bargaining power* is that it coincides with *power* itself in the fully non-cooperative scenario. In order to model this, we need to incorporate more data from the bargaining situation into a (generalized) bargaining problem. Instead of limiting the transmitted information to the solution of the disagreement game, we incorporate into the description a reduced form of the disagreement game itself. As it turns out, all the relevant information can be summarized by the specification of how the outcome of conflict, d, varies as a function of the stakes, S. This is the game-specific *disagreement function*, $D(.)$, which maps sets of payoffs into the corresponding equilibrium of the

1

disagreement game.[1] A generalized bargaining problem is thus completely described by a pair *(S, D(.))*: the set of the pay-offs initially available and the disagreement function providing the outcome of disagreement for any subset of these pay-offs.

The nature of disagreement games depends on the problem at hand. In some situations, the underlying disagreement game is so rudimentary that players do not even have a choice over alternative strategies. Consider, for instance, bargaining over the price of an object in the middle of a bazaar. If the players do not reach an agreement, the potential buyer walks out and goes to the next shop. However, such extremely simple situations seem the exception rather than the rule. In the previous case, it is essential that players terminate any future relationship after reaching disagreement. Whenever players do not cease to interact, the disagreement game is necessarily richer. Social relationships are of this type. The fact that we may fail to achieve a particular collective agreement simply means that the future relationship among agents will be non-cooperative. The same can be said of oligopolistic markets, industrial disputes, or simply of individuals litigating over a particular issue of their concern. This is also the case in the international arena, where one cannot modify who are one's neighbors. In all these important instances, players have a menu of potential strategies to follow in the non-cooperative mode.

This paper deals with bargaining situations characterized by truly strategic disagreement games. For this class of bargaining situations one simple axiom – essentially positing individual rationality – permits the characterization of a unique and efficient solution –the agreement in the shadow of conflict, ASC. The axiom of the Independence of Individually Irrational Alternatives, IIIA, simply states that the agreement should not depend on the availability of alternatives that give to at least one player strictly less than what she would get in disagreement and hence are not individually rational. This axiom is in fact a weakening of Nash's axiom of Independence of Irrelevant Alternatives, though here it is applied in the context of a generalized bargaining game, what increases its bite.

---

[1] Note that this additional information was already required for the determination of *d* in the standard context, since the knowledge of the game is necessary to find its equilibrium. Once the game is well defined, it is straightforward to calculate its equilibria under different hypotheses (*S*'s).

The key observation driving our result is that once we eliminate the individually non-rational agreements, the bargaining problem becomes a different one – with a new bargaining set reduced to the remainder. Via the disagreement function, the new bargaining set yields a new threat point as well. Since our axiom applies to all bargaining problems, it also applies to this new (continuation) one, and further reduces the set of feasible agreements. What we show is that the repeated application of the axiom to the resulting sequence of bargaining games converges to a situation where the disagreement outcome is efficient, thus pinpointing a unique solution.

To fix ideas, consider the simple example of splitting an inheritance of, say, ten euros, between two siblings (who do not fancy each other). The siblings can either agree on a particular split at no cost, or disagree and engage in a costly dispute over the money. Suppose that, if players engaged in conflict, in equilibrium seven euros would be wasted (on, say, lawyers' fees), while of the remaining three euros one player would expect to obtain two and the other one. This allocation may reflect the fact that, for instance, one's lawyer is "twice" as influential as the other's.[2] As a result of the expected outcome of conflict, any agreement must give to the siblings at least two and one euros, respectively. Recognising this, they are willing to get to a partial agreement, which guarantees them these outside payoffs. Consequently, the effective area of dissent shrinks to the remaining seven euros, which are precisely the benefits from cooperation. On the division of these seven euros the siblings may again either agree or disagree and engage in a dispute. In the dispute, say, four euros would be wasted and the strong sibling would obtain two and the weak one. Notice that even if they disagree, both siblings are better off by respecting their partial agreement and restricting the dispute to the distribution of the seven-euro surplus. It thus follows that any agreement must give to the siblings at least four and two euros, respectively. This observation generates a new partial agreement. Applying the argument

---

[2] For example, the expected division ruled by the court may be 7:3, but the cost of the better lawyer is 5 why the cost of the worse one is only 2.

repeatedly, we reach the final agreement, where the ten euros are distributed according to the power of the parties in the conflict game:[3] 20/3 and 10/3.

The argument above provides an attractive interpretation of the negotiation as a *process*[4] where, driven by the fear of a conflictual resolution, the parties accept to gradually narrow down the extent of their dissent.[5] Along each step of this process, it is the relative power of the players, as embodied in the disagreement function, what shapes the solution. We prove that for a very rich class of games, perfectly informed, rational agents will accept to reduce the area of their dissent completely: they will reach an agreement.

We also clarify the connection between the ASC, the generalized Nash and the Rubinstein solutions. Restricting attention to proportional disagreement functions (what corresponds to fixed discount factors in the Rubinstein-type models) and to the unit simplex as the Pareto frontier, we show that the ASC solution coincides with the asymmetric Nash solution, where the ratio of bargaining weights is equal to the proportion of the disagreement utilities.

The paper is structured as follows. In the next section, we develop our theory of bargaining in the shadow of conflict. We discuss the properties of disagreement

---

[3] In contrast, both the Nash (1950) and the Kalai-Smorodinsky (1975) solutions would predict that the seven-euro surplus over and above the (total) disagreement point would be brotherly shared by the two players. They would obtain 5.5 and 4.5 pounds in total, respectively.

[4] This process may be an actual one or just a thought process, which directly leads the players to agreement.

[5] Indeed, we observe that even in the cases in which players do not reach agreement and go into playing the conflict game –think of the extreme case of wars– they do accept restricting its amplitude. Thus, countries accept not to bomb civilian targets or to abstain from the use of particularly harmful weapons. In some cases – think of India and Pakistan – they tacitly agree to keep the conflict as mere border skirmishes. What keeps the conflict from escalation is the separation between the agreement and conflict *games:* not respecting a (partial) agreement is not a unilateral deviation *in* the conflict game; instead it is a unilateral deviation provoking a transition *to* the conflict game. This way such a deviation is observable: the countries foresee each other's reaction to a unilateral deviation. For example, according to our solution, in a complete information Cournot model, two identical firms would each agree to produce half the monopoly quantity, which is indeed the optimal colluding outcome (for them). The Nash equilibrium would correspond to unrestricted conflict (that is, competition) in this case.

4

functions, present the IIIA axiom, characterize the bargaining solution and prove that it is unique. We also compare the ASC and Nash solutions for feasible sets with a linear Pareto frontier when disagreement functions belong to a standard class. In section 3 we establish the connection between "bargaining power" in asymmetric Nash solutions and the power of the players in the disagreement game. In section 4 we use a family of conflict games to provide an application, which also yields an alternative characterization. Finally, we provide a discussion of related bargaining literature, followed by some concluding remarks.

2. A disagreement theory of bargaining

Suppose that there are *N* players, who wish to reach an agreement in $S^0 \in \Sigma$, where Σ is the set of convex, compact subsets of the utility[6] space, $\Re_+^N$. Assume further the existence of a *disagreement function*, *D(.)*, which assigns a disagreement point, *d,* to every convex,, compact subset of S^0. That is, if the set of alternatives considered were *S*, the outcome of disagreement would be *d = D(S)∈ S*. This function is to be interpreted as shorthand for the solution[7] to the underlying *non-cooperative game* that agents will play in case they fail to reach an agreement. We would like to stress that *D(S)* may depend on additional parameters, especially those related to the players' "strength", which form part of the description of this conflict game. A *bargaining problem in the shadow of conflict* (BPSC) is then completely described by the pair $\left(S^0, D(.)\right)$.

  We need not impose any structure on *D(.)*, since it is meant to be a positive description of some real underlying conflict situation and therefore it cannot be freely chosen by the modeler.

---

[6]   Actually, for our analysis it is not necessary that preferences satisfy the von Neumann-Morgenstern axioms. We could directly phrase our model in terms of money, prestige or the like. We elaborate on this issue in the Conclusions.

[7]   This solution maybe a unique Nash (subgame-perfect?) equilibrium, but uniqueness of equilibrium is not necessary. In case of multiplicity, the "disagreement outcome" can be defined as the meet of the utilities gained at the different equilibria.

Let B denote the set of all BPSCs. A *solution* for a BPSC is then a mapping, *f: B* $\rightarrow \Sigma$, satisfying $f(S,D(.)) \subseteq S$ for all $S \in \Sigma$. That is, the solution selects a subset of the alternatives as acceptable.

Let $S_x = \{s \in S \mid s \geq x\}$. That is, S_x is the subset of S which weakly Pareto dominates x.

As long as players act rationally, any solution should weakly Pareto dominate the disagreement outcome, since otherwise at least one player would prefer to provoke disagreement. In other words, from the knowledge that players are rational we can deduce that any agreement on S should be a member of the set S_d. Our only assumption is that any eventual agreement on S should not be altered if we eliminate all the alternatives that cannot be candidate solutions under individual rationality (i.e. the complement of the set S_d).

We thus impose the following axiom:

Axiom 1: Independence of Individually Irrational Alternatives (IIIA):

$f(S,D(.)) = f(S_{D(S)},D(.))$ *for all* $(S,D(.)) \in B.$

Conceptually, IIIA is much weaker than Nash's Independence of Irrelevant Alternatives (IIA) axiom, since it only eliminates a subset of his "irrelevant alternatives" and the definition of this subset makes no reference to the final solution. Placed on a standard bargaining problem (SBP), IIIA would simply eliminate the alternatives that do not weakly dominate the disagreement point. However, when applied to a BPSC, IIIA has a recursive effect: once we eliminate the individually irrational alternatives, the application of the disagreement function to the remaining set results, in general, in a different disagreement point than before. To this new BPSC the axiom also applies (note that, if $(S,D(.)) \in B$, then $(S_{D(S)},D(.)) \in B$ as well). Thus, as long as $D(.)$ is not constant (as in a SBP), the application of IIIA generates new BPSCs which, in turn, also have to satisfy the axiom.

In view of its recursive implications, should we still find IIIA a plausible axiom? We certainly think so. The point of all "irrelevant alternatives" type axioms is to provide

some consistency between solutions of the same underlying bargaining situation but with different sets of available agreements. In our view, the appropriate description of the bargaining situation should not be confined to a fixed disagreement point, since the outcome of disagreement is likely to depend on the alternatives available. Therefore, what should be kept fixed when carrying out the "consistency check" is the disagreement *function*, just as it is done in IIIA. That is, our assumption compares bargaining situations where the same set of players are bargaining in the shadow of the same conflict game but with different sets of feasible utility pay-offs.

Definition 1 *The agreement in the shadow of conflict solution (ASC) assigns to each BPSC the maximal set that is consistent with IIIA.*

Our first result shows that the requirement imposed on the solution is not too stringent –that is, for every BPSC there exists a non-empty set of agreements consistent with IIIA.

Proposition 1 *For every BPSC there exists a well-defined ASC bargaining solution.[8] Moreover, the set of agreements in the shadow of conflict is always non-empty.*

Proof: Let us look at the implications of Definition 1. Let $f^*(.,.)$ be a bargaining solution, $S^0 \in \Sigma$ an arbitrary bargaining set and $D(.)$ a disagreement function. IIIA implies that f^* $(S^0,D) = f^*(S^o_{d^0},D)$, where $d^0 = D(S^0)$. The disagreement point corresponding to the set $S^o_{d^0}$, however, is not d^0 but it is given by $d^1 = D(S^o_{d^0})$. Thus, the application of IIIA results in a new set, S^1. Repeatedly eliminating the individually irrational alternatives, for the t-th iteration we will have $S^t = \{u \in S^{t-1} | \ u \geq d^{t-1}\}$. IIIA requires exactly that for all the sets of this sequence, when coupled with $D(.),$ the solution be the same. In other words, a bargaining solution satisfies IIIA if and only if $f^*(S^0,D) \subseteq$

$$S^* = lim_{T \to \infty} \ \prod_{t=0}^{T} \ S^t \ .[9]$$ Thus the ASC solution is defined as the limit set, S^*.

Note now that the sets S_t are compact and nested. Therefore their intersection is uniquely defined and, by Tychonov's theorem, it is non-empty as well. Q.E.D.

Proposition 1 shows that, independently of the exact form it takes, just the conceptual increase in the informational content of the description of the bargaining problem is sufficient to provide us with a set of "acceptable" agreements. In general, these agreements need not be unique. Whether the solution is determinate or not depends on the nature of the disagreement game. We shall now prove that for strategic disagreement games the above result can be strengthened: the ASC solution singles out a unique, Pareto efficient agreement.

The amount of information contained in the disagreement function depends on whether the non-cooperative game allows for players to choose over alternative strategies. In the games in which no player has any choice or all players are indifferent over all choices, the equilibrium pay-offs convey little, if any, information on the characteristics of the players. We wish to single out the class of strategic games in which at least one player can choose over at least two strategies and thus avoid the worst outcome possible. Formally, this amounts to the following assumption.

Assumption 1 *Unless S is singleton, the disagreement outcome is strictly preferred to her worst agreement in S by at least one player: for all $S \subseteq S^0$ such that $S \in \Sigma$, there exists $z \in S$ such that $z_i < D_i(S)$ for some $i \in \{1,2,...,N\}$.*

Assumption 1 is satisfied by all non-cooperative games in which there is at least one player that has a choice over a set of possible strategies and that in equilibrium is not indifferent to all of them.[10] Examples of this class of games abound: in pre-trial bargaining the lawyer's fees are often set as a percentage of the amount under dispute; in collusive agreements in a market setting, even if there is cut-throat Bertrand competition, unless the firms are identical, there is always positive profits for the more efficient firm;

---

[10] Notice that the familiar case of bargaining over the price to be paid for an object to be traded violates in principle Assumption 1.

in conflict models with endogenous choice of effort there is usually a unique interior Nash equilibrium, etc. [11]

Next, we assume continuity of the equilibrium pay-offs of the disagreement game relative to small perturbations in the set of feasible allocations S: small changes should not provoke major changes in the outcome of disagreement.

Assumption 2 *D is continuous in the Hausdorff topology: if a sequence of elements of Σ converges to S in the Hausdorff topology, then the corresponding sequence of disagreement points converges to D(S).*

Without continuity the ASC solution could be set valued, since the sequence of disagreement points starting from d_0, might converge to an interior point of S. In this case, if we imposed that IIIA had to apply to the limit set S* as well, we could drop the continuity assumption and still obtain a unique and efficient agreement as long as Assumption 2 is satisfied. However, as a principle, we prefer to put more structure on the (empirically testable) disagreement game rather than to increase our normative requirements (no matter how reasonable) on the solution.

Proposition 2 *For every BPSC with D(.) satisfying Assumptions 2 and 3 the ASC bargaining solution selects a unique and efficient agreement.*

<u>Proof:</u> To see that S^* has a unique element, note that, by the continuity of $D(.)$, $\lim_{t\to\infty} D(St) = D(S^*)$, and thus $S^* = S^*_{D(S^*)}$. Suppose that S^* is not a singleton. Then, by Assumption 3, $D(S^*)$ does dominate some points in S^*. Contradiction.

By construction, each set St contains the points of the weak Pareto frontier of S_0 that dominate $D(S_{t-1})$. Therefore, the point S^* is on the frontier of S_0. This proves the efficiency of the solution. Q.E.D.

For a rich class of disagreement functions we have that a mild assumption on player's rationality is enough to identify a determinate agreement. The essential point is

---

[11] Esteban and Ray (1999) show that for a generalized version of the rent-seeking model, there always exists a *unique* Nash equilibrium and at this equilibrium each contending party expends strictly positive amounts of resources. It is straightforward to show that the disagreement point generated by the Nash equilibrium satisfies our Assumptions 1 and 2.

that by using more information on the player's characteristics, as revealed by the outcomes of the disagreement game, we can dispense with most axioms, except rationality. This argument is similar to how Economics deals with exchange. If we know the supply and demand schedules (at different prices) we can determine the equilibrium price, without having to axiomatize on what would be the "just" price.

2.1. The ASC for feasible sets with linear Pareto frontier

We now restrict to sets S that can be written as

$$S = \left\{ s \geq 0, \sum_i \beta_i s_i \leq R \right\} \quad \text{for some } R > 0 \text{ and some } \beta \text{ in the unit simplex.}$$

We shall also focus on the class of disagreement functions that satisfy the following property:

Assumption 3 *The disagreement function D(.) is invariant to affine transformations, that is, $D(\alpha + \lambda S) = \alpha + \lambda D(S)$ for all α and all $\lambda > 0$.*[12]

We can now easily compute the ASC solution to any such BPSC problem.

Proposition 3 *Let* $S^0 = \left\{ s \geq 0, \sum_i \beta_i s_i \leq R \right\}$ *and D(.) satisfy Assumptions 1 to 3. Then, the*

ASC solution satisfies

$$\frac{f_i^*(S^0, D)}{f_j^*(S^0, D)} = \frac{D_i(S^0)}{D_j(S^0)} \quad \text{for } i,j=1,2,...,N.$$

<u>Proof:</u> We know that the solution is $S^* = lim_{T \to \infty} \prod_{t=0}^{T} S^t$, where $S^t = \{u \in S^{t-1} | \ u \geq d^{t-1}\}$.

Therefore, to obtain S^* we need only to compute $d^* = lim_{t \to \infty} d^t$. Note now that any set S^{t+1} in the sequence satisfies $S^t = d^{t-1} + \lambda^t S^0$, for some appropriate λ^t. By Assumption 4 we

---

[12] It is easy to show that the endogenous contest model of Esteban and Ray (1999) mentioned earlier, satisfies this assumption whenever the Pareto frontier of the bargaining set is linear.

have that $d^t = D(S^t) = d^{t-1} + \lambda^t D(S^0)$. Therefore, $d^t = D(S^0) \sum_{r=1}^{t} \lambda^r$ and the equality in the statement of the Proposition follows immediately. Q.E.D.

In order to illustrate the differences between ASC and Nash's solutions, let us examine the case of splitting one euro. The disagreement game is as follows. Whatever the amount of money at stake, x, a fraction γ, $0 < \gamma < 1$, is lost and of the remainder $(1-\gamma)x$, a fraction α_i goes to each player $i=1,...,N$. It is straightforward that the ASC solution is to give $s_i = \alpha_i$ to each player. Clearly, α_i can be interpreted and the relative *power* of player i and γ the degree of inefficiency induced by playing the disagreement game. The ASC solution does not depend on γ because the sharing of the surplus of cooperation respects the power of the players as reflected in the disagreement game. The Nash solution instead allocates the euro as $s_i^N = (1 - \gamma)\alpha_i + \dfrac{\gamma}{N}$. The Nash solution is arrived at by combining the outcome of the disagreement game (biased by the power of the players) and the brotherly sharing of the surplus on the basis of equal division. Parameter γ is now the weight assigned by the Nash solution to each sharing rule. The Nash solution thus has the unappealing feature that the more destructive the disagreement game is, the closer the solution will be to equal division, irrespective of the relative power of the players. Further, the more biased the disagreement game the greater the discrepancy between the Nash and ASC solutions.

3. ASC as an asymmetric Nash solution

Our previous example makes it clear that ASC is a solution based on the asymmetric treatment of the players. It seems natural to explore the relationship between ASC and asymmetric Nash solutions. Recall that the asymmetric Nash solution (see Harsányi and Selten, 1972) results from the constrained maximization of a social welfare function where the individual welfare weights are supposed to embody the differential (bargaining) power of the players: $W(x,d) = \prod_{i=1}^{N}(x_i - d_i)^{\gamma_i}$. We shall now discuss the

relationship between the vector γ and the power of the parties as embodied in the disagreement function.

Recall that the asymmetric Nash solution can be characterized as the point on the Pareto frontier, where the pair-wise elasticity of this frontier is equal to the corresponding ratio of the bargaining weights.

Let us assume that the Pareto frontier of S is the unit simplex and that $D(.)$ satisfies Assumptions 1 to 3. By Proposition 3 the players distribute utilities (efficiently) in the same proportion as the disagreement function does. Consequently, the pair-wise ratio of bargaining weights corresponds to the elasticity of the Pareto frontier at the point where the utilities are distributed in the same proportion as in the disagreement point.

Proposition 4 *When the Pareto frontier of S is the unit simplex and the disagreement function satisfies Assumptions 1 to 3, the bargaining weights corresponding to the ASC solution are* $\gamma_i = D_i(S)$*, i = 1,2.*

Proof: When the Pareto frontier is the unit simplex, the marginal rate of substitution is 1, everywhere. Consequently the elasticity of the Pareto frontier is equal at every point to the ratio of the utilities at that point. By Proposition 3, this ratio is equal to the ratio of the disagreement utilities. Q.E.D.

Note that this scenario is equivalent in "richness" to the one analyzed by Rubinstein (1982), in the sense that in both models at each step of the process, the pie remaining in dispute decreases at some given proportion. In Rubinstein's alternating-offer bargaining model the unique subgame-perfect equilibrium yields an agreement as a function of the discount factors (δ_i) and the selection of the first mover. As the time between offers shrinks to zero this solution converges to the same outcome as the asymmetric Nash solution − with bargaining weights[13] $\gamma_1 = \log \delta_2$ and $\gamma_2 = \log \delta_1$ − independently of the identity of the first mover. Assuming that the Pareto frontier is the unit simplex, we can prove a similar result for the ASC solution, without having to resort

---

[13] See Binmore (1987a,b) and Binmore et al. (1986). Wilson (2001), has obtained the same result in a model with a mediator who makes random proposals.

to taking limits. That is, the ASC solution will exactly coincide with the asymmetric Nash solution, while Rubinstein's does so only in an approximate sense.

The following corollary is now immediate.

Corollary 1 *Under the Rubinstein assumptions (including the proportionality of the disagreement function), the ASC solution and the Rubinstein solution coincide if and only if* $\dfrac{D_1(S)}{D_2(S)} = \dfrac{\log \delta_2}{\log \delta_1}$.

When the disagreement function is not restricted to be proportional, our model still resembles somewhat a Rubinstein-like model, where the discount rates are not stationary (see Binmore, 1987b, for a detailed discussion of these games). Both models are still equivalent to some asymmetric Nash solution. However, the bargaining weights − just as the actual solutions − are no longer easily computable. In terms of computability, the ASC solution has a significant advantage over the Rubinstein-like one: each step in the calculation of the ASC solution improves the precision of the current estimate, and this precision is known. In contrast, to calculate the subgame-perfect equilibrium of a Rubinstein-like game, one has to work backwards from the solution, trying to end up at the disagreement point. At no point in the process, can one have a precise idea about how good the approximation is.

4. An application leading to an alternative characterization

Assume that the disagreement game is a contest with endogenous choice of effort, as in Esteban and Ray (1999). To be specific let there be two players, whose probability that Player i wins the contest (the entire pie) is given by

$$p_i = \frac{n_i r_i}{n_i r_i + n_j r_j},$$

where $n \in [0,1]$ is the intrinsic power of each party (capturing the efficiency with which she can convert effort into winning probability), while r is the effort invested in conflict. For simplicity, assume that power is normalized, so that $n_i + n_j = 1$. The cost of effort is given by $c(r) = \dfrac{r^\alpha}{\alpha}$, where $\alpha \geq 2$.

13

Suppose now that a particular sharing rule is used, giving shares β and (1-β) to each player over whatever is to be distributed. Assume that agents may decide to distribute some fraction, λ, of what is at stake accordingly with the given distributional rule and play the conflict game over the remainder. Notice that any $λ > 0$ would produce an outcome that dominates playing the conflict game over the entire cake. Clearly, a full agreement is reached when both players agree on bringing λ to unity.

We can now address the following question: are there restrictions on the values of β that guarantee that players will unanimously choose λ?

Let us compute the expected utility for both players of distributing a fraction λ under shares β and (1-β) . As it is shown in the Appendix, $u_i(\lambda) = \lambda\beta + \dfrac{(1-\lambda)n_i(\alpha - n_j)}{\alpha}$

and $u_j(\lambda) = \lambda(1-\beta) + \dfrac{(1-\lambda)n_j(\alpha - n_i)}{\alpha}$.

Observe that utility is linear in λ for both players. Therefore, it is immediate that for the two players to accept a full agreement with $λ = 1$, the necessary and sufficient conditions are that $\beta \geq \dfrac{n_i(\alpha - n_j)}{\alpha}$ and $\beta \leq 1 - \dfrac{n_j(\alpha - n_i)}{\alpha}$. Note that the RHS of both inequalities run between zero and unity as the parties power varies. Therefore, we have that there is no constant sharing rule β such that players prefer to avoid conflict altogether in all circumstances.

Let us thus consider whether there are rules in which the shares depend on the power of the parties, β(n), such that players always prefer full agreement. Letting now α tend to infinity, we observe that the RHS of both inequalities converge to n_i. Hence, the only sharing rule that makes the full agreement acceptable in all circumstances is setting $\beta(n) = n_i$. We have thus obtained that the only agreement that always avoids conflict is the one that fully recognizes the power of the parties and share what is at stake accordingly.

5. A comparative analysis

In this section, we clarify our theory by contrasting it to the most related literature.

i) Endogenous determination of the disagreement point.

In his 1953 paper, Nash proposed a generalization of his original model of 1950. In this game, known as the "variable threat" model of bargaining, the players choose threats before the actual bargaining phase, of which they serve as the disagreement point. At first blush, our model may seem just like Nash's one, with a specific, well-motivated threat game. Actually, however, our contribution goes well beyond that. There are two important differences between the models that we would like to underline:

a) Nash needs to employ an "umpire" to oblige the players to carry out their threats (in case of disagreement). We do without a $n+1^{st}$ party. The underlying reason for this is quite relevant. Nash thinks of the threat phase as one preceding the Nash bargaining game. Therefore, this phase has no interpretation on its own, it is simply a – perhaps realistic – way to make the bargaining game more detailed. In contrast, we think of our conflict subgame as one posterior to bargaining. By invoking sequential rationality, we can then analyze the players' optimal behavior in that subgame without any additional commitment device. Apart from the obvious difference in philosophy, the technical difference is also apparent, since in Nash's game by a well-chosen threat (which she would prefer not to carry out) a player can improve her share, without her bluff ever being called. Thus, even if we used our conflict game as the threat game, the equilibria would differ, since the players, in general, would not use a threat that forms part of an equilibrium of the conflict game.

b) When Nash's players generate a disagreement point, he considers the bargaining problem properly defined and proceeds to its solution (according to his 1950 paper). In contrast, we argue that they have simply arrived at a new bargaining situation, where they might wish to employ different threats than before. To put it another way: while in the Nash model the demand phase depends on the outcome of the threat phase, in our model the conflict game is supposed to depend on the demands made (when they are not compatible).

ii) Step-by-step resolution.

Kalai (1977) introduced the axiom of decomposability. This assumption requires that if we break up the set of available agreements, *S*, into two subsets, *X* and *Y*, then using the solution of (either) one of these as a partial agreement to subsequently bargain over the rest, *(S-f(X, d))* ▌ $\Re_+^N$*,* should give the same result as applying the solution directly. Note that Kalai's model agrees with ours in the idea that partial agreements are only renegotiated if this yields a Pareto improvement. On the other hand, Kalai does not propose a well-defined solution: he only establishes that the solution should be "proportional," without identifying what should these proportions be. In addition, Kalai's model has two caveats, first pointed out by Ponsati and Watson (1997). The first of these is that when agreeing on the first sub-problem, the bargainers of Kalai are not supposed to take into account the effect of today's agreement on tomorrow's one. This is not true in our model. Second, there seems to be an inconsistency between the assumption that the agreement on the first subproblem is binding, but at the same time can be renegotiated –since the second sub-problem is not *S\X = Y* but *(S-f(X, d))* ▌ $\Re_+^N$. In our model, however, these two sets coincide so we avoid any confusion.

Wiener and Winter (1999) (see O'Neill et al., 2001 for the published version) propose a solution for bargaining problems where the feasible set is exogenously divided up into smaller pieces. Their solution is equivalent to agreeing step-by-step on each "crumb" according to the Nash solution, using the result of the previous step as the new disagreement point. This procedure is similar to ours, but we use the disagreement function to determine the new *status quo* and we do not need the arbitrary division.

iii) Bargaining under the threat of some outside enforcement mechanism.

This topic has been extensively dealt with in the applied literature (pre-trial negotiations, strikes, arbitration etc.). Perhaps, the piece closest to our approach is Powell (1996). Powell sets up a non-cooperative bargaining game where the players can choose to force a (probabilistic) settlement at some cost. The important difference with respect to our approach is that, in his model, forcing the settlement is equivalent to taking an *outside option*. However, outside options do not determine, in general, the outcome of a bargaining game. Therefore, Powell needs to rely on the solution to the bargaining game,

16

which would come about in the absence of outside options. In our case, in contrast, the solution of the game cannot be dissociated from the underlying conflict situation.

iv) Recursive solutions.

We are not the first ones to use a recursive application of some rule in bargaining theory. Let us mention just a couple. Raiffa (1953) proposes a method where the players first pocket half of their most preferred allocation, then half of their most preferred allocation in the remainder... etc. While in (its recursive) structure his procedure is very much like ours, the important difference is that he has no justification other than some vague consideration of "fairness" for the fifty percent rule. van Damme (1986) considers a recursivity axiom which imposes that if the players are making demands according to some individual theories, then in every step of the iteration, as a function of these demands some subset of S is to be discarded, and the negotiation resumed. Technically, the IIIA assumption is very similar, with the important difference that we only invoke individual rationality for discarding "irrelevant alternatives."

v) Disagreement modelled as a non-cooperative game.

Lundberg and Pollak (1993) replace divorce by a non-cooperative equilibrium within marriage, as the disagreement point in a model of marital bargaining. While they implicitly recognize that the forces determining the threat point are the same ones that influence the bargaining process, they do not make this connection explicit, and simply use the Nash solution.

6. Concluding remarks

In this paper we have presented a new approach to the theory of negotiation and have introduced the corresponding agreement concept. The cornerstone of our theory is the more efficient use of information that was already necessary for the standard theory: the description of the non-cooperative resolution of conflict. Indeed, we use not only the utility allocation in a particular equilibrium (the disagreement point), but we make full use of the primitives behind this equilibrium. In fact, we have shown that the disagreement function contains sufficient information to derive a unique agreement when coupled with a mild generalization of individual rationality.

We consider our theory to be complementary to the one based on time preferences. In scenarios where delay costs (and the risk of breakdown) are negligible with respect to the stakes of negotiation, like political disputes; or where disagreement leads into conflict which generates inefficiencies that are not related to delay, our approach seems to be more appropriate. In addition, the ASC solution yields a unique solution for an arbitrary number of negotiators, while the alternating-offers models usually generate multiple equilibria for more than two players.

Our theory carries with it a conceptual novelty as well. This insight relates to the interpretation of the terms: *agreement* and *disagreement*. Recall that the general idea of offer-counteroffer models is that disagreement is temporary − in the sense that the rejection of an offer does not end the negotiation − and that agreement is total −in the sense that at each point in time the players are either in agreement or not, no intermediate possibility is considered. Instead, we make the "dual" assumption: we posit that disagreement is final but possibly partial, while agreements can be temporary, and therefore partial as well. That is, we allow for the possibility that the players agree on the sharing of part of the surplus and either postpone agreement or disagree on the rest. The important observation is that the fact that they did not get to full agreement is not interpreted as a complete failure of the negotiation: the partial agreement can be implemented and the extent (and the efficiency cost) of disagreement is reduced.

To appreciate the degree of the meta-similarity of the dual approaches, note that our enrichment of the bargaining problem with the disagreement function merely corresponds to the incorporation of an exogenous cost of disagreement −over the surplus remaining, conditional on any partial agreement. This is completely parallel to the case where the description of the bargaining problem is augmented with the exogenous parameters of the cost of (temporary) disagreement to each party − following any length of past disagreement. Similarly, our ruling out of a trivial disagreement game corresponds to Rubinstein's ruling out perfectly patient players. Finally, in both cases the sequential story behind the solution is not meant to be actually followed in real time. Rational, fully informed agents will immediately identify which is the unique solution.

Finally, we should emphasize that we have presented our model based on cardinal preferences only to minimize our departure from standard theory. It is easy to see that we need not restrict attention to the utility space in order to derive our results. Any underlying space of bargaining outcomes, together with a complete preference relation, would suffice. In other words, our theory is one based on ordinal preferences, an unreachable goal for solutions to the standard bargaining problem.

Appendix

References

Binmore, K. (1987a), "Nash Bargaining Theory II," Chapter 4 in *The Economics of Bargaining* (eds. Binmore and Dasgupta) Basil Blackwell, Oxford.

Binmore, K. (1987b), "Perfect Equilibria in Bargaining Models," Chapter 5 in *The Economics of Bargaining* (eds. Binmore and Dasgupta) Basil Blackwell, Oxford.

Binmore, K., Rubinstein, A. and A. Wolinsky (1986), "The Nash Bargaining Solution and Economic Modelling," *RAND Journal of Economics* 17(2), 176-188.

Chen, M. and E. Maskin (1999), "Bargaining, Production, and Monotonicity in Economic Environments," *Journal of Economic Theory* 89, 140-147.

van Damme, E. (1986), "The Nash Bargaining Solution is Optimal," *Journal of Economic Theory* 38, 78-100.

Esteban, J. and D. Ray (1999), "Conflict and Distribution," *Journal of Economic Theory* 87, 379-415.

Harsányi, J. and R. Selten (1972), "A Generalized Nash Solution for Two-Person Bargaining Games with Incomplete Information," *Management Science* 18, 80-106.

Kalai, E. and M. Smorodinsky (1975), "Other Solutions to Nash's Bargaining Problem," *Econometrica* 43, 513-518.

Kalai, E. (1977), "Proportional Solutions to Bargaining Situations: Interpersonal Utility Comparisons," *Econometrica* 45(7), 1623-1630.

Lundberg, S. and R. Pollak (1993), "Separate Spheres Bargaining and the Marriage Market," *Journal of Political Economy* 101(6), 988-1010.

Nash, J. (1950), "The Bargaining Problem," *Econometrica* 18, 155-162.

Nash, J. (1953), "Two Person Cooperative Games," *Econometrica* 21, 128-140.

O'Neill, B., Samet, D., Wiener, Z. and E. Winter (2001), "Bargaining with an Agenda", mimeo, forthcoming in *Games and Economic Behavior*.

Ponsati, C. and J. Watson (1997), "Multiple-Issue Bargaining and Axiomatic Solutions," *International Journal of Game Theory* 26, 501-524.

Powell, R. (1996), "Bargaining in the Shadow of Power," *Games and Economic Behavior* 15, 255-289.

Raiffa, H. (1953), "Arbitration Schemes for Generalized Two-Person Games," in Kuhn and Tucker (eds.) *Contributions to the theory of games II*, Annals of Mathematics Studies #28. Princeton University Press.

Roemer, J. (1988), "Axiomatic Bargaining Theory on Economic Environments," *Journal of Economic Theory* 45, 1-31.

Rubinstein, A. (1982), "Perfect Equilibrium in a Bargaining Model," *Econometrica* 50, 97-109.

Smith, A. (1776), *The Wealth of Nations*.

Svejnar, J. (1986), "Bargaining Power, Fear of Disagreement, and Wage Settlements: Theory and Evidence from U.S. Industry," *Econometrica* 54(5), 1055-1078.

Wiener, Z. and E. Winter (1999), "Gradual Bargaining," mimeo, Hebrew University, Jerusalem, March.

Wilson, C. (2001), "Mediation and the Nash Bargaining Solution," *Review of Economic Design*, 6, 353-370.