Centre de Referència en Economia Analítica

Barcelona Economics Working Paper Series

Working Paper nº 255

A Theory of Agreements in the Shadow of Conflict

Joan Esteban and József Sákovics

February 2005

A theory of agreements in the shadow of conflict*

Joan Esteban[†]

Institut d'Anàlisi Econòmica (CSIC) and Universitat Pompeu Fabra

and

József Sákovics[‡]

Edinburgh School of Economics

February 2005

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. Adding this function to the description of a bargaining problem, a weak axiom based on individual rationality leads to 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 disagreement scenario. We provide a non-cooperative implementation as well.

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

«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 among the 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 these two steps in a consistent manner.

Specifically, we claim that the differential power of the players in the underlying non-cooperative game supporting the disagreement point also influences the bargaining process itself and, in fact, it actually determines its solution. In order to capture the differential power in the non-cooperative game we need more information on the disagreement game than just the equilibrium payoffs. We need a – reduced-form – description 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 disagreement game.[1] An *extended bargaining problem* is thus described by a pair $(S, D(.))$: the set of the payoffs initially available and the disagreement function providing the outcome of disagreement for any subset of these payoffs. We wish to stress that the disagreement function is NOT freely chosen by the modeler, rather it forms part of the description of a bargaining situation and therefore it is exogenously given.

The nature of the disagreement game depends on the problem at hand. In some situations, it 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. Thus, the disagreement payoffs are simply the outside options of the parties. However, such extremely simple situations

---

[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 hypothetical $S' \subseteq S$.

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.

For our purposes, the only condition we need the disagreement game/function to satisfy is that disagreement does not result in the loss of the entire surplus. That is, either the surplus should not be fully relationship specific, or upon disagreement the players should continue the interaction and reach a non-negotiated settlement which does not destroy the full surplus. For this class of bargaining situations two simple axioms – essentially positing individual rationality and internal consistency – permit the characterization of a unique and efficient solution: the *agreement in the shadow of conflict*, or 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 are not individually rational –that is, they are dominated by the outcome of the disagreement game for at least one player. This axiom is conceptually a weakening of Nash's axiom of Independence of Irrelevant Alternatives (since it requires the solution's independence of fewer alternatives), though here it is applied in the context of an extended bargaining game, what increases its bite.

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

4

reduces the set of feasible agreements. Using our second axiom –whic h only requires that the solution to a bargaining game should remain unaltered if we restart the bargaining game using the solution as the original bargaining set – we show that the repeated application of the first axiom to the resulting sequence of bargaining games converges to a situation where the disagreement outcome is efficient, thus pinpointing a unique solution.[2]

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.[3] As a result of the expected outcome of conflict, any agreement must give to the siblings at least two and one euros, respectively. Recognizing 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.

---

[2]Note that we for our argument to work we need not assume that the bargaining set is defined in utility space, it could be defined directly in outcome space. Consequently it can be construed as an ordinal solution.

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

Applying the argument repeatedly, we reach the final agreement, where the ten euros are distributed according to the power of the parties in the conflict game:[4] 20/3 and 10/3.

The argument above provides an attractive interpretation of negotiation as a *process*[5] where, driven by the fear of a conflictual resolution, the parties accept to gradually narrow down the extent of their dissent.[6] 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.

As no theory of bargaining is complete without a non-cooperative implementation, we also exhibit an extensive form game, which (asymptotically) implements the ASC. This game is a finite horizon bargaining game with randomly selected proposers, where unanimity is required for agreement. We show that as the number of periods tends to infinity the unique expected subgame perfect outcome of the

---

[4]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 euros in total, respectively.

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

[6]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. Likewise, plaintiffs may reach a *partial* pre-trial agreement still leaving part of the dispute unresolved. 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.

game converges to immediate agreement at the ASC.

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 a 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. In this, linear, set-up we are also able to provide an alternative extensive form (concession) game – from the class analyzed by Compte and Jehiel (2004) – which has a subgame perfect equilibrium, which reaches the ASC solution via a sequence of partial agreements.

The paper is structured as follows. In the next section, we discuss the conceptual forbearers of our approach. In section 3 we develop our theory of bargaining in the shadow of conflict for general environments. We discuss the properties of disagreement functions, present the IIIA axiom, characterize the bargaining solution and prove that it is unique. In section 4 we provide our strategic implementation of the ASC solution. In section 5 we analyze the linear model. Finally, we provide a discussion of the related bargaining literature, followed by some concluding remarks.

2   Might and right in social agreements

The view we develop here is to a good extent inspired by Hobbes' theory of social agreements. Well before Economics developed the theory of bargaining, Political Philosophy had addressed the question of social agreements in its inquiry about the foundations of the state. Thomas Hobbes (1588-1679) was possibly the first modern political philosopher who formulated an articulated theory of the *social contract*.[7] He viewed the possibility of a collective agreement as a case of "conditional cooperation" (in Taylor's, 1987, words), constrained by what individuals can obtain in

---

[7]See Taylor (1987) and Gauthier (1990).

the state of nature. The state of nature is the outcome that would ensue from a non-cooperative, rule-free interaction among utility maximizing, selfish individuals (Hobbes' *first axiom*). The outcome of this interaction is resource consuming and is governed by the differences in endowments across individuals. His second axiom asserts that there exist agreements that Pareto dominate the allocation achieved under the state of nature. Finally, according to his third axiom, agreements should be conditioned by the allocation resulting in the state of nature: «*. . . it is a precept, or generall rule of Reason, That every man, ought to endeavour Peace, as farre forth as he has hope of obtaining it; and when he cannot obtain it, that he may seek, and use, all helps, and advantages of Warre*» (Leviathan, 100, as cited in Taylor, 1987, 131). Therefore, in Hobbes' view, social agreements are not the outcome of an idealistic introspection on how things ought to be, but rather the viable outcome of a process conditioned by the might of the parties.

This view was largely shared by Jean-Jacques Rousseau (1782), for whom the formation of the political society and the establishment of laws «*gave new constraints to the weak and new forces to the rich, irreversibly destroyed natural freedom, established forever property law and inequality*» (p.170-1). Adam Smith (1776) also conceived the state as the creature of the mighty, specifically designed to give stability to the unequal distribution of wealth. In his own words: «*The rich, in particular, are necessarily interested to support that order of things, which can alone secure them in the possession of their own advantages. (. . .) Civil government, so far as it is instituted for the security of property, is, in reality, instituted for the defence of the rich against the poor, of those who have some property against those who have none at all*». (Book V, Chap. 1, Part II)

Despite the dramatically different normative positions as to what a "social contract" ought to be, all of them coincide in the positive analysis: in actual social agreements the mighty obtain a preferential treatment.[8] That actual social agree-

---

[8]One of the lines along which the position held by Rousseau (and by Smith) departs from Hobbes' views is on whether the inequality that forces a biased social contract is innate to humans or it is

ments will, at least partly, reflect the distribution of power is to be expected as long as a social contract is to be found acceptable by all parties. Therefore, and this is one of Hobbes' characteristic themes, we cannot develop a theory of social contracts without reference to the power of the parties in the non-cooperative scenario. The state of nature not only determines the size of the potential surplus to be shared, but also the shares themselves.

It is our opinion that standard bargaining theory has been driven to the use of normative axioms because the description of the bargaining problem was so stylized that there were no bases left for a positive derivation of the corresponding agreement.[9] We develop a positive theory of agreements, adopting Hobbes'p osition that takes the initial conditions as given and focuses on *reachable* social agreements, quite independently of the moral judgement they might deserve. We reserve normative considerations for the "state of nature," the initial conditions under which a particular agreement has been reached.[10] This view is consistent with Roemer's (1996) reservations about the moral content of a bargaining agreement obtained without a prior redistribution of the initial endowments.

We explore whether a solution can be characterized saving on axioms and making a more intensive use of the information contained in the description of the background game. This approach is in line with the growing literature on the explicit modeling of the conflictual resolution of opposing interests. The works by Becker (1983) on pressure groups and Tullock (1980) on rent-seeking, are the pre-

---

acquired.

[9]Svejnar (1986), Roemer (1988) and, more recently, Chen and Maskin (1999) have also expressed their reservations about the standard description of a bargaining problem, pointing out that Nash's abstraction might be dispensing with essential information.

[10]Consider the parallel case of assigning the gains from exchange. Economics takes a positive stand and investigates the terms of trade that will actually take place, resulting from different market structures and characteristics of the traders. It does not inquire about which would have been the "fair" terms of trade. The normative valuations are reserved for the comparison of the distribution of the characteristics that condition the trade (distribution of endowments, for instance).

decessors of the more recent papers by Esteban and Ray (1999), Grossman (1991, 1994), Grossman and Kim (1995), Hirshleifer (1991, 1995), Horowitz (1993), and Skaperdas (1992) among many others.[11] The common feature of all these models is that the opposition of interests is resolved via conflict. Players expend resources into trying to make their preferred option prevail. The equilibrium outcome entails waste of resources and the particular allocation reached critically depends on what is at stake as well as on the relative power, among other relevant characteristics, of the players. It seems only natural to enquire why there is conflict to start with and whether there exist plausible conflict-avoiding agreements in this scenario.

3 Agreements in the shadow of conflict: an axiomatic characterization

Suppose that there are N players, who wish to reach an agreement in $S^0 \in \Sigma$, where Σ is the set of non-empty, convex, compact subsets of the utility[12] space, $\Re_+^N$. In case of disagreement, during any point of the negotiation, the payoffs are given by (the solution[13] to) a disagreement game played for the currently available stakes (some subset of S^0). Even though this game may be elaborate and may depend on a number of parameters, for our purposes the only relevant information is the relationship between the stakes and the outcome. Hence, we posit the existence of a

---

[11]Models of the conflictual resolution of opposing interests have also been developed in areas such us growth, international trade, industrial organization, organizational design, patent races, or economics of litigation, to mention just a few. Conflict models have also been developed for boundedly rational individuals (see, for example, Anderson et al., 1998).

[12]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.

[13]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.

disagreement function, $D(.)$, which assigns a disagreement point, d, to every $S \in \Sigma$, (an therefore, to every non-empty convex, compact subset of S^0). That is, if the set of alternatives currently considered were S, the outcome of disagreement would be $d = D(S) \in S$.

It is important to stress that 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. $D(.)$ may depend on additional parameters, especially those related to the players' "strength", which form part of the description of the conflict game.

A *bargaining solution* is a mapping, f_D: $\Sigma \to \Sigma$, satisfying $f_D(S) \subseteq S$ for all $S \in \Sigma$ and $D(.)$. That is, given $D(.)$, the solution selects a non-empty convex, compact subset of the alternatives as acceptable.

Define $S_x = \{s \in S | s \geq x \in S\}$. That is, S_x is the subset of S which weakly Pareto dominates x. Note that, if $S \in \Sigma$, then $S_{D(S)} \in \Sigma$ as well.

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 in S).

We thus impose the following axiom on the bargaining solution:

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

$$f_D(S) = f_D(S_{D(S)}) \quad \text{for all } S \in \Sigma.$$

Conceptually, IIIA is much weaker than Nash's Independence of Irrelevant Alternatives (IIA) axiom, since it only eliminates a subset of his "irrelevant alternatives". Imposed on a standard bargaining problem, IIIA would simply eliminate

11

the alternatives that do not weakly dominate the disagreement point. However, in our context, 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. Consequently, the axiom applies again, eliminating further possible agreements. In fact, as long as $D(S)$ is interior to S, the application of IIIA keeps generating new – smaller and smaller – sets of acceptable agreements.

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 payoffs.

Our first result shows that the requirement imposed on the solution is not too stringent.

Lemma 1 *For any $D(.)$ there exists a bargaining solution consistent with IIIA.*

Proof.   For any $S^0 \in \Sigma$, define the sequences, $d^t = D(S^t)$ and $S^{t+1} = S^t_{d^t}$, $t = 0, 1, 2, ...$. Note now that the sets S^t are compact and nested. Therefore, their intersection is uniquely defined and it is non-empty as well. Let us now consider the bargaining solution $f^*_D(S^0) = \bigcap\limits_{t=0}^{\infty} S^t$.

To see that this solution satisfies IIIA, take any $\widehat{S}^0 \in \Sigma$. We have that $f^*_D(\widehat{S}^0) = \bigcap\limits_{t=0}^{\infty} \widehat{S}^t$. At the same time $f^*_D(\widehat{S}^0_{D(\widehat{S}^0)}) = \bigcap\limits_{t=1}^{\infty} \widehat{S}^t$. Thus, since $\widehat{S}^0_{D(\widehat{S}^0)} \subseteq \widehat{S}^0$, $\bigcap\limits_{t=0}^{\infty} \widehat{S}^t = \bigcap\limits_{t=1}^{\infty} \widehat{S}^t$ and therefore $f^*_D(\widehat{S}^0) = f^*_D(\widehat{S}^0_{D(\widehat{S}^0)})$. ∎

A further nice property of bargaining solutions satisfying IIIA is that for any bargaining set, the union of all solutions is itself a solution satisfying IIIA, and in fact it is $f_D^*(.)$, the solution displayed in the existence proof above.

Lemma 2 *For any $D(.)$ and for any $S \in \Sigma$, if $f_D(.)$ satisfies IIIA then $f_D(S) \subseteq f_D^*(S)$.*

Proof.   Suppose otherwise. Then for some $S^0 \in \Sigma$, $f_D(S^0) \setminus f_D^*(S^0)$ is non-empty. By IIIA, using the construction in the proof of Lemma 1, we have that $f_D(S^0) = f_D(S^t) \subseteq S^t$ for all t. However, since $\lim_{t \to \infty} S^t = f_D^*(S^0)$, for t large enough it must be the case that $f_D(S^0) \setminus S^t$ is non-empty. Contradiction. ∎

As a result of Lemma 2, we can meaningfully talk about $f_D^*(.)$ as the "maximal" solution satisfying IIIA.

We are now set for the formal definition of our solution concept.

Definition 1 *The agreement in the shadow of conflict solution (ASC) is $\bigcap_{t=0}^{\infty} S^t$, the maximal bargaining solution consistent with IIIA.*

The increase in the informational content of the description of the bargaining problem – together with our axiom – is sufficient to provide us with a unique set of "acceptable" agreements. In general, this set of agreements needs not be a singleton. Whether the solution is determinate or not depends on the nature of the disagreement game. We shall now prove that – provided that the solution is internally consistent – for disagreement games where at least some of the players can obtain a payoff which is superior to the worst possible agreement for them, the above result can be strengthened: the ASC solution singles out a unique, Pareto efficient agreement.

Internal consistency of the solution simply means that if we apply the bargaining solution to the solution set itself, the solution remains invariant.

Axiom 2 *Internal Consistency (IC): $f_D(f_D(S)) = f_D(S)$ for all $S \in \Sigma$.*

13

This axiom is obviously satisfied whenever the solution is a singleton. Intuitively it simply says that if starting from a larger set the solution can only be refined to S^*, then starting from S^* it should not be possible to obtain a more refined solution, since then the solution from the original set should also get refined.

The formal requirement we impose on the disagreement function 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 $i \in \{1, 2, ..., N\}$.*

The amount of information contained in the disagreement function depends on whether the players can affect the status quo (the worst point in S). If they can't, the disagreement payoffs convey little, if any, information about the characteristics of the players, so we should not be surprised if it does not pinpoint a unique solution. 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. Alternatively, if the surplus bargained over is not fully relationship specific, some players must be better off "opting out" than with the status quo.[14] 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; in conflict models with endogenous choice of effort there is usually a unique interior Nash equilibrium, etc.[15]

---

[14]Notice that the familiar case of bargaining over the price to be paid for an object to be traded would appear to violate Assumption 1, since the payoffs upon no trade are zero. However, note that in any realistic application, the players are likely to have outside opportunities, bringing us back to the realm of Assumption 1.

[15]Esteban and Ray (1999) show that for a generalized version of the rent-seeking model, there always exists a *unique* Nash equilibrium and that at this equilibrium each contending party expends

Just as with IIIA, we have to ask ourselves the question, whether Assumption 1 remains reasonable when used in a recursive context. The question boils down to the interpretation of disagreement over a set, which Pareto dominates the status quo. What do we mean by disagreement over a set whose worse element gives the players more utility than the status quo? We believe that this should be understood as a partial agreement: the players agree to distribute some of the surplus in a certain way, thereby reducing the set over which they disagree. As long as this partial agreement is honored even in case of a later break-up, or conflict, the assumption is as valid as when it is made about the original bargaining set.

We now can reap the benefits of Assumption 1:

Proposition 3 *For every $D(.)$ satisfying Assumption 1, the internally consistent ASC bargaining solution selects a unique and efficient agreement.*

Proof. Suppose that $ASC_D(S^0) = S^*$ is not a singleton. Then, by Assumption 1, $D(S^*)$ dominates some points in S^*. This, together with IIIA implies that $ASC_D(S^*) = ASC_D(S^*_{D(S^*)}) \subset S^*$, violating IC. Therefore, S^* must be a singleton.

By the construction of the solution in the proof of Lemma 1, each set S^t 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, proving the efficiency of the solution. ∎

Thus, we have shown that whenever the disagreement functions are informative, a mild axiom about the players' rationality is sufficient to identify a determinate agreement. The essential point is 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 what the "just" price would be.

_____

strictly positive amounts of resources. It is straightforward to show that the disagreement point generated by the Nash equilibrium satisfies Assumption 1.

4 A strategic implementation

In the spirit of the Nash Program, we wish to provide a strategic implementation as well for our solution. To that effect, consider the following bargaining game in extensive form first:

There are T periods. In each period, t, one player is randomly selected to be a proposer. For all $i \in N$, Player i is chosen with probability $\lambda_i^t \geq 0$, where $\sum_i \lambda_i^t = 1$, for all $1 \leq t \leq T$. The proposer makes an offer to her adversaries, consisting of a utility vector $x \in S$. If the offer is unanimously accepted, the game is over and x is implemented. If at least one responder rejects the offer the game proceeds to the next period (without discounting), where the same procedure is repeated. If there is no agreement by the end of period T, the players collect their disagreement payoffs, given by $D(S)$.

Note that the driving force towards agreement is neither impatience nor the fear of a possible (exogenous) breakdown – the usual arguments in favor of stationary game forms (c.f. Binmore et al. 1986) – rather it is the existence of a deadline.

This game is a variant of the one proposed in Gomes et al. (1999),[16] with the differences that in our game no player is eliminated at any stage and that the probabilities to be selected as a proposer are neither symmetric nor constant. For the case of two players, Sjostrom (1991) had showed that – when $\lambda_i^t \equiv 1/2$ – the unique SPE of this game converges to the Raiffa solution (see Raiffa, 1953, Luce and Raiffa, 1957) as T tends to infinity.

Let us now enrich the above game by the following. In every period, bargaining (or communication) is only possible with probability μ^t (with $\mu^1 = 1$), otherwise the game directly proceeds to the following period.[17]

The motivation for this feature of the model is that in every period, after the

---

[16]Hart and Mas-Colell (1996) intoduced the infinite horizon version of the game previously, but as the type of horizon is crucial to the solution, our model is much closer to Gomes et al. (1999).

[17]If the game reaches period T but bargaining is not possible, the players collect their disagreement payoffs.

start of negotiation, some exogenous event may occur, which makes it impossible for all the players to show up at the negotiating table. It is similar in spirit to the random breakdown model except that here breakdown is only temporary.

Lemma 4 *The enriched T-period bargaining game has a unique expected outcome implementable by a subgame-perfect equilibrium, denoted by $A^T(S)$.*

Proof. We proceed by backward induction. In period T, if Player i is the proposer she can ask for (and obtain, if communication is possible) the highest utility in S, which leaves her opponents their disagreement payoff: $\max_i D_{-i}(S)$. Therefore, in period $T-1$ the expected continuation vector upon disagreement is

$$\Delta^{T-1} = \mu^T \sum_{i \in N} \lambda_i^T \left(D_1(S), D_2(S), ..., \max_i D_{-i}(S), ..., D_N(S) \right) + \left(1 - \mu^T\right) D(S).$$

For an arbitrary period, t, we then have

$$\Delta^t = \mu^{t+1} \sum_{i \in N} \lambda_i^{t+1} \left(\Delta_1^{t+1}, \Delta_2^{t+1}, ..., \max_i \Delta_{-i}^{t+1}, ..., \Delta_N^{t+1} \right) + \left(1 - \mu^{t+1}\right) \Delta^{t+1}. \quad (1)$$

Consequently, in period t, the unique equilibrium action of a proposer is to propose

$$\left(\Delta_1^t, \Delta_2^t, ..., \max_i \Delta_{-i}^t, ..., \Delta_N^t \right)$$

whenever she is chosen. Thus the period 1 proposals will be

$$\left(\Delta_1^1, \Delta_2^1, ..., \max_i \Delta_{-i}^1, ..., \Delta_N^1 \right)$$

and will always be accepted in equilibrium. Therefore the unique expected equilibrium outcome is[18]

$$A^T(S) = \sum_{i \in N} \lambda_i^1 \left(\Delta_1^1, \Delta_2^1, ..., \max_i \Delta_{-i}^1, ..., \Delta_N^1 \right).$$

∎

We then have the following:

---

[18]Recall that $\mu^1 = 1$.

Proposition 5 *For any $D(S)$ satisfying Assumption 1, there exist a sequence of probability vectors $\boldsymbol{\lambda}^t$ and non-negative numbers μ^t, such that the corresponding $\lim_{T\to\infty} A^T(S) = ASC(S)$.*

Proof.   Note that if we can choose $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ so that $\Delta^t = D(S^{T-t})$, for all $t \in 1, 2, ..., T$, – recalling that the S^t refer to the sequence of bargaining sets generated in the proof of Proposition 1 – then the claim is true. In order for that to hold (c.f. Equation (1))we need the following relationship to hold (for all i and t):

$$\mu^t \left[\lambda_i^t \max_i D_{-i}(S^{T-t}) + \left(1 - \lambda_i^t\right) D_i(S^{T-t}) \right] + \left(1 - \mu^t\right) D_i(S^{T-t}) = D_i(S^{T-t+1}). \quad (2)$$

Solving for λ_i^t,

$$\lambda_i^t = \frac{D_j(S^{T-t+1}) - D_j(S^{T-t})}{\mu^t \left[\max_j D_{-j}(S^{T-t}) - D_j(S^{T-t})\right]}$$

and observing that the λ_i^t must add up to one (for each t) we have

$$\sum_{j \in N} \frac{D_j(S^{T-t+1}) - D_j(S^{T-t})}{\mu^t \left[\max_j D_{-j}(S^{T-t}) - D_j(S^{T-t})\right]} = 1.$$

Solving now for μ^t we obtain

$$\mu^t = \sum_{j \in N} \frac{D_j(S^{T-t+1}) - D_j(S^{T-t})}{\max_j D_{-j}(S^{T-t}) - D_j(S^{T-t})}. \quad (3)$$

∎

Since the convergence is monotonic, an alternative expression of this result is that for any $\epsilon > 0$, there exists a high enough T, so that $\left| A^T(S) - ASC(S) \right| < \epsilon$. In the same spirit, if we were to take the set of feasible agreements to be finite (c.f. Stahl 1994), there would be an upper bound on the periods needed for (full) convergence.

Unfortunately, the proposition does not guarantee that $\mu^t \leq 1$. In order to ensure that, we need to make an additional assumption. Recall from the previous proof that we denote the best payoff vector or "blisspoint" for Player i in S_d by $\max_i d_{-i}$.

Assumption 2 *$D(S_d)$ lies (weakly) below the hyper-plane defined by $\max_i d_{-i}$, i= 1,2,...,N.*

Note that the assumption simply posits that the efficiency cost of disagreement is sufficiently large so that there exists some lottery over the players' blisspoints which Pareto dominates disagreement.

Corollary 6 *When Assumption 2 is satisfied, Proposition 5 yields* $(0 \leq) \mu^t \leq 1$.

Proof. Note that by (3) $\mu^t = 1$ would imply that the disagreement point is exactly on the hyper-plane defined by $\max_i d_{-i}$, i= 1,2,...,N. ∎

The converse of Proposition 5 holds as well, so that we have a full equivalence between the strategic and axiomatic models:

Proposition 7 *For any sequence of probability vectors* $\boldsymbol{\lambda}^t$ *and probabilities* μ^t *there exists a* $D(S)$ *satisfying Assumption 1, such that* $ASC(S) = \lim_{T \to \infty} A^T(S)$.

Proof. Let

$$
\begin{aligned}
D(S^t) &= \mu^{T-t+1} \sum_{i \in N} \lambda_i^{T-t+1} \left(D_1(S^{t-1}), D_2(S^{t-1}), ..., \max_i D_{-i}(S^{t-1}), ..., D_N(S^{t-1}) \right) + \\
&\quad + \left(1 - \mu^{T-t+1} \right) D(S^{t-1}).
\end{aligned}
$$

By (1) and the definition of $A^T(S)$, this implies that $A^T(S)$ and $D(S^T)$ converge to the same limit. To see that such a $D(.)$ satisfies Assumption 1, note that unless $\max_i D_{-i}(S^{t-1}) = D_i(S^{t-1})$ for all i, this is clearly the case, while otherwise S^t must be a singleton. ∎

5   The linear case

In this section we restrict our attention to bargaining situations where the sub-problems considered are affine transformations of the original one. In other words, we assume that the bargaining set has a linear Pareto frontier (corresponding to risk neutrality), and that the disagreement function is homogeneous. These simplifications render the calculation of ASC easy and it also facilitates its comparison with other solutions.

Therefore, in this section we assume that S can be written as $S = \left\{ s \geq 0, \sum_i \beta_i s_i \leq R \right\}$ for some $R > 0$ and some β 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* homogeneous, *that is,* $D\left(\alpha + \lambda S\right) = \alpha + \lambda D\left(S\right)$ *for all* $\alpha \in \Re_+^N$ *and all* $\lambda > 0$.[19]

We can now easily compute the ASC solution.

Proposition 8 *Let* $S^0 = \left\{ s \geq 0, 0 \leq \sum_i \beta_i s_i \leq R \right\}$ *and let $D(.)$ satisfy Assumptions 1 and 3. Then, the ASC solution satisfies*

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

Proof. We know that the solution is $S^* = \lim_{T \to \infty} \bigcap_{t=0}^{T} S^t$, where $S^t = \left\{ u \in S^{t-1} \mid u \geq d^{t-1} \right\}$. Therefore, to obtain S^* we need only to compute $d^* = \lim_{t \to \infty} d^t$. Note that any set S^t in the sequence satisfies $S^t = d^{t-1} + \lambda^t S^0$, with $\lambda^t = \frac{R - \sum_i d_i^{t-1}}{R}$. By Assumption 3 we have that $d^t = D\left(S^t\right) = d^{t-1} + \lambda^t D\left(S^0\right)$. Therefore, $d^t = D\left(S^0\right) \sum_{r=1}^{t} \lambda^r$ and the equality in the statement of the Proposition follows immediately. ■

5.1 A comparison with the Nash Bargaining Solution

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 $i = 1, \ldots, N$. It is straightforward that the ASC solution is to give $s_i = \alpha_i$ to each player. Clearly, α_i can

---

[19]It is easy to show that the endogenous contest model of Esteban and Ray (1999) mentioned earlier, satisfies that $D(\lambda S) = \lambda D(S)$ whenever the Pareto frontier of the bargaining set is linear and the utility cost of effort is isoelastic.

be interpreted and the relative *power* of player i and β the degree of inefficiency induced by playing the disagreement game. The ASC solution does not depend on β 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 $s_i^N = (1 - \beta)\,\alpha_i + \frac{\beta}{N}$. 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 β is now the weight assigned by the Nash solution to the equal 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.

Take now the construction of the Nash solution – that is a part of the surplus is shared in proportion to the disagreement payoffs and the remainder according to some other sharing rule, but assume that the parameter β is not constrained to coincide with the loss from conflict, rather it can be freely agreed among the players. It is easy to see that the only situation in which the players would agree is when the "other" sharing rule coincides with the one which is proportional to the disagreement payoffs, resulting in the entire surplus being shared in those proportions (irrespective of β).

The example makes it clear that ASC is a solution based on the asymmetric treatment of the players. Thus, it is 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: $W\left(x, d\right) = \prod\limits_{i=1}^{N} \left(x_i - d_i\right)^{\gamma_i}$. We shall now discuss the relationship between the vector $\boldsymbol{\gamma}$ 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.

We the have the following:

Proposition 9 *When the Pareto frontier of S is the unit simplex and the disagreement function satisfies Assumptions 1 and 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 4, this ratio is equal to the ratio of the disagreement utilities. ∎

5.2 The non-cooperative implementation revisited

Note that the current set-up 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[20] – with bargaining weights $\gamma_1 = \log \delta_2$ and $\gamma_2 = \log \delta_1$ – independently of the identity of the first mover. Assuming that the Pareto frontier is the unit simplex, we have proved a similar result for the ASC solution, without having to resort to taking limits. That is, the ASC solution will exactly coincide with the asymmetric Nash solution, while Rubinstein's does so only

---

[20]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.

in an approximate sense.[21]

The linear set-up allows us to improve upon our non-cooperative implementation in two distinct ways. On one hand, we overcome the non-stationary nature of our extensive form (apart from the deadline). That is, the probabilities of a given player making the offer and that bargaining is possible are now constant over time:

Corollary 10 *In the linear set-up the ASC can be (asymptotically) implemented by the game presented in Section 4, with*

$$\lambda_i = \frac{D_j(S_D) - D_j(S)}{\mu \left[\max_j D_{-j}(S) - D_j(S) \right]}$$

and

$$\mu = \sum_{j \in N} \frac{D_j(S_D) - D_j(S)}{\max_j D_{-j}(S) - D_j(S)}.$$

Proof. Homogeneity of D(.), together with the linear Pareto frontier, implies that the bargaining situation is qualitatively the same following any partial agreement. Consequently, we can drop the time indices. ∎

The other advantage of the linear set-up is that we can now propose an alternative extensive form game, which has a subgame perfect equilibrium leading to the ASC *via a sequence of partial agreements*. Therefore, this game is a better illustration of the motivation provided for the axiomatic solution.[22]

---

[21]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.

[22]On the other hand, the ASC is not the only outcome supportable, so this game does not serve as implementation in the classical sense.

Consider the following non-cooperative bargaining game. In each period the players first simultaneously choose whether to provoke conflict over the current stakes (that is, the part of the surplus that has not yet been conceded). If no one provokes conflict then the players simultaneously[23] make – irreversible – concessions and the game continues to the following period. The game ends when either conflict is generated, or there remains nothing to distribute.

Proposition 11 *As long as conflict is strictly inefficient, the players conceding up along the ray leading from the origin to the ASC is a subgame perfect equilibrium path for a sufficiently slow rate of concession.*

Proof. Note that we can always require that everyone provokes conflict following a deviation, what – by the simultaneity – will be self-enforcing. Therefore, the only thing to worry about is that not making the current period concession and provoking conflict should not dominate the ASC payoff. However, this can always be avoided by making the required concession sufficiently low, since the loss due to conflict is bounded away from zero. ∎

6   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 (like Anbarci et al. 2002). Actually, however, our contribution goes well beyond that. There are two important differences between the models that we would like to underline:

---

[23]It is not crucial for the proposals to be simultaneous.

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 ommitment 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, $(S - f(X, d)) \cap \Re_+^N$, 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 $S \backslash X = Y$ but $(S - f(X,d)) \bigcap \Re_+^N$. In our model, however, these two sets coincide so we avoid any confusion.

Wiener and Winter (1999) (see O'Neill et al., 2004 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.

Fearon (1996) proposes a model with flow payoffs where the current distribution of the assets determines the current power in case of (costly) conflict. Assuming one-sided offers, he obtains a unique path of gradually increasing demand.

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

26

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 ercent rule. van Damme (1986) considers a ecursivity 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.

vi) *Endogenous bargaining power.*

Spindler (1974, 1976) proposed a definition of bargaining power which is not fixed, rather it is a function of the agreement considered (and of the status quo). Then the solution can be calculated as the one which equates these endogenous bargaining powers. While his method is very different from hours, it shares the idea of trying to read more out of the bargaining situation than Nash's bargaining set.

7   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 solution. 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 elusive[24] goal for solutions to the standard bargaining problem.

---

[24]Until recently (see Sákovics (2004)) it was thought that –for two-player games – no (meaningful) such solution was possible at all.

References

[1] Anbarci, N., Skaperdas, S. and C. Syropoulos (2002), "Comparing Bargaining Solutions in the Shadow of Conflict: How Norms against Threats Can Have Real Effects," *Journal of Economic Theory* 106(1), 1-16.

[2] Anderson, S., Goeree, J. and C. Holt (1998), "Rent Seeking with Bounded Rationality: An Analysis of the All-Pay Auction," *Journal of Political Economy* 106(4), 828-853.

[3] Becker, G. (1983), "A Theory of Competition among Pressure Groups for Political Influence," *Quarterly Journal of Economics* 98, 371-400.

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

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

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

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

[8] Compte, O. and P. Jehiel (2004), "Gradualism in Bargaining and Contribution Games," *Review of Economic Studies*, 71(4), 975-1000.

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

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

[11] Fearon, J. (1996), "Bargaining over Objects that Influence Future Bargaining Power," mimeo, October. http://www.stanford.edu/~jfearon/papers/booifbp.pdf

[12] Gauthier, D. (1990), *Moral Dealing. Contracts, Ethics and Reason*, Cornell University Press, Ithaca N.Y..

[13] Gomes, A., Hart, S. and A. Mas-Colell (1999), "Finite Horizon Bargaining and the Consistent Field," *Games and Economic Behavior* 27, 204-228.

[14] Grossman, H. (1991), "A General Equilibrium Model of Insurrections," *American Economic Review* 81, 912-921.

[15] Grossman, H. (1994), "Production, Appropriation and Land Reform," *American Economic Review* 84, 705-712.

[16] Grossman, H. and M. Kim (1995), "Swords or Plowshares? A Theory of the Security of Claims to Property," *Journal of Political Economy* 103(6), 1275-1288.

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

[18] Hart, S. and A. Mas-Colell (1996), "Bargaining and Value," *Econometrica* 64(2), 357-380.

[19] Hirshleifer, J. (1991), "The Paradox of Power," *Economics and Politics* 3, 177-200.

[20] Hirshleifer, J. (1995), "Anarchy and its Breakdown," *Journal of Political Economy* 103, 26-52.

[21] Horowitz, A. (1993), "Time Paths of Land Reform: A Theoretical Model of Reform Dynamics," *American Economic Review* 83(4), 1003-1010.

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

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

[24] Luce, R. and H. Raiffa (1957), *Games and Decisions*, Wiley, New York.

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

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

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

[28] O'Neill, B., Samet, D., Wiener, Z. and E. Winter (2004), "Bargaining with an Agenda," *Games and Economic Behavior* 48(1), 139-153.

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

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

[31] 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.

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

[33] Roemer, J. (1996), *Theories of Distributional Justice*, Harvard University Press, Cambridge (Mass.).

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

[35] Sákovics, J. (2004), 'A Meaningful Two-Person Bargaining Solution Based on Ordinal Preferences," *Economics Bulletin*, 3(26), 1-6.

[36] Sjostrom, T. (1991), "Ståhl's Bargaining Model," *Economics Letters* 36:153-157.

[37] Skaperdas, S. (1992), "Cooperation, Conflict, and Power in the Absence of Property Rights," *American Economic Review* 82, 720-739.

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

[39] Spindler, Z. (1974), "Endogenous Bargaining Power in Bilateral Monopoly and Bilateral Exchange," *Canadian Journal of Economics*, 7(3), 463-474.

[40] Spindler, Z. (1976), "Endogenous Bargaining Power and Small Group Collective Choice," *Public Choice*, 28, 67-78.

[41] Stahl, I. (1994), "The Rubinstein and Stahl Bargaining Models: A Comparison and an Attempt at a Synthesis" EFI Research paper #6535, Stockholm School of Economics, April.

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

[43] Taylor, M. (1987), *The Possibility of Cooperation*, Cambridge University Press.

[44] Tullock, G. (1980), "Efficient Rent Seeking," in J.M. Buchanan, R.D. Tollison and G. Tullock (eds.) *Toward a Theory of the Rent-Seeking Society*, College Station: Texas A&M University Press, 97-112.

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

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