Bargaining and Destructive Power*

Partha Dasgupta

University of Cambridge

and

Eric S. Maskin

Institute for Advanced Study, Princeton E-mail: Maskin@IAS.edu

We explore the effect that the power to destroy the feasible set has on two-person bargaining outcomes.

Key Words: Bargaining; Axioms; Destruction. JEL Classification Number: C78.

1. MOTIVATION

The theory of bargaining as formulated by Nash (1950, 1953) has developed along two routes. One is axiomatic (e.g., Nash 1950; Kalai and Smorodinsky 1975; Roemer 1988). Here, the negotiation process underlying the bargaining is only implicit. The idea is to try to characterize the negotiated outcome (the solution) through a set of axioms without formally modeling the process.

The advantages of the axiomatic route are clear enough. A fruitful solution concept will be applicable to a wide class of negotiation procedures. Partly for this reason the axiomatic approach has been adopted not only in the classical bargaining problem but more generally in cooperative game theory, where axiomatizations have been devised for such solution concepts as the Shapley value and the nucleolus.

The second route formulates the problem of bargaining in strategic terms. Here, the negotiation procedure is described explicitly as a noncooperative game, and its equilibrium points are studied (see, e.g., Nash 1953). This

* We thank the National Science Foundation for research support.

217

1529-7373/2007 All rights of reproduction in any form reserved. approach has become increasingly popular in recent years. But it has limitations. Because the negotiation procedure is modeled explicitly, one may not always be sure which properties of equilibrium are sensitive to particular (and perhaps arbitrary) features of the game, and which are not. Accordingly, Nash (1950, 1953) regarded the two avenues as complementary. In this view, modeling the negotiation process explicitly is a way of testing a solution concept. This two-sided perspective on cooperative games is sometimes referred to as the Nash program.¹

The focal point of the bargaining theory literature continues to be Nash's own solution. However, the perspective has shifted. For years, Nash's strategic (1953) formulation was largely ignored. In that model bargainers simultaneously make utility demands. They obtain these utilities if the demands are jointly feasible and otherwise remain at the status quo (threat point). Nash showed that any noncooperative equilibrium of a smooth approximation of this game approximates the utilities of his bargaining solution. But, until recently, the profession concentrated almost entirely on his (1950) axiomatic approach, perhaps in part because of the Luce-Raiffa (1957) interpretation of Nash's axioms as normative properties of an arbitration scheme (an interpretation that Nash himself did not put forward).

The strategic approach to bargaining was revived in the early 1980's (see, e.g., Fudenberg-Tirole 1983, Rubinstein 1982, and Sobel-Takahashi 1983). In particular, Rubinstein considered bargaining without time limit over a "cake." Players make alternating proposals about the cake's division until some proposal is accepted.² Rubinstein showed that, if there is a finite time between offers and players discount the future, such a game has a unique subgame perfect equilibrium. In fact, if one reinterprets the cake as a feasible set of utility pairs, the equilibrium outcome approximates the Nash bargaining solution if the time between successive proposals is small and the players have nearly the same discount rate (see Binmore 1987). Rubinstein's game is therefore another negotiation procedure that (approximately) yields the Nash bargaining solution as an equilibrium outcome.

In Nash's and Rubinstein's models, as well as in other strategic bargaining work that we know of, a bargainer exercises power over the other party only by the threat that a deal will not occur. Even in Nash's variable-threat version (1953) of his model, the threat is carried out only if a bargain is never struck. The emphasis on delay and recalcitrance as a bargainer's

¹For extensive discussion of these complementary approaches to cooperative game theory in general and to the theory of bargaining in particular, see the essays in Binmore and Dasgupta, ed. (1987).

 $^{^{2}}$ Stahl (1972) analyzed the same model but with a finite time horizon.

only instruments ignores the common possibility that he might take active steps to affect the bargain by destroying part of the feasible set.

In practice, bargaining frequently entails destruction in this sense and sometimes even violence: A terrorist threatens to kill one of his hostages if his demands are not met, and then does so when the deadline has passed; the United States calls for Japan to surrender and then bombs Hiroshima and Nagasaki before Japan concedes. Indeed, war, in general, can be viewed as the exercise of destructive power within a framework of political bargaining.

In this paper we explore the effect of such power on bargaining. In Section 2 we provide an example of negotiation between a firm (management) and its (unionized) workers in which each party is capable of inflicting some damage on the other party's interest during bargaining: the workers, by neglecting to maintain capital and equipment; the management, by replacing the existing production technology by ones that are less advantageous to the union. The example is designed to illustrate that in many situations each party has the ability to direct its destructive power toward the other party; that is, in damaging the other party it need not hurt itself. In Section 3 we present the strategic form of a negotiation process that idealizes the destructive abilities in this example. We show that if the parties have equal power to delete portions of the feasible set outcomes, the negotiation process, although otherwise similar to that of Rubinstein (1982) results in a unique subgame perfect equilibrium quite different from the Nash bargaining solution. In Section 4 we propose a set of axioms for the classical bargaining problem that yield the equilibrium outcome of our negotiation process as the unique solution. In brief, while retaining Nash's other axioms, we replace his "independence of irrelevant alternatives" (or the Kalai-Smorodinsky 1975 "monotonicity" axiom) by a "deletion" axiom that formalizes the idea that parties have the power to affect the size and shape of the 'cake' being bargained over. Finally, in Section 5 we discuss the robustness of our solution concept to alterations in the negotiation process, extend the analysis of Section 3 to cases where the parties differ in their potential to destroy, and briefly examine how introducing imperfect information would change the results.

2. UNION-MANAGEMENT NEGOTIATION: AN EXAMPLE

Consider negotiation between a firm (or its management) and its union workers. The negotiation is over hours of work (L) and the wage bill (W). Assume that the (von Neumann-Morgenstern) utility function of the union, U, is

$$U = W^{\frac{1}{2}} (1 - L)^{\frac{1}{2}}$$
, where $0 \le L \le 1$ and $W \ge 0$, (1)

and that of the management, V, is

$$V = \pi - W, \tag{2}$$

where π is the firm's revenue. The firm's revenue (output) is determined by labor hours and installed capital (K) such that

$$\pi = \begin{cases} \frac{L}{K}, & \text{if } L \le \overline{K} \\ \overline{K}, & \text{if } L > \overline{K} \end{cases}$$
(3)

where \overline{K} is given initially. Let us assume that $\overline{K} \geq 1$. Then, the set of efficient (u, v) pairs is given by the straight line

$$v = 1 - 2u. \tag{4}$$

Let the status-quo point be the origin, i.e., the firm earns zero profit and the union zero utility if a bargain is not struck. It follows that negotiation between management and union is aimed at agreeing on a point (u, v) such that $2u + v \leq 1, u \geq 0$, and $v \geq 0$. It is also simple to check that, at any point on the efficient frontier,

$$L = 1 - u. \tag{5}$$

Therefore, as we move along the efficient frontier toward higher union welfare, the number of hours the union works falls.

We wish to consider situations where both parties can affect the set of feasible utility pairs (u, v) during negotiation. To see how the union might affect the set of feasible set, suppose that the capital equipment requires maintenance. Then, during negotiation the union could refrain from keeping it properly maintained. This can be modeled by allowing the value of (K) to fall through time. When $\overline{K} < 1$ the efficient set is no longer given by (4), but rather by the equation

$$V = \begin{cases} \overline{K} - \frac{u^2}{1 - \overline{K}}, & u \le u^* \\ 1 - 2u, & u > u^*, \end{cases}$$

where $\overline{K} - \frac{(u^*)^2}{1-\overline{K}} = 1 - 2u^*$. By not maintaining equipment, therefore, the union deletes a portion of the utility set most favorable to management.

What of the power that management can exercise? Let us assume that management can slowly modify the technology of production—say, by gradually trading its existing technology for a different one. Suppose that it does so in such a way that (3) continues to hold, except that π is now zero

220

for L less than some minimum labor requirement $\underline{L} \geq \frac{1}{2}$. Equation (5) shows that such a modification is tantamount to deleting an area from the utility set most favorable to the union, so that the efficient frontier becomes

$$V = \begin{cases} 1 - \underline{L} - \frac{u^2}{1 - \underline{L}}, & u \ge u^{**} \\ 1 - 2u, & u < u^{**}, \end{cases}$$

where $1 - \underline{L} - \frac{(u^{**})^2}{1 - \underline{L}} = 1 - 2u^{**}$.

This sort of example, where each party has the power to destroy a part of the feasible set most favored by the other side, provides the motivation for the negotiation game of the next section.

3. NEGOTIATION IN STRATEGIC FORM

The negotiation is between two parties (i = 1, 2), and a negotiated outcome is a pair of utilities (u, v). Let R_0 denote the set of utility pairs that are feasible before negotiation begins. Among the points in R_0 is the status quo.³ We assume that is convex⁴ and compact. Each party's utility function is a representation of preferences satisfying the von Neumann-Morgenstern axioms. Without loss of generality we translate the players' utilities so that the status quo point is the origin. It will be convenient to assume that R_0 lies in the "north-east" quadrant (i.e., if $(u, v) \in R_0$, then $u \ge 0$ and $v \ge 0$) and that it is comprehensive (i.e., if (u, v) belongs to R_0 then so do $(u - \alpha, v)$ and $(u, v - \beta)$, where $0 \le \alpha \le u$ and $0 \le \beta \le v$). Out results can easily be modified in the absence of these last two assumptions about R_0 , but the argument is then more complicated.

The players move alternately. Let R_t be the feasible set of utility pairs before move t (t = 0, 1, 2, ...). A move consists of deleting⁵ a portion of R_t of any size or shape up to a maximum area $\overline{\delta}$ (where $\overline{\delta}$ is understood to be small relative to R_0), and simultaneously proposing a point in the set of utility pairs that remains. The proposal is either accepted or rejected by the other party. If accepted, it is implemented and negotiation ends. If it

 $^{^{3}}$ By the status quo point we mean the utility pair that would result if a deal were never struck. We assume that this point is unaffected by any threats carried out during negotiation. This assumption is, in a sense, the opposite of that made in Nash's (1953) variable-threat model, where threats affect only the status-quo.

⁴If the set failed to be convex, the parties could convexify it by randomization.

⁵The deletion must be such that the remaining set, R_{t+1} , is compact, convex (since randomization is possible) and comprehensive. We allow bargainers to remove areas from any part of R_t . However, as we shall see, player 1 always finds it advantageous to delete from player 2's favorite corner of the feasible set (and vice versa). Hence, we are capturing in idealized form the kind of deletion the parties in the example of section 2 could undertake. Because $\overline{\delta}$ is the same for both bargainers we are assuming that they have equal bargaining power (see Section 5 where this assumption is relaxed).

is rejected, negotiation continues, with the other party now moving. And so forth.

We shall suppose that parties discount the future in the following weak sense: Given a choice between the same utility level at an earlier or later time a bargainer strictly prefers the former. We shall examine the role of this assumption in Section 5. Finally, we assume that all the above is common knowledge to the players.

Our central result is:

THEOREM 1. A subgame perfect equilibrium of the negotiation process outlined above exists. Moreover, if $\overline{\delta}$ is small, any equilibrium pair of utilities approximately equals the intersection of the Pareto frontier of R_0 and the ray from the origin that bisects R_0 , i.e., divides R_0 into subsets of equal area. In equilibrium, negotiation is concluded in the first round (t = 0).

Proof. As is standard, we argue backwards from the end. Suppose without loss of generality that it is player 1's turn to move. Assume first that the remaining feasible set R_T (where T is an even number because it is 1's move) is the rectangle $(0, 0), (u^*, 0), (0, v^*)$ and (u^*, v^*) . (Call $(R)_T$ the class of all such rectangles.) We claim that is soptimal for 1 to delete nothing and propose the "north-east" corner (u^*, v^*) of R_T . This proposal will be accepted by player 2 from our weak discounting assumption, since he can never obtain more than v^* . Thus, since there is no way that 1 can obtain more than u^* , he might as well act as we have asserted.⁶ We conclude that negotiation ends when the remaining feasible set is a rectangle. The proposal will be accepted by 2, from our weak discounting assumption, since he cannot gain from delay.

Suppose next that it is player 2's move and that the feasible set R_{T-1} belongs to the class of all (convex, compact, comprehensive) sets that can be reduced to a rectangle with a single deletion of area not exceeding $\overline{\delta}$ from the lower right-hand corner. Call this class \mathcal{R}_{T-1} . Confronted with R_{T-1} , player 2 can do no better than to make the minimal deletion γ_1 that reduces the feasible set to a rectangle and to propose the north-east corner (u^*, v^*) of that rectangle. To see this, note that, from the argument of the previous paragraph and weak discounting, player 1 will accept the

222

⁶Player 1's optimal move is not unique. He could, alternatively, delete a rectangular horizontal strip of area δ (with $0 \le \delta \le \overline{\delta}$) from the top of the feasible set and offer 2 the north-east corner of the remaining rectangle. This would be accepted by 2 because he cannot feasibly do better. Player 1 is indifferent among these strategies because they all guarantee him u^* . As we will see below, this multiplicity does not affect the characterization of the equilibrium payoffs.

proposal. Moreover, player 2 can never obtain more than v^* , and so cannot do better with a different move.⁷

Next, suppose that player 1 is to move and that the feasible set R_{T-2} belongs to the class of sets that can be reduced to the class \mathcal{R}_{T-1} with a single deletion γ_2 of area $\overline{\delta}$ from the upper left hand corner (i.e., R_{T-2} can be reduced to a rectangle with two deletions). Call this class \mathcal{R}_{T-2} . We claim that player 1's unique best move is to delete γ_2 and propose the northeast corner (u^*, v^*) of the ultimate rectangle. Notice that if he does so, player 2 will accept the proposal (from weak discounting and the argument of the previous paragraph), and thus player 1 obtains utility u^* . But if 1 deletes less than γ_2 (or deletes an area from some other part of R_{T-2}), player 2 can delete more than γ_1 (this assumes that the area of γ_1 is less than $\overline{\delta}$)⁸ and thereby get a payoff higher than v^* (implying that 1's payoff is less than v^* for 2), player 2 will reject the proposal since, from previous argument, he can guarantee himself at least v^* . Hence player 1's unique best move is as specified.

Continuing inductively, let us suppose that it is player 2's turn and that the feasible set R_{T-t} (where t is odd; the argument is very similar if t is even) belongs to the class \mathcal{R}_{T-t} of sets that can be reduced to the class \mathcal{R}_{T-t+1} with a single deletion γ_t of area $\overline{\delta}$ from the lower right-hand corner. Let (u^*, v^*) be the north-east corner of the ultimate rectangle when R_{T-t} is reduced successively to elements of $\mathcal{R}_{T-t+1}, \mathcal{R}_{T-t+2}, \ldots, \mathcal{R}_T$. It lies on the efficient frontier of R_{T-t} . From above argument, player 2's unique best move is to delete γ_t and propose (u^*, v^*) , and if he does so the proposal will be accepted.

Hence, a subgame-perfect equilibrium exists, and in any such equilibrium negotiation ceases in the first round. Moreover, in the reduction of R_{T-t} to the class of rectangle \mathcal{R}_T , player 1 makes t - 1/2 deletions of area $\overline{\delta}$ and one of area no greater than $\overline{\delta}$ from the lower right-hand part. (Once a rectangle is reached, player 1 might delete a further area of no more than $\overline{\delta}$, see footnote 6. However, when $\overline{\delta}$ is small this multiplicity of optimal actions does not invalidate the following sentence.) Hence, when $\overline{\delta}$ is small, the total areas deleted by each are nearly the same, and so the equilibrium payoffs (u^*, v^*) lie near the ray from the origin bisecting R_0 .

⁷Again, the move we described is not uniquely best for 2 if the area of γ_1 is less than $\overline{\delta}$. He could equally well delete more than γ_1 since he will still wind up with v^* . However, he must not delete less than γ_1 , since otherwise player 1 will make a deletion that forces him below v^* .

⁸In the (nongeneric) case where the area of γ_1 equals $\overline{\delta}$, player 1's best move is not unique.

4. AXIOMATICS

We now turn to an axiomatic development. Suppose that the two bargainers' preferences satisfy the von Neumann-Morgenstern axioms. Utilities are therefore unique up to positive affine transformations. Given a utility representation for each party let (u_0, v_0) be the status quo point and R the set of feasible utility paris. As before, we assume that R lies in the north-east quadrant relative to (u_0, v_0) , and is convex, compact and comprehensive. A bargaining game can therefore be denoted as $[R, (u_0, v_0)]$. A solution concept is a function, F, that operates on $[R, (u_0, v_0)]$ to give a point $(\overline{u}, \overline{v})$ in R, i.e., $F[R, (u_0, v_0)] = (\overline{u}, \overline{v})$. The following three axioms have often been required of solution concepts (see e.g., Nash 1950; Kalai and Smorodinsky 1975).

(A1) Invariance: Let $[R, (u_0, v_0)]$ and $[R', (u'_0, v'_0)]$ be two versions of the same bargaining game; that is, they differ only in the units and origins of the utility function. Then $F[R, (u_0, v_0)]$ and $F[R', (u'_0, v'_0)]$ are related by the same two utility transformations.

(A2) Weak Pareto Efficiency: There is no $(u, v) \in R$ such that $u > \overline{u}$ and $v > \overline{v}$.

(A3) Symmetry: Suppose that $[R, (u_0, v_0)]$ satisfies the properties:

i) $u_0 = v_0$

ii) $(u, v) \in R$ if and only if $(v, u) \in R$. Then $\overline{u} = \overline{v}$.

The Nash bargaining solution satisfies (A1)-(A3). Nash (1950) also imposed the more the controversial axiom, "independence of irrelevant alternatives." Kalai and Smordinsky (1975) replaced IIA with a monotonicity axiom. In what follows we will refer to a horizontal strip as that part of R above a horizontal line and to a vertical strip as that part of R to the right of a vertical line. We can now state

(A4) Deletion: Let [R, (0, 0)] be a bargaining game. If R is what remains of R when horizontal and vertical strips of equal area have been deleted, then neither component of $F[\overline{R}, (0, 0)]$ exceeds the corresponding component of F[R, (0, 0)].⁹

We may now state and prove:

THEOREM 2. A solution concept F satisfies (A1)-(A4) if and only if it selects the Pareto efficient point on the ray from the origin that bisects R.

Proof. We have already used (A1) once to translate the status quo point to the origin. Bisect R with a ray through the origin. Next, use (A1) again to rescale utilities so that the intersection of the Pareto efficient frontier and the ray bisecting R is transformed to the point (1, 1). Call the

 $^{^{9}}$ We comment on (A4) in section 5.

transformed bargaining region R', and let the points (0, 0), (0, 1), (1, 0), and (1, 1) be O, A, B, and C, respectively. Notice that the horizontal strip above AC has the same area as the vertical strip to the right of CB. If the unit square OACB were the bargaining region, (A2) and (A3) would imply that C is the solution of this game. We claim that C is the solution of the game [R', (0, 0)]. To see this, note that by (A2) the solution of [R', (0, 0)] is on the Pareto frontier of R'. Any point on this frontier other than C is better for one of the players. So, if the solution of [R', (0, 0)] were any point other than C, (A4) would be violated. To establish the converse is a straightforward verification.

5. COMMENTARY

The theory of economic externalities provides a natural framework for the analysis of individual and group power. In their well-known contributions, Shapley and Shubik (1969) and Starrett (1973) explored the existence of core allocations in an environment where individuals and coalitions have the power to inflict damage on the rest of society in the form of "external diseconomies" (in their example, the power to dump their garbage on the rest). Another formulation of power is that which derives from "self-destruction." Aumann and Kurz (1977) studied the structure of redistributive taxation associated with the Shapley value of an economy in which taxes and subsidies are chosen by majority rule and where individuals and groups can destroy their own endowments to prevent majority tyranny.

Each of these studies examined the implications of the power to destroy for a specific cooperative solution concept. In this paper we have set ourselves a somewhat different task. Rather than taking the solution concept as given, we have instead derived it, first from an explicit negotiation process (Theorem 1), and second from a set of axioms (Theorem 2).

In fact, Theorem 2 can be viewed as a generalization of Theorem 1. The negotiation process of section 3 is clearly invariant to positive affine transformations of the utility functions and results in Pareto optima. Moreover, when $\overline{\delta}$ is small, it is nearly symmetric. As for the deletion axiom (A4), notice that it stipulates that if horizontal and vertical strips of equal area are deleted from the feasible set, then neither bargainer should be better off than before. But this property is certainly satisfied by our negotiation process; as we have seen, if the bargainers each delete a portion of area $\overline{\delta}$, the equilibrium outcome remains the same as before.

There is a degree of symmetry embodied in the deletion axiom, but this is not enough to establish Theorem 2 in the absence of (A3). For example, a dictatorship by bargainer 1, in which the solution is always the intersection of Pareto boundary of the feasible set with the horizontal axis satisfies axioms (A1), (A2) and (A4).

Theorem 2 suggests that there should be many negotiation processes that lead to the bargaining outcome of Theorem 1. Indeed, we can vary the details of the section 3 negotiation process quite a bit without affecting the nature equilibrium. In particular, just as that process was analogous (in its move structure) to Rubinstein's bargaining game, we could work instead with the counterpart to Nash's (1953) game.

In practice, parties differ according to the power they wield. But in the strategic game of Section 3 we assumed that the players possess equal power to damage the other party; specifically, each player can delete an area of maximum size $\overline{\delta}$. Suppose instead that at each move, player i, i = 1, 2, can delete at most an area of size $\overline{\delta_i}$. We then have:

THEOREM 3. Let $\overline{\delta_1}/\overline{\delta_2} \to \alpha$ as $\overline{\delta_1}, \overline{\delta_2} \to 0$. Then, the limit of the equilibrium utilities corresponds to the intersection of the Pareto frontier and the ray from the origin that divides R_0 into sets for which the ratio of the upper to lower areas is α .

Proof. Simple adaptation of the proof of Theorem 1.

Theorem 3 captures the sense in which the relative power that players have to damage the interests of their rivals influences the outcome of negotiation. As α varies from 0 to infinity, the equilibrium outcome sweeps across the entire Pareto-efficient frontier.

In the games of Theorems 1 and 3, negotiation ends immediately in equilibrium, so that destructive power is never exercised. This unrealistic feature depends, however, on our perfect information assumption. If parties knew each other's destructive capabilities only imperfectly, negotiation would not ordinarily cease in the first period and some destruction would occur. In fact such a model of imperfect information may help resolve a theoretical puzzle. Models where delay is the only threat can explain a negotiation period of more than zero length if there is asymmetric information (c.f. Grossman-Perry (1986)). However as Gul-Sonnenschein-Wilson (1986) point out, such models do not readily predict delays of significant length, if, as often seems plausible, proposals and counterproposals can be made very quickly. By contrast, destruction often takes some time, and so perhaps we can interpret strikes and other breakdowns in negotiation as, in part, vehicles by which parties can learn their fellow bargainers' destructive capabilities.

REFERENCES

Aumann, R. and M. Kurz, 1977, Power and taxes. Econometrica 45.

Binmore, K., 1987, Perfect equilibria in bargaining models. In: K. Binmore and P. Dasgupta, eds.

Binmore, K., and P. Dasgupta, eds., 1987, The Economics of Bargaining, Basil Black-well (Oxford).

Fudenberg, D., and J. Tirole, 1983, Sequential Bargaining with Incomplete Information. *Review of Economic Studies* **50**.

Grossman, S. and M. Perry, 1986, Sequential Bargaining under Asymmetric Information. *Journal of Economic Theory* **39**.

Gul, F., H. Sonnenschein, and R. Wilson, 1986, Foundations of Dynamic Monopoly and the Coase Conjecture. *Journal of Economic Theory* **39**, 155-190.

Kalai, E. and M. Smorodinsky, 1975, Other Solutions to Nash's Bargaining Problem. $Econometrica~{\bf 43}.$

Luce, R.J. and H. Raiffa, 1957, Games and Decisions, Wiley (New York).

Nash, J., 1950, The Bargaining Problem. Econometrica 18.

Nash, J., 1953, Two-person Cooperative Games. Econometrica 21.

Roemer, J., 1988, Axiomatic Bargaining Theory on Economic Environments. *Journal of Economic Theory* **45**.

Rubinstein, A., 1982, Pefect Equilibrium in a Bargaining Model. Econometrica 50.

Shapley, L. and M. Shubik, 1969, On the Core of an Economic System with Externalities. *American Economic Review* **59**.

Sobel, J. and I. Takahashi, A Multi-Stage Model of Bargaining. *Review of Economic Studies* **50**.

Stahl, I., 1972, Bargaining Theory, Economic Research Institute, Stockholm.

Starrett, D., 1973, A Note on Externalities and the Core. Econometrica 44.