Center for Research on Economic and Social Theory

CREST Working Paper

Noncooperative Models of Bargaining

Ken Binmore
Martin J. Osborne
Ariel Rubinstein

1989
Number 89-26

DEPARTMENT OF ECONOMICS
University of Michigan
Ann Arbor, Michigan 48109

NONCOOPERATIVE MODELS OF BARGAINING*

by

Ken Binmore**
Department of Economics
University of Michigan
Ann Arbor, MI 48109, USA

Martin J. Osborne
Department of Economics
McMaster University
Hamilton L8S 4M4, Canada

and

Ariel Rubinstein**
Department of Economics
The Hebrew University
Jerusalem, Israel

*Parts of this chapter use material from Rubinstein [1987], a survey of sequential bargaining models. Parts of Sections 5, 6, 7 and 8 are based on a draft of parts of Osborne and Rubinstein [1990].

**The first author wishes to thank Avner Shaked and John Sutton and the third author wishes to thank Asher Wolinsky for long and fruitful collaborations.

1. Introduction

John Nash's [1950] path-breaking paper introduces the bargaining problem as follows:

> "A two-person bargaining situation involves two individuals who have the opportunity to collaborate for mutual benefit in more than one way." (p. 155)

With such a definition, nearly all human interaction can be seen as bargaining of one form or another. To say anything meaningful on the subject, it is necessary to narrow the scope of the inquiry. We follow Nash in assuming that

> "the two individuals are highly rational,...each can accurately compare his desires for various things,...they are equal in bargaining skill."

In addition we assume that the procedure by means of which agreement is reached is both clear-cut and unambiguous. This allows the bargaining problem to be modeled and analyzed as a noncooperative game.

The target of such a noncooperative theory of bargaining is to find theoretical predictions of what agreement, if any, will be reached by the bargainers. One hopes thereby to explain the manner in which the bargaining outcome depends on the parameters of the bargaining problem and to shed light on the meaning of some of the verbal concepts which are used when bargaining is discussed in ordinary language. However, the theory has only peripheral relevance to such questions as:

What is a just agreement?
How would a reasonable arbiter settle a dispute?
What is the socially optimal deal?

Nor is the theory likely to be of more than background interest to those who write manuals on practical bargaining techniques. Such questions as

How can I improve my bargaining skills?
How do bargainers determine what is jointly feasible?

are psychological issues that the narrowing of the scope of the inquiry is designed to exclude.

Cooperative bargaining theory (Chapter ???) differs mainly in that the bargaining <u>procedure</u> is left unmodeled. Cooperative theory therefore has to operate from a poorer informational base and hence its fundamental assumptions are necessarily abstract in character. As a consequence, cooperative solution concepts are often difficult to evaluate. Sometimes they may have more than one viable interpretation, and this can lead to confusion if distinct interpretations are not clearly separated. In this chapter we follow Nash in adopting an interpretation of cooperative solution concepts that attributes the same basic aims to cooperative as to noncooperative theory. That is to say, we focus on interpretations in which, to quote Nash [1953],

> "the two approaches to the [bargaining] problem...are complementary; each helps to justify and clarify the other." (p. 129)

This means in particular that what we have to say on cooperative solution concepts is not relevant to interpretations which seek to address questions like those given above which are specifically excluded from our study.

Notice that we do not see cooperative and noncooperative theory as rivals. It is true that there is a sense in which cooperative theory is "too general"; but equally there is a sense in which noncooperative theory is "too special". Only rarely will the very concrete procedures studied in noncooperative theory be observed in practice. As Nash [1953] observes,

> "Of course, one cannot represent all possible bargaining devices as moves in the noncooperative game. The negotiation process must be formalized and restricted, but in such a way that each participant is still able to utilize all the essential strengths of his position." (p. 129)

Even if one makes good judgments in modeling the essentials of the bargaining process, the result may be too cumbersome to serve as a tool in applications, where what is required is a reasonably simple mapping from the parameters of the problem to a solution outcome. This is what cooperative theory supplies. But which of the many cooperative solution concepts is appropriate in a given context, and how should it be applied? For answers to such questions, one may look to noncooperative theory for guidance. It is in this sense that we see cooperative and noncooperative theory as complementary.

2. A sequential bargaining model

The archetypal bargaining problem is that of "dividing the dollar" between two players. However, the discussion can be easily interpreted broadly to fit a large class of bargaining situations. The set of feasible agreements is identified with $A = [0,1]$. The two bargainers, players 1 and 2, have opposing preferences over A. When $a > b$, 1 prefers a to b and 2 prefers b to a. Who gets how much?

The idea that the information so far specified is not sufficient to determine the bargaining outcome is very old. For years, economists tended to agree that further specification of a bargaining solution would need to depend on the vague notion of "bargaining ability". Even von Neumann and Morgenstern [1944] suggested that the bargaining outcome would necessarily be determined by unmodeled psychological properties of the players.

Nash [1950, 1953] broke away from this tradition. His agents are fully rational. Once their preferences are given, other psychological issues are irrelevant. The bargaining outcome in Nash's models is determined by the players' attitudes towards risk--i.e. their preferences over lotteries in which the prizes are taken from the set of possible agreements together with a pre-determined "disagreement point".

A sequential bargaining theory attempts to resolve the indeterminacy by explicitly modeling the bargaining procedure as a sequence of offers and counter-offers. In the context of such models, Cross [1969] remarks, "If it did not matter when the parties agreed, it would not matter if they agreed at all". This suggests that the players' time preferences may be highly relevant to the outcome. In what follows, who gets what depends exclusively on how patient each player is.

The following procedure is familiar from street markets and bazaars all over the world. The bargaining consists simply of a repeated exchange of offers. Formally, we study a model in which all events take place at one of the times t in a prespecified set $T = \{0, t_1, t_2, \ldots\}$, where (t_n) is strictly increasing. The players alternate in making offers, starting with player 1. An offer x, made at time t_n, may be accepted or rejected by the other player. If it is accepted, the game ends with the agreed deal being implemented at time t_n. This outcome is denoted by (x,t_n). If the offer is rejected, the rejecting player makes a counter-offer at time t_{n+1}. And so on. Nothing binds the players to offers they have made in the past, and no predetermined limit is placed on the time that may be expended in bargaining. In principle, a possible outcome of the game is therefore perpetual disagreement or impasse. We denote this outcome by D.

Suppose that, in this model, player 1 could make a commitment to hold out for a or more. Player 2 could then do no better than to make a commitment to hold out for 1-a or better. The result would be a Nash equilibrium sustaining an agreement on a. The indeterminacy problem would therefore remain. However, we follow Schelling [1960] in being skeptical about the extent to which such commitments can genuinely be made. A player may make threats about his last offer being final, but the opponent can dismiss such threats as mere bombast unless it would actually be in the interests of the threatening player to carry out his threat if his implicit ultimatum were disregarded. In such situations, where threats need to be credible to be effective, we replace Nash equilibrium by Selten's notion of subgame-perfect equilibrium (Chapter ???).

The first to investigate the alternating offer procedure was Ståhl [1967, 1972, 1988]. He studied the subgame-perfect equilibria of such time-structured models by using backwards induction in finite horizon models. Where the horizons in his models are infinite, he postulates non-stationary time preferences which lead to the existence of a "critical period" at which one player prefers to yield rather than to continue, independently of what might happen next. This creates a "last interesting period" from which one can start the backwards induction. (For further comment, see Ståhl [1988].) In the infinite horizon models studied below, different techniques are required to establish the existence of a unique subgame-perfect equilibrium.

Much has been written on procedures in which all the offers are made by only one of the two bargainers. These models assign all the bargaining power to the party who makes the offers. Such an asymmetric set-up does not fit very comfortably within the bargaining paradigm as usually understood and so we do not consider models of this type in the current paper.

Impatience

Players are assumed to be impatient with the unproductive passage of time. The times in the set T at which offers are made are restricted to $t_n = n\tau$ ($n = 0,1,2,\ldots$), where $\tau > 0$ is the length of one period of negotiation. Except where specifically noted, we take $\tau = 1$ to simplify algebraic expressions. Rubinstein [1982] imposes the following conditions on the players' (complete, transitive) time preferences. For a and b in A, for s and t in T, and for $i = 1, 2$:

(TP1) $a < b$ implies $(a,t) >_i (b,t)$ and $(b,t) >_2 (a,t)$
(TP2) $0 < a < 1$ and $s < t$ imply that $(a,s) >_i (a,t) >_i D$
(TP3) $(a,s) \geq_i (b,s+\tau)$ if and only if $(a,t) \geq_i (b,t+\tau)$
(TP4) the graphs of the relations $\geq_i$ are closed.

These conditions are sufficient to imply the existence of utility function representations in which $\Phi_1(D) = \Phi_2(D) = 0$, $\Phi_1(a,t) = \phi_1(a)\delta_1^t$ and $\Phi_2(a,t) = \phi_2(1-a)\delta_2^t$, where the functions ϕ_i: $[0,1] \to [0,1]$ are strictly increasing and continuous, and $0 < \delta_i < 1$ (see Fishburn/Rubinstein [1982]). Sometimes we may take as primitives the "discount factors" δ_i. However, note that if we start, as above, with the preferences as primitives, then the number δ_i may be chosen arbitrarily in the range (0,1). The associated discount rates ρ_i are then given by $\delta_i = e^{-\rho_i}$.

To these conditions, we add the requirement:

(TP0) for each $a \in A$ there exists $b \in A$ such that $(b,0) =_i (a,\tau)$.

By (TP0) we have $\phi_i(0) = 0$; without loss of generality, we take $\phi_i(1) = 1$.

The function f: $[0,1] \to [0,1]$ defined by $f(u_1) = \phi_2(1 - \phi_1^{-1}(u_1))$ is useful. A deal reached at time 0 that assigns utility u_1 to player 1 assigns $u_2 = f(u_1)$ to player 2. More generally, the set U^t of utility pairs available at time t is

(1) $$U^t = \{(u_1\delta_1^t,\ f(u_1)\delta_2^t) : 0 \leq u_1 \leq 1\}.$$

Note that a feature of this model is that all subgames in which a given player makes the first offer have the <u>same</u> strategic structure.

Our goal is to characterize the subgame-perfect equilibria of this game. We begin by examining a pair of stationary strategies in which both players always plan to do the same in strategically equivalent subgames, regardless of the history of events that must have taken place for the subgame to have been reached. Consider two possible agreements a^* and b^*, and let u^* and v^* be the utility pairs that result from the implementation of these agreements at time 0. Let s_1 be the strategy of 1 which requires him always to offer a^* and to accept an offer of b if and only if $b \geq b^*$. Similarly, let s_2 be the strategy of 2 which requires him always to offer b^* and to accept an offer of a if and only if $a \leq a^*$. The pair (s_1,s_2) is a subgame-perfect equilibrium if and only if

(2) $$v_1^* = \delta_1 u_1^* \quad \text{and} \quad u_2^* = \delta_2 v_2^*.$$

In checking that (s_1,s_2) is a subgame-perfect equilibrium, observe that each player is always offered precisely the utility that he will get if he refuses the offer and s_1 and s_2 continue to be used in the subgame that ensues.

Notice that (2) admits a solution if and only if the equation

(3) $\quad f(x) = \delta_2 f(x\delta_1)$

has a solution. This is assured under our assumptions because f is continuous, $f(0) = 1$ and $f(1) = 0$.

Each solution to (2) generates a different subgame-perfect equilibrium. Thus, the uniqueness of a solution to (2) is a necessary condition for the uniqueness of a subgame-perfect equilibrium in the game.

In the following we will assume that

(TP5) (2) has a unique solution.

In Rubinstein [1982], this is ensured with the condition

(TP5*) $(a + \alpha, \tau) =_i (a,0)$, $(b + \beta, \tau) =_i (b,0)$ and $a < b$ implies that $\alpha \leq \beta$.

This has the interpretation that the more you get, the more you have to be compensated for delay in getting it. A weak sufficient condition for the uniqueness of the solution for (2) is that ϕ_1 and ϕ_2 be concave. (This condition is far from necessary. It is enough, for example, that $\log \phi_1$ and $\log \phi_2$ be concave.)

Result 1 (Rubinstein, 1982) Under assumptions (TP0)–(TP5) the bargaining game has a unique subgame-perfect equilibrium. In this equilibrium, agreement is reached immediately, and the players' utilities satisfy (2).

Alternative versions of Rubinstein's proof appear in Binmore [1987b] and Shaked/Sutton [1984]. The following proof of Shaked and Sutton is especially useful for extensions and modifications of the theorem.

Proof: Without loss of generality, we take $\tau = 1$.

Let the supremum of all subgame-perfect equilibrium payoffs to player 1 be M_1 and the infimum be m_1. Let the corresponding quantities for player 2 in the companion game, in which the roles of 1 and 2 are reversed, be M_2 and m_2. We will show that $m_1 = u_1^*$ and $M_2 = v_2^*$, where u_1^* and v_2^* are uniquely defined by (2). An analogous argument shows that $M_1 = u_1^*$ and $m_2 = v_2^*$. It follows that the equilibrium <u>payoffs</u> are uniquely determined. To see that this implies that the equilibrium <u>strategies</u> are unique, notice that, after every history, the proposer's offer must be accepted in equilibrium. If, for example, player 1's demand of u_1^* were refused, he would get at most $\delta_1 v_1^* < u_1^*$.

Since u^* is a subgame-perfect equilibrium outcome by the construction given above, $m_1 \leq u_1^*$ and $M_2 \geq v_2^*$. We now show that (i) $\delta_2 M_2 \geq f(m_1)$ and (ii) $M_2 \leq f(\delta_1 m_1)$.

(i): Observe that if 2 refuses the opening offer, then the companion game will be played from time $\tau = 1$. If equilibrium strategies are played in this game, 2 gets no more than $\delta_2 M_2$. Player 2's equilibrium behavior at time $t = 0$ must therefore include accepting any offer that assigns 2 a payoff strictly greater than $\delta_2 M_2$. Player 1 can therefore guarantee himself any payoff less than $f^{-1}(\delta_2 M_2)$. Thus $m_1 \geq f^{-1}(\delta_2 M_2)$.

(ii): In the companion game, player 1 can guarantee himself any payoff less than $\delta_1 m_1$ by refusing player 2's opening offer (provided equilibrium strategies are used thereafter). Thus $M_2 \le f(\delta_1 m_1)$.

The uniqueness of (u_1^*, v_2^*) satisfying (2) is expressed in Figure 1 by the fact that the curves $f(\delta_1 u_1) = u_2$ and $f^{-1}(\delta_2 u_2) = u_1$ intersect only at (u_1^*, v_2^*). From (i) and $m_1 \le u_1^*$, (m_1, M_2) lies in region (i). From (ii) and $M_2 \ge v_2^*$, (m_1, M_2) lies in region (ii). Hence $(m_1, M_2) = (u_1^*, v_2^*)$. Similarly $(M_1, m_2) = (u_1^*, v_2^*)$. ■

Figure 1

Shrinking cakes

Binmore's [1985] geometric characterization (see Figure 2) applies to any preferences, whether or not they satisfy the stationarity assumption (TP3). The "cake" available at time t is identified with a set U^t of utility pairs which is assumed to be closed, bounded above and to have a connected Pareto frontier. It is also assumed to shrink over time. This means that if $s \le t$, then, for each $y \in U^t$, there exists $x \in U^s$ satisfying $x \ge y$.

Figure 2

The construction begins by truncating the game to a finite number n of stages. Figure 2 shows how a set E_{t_n} of payoff vectors is constructed from the truncated game in the case when n is odd. The construction when n is even is similar. The set of all subgame-perfect equilibrium payoff vectors is shown to be the intersection of all such E_{t_n}. Since the sets E_{t_n} are nested, their intersection is also their limit as $n \to \infty$. The methodology reveals that, when there is a unique equilibrium outcome, this must be the limit of the equilibrium outcomes in the finite horizon models obtained by calling a halt to the bargaining process at some pre-determined time t_n. In fact, the finite horizon equilibrium outcome in Figure 2 is the point m.

Discounting

A very special case of the time preferences covered by Result 1 occurs when $\phi_1(a) = \phi_2(a) = a$ $(0 \le a \le 1)$. Reverting to the case of an arbitrary $\tau > 0$, we have $u_1^* = (1-\delta_2^\tau)/(1 - \delta_1^\tau \delta_2^\tau) \to \rho_2/(\rho_1 + \rho_2)$ as $\tau \to 0+$. When $\delta_1 = \delta_2$, it follows that the players share the available surplus of 1 equally in the limiting case when the interval between successive proposals is negligible.

If δ_1 decreases, so does player 1's share. This is a general result in the model: it always pays to be more patient. More precisely, define $\succeq_1$ to be at least as patient preference relation than $\succeq_1'$ if whenever $(y,0) \succeq_1' (x,1)$ we have $(y,0) \succeq_1 (x,1)$. Then player 1 always gets at least as much in equilibrium when his preference relation is $\succeq_1$ than when it is $\succeq_1'$ (Rubinstein [1987]).

Fixed costs

Rubinstein [1982] characterizes the subgame-perfect equilibrium outcomes in the alternating offers model under the hypotheses (TP1) – (TP4) and a weak version of (TP5*). These conditions also cover another interesting case when each of the players incurs a fixed cost $c_i > 0$ for each unit of time that elapses without an agreement being reached. Suppose, in particular, that their respective utilities for the outcome (a,t) are $a - c_1t$ and $1 - a - c_2t$. It follows from Rubinstein [1982] that, if $c_1 < c_2$, the only subgame-perfect equilibrium assigns the whole surplus to 1. If $c_1 > c_2$, then 1 obtains only c_2 in equilibrium. If $c_1 = c_2 = c < 1$, then <u>many</u> subgame-perfect equilibria exist. If c is small ($c < 1/3$), some of these equilibria involve <u>delay</u> in agreement being reached. That is, equilibria exist in which one or more offers get <u>refused</u>. It should be noted that, even when the interval τ between successive proposals becomes negligible ($\tau \to 0+$), the equilibrium delays do not necessarily become negligible.

Stationarity, Efficiency and Uniqueness

We have seen that, when (2) has a unique solution, the game has a <u>unique</u> subgame-perfect equilibrium which is <u>stationary</u> and that its use results in the game ending <u>immediately</u>.

The efficiency of the equilibrium is not a consequence of the requirement of perfection by itself. As we have just seen, when multiple equilibria exist, (that is when (2) has more than one solution) some of these may call for some offers to be refused before agreement is reached so that the final outcome need not be Pareto-efficient. It is sometimes suggested that rational players with complete information must necessarily reach a Pareto-efficient outcome when bargaining costs are negligible. This example shows that the suggestion is questionable.

Some authors consider it adequate to restrict attention to stationary equilibria on the grounds of simplicity. We do <u>not</u> make any such restriction, since we believe that, for the current model, such a restriction is hard to justify. Strategies in sequential games are more than plans of how to play the game. They include also a description of the beliefs the players have about what a player would do if he were to deviate from his own plan of action. (We are not talking here about beliefs as formalized in the notion of sequential equilibrium, but of the beliefs built into the definition of a strategy in an extensive form game.) Therefore, a stationarity assumption does more than attribute simplicity of behavior to the players: it also makes players' beliefs insensitive to past events. For example, stationarity requires that, if player 1 is supposed to offer a 50:50 split in equilibrium, but has always demanded an out-of-equilibrium 60:40 split in the past, then player 2 still continues to hold the belief that player 1 will offer the 50:50

split in the future. For a more detailed discussion of this point see Rubinstein [1988].

Finally, it should be noted that the uniqueness condition of Result 1 can fail if the set A from which players choose their offers is sufficiently restricted. Suppose, for example, that the dollar to be divided can only be split into a whole number of cents, so that A = {0, .01,..., .99, 1}. If $\phi_1(a) = \phi_2(a) = a$ and $\delta_1 = \delta_2 = \delta > .99$, then any division of the dollar can be supported as the outcome of a subgame-perfect equilibrium. Does this conclusion obviate the usefulness of Result 1? This depends on the circumstances in which it is proposed to apply the result. If the grid on which the players locate values of δ is finer than that on which they locate values of a, then the bargaining problem remains indeterminate. Our judgment, however, is that the reverse is usually the case.

Outside options

When bargaining takes place it is usually open to either player to abandon the negotiation table if things are not going well, to take up the best option available elsewhere. This feature can easily be incorporated into the model analyzed in Result 1 by allowing the players to opt out whenever they have just refused an offer. Let us suppose that opting out at time t yields a payoff of $e_i\delta_i^t$ to player i. The important point is that, under the conditions of Result 1, the introduction of such exit opportunities is irrelevant to the equilibrium bargaining outcome when $e_1 < \delta_1 u_1^*$ and $e_2 < \delta_2 v_2^*$. In this case the players always prefer to continue bargaining rather than to opt out. The next result exemplifies this point.

Result 2 (Binmore/Shaked/Sutton, 1988) Take $\phi_1(a) = \phi_2(a) = a$ $(0 \le a \le 1)$ and $\delta_1 = \delta_2 = \delta$. If $e_1 + e_2 < 2\delta/(1 + \delta)$ then there exists a unique subgame-perfect equilibrium in which neither player exercises his outside option. It assigns $1/(1 + \delta)$ to player 1 unless $e_1 > \delta/(1 + \delta)$ or $e_2 > \delta/(1 + \delta)$. If $e_2 > \delta/(1 + \delta)$, player 1 gets $1 - e_2$. Otherwise he gets $1 - \delta(1 - e_1)$.

As modeled above, a player cannot leave the table without first listening to an offer from his opponent, who therefore always has a last chance to save the situation. This seems to capture the essence of traditional face-to-face bargaining. Shaked [1987] finds multiple equilibrium if a player's opportunity for exit occurs, not after a refusal by himself, but after a refusal by his opponent. He has in mind "high tech" markets in which binding deals are made quickly over the telephone. Intuitively, a player then has the opportunity to accompany the offer with a threat that the offer is final. Shaked shows that equilibria exist in which the threat is treated as credible and others in which it is not. When outside options are mentioned later, it is the face-to-face model that is intended. But it is important to bear in mind how sensitive the model can be to apparently minor changes in the structure of the game. For a further discussion of the "outside option" issue see Sutton [1986] and Bester [1988].

Finally, reference needs to be made to the work of Harsanyi and Selten [1988, p.243]. Although the bargaining model they consider is not strictly comparable with the model used in deriving Result 2, it is worrying that they should obtain such strikingly different conclusions about how outside options should be expected to influence the bargaining outcome. Clearly further research on the many possible bargaining models that can be constructed in this context is much needed.

Risk

Binmore/Rubinstein/Wolinsky [1986] reinterpret the alternating offers model by supposing that the players are indifferent to the passage of time but face a probability p that any refused offer will be the last that can be made. The fear of getting trapped in a bargaining impasse is then replaced by the possibility that intransigence will lead to a breakdown of the negotiating process due to the intervention of some external factor. The extensive form in the new situation is somewhat different from the one described above: at the end of each period the game may end with a breakdown outcome with probability p. Moreover, the functions ϕ_1 and ϕ_2 need to be reinterpreted as von Neumann and Morgenstern utility functions. That is to say, they are derived from the players' attitudes to risk rather than from their attitudes to time. The conclusion is essentially the same as in the time-based model. We denote the breakdown payoff vector by b and replace the discount factors by $1 - p$. Since account needs to be taken of the fact that b may not be located at the origin, (2) is replaced by

(4) $\quad v_1^* - b_1 = (1 - p)(u_1^* - b_1)$ and $u_2^* - b_2 = (1 - p)(v_2^* - b_2)$,

where u^* is the agreement payoff vector when player 1 makes the first offer and v^* is its analog for the case in which it is 2 who makes the first offer.

More than two players

Result 1 does not extend to the case when there are more than two players, as the following three-player example of Shaked demonstrates.

Three players rotate in making proposals $a = (a_1, a_2, a_3)$ on how to split a cake of size one. We require that $a_1 + a_2 + a_3 = 1$ and $a_i \geq 0$ $(i = 1,2,3)$. A proposal a accepted at time t is evaluated as worth $a_i\delta^t$ by the i^{th} player. A proposal a made by the j^{th} player at time t is first considered by player $j + 1$ (mod 3) who may accept or refuse. If he accepts, player $j + 2$ (mod 3) may accept or refuse. If both accept, the game ends and the proposal a is implemented. Otherwise player $j + 1$ (mod 3) makes the next proposal at time $t + 1$.

Let $1/2 \leq \delta < 1$. Then, for every proposal a, there exists a subgame-perfect equilibrium in which a is accepted immediately. We describe the equilibrium in terms of the four commonly held "states (of mind)" a, e^1, e^2 and e^3, where e^i is the i^{th} unit vector. In state y, each player i makes the proposal y and accepts the proposal z if and only if $z_i \geq \delta y_i$. The initial state is a. Transitions occur only after a proposal has been made, before the response. If, in state y, player i proposes z with $z_i > y_i$ then the state becomes e^j, where $j \neq i$ is a player for whom $z_j < 1/2$. Such a player j exists, and the requirement that $\delta \geq 1/2$ guarantees that it is optimal for him to reject player i's proposal.

Efforts have been made to reduce the indeterminacy in the n-player case by changing the game or the solution concept. One obvious result is that, if attention is confined to stationary (one-state) strategies, then the unique subgame-perfect equilibrium assigns the cake in the proportions $1:\delta:\ldots:\delta^{n-1}$. The same result follows from restricting the players to have continuous expectations about the future (Binmore [1986]).

3. The Nash program

The ultimate aim of what is nowadays called the "Nash program" (see Nash [1953]) is to classify the various institutional frameworks within which negotiation takes place and to provide a suitable "bargaining solution" for each class. As a test of the suitability of a particular solution concept for a given type of institutional framework, Nash proposed that attempts be made to reduce the available negotiation ploys within that framework to moves within a formal bargaining game. If the rules of the bargaining game adequately capture the salient features of the relevant bargaining institutions, then a "bargaining solution" proposed for use in the presence of these institutions should appear as an equilibrium outcome of the bargaining game.

The leading solution concept for bargaining situations is the Nash bargaining solution (see Nash [1950]). The idea belongs in cooperative game theory. A "bargaining problem" is a pair (U,q) in which U is a set of pairs of von Neumann and Morgenstern utilities representing the possible deals available to the bargainers, and q is a point in U interpreted by Nash as the <u>status quo</u>. The Nash bargaining solution is a point at which the Nash product

$$(u_1 - q_1)(u_2 - q_2) \tag{5}$$

is maximized subject to the constraints $u \in U$ and $u \geq q$. Usually it is assumed that U is convex, closed and bounded above to ensure that the Nash bargaining solution is uniquely defined, but convexity is not strictly essential in what follows.

When is such a Nash bargaining solution appropriate for a two-player bargaining environment involving alternating offers? Consider the model we studied in the previous section, in which there is a probability p of breakdown after any refusal. We have the following result. (See also McLennan [1982], Moulin [1982], Binmore/Rubinstein/Wolinsky [1986].)

Result 3 (Binmore, 1987a) When a unique subgame-perfect equilibrium exists for each p sufficiently close to one, the bargaining problem (U,q), in which U is the set of available utility pairs at time 0 and $q = b$ is the breakdown utility pair, has a unique Nash bargaining solution. This is the limiting value of the subgame-perfect equilibrium payoff pair as $p \to 0+$.

Proof. To prove the concluding sentence, it is necessary only to observe from (4) that $u^* \in U$ and $v^* \in U$ lie on the same contour of $(u_1 - b_1)(u_2 - b_2)$ and that $u^* - v^* \to (0,\ 0)$ as $p \to 0+$.

■

We can obtain a similar result in the time-based alternating offers model when the length τ of a bargaining period approaches 0. One is led to this case by considering two objections to the alternating offers model. The first is based on the fact that the equilibrium outcome favors player 1 in that $u_1^* > v_1^*$ and $u_2^* < v_2^*$. This reflects player 1's first-mover advantage. The .

objection evaporates when τ is small and so "bargaining frictions" are negligible. It then becomes irrelevant who goes first. The second objection concerns also the reasons why players abide by the rules. Why should a player who has just refused an offer patiently wait until a period of length $\tau > 0$ before making a counter-offer? If he were able to abbreviate the waiting time, he would respond immediately. Considering the limit as $\tau \to 0+$, removes some of the bite of the second objection in that the players need no longer be envisaged as being constrained by a rigid, exogenously determined timetable.

Figure 3

Figure 3 illustrates the solutions u^* and v^* of equations (2) in the case when δ_1 and δ_2 are replaced by δ_1^τ and δ_2^τ and $\rho_i = -\log \delta_i$. It is clear from the figure that, when τ approaches zero, u^* and v^* both approach the point n^* in U at which $u_1^{1/\rho_1} u_2^{1/\rho_2}$ is maximized. Although we are not dealing with von Neumann and Morgenstern utilities, it is convenient to describe n^* as being located at an <u>asymmetric</u> Nash bargaining solution of U relative to a status quo q located at the impasse payoff pair (0, 0). (See Chapter ??? and Roth [1978].)

Such an interpretation should not be pushed beyond its limitations. In particular, with our assumptions on time preferences, it has already been pointed out that, for <u>any</u> δ in (0,1), there exist functions w_1 and w_2 such that $w_1(a)\delta^t$ and $w_2(1 - a)\delta^t$ are utility representations of the players' time preferences. Thus if the utility representation is tailored to the bargaining problem, then the equilibrium outcome in the limiting case as $\tau \to 0+$ is the <u>symmetric</u> Nash bargaining solution for the utility functions w_1 and w_2.

This discussion of how the Nash bargaining solution may be implemented by considering limiting cases of sequential noncooperative bargaining models makes it natural to ask whether other bargaining solutions from cooperative game theory can be implemented using parallel techniques. We mention only Moulin's [1984] work on implementing the Kalai-Smorodinsky solution. (See Chapter ??? and Kalai and Smorodinsky [1975].) Moulin's model begins with an auction to determine who makes the first proposal. The players simultaneously announce probabilities p_1 and p_2. If $p_1 > p_2$, then player 1 begins by proposing an outcome a. If player 2 rejects a, then he makes a counter-proposal b. If player 1 rejects b, then the status quo q results. If player 1 accepts b, a referee organizes a lottery which yields b with probability p_1 and q with probability $1 - p_1$. (The tie-breaking device used in the auction is irrelevant.) The natural criticism is that it is not clear to what extent such an "auctioning of fractions of a dictatorship" qualifies as bargaining in the sense that this is normally understood.

Economic modeling

The preceding section provides some support for the use of the Nash bargaining solution in economic modeling. One advantage of a noncooperative approach is that it offers some insight into how the various economic parameters that may be relevant should be assimilated into the bargaining model when the environment within which bargaining takes place is complicated (Binmore/Rubinstein/Wolinsky [1986]). In the following, we draw together some of the relevant considerations.

Assume that the players have von Neumann and Morgenstern utilities of the form $u_i(a)\delta^t$. Consider the placing of the status quo. In cooperative bargaining theory this is interpreted as the utility pair that results from a failure to agree. But such a failure to agree may arise in more than one way. We shall, in fact, distinguish three possible ways:

a. A player may choose to abandon the negotiations at time t. Both players are then assumed to seek out their best outside opportunities, thereby deriving utilities $e_i\delta_i^t$. Notice that it is commonplace in modeling wage negotiations to ignore timing considerations and to use the Nash bargaining solution with the status quo placed at the "exit point" e.
b. The negotiations may be interrupted by the intervention of an exogenous random event which occurs in each period of length τ with probability $\lambda\tau$. If the negotiations get broken off in this way at time t, each player obtains utility $b_i\delta_i^t$.
c. The negotiations may continue for ever without interruption or agreement, which is the outcome denoted by D in Section 2. As in Section 2, utilities are normalized so that each player then gets $d_i = 0$.

Assume that the three utility pairs e, b and d, satisfy $0 = d < b < e$.

When contemplating the use of an asymmetric Nash bargaining solution in the context of an alternating offers model for the "frictionless" limiting case when $\tau \to 0+$, the principle is that the status quo is placed at the utility pair q that results from the use of impasse strategies. Thus, if we ignore the exit point e then

$$q_i = \lim_{\tau\to 0+}\Sigma_{j=0}^{\infty} b_i\delta_i^{\tau j}\lambda\tau(1 - \lambda\tau)^j = \lambda b_i/(\lambda + \rho_i) \quad (i = 1,2).$$

The (symmetric) Nash bargaining solution of the problem in which (q_1, q_2) is the status quo point is the maximizer of $u_1^{\alpha_1}u_2^{\alpha_2}$, where $\alpha_i = 1/(\lambda + \rho_i)$ (i.e. it is the asymmetric Nash bargaining solution in which the "bargaining power" of player i is α_i). This reflects the fact that both time and risk are instrumental in forcing an agreement.

It is instructive to look at two extremal cases. The first occurs when λ is small compared with the discount rates ρ_1 and ρ_2 so that it is the time costs of disagreement that dominate. The status quo goes to d and the bargaining powers become $1/\rho_i$. The second case occurs when ρ_1 and ρ_2 are both small compared with λ so that risk costs dominate. This leads to a situation closer to that originally envisaged by Nash [1950]. The status quo goes to the breakdown point b and the bargaining powers approach equality so that the Nash bargaining solution becomes symmetric.

As to the exit point, the principle is that its value is irrelevant unless at least one player's outside option e_i exceeds the appropriate Nash bargaining payoff. There will be no agreement if this is true for both players. When it is true for just one of the players, he gets his outside option and the other gets what remains of the surplus.

Note finally that the above considerations concerning bargaining over stocks translate immediately to the case of bargaining over flows. In bargaining over the wage rate during a strike, for example, the status quo payoffs should be the impasse flows to the two parties during the strike.

4. Commitment and concession

A commitment is understood to be an action available to an agent that constrains his choice set at future times in a manner beyond his power to revise. Schelling [1960] has emphasized, with many convincing examples, how difficult it is to make genuine commitments in the real world to take-it-or-leave-it bargaining positions. It is for such reasons that subgame-perfect equilibrium and other refinements now supplement Nash equilibrium as the basic tool in noncooperative game theory. However, when it is realistic to consider take-it-or-leave-it offers or threats, these will clearly be overwhelmingly important. Nash's [1953] demand game epitomizes the essence of what is involved when both sides can make commitments.

In this model, the set U of feasible utility pairs is assumed to be convex, closed, and bounded above, and to have a non-empty interior. A point $q \in U$ is designated as the status quo. The two players simultaneously make take-it-or-leave-it demands u_1 and u_2. If $u \in U$, each receives his demand. Otherwise each gets his status quo payoff.

Any point of $V = \{u \geq q\text{: } u \text{ is Pareto-efficient in } U\}$ is a Nash equilibrium. Other equilibria result in disagreement. Nash [1953] dealt with this indeterminacy by introducing a precursor of modern refinement ideas. He assumed some shared uncertainty about the location of the frontier of U embodied in a quasi-concave, differentiable function $p\colon \mathbb{R}^2 \to [0,1]$ such that $p(u) > 0$ if u is in the interior of U and $p(u) = 0$ if $u \notin U$. One interprets $p(u)$ as the probability that the players commonly assign to the event $u \in U$. The modified model is called the smoothed Nash demand game. Interest centers on the case in which the amount of uncertainty in the smoothed game is small. For all small enough $\epsilon > 0$, choose a function $p = p^\epsilon$ such that $p^\epsilon(u) = 1$ for all $u \in U$ whose distance from V exceeds ϵ. The existence of a Nash equilibrium which leads to agreement with positive probability for the smoothed Nash demand game for $p = p^\epsilon$ follows from the observation that the maximizer of $u_1 u_2 p^\epsilon(u_1, u_2)$ is such a Nash equilibrium.

Result 4 (Nash, 1953) Let u^ϵ be a Nash equilibrium of the smoothed Nash demand game associated with the function p^ϵ which leads to agreement with positive probability. When $\epsilon \to 0$, u^ϵ converges to the Nash bargaining solution for the problem (U,q).

Proof. The following sketch follows Binmore [1987c]. Player i seeks to maximize $u_i p^\epsilon(u) + q_i(1 - p^\epsilon(u))$. The first order conditions for $u^\epsilon > q$ to be a Nash equilibrium are therefore

$$(6) \qquad (u_i^\epsilon - q_i)p_i^\epsilon(u^\epsilon) + p^\epsilon(u^\epsilon) = 0 \qquad (i = 1,2)$$

where p_i^ϵ is the partial derivative of p^ϵ with respect to u_i. Suppose that $p(u^\epsilon) = c > 0$. From condition (6) it follows that the vector u^ϵ must be a maximizer of $H(u_1,u_2) = (u_1 - q_1)(u_2 - q_2)$ subject to the constraint that $p(u) = c$. Let w^ϵ be the maximizer of $H(u_1,u_2)$ subject to the constraint that $p(u) = 1$. Then $H(u^\epsilon) \geq H(w^\epsilon)$. By the choice of p^ϵ the sequence w^ϵ converges to the Nash bargaining solution and therefore the sequence u^ϵ converges to the Nash bargaining solution as well.

■

There has been much recent interest in the Nash demand game with incomplete information, in which context it is referred to as a "sealed-bid, double-auction" (see for example Wilson [1987a,b], Leininger/Linhart/Radner [1986], Matthews/Postlewaite [1987]). It is therefore worth noting that the smoothing technique carries over to the case of incomplete information and provides a noncooperative defense of the Harsanyi and Selten [1972] axiomatic characterization of the (M + N)-player asymmetric Nash bargaining solution in which the bargaining powers β_i $(i = 1,..,M)$ are the (commonly known) probabilities that player 2 attributes to player 1's being of type i and β_j $(j = M + 1,...,M + N)$ are the probabilities attributed by player 1 to player 2's being of type j. If attention is confined to pooling equilibria in the smoothed demand game, the predicted deal $a \in A$ is the maximizer of $\Pi_{i=1,..,M}(\phi_i(a))^{\beta_i}\Pi_{j=M+1,..,M+N}(\phi_j(a))^{\beta_j}$. Here $\phi_i: A \rightarrow \mathbb{R}$ is the von Neumann and Morgenstern utility function of the player of type i (Binmore, [1987c]).

Nash's threat game

In the Nash demand game, the status quo q is given. Nash [1953] extended his model in an attempt to endogenize the choice of q. In this later model, the underlying reality is seen as a finite two-person game G. The bargaining activity begins with each player making a binding threat as to the (possibly mixed) strategy for G that he will use *if* agreement is not reached in the negotiations that follow. The ensuing negotiations consist simply of the Nash demand game being played. If the latter is appropriately smoothed, the choice of threats t_1 and t_2 at the first stage serves to determine a status quo $q(t_1,t_2)$ for the use of the Nash bargaining solution at the second stage. The players can write contracts specifying the use of lotteries, and hence we identify the set U of feasible deals with the convex hull of the set of payoff pairs available in G when this is played noncooperatively. This analysis generates a reduced game in which the payoff pair n(t) that results from the choice of the strategy profile t is the Nash bargaining solution for U relative to the status quo q(t).

Result 5 (Nash, 1953) The Nash threat game has an equilibrium, and all equilibria yield the same agreement in U.

The threat game is strictly competitive in that the players' preferences over the possible outcomes are diametrically opposed. The result is therefore related to von Neumann's maximin theorem for two-person zero-sum games. In particular, the equilibrium strategies are the players' security strategies and the equilibrium outcome gives each player his security level. For a further discussion of the Nash threat game, see Owen [1982].

The model described above, together with Nash's [1953] axiomatic defense of the same result, is often called his *variable threats* theory. The earlier model, in which q is given, is then called the <u>fixed threat</u> theory and q itself is called the <u>threat point</u>. It needs to be remembered, in appealing to either theory, that the threats need to have the character of conditional <u>commitments</u> for the conclusions to be meaningful.

The Harsanyi-Zeuthen model

In what Harsanyi [1977] calls the "compressed Zeuthen model", the first stage consists of Nash's simple demand game (with no smoothing). If the opening demands are incompatible, a second stage is introduced in which the players simultaneously decide whether to concede or to hold out. If both concede, they each get only what their opponent offered them. If both hold out, they get their status quo payoffs, which we normalize to be zero.

The concession subgame has three Nash equilibria. Harsanyi [1977] ingeniously marshals a collection of "semi-cooperative" rationality principles in defense of the use of Zeuthen's [1930] principle in making a selection from these three equilibria. Denoting by r_i the ratio between i's utility gain if j concedes and i's utility loss if there is disagreement, Zeuthen's principle is that, if $r_i > r_j$, then player i concedes. When translated into familiar terms, this calls for the selection of the equilibrium at which the Nash product of the payoffs is biggest. When this selection is made in the concession subgames, the equilibrium pair of opening demands is then simply the Nash bargaining solution.

The full Harsanyi-Zeuthen model envisages, not one sudden-death encounter, but a sequence of concessions over small amounts. However, the strategic situation is very similar and the final conclusion concerning the implementation of the Nash bargaining solution is identical.

Making commitments stick.

Crawford [1982] offers what can be seen as an elaboration of the compressed Harsanyi-Zeuthen model with a more complicated second stage in which making a concession (backing down from the "commitment") is costly to an extent that is uncertain at the time the original demands are made. He finds not only that impasse can occur with positive probability, but that this probability need not decrease as commitment is made more costly.

More recent work has concentrated on incomplete information about preferences as an explanation of disagreement between rational bargainers (see Section 7). In consequence, Schelling's [1960] view of bargaining as a "struggle to establish commitments to favorable bargaining positions" remains largely unexplored as regards formal modeling.

5. Pairwise bargaining with few agents

In many economic environments, the parameters of one bargaining problem are determined by the forecast outcomes of other bargaining problems. In such situations, the result of the bargaining is highly sensitive to the detailed structure of the institutional framework that governs how and when agents can communicate with each other. The literature on this topic remains exploratory at this stage, concentrating on a few examples with a view to isolating the crucial institutional features. We examine subgame-perfect equilibria of some elaborations of the model of Section 2.

One seller and two buyers.

An indivisible good is owned by a seller S whose reservation value is $v_S = 0$. It may be sold to one and only one of two buyers, H and L, with reservation values $v = v_H \geq v_L = 1$. In the language of cooperative game theory, we have a three-player game with value function V satisfying $V(\{S,H\}) = V(\{S,H,L\}) = v$, $V(\{S,L\}) = 1$, and $V(C) = 0$ otherwise. The game has a non-empty core in which the object is sold to H for a price $p \geq 1$ when $v > 1$. (When $v = 1$, it may be sold to either of the buyers at the price $p = 1$.) The Shapley value is $(1/6 + v/2,\ v/2 - 1/3,\ 1/6)$ (where the payoffs are given is the order S, H, L).

How instructive are such conclusions from cooperative theory? The following noncooperative models are intended to provide some insight. In these models, if the object changes hands at price p at time t, then the seller gets $p\delta^t$ and the successful buyer gets $(v_B - p)\delta^t$, where v_B is his valuation and $0 < \delta < 1$. An agent who does not participate in a transaction gets zero. Information is always perfect.

a. **"Auctioning"** [Binmore, 1985; Wilson, 1985]

The seller begins at time 0 by announcing a price, which both buyers hear. Buyer H either accepts the offer, in which case he trades with the seller and the game ends, or rejects it. In the latter case, buyer L then either accepts or rejects the seller's offer. If both buyers reject the offer, then there is a delay of length τ, after which both buyers simultaneously announce counter-offers; the seller may either accept one of these offers or reject both. If both are rejected then there is a delay of length τ, after which the seller makes a new offer; and so on.

b. **"Telephoning"** [Binmore, 1985]

The seller begins by choosing a buyer to call. During their conversation, the seller and buyer alternate in making offers, a delay of length τ elapsing after each rejection. Whenever it is the seller's turn to make an offer, he can hang up, call the other buyer and make an offer to him instead. An excluded buyer is not allowed to interrupt the seller's conversation with the other buyer.

c. **"Random matching"** [Rubinstein/Wolinsky, 1986])

At the beginning of each period, the seller is randomly matched with one of the two buyers with equal probability. Each member of a matched pair then has an equal chance of getting to make a proposal which the other can then accept or refuse. If the proposal is rejected, the whole process is repeated after a period of length τ has elapsed.

d. **"Acquiring property rights"** [Gul, 1989]

The players may acquire property rights originally vested with other players. An individual who has acquired the property rights of all members of the coalition C enjoys an income of V(C) while he remains in possession. Property rights may change hands as a consequence of pairwise bargaining. In each period, any pair of agents retaining property rights has an equal chance of being chosen to bargain. Each of the matched pair then has an equal chance of getting to make a proposal to the other about the rate at which he is willing to rent the property rights of the other. If the responder agrees, he leaves the game and the remaining player enjoys the income derived from coalescing the property rights of both. If the responder refuses, both are returned to the pool of available bargainers. In this model, a strategy is to be understood as stationary if the behavior for which it calls depends only on the current distribution of property rights and not explicitly on time or other variables.

Result 6

(a) **[Binmore, 1985]** If, in the <u>auctioning</u> model, $1 < v \le 2$ and $\delta^\tau > 1/v$, or $v = 1$, then there is a subgame-perfect equilibrium, and in all such equilibria the good is sold immediately (to H if $v > 1$) at the price $\delta^\tau + (1 - \delta^\tau)v$. If $v > 2$ and τ is sufficiently small, then the only subgame-perfect equilibrium outcome is that the good is sold to H at the bilateral bargaining price (of approximately $v/2$) that would obtain if L were absent altogether.

(b) **[Binmore, 1985]** For the <u>telephoning</u> model, if $v \ge 1$, in any subgame-perfect equilibrium immediate agreement is reached on the bilateral bargaining price (approximately $v/2$ when τ is small) that would obtain if L were absent altogether. If $v > 1$ then the good is sold to H.

(c) **[Rubinstein/Wolinsky, 1986]** For the <u>random matching</u> model when $v = 1$, there is unique subgame-perfect equilibrium in which the good is sold to the first matched buyer (an inefficient outcome) for a price that is approximately 1 when τ is small. (The case $v > 1$ has not been analyzed yet).

(d) **[Gul, 1989]** For the <u>acquiring property rights</u> model, among the class of stationary subgame-perfect equilibria there is a unique equilibrium, in which all matched pairs reach immediate agreement. When τ is small, this equilibrium assigns each player an expected income approximately equal to his Shapley value allocation.

Remark Some of the above results are valid for more general cooperative games (see for example Jun [1987]). Gul's result is valid for all cooperative games with a strictly superadditive value function. For a distinct but related implementation of the Shapley value, see Dow [1989]. For a very different "semi-cooperative" interpretation of the Shapley value, see Harsanyi [1977]. Horn and Wolinsky [1986, 1988] have analyzed a three player cooperative game in which $V(1,2,3) > V(1,2) = V(1,3) > 0$ and for all other coalitions C we have $V(C) = 0$. In this case, the game does not end as soon as one agreement is reached and the question of whether the first agreement is implemented immediately becomes an important factor. (Related work appears in Jun [1989] and Chae/Yang [1988]).

6. Bargaining in markets

Bargaining theory provides a natural framework within which to study price formation in markets where transactions are made in a decentralized manner via interaction between pairs of agents rather than being organized centrally through the use of a formal trading institution like an auctioneer. One might describe the aim of investigations in this area as that of providing "mini-micro" foundations for the micro-economic analysis of markets and, in particular, to determine the range of validity of the Walrasian paradigm. Such a program represents something of a challenge for game theorists in that its success will presumably generate new solution concepts for market situations intermediate between those developed for bilateral bargaining and the notion of a Walrasian equilibrium.

Early studies of matching and bargaining models are Diamond [1981], Mortensen [1982a, b] and Diamond/Maskin [1979] in which bargaining is modeled using cooperative game theory. This approach is to be contrasted with the noncooperative approach of the models that follow. A pioneering paper in this direction is that of Butters [1977].

The models that exist differ in their treatment of several key issues. First, there is the information structure. What does a player know about the events in other bargaining sessions? Second, there is the question of the detailed structure of the pairwise bargaining games. In particular, when can a player opt out? Third, there is the modeling of the search technology through which the bargainers get matched. Finally, there is the nature of the data given about agents in the market. Sometimes, for example, it relates to stocks of agents in the market, and sometimes to flows of entrants or potential entrants.

Markets in steady state [Rubinstein/Wolinsky, 1985]

Most of the literature has concentrated on a market for an indivisible good in which agents are divided into two groups, sellers and buyers. All the sellers have reservation value 0 for the good and all the buyers have reservation value 1. A matched seller and buyer can agree on any price p with $0 \le p \le 1$. If agreement is reached at time t, then the seller leaves the market with a von Neumann and Morgenstern utility of $p\delta^t$ and the buyer leaves with $(1 - p)\delta^t$.

The first event in each period is a matching session in which all agents in the market participate, including those who may be matched already. Any seller has a probability σ of being matched with a buyer and any buyer has a probability β of being matched with a seller. The numbers σ and β are assumed to be constant so that the economic environment remains in a steady state.

Bargaining can take place only between individuals in a matched pair. After the matching session, each member of a matched pair is equally likely to be chosen to make the first offer. This may be accepted or rejected by the proposer's partner. If it is accepted, both leave the market. In either case, the next period commences after time τ has elapsed.

Pairs who are matched at time t but do not reach agreement remain matched at time $t + \tau$, unless one or both partners get matched elsewhere. In the model, an agent <u>must</u> abandon his old partner when matched with a new one. Thus, for example, a seller with a partner at time t who does not reach agreement at time t, has probability $(1 - \sigma)\beta$ of being without a partner at time $t + \tau$. (A story can be told about the circumstances under which it would always be optimal to abandon the current partner if this decision were the subject of strategic choice, but this issue is neglected here.)

The model is not a game in the strict sense. For example, the set of players is not specified. Nevertheless, a game-theoretic analysis makes sense using a solution concept that is referred to as a "market equilibrium". This is a pair of strategies, one for buyers and one for sellers, that satisfies:

1. semi-stationarity The strategies prescribe the same bargaining tactics for all buyers (or sellers) independently of their personal histories.
2. sequential rationality The strategies are optimal after all possible histories.

Result 7 (Rubinstein/Wolinsky, 1985) There is a unique market equilibrium. As $\tau \to 0+$, the price at which the good changes hands converges to $\sigma/(\sigma + \beta)$.

The probabilities σ and β depend on the matching technology which depends in turn on how search is modeled. Let S and B be the steady-state measures of sellers and buyers respectively, and consider the most naïve of search models in which $\sigma = c\tau[B/(S + B)]$ and $\beta = c\tau[S/(B + S)]$, where the constant c represents a "search friction". In the limit as $\tau \to 0+$, the market equilibrium price approaches $B/(B + S)$. Thus, for example, if there are few sellers and many buyers, the price is high.

Notice that the short side of the market does not appropriate the entire surplus even in the case when search frictions become negligible. Gale [1987] points out that, if this conclusion seems paradoxical, it is as a consequence of thinking of supply and demand in terms of the stocks S and B of agents in the market at any time. To keep the market in a steady-state, the flows of buyers and sellers into the market at any time have to be equal. If supply and demand are measured in terms of these flows, then any selling price is Walrasian. For further discussion, see Rubinstein [1987, 1989].

Unsteady states [Binmore/Herrero, 1988a, b]

Binmore and Herrero [1988a,b] generalize the preceding model in two directions. The informational difficulties finessed by Rubinstein and Wolinsky's "semi-stationarity" condition are tackled by observing that subgame-perfect equilibria in alternating-offers models can be replaced by "security equilibria" without losing the uniqueness conclusion. A security equilibrium is related to the notion of "rationalizability" introduced by Bernheim [1984] and Pearce [1984]. Their requirement about its being common knowledge that strictly dominated strategies are never played, is replaced by a similar requirement concerning security levels. It is assumed to be common knowledge that no player takes an action under any contingency that yields less than he calculates his security level to be given the occurrence of the contingency. Any equilibrium notion normally considered is also a security equilibrium. A proof of uniqueness for security equilibria therefore entails uniqueness for more conventional equilibria also. However, in markets with a continuum of traders, security equilibria are insensitive to the players' personal histories. The immediate point is that stationarity restrictions on the equilibrium concept used in the Rubinstein/Wolinsky model and its relatives are not crucial in obtaining a uniqueness result (provided $\delta < 1$).

The second generalization of the Rubinstein/Wolinsky model results from applying the technique to markets which are not necessarily in a steady-state in that the equilibrium measures of traders may vary with time as a consequence of satisfied traders leaving the market without there being an exactly counterbalancing inflow of new traders. Closed-form conclusions are obtained for the continuous time case obtained by considering the limit as $\tau \to 0+$. In particular, the equilibrium deal can be expressed as an integral involving the equilibrium probabilities that a buyer or a seller is matched at all future times.

Aside from the steady-state model, the simplest special case occurs when no new traders enter the market after time 0. There is then no replacement of those traders present at time 0 when they finally conclude a successful deal and leave the market. With the naive search technology considered in the Rubinstein/Wolinsky model, the following Walrasian conclusion may be obtained:

Result 8 (Binmore/Herrero, 1988a,b) There is a unique security equilibrium. As search frictions become negligible, the equilibrium deal approximates that in which the entire surplus is assigned to agents on the short side of the market.

Among many other results, Gale [1987] has extended versions of both Results 7 and 8 to the case in which there is a spectrum of reservation prices on both sides of the market.

Divisible goods with multiple trading [Gale, 1986c]

Gale [1986a, b, c] studies traditional barter markets in which many divisible goods are traded and agents can transact many times before leaving the market. We now describe one of the models from Gale [1986c] (which is a simplification of the earlier paper Gale [1986a]). The existence of market equilibrium was established in Gale [1986b]. The relation between Gale's work and general equilibrium theory is explored in McLennan/Sonnenschein [1989].

All agents, of which there are K types, enter the market at time zero. Initially, there is a measure n_k of agents of each type $k = 1,2,\ldots,K$. An agent of type k is characterized by his initial commodity bundle w_k and his utility function $u_k:\mathbb{R}_+^m \cup \{D\} \to \mathbb{R} \cup \{-\infty\}$, where $\mathbb{R}_+^m$ is the space of commodity bundles with which he might leave the market and D is the event of his remaining in the market for ever. Agents are not impatient ($\delta = 1$) and bundles may be stored costlessly.

Each period begins with a matching session which operates independently of past events. In particular, no matches survive from previous periods. The probability of a given agent getting matched with an agent with specified characteristics is proportional to the current measure of such agents in the population. Once a match is established, each of the paired players learns the type of his partner and his partner's commodity bundle. Bargaining then begins. Each member of a matched pair is equally likely to be chosen to make a proposal. This must consist of a vector representing a feasible transfer of goods from himself to his bargaining partner. This proposal may be accepted or rejected. If it is rejected, the responding player then decides whether or not to leave the market. An important assumption is that players do not leave the market except after such a rejection.

As trade occurs, the bundle held by each agent changes. Given the restrictions on strategies imposed below, the number of different bundles held is always finite. Thus, in any period the state of the market can be characterized by a finite list (c_i, k_i, ν_i), where c_i is a feasible holding and ν_i is the measure of agents of type k_i holding c_i.

A market equilibrium is defined to be a K-tuple of strategies, one for each type, which satisfies the following conditions:

1. semi-stationarity The bargaining tactics prescribed by the strategy depend only on time, the agent's current bundle and the opponent's type and current bundle.
2. sequential rationality Whenever an agent makes a decision, the strategy calls for an optimal decision, given the strategies of the other types and the current state of the market.

The definition is closely related to the notion of sequential equilibrium with "passive conjectures".

A K-tuple of bundles $(x_1,\ldots,x_K)$ is an allocation if $\Sigma n_k x_k = \Sigma n_k w_k$. If there exists a price vector p such that, for all k, the bundle x_k maximizes u_k subject to the budget constraint $px \le pw_k$, then the allocation is Walrasian. Gale's concern is with the circumstances under which the equilibrium outcome is Walrasian.

For technical reasons, Gale restricts the utility functions to be considered. Here (as in the presentation in Osborne and Rubinstein [1990]) we require the existence of an increasing and continuous function $\phi_k:\mathbb{R}^m_+ \to \mathbb{R}$ which is zero on the boundary of $\mathbb{R}^m_+$ and strictly concave in its interior. For a given $\phi > 0$, it is then required that

$$u_k(x) = \begin{cases} \phi_k(x) & \text{if } \phi_k(x) \ge \phi \\ -\infty & \text{otherwise.} \end{cases}$$

In addition, some regularity conditions have to be imposed on the indifference curves. It is sufficient, for example, that ϕ_k have a uniformly continuous derivative.

Result 9 (Gale 1986a, b) For every market equilibrium, there is a Walrasian allocation $(x_1,\ldots,x_K)$ such that each agent of type k leaves the market holding bundle x_k with probability one.

The constraint that the strategies have to be semi-stationary may reflect an assumption about the information available to the players. The role of the informational structure in such models is explored in Rubinstein/Wolinsky [1986]. In particular, it is shown that, in a model with $\delta = 1$ and a finite number of traders, any price can be supported as a sequential equilibrium, provided that players are permitted perfect knowledge of the events in the market or even if the players are able to recall only their personal histories.

Notes Other related works include Shaked/Sutton [1984] which aims to explain involuntary unemployment with a stylized model of the labor market. Wolinsky

[1987a] makes the intensity with which agents search an endogenous variable. (See also Bester [1988], Chikte/Deshmukh [1985] and Muthoo [1988] on this latter topic.) Wolinsky [1988] analyzes the case in which transactions are made by auction, rather than by matching and bargaining.

A very challenging issue is that of extending the models to include asymmetric information. A pioneering work in this direction is Wolinsky [1987b]. In Wolinsky's model the equilibrium outcome of a decentralized trading process may not approximate the rational expectations equilibrium of the corresponding trading process, even when the market is "approximately frictionless". For related models see Rosenthal/Landau [1981] and Samuelson [1987].

7. Bargaining with incomplete information

This section presents some attempts to build theories of bargaining when the information available to the bargainers about their opponents is incomplete. The proposals and responses in an alternating-offers model then do more than register a player's willingness to settle on a particular deal: they also serve as signals by means of which the players may communicate information to each other about their private characteristics. Such signals need not be "truthful". A player in a weak bargaining position may find it worthwhile to imitate the bargaining behavior that he would use if he were strong with a view to getting the same deal as a strong player would get. A strong player must therefore consider whether or not to choose a bargaining strategy that it would be too costly for a weak player to imitate lest the opponent fail to recognize that he is strong. Such issues are studied in the literature on signalling games (Chapter ?) which is therefore central to what follows.

A central goal in studying bargaining with incomplete information is to explain the delays in reaching agreement that we observe in real-life bargaining. (Recall that the alternating-offers model of Result 1, in which information is complete, predicts no delay at all). Much has been learned in pursuing this goal, but its attainment remains elusive. In this section, we propose to do no more than indicate the scope of the difficulties as currently seen.

The literature uses the Kreps/Wilson [1982] notion of a sequential equilibrium after reducing the bargaining situation with incomplete information to a game with imperfect information in accordance with Harsanyi's [1967/68] theory, within which each player is seen as being chosen by a chance move from a set of "types" of player that he might have been. Although subgame-perfection is a very satisfactory concept for complete information bargaining games, the set of sequential equilibria for bargaining games with incomplete information is typically enormously large. It is therefore necessary, if informative results are to emerge, to refine the notion of sequential equilibrium. Progress in the study of bargaining games of incomplete information, as with signalling games in general, is therefore closely tied to developments in the literature on refinements of sequential equilibrium. It should be noted, however, that advances in refinement theory have only a tentative character. Although one idea or another may seem intuitively plausible in a particular context, the theory lacks any firmly grounded guiding principles. Until these problems in the foundations of game theory are better understood, it therefore seems premature to advocate any of the proposed resolutions of the problem of bargaining under incomplete information for general use in economic theory.

An alternating-offers model with incomplete information [Rubinstein, 1985a,b]

We return to the problem of "dividing the dollar" in which the set of feasible agreements is identified with $A = [0,1]$. For simplicity, we confine attention to the case of fixed costs per unit time of delay. Recall that the players' preferences over the possible deals (a,t), in which 1 gets a and 2 gets $1 - a$ at time t, may then be represented by $a - c_1 t$ and $1 - a - c_2 t$, where $c_i > 0$ $(i = 1,2)$. Player 1's cost $c_1 = c$ per unit time of delay is taken to be common knowledge, but 2's cost c_2 is known for certain only by 2. It is common knowledge only that player 1 initially believes that c_2 must take one of the two values c_W or c_S and that the probability of the former is π_W. It is assumed that $c_S < c < c_W$ and the costs are small enough that $c + c_W + c_S < 1$. The interval between successive proposals is fixed at $\tau = 1$ except where otherwise noted.

Having a high cost rate is a source of weakness in one's bargaining position. For example, if $\pi_W = 1$, so that it is certain that 2 has a higher cost rate than 1, then we have seen that 1 gets the entire surplus in equilibrium. On the other hand, if $\pi_W = 0$, so that it is certain that 2 has a lower cost rate than 1, then 1 gets only c_S. For this reason, a high cost type of 2 is said to be <u>weak</u> and a low cost type to be <u>strong</u>.

In the context of this model, a sequential equilibrium is a strategy triple, one for player 1 and one each for the two possible types of player 2, combined with a belief function which assigns, to every possible history after which player 1 has to move, the probability that player 1 attaches to the event that player 2 is weak. The beliefs have to be updated using Bayes' Rule whenever this is possible, and the initial belief has to be π_W. The strategy of each player must be optimal after <u>every</u> history (sequential rationality). We impose two auxiliary requirements. First, if the probability that player 1 attaches to the event that player 2 is weak is zero (one) for some history, it remains zero (one) subsequently. Thus, once player 1 is convinced of the identity of his opponent, he is never dissuaded of this view. Second, player 1 never changes his belief in response to a deviation by himself.

As is shown in Rubinstein [1985a,b], many sequential equilibria may exist:

1. If $2c/(c + c_W) < \pi_W$ then in all sequential equilibria player 1's expected payoff is at least $\pi_W + (1 - \pi_W)(1 - c_W - c)$.
2. If $2c/(c + c_W) \geq \pi_W$, then, for any a^* between c and $1 - c - c_S$, there exists a ("pooling") sequential equilibrium in which both a weak and a strong player 2 accept an opening demand of a^* by player 1.
3. If $2c/(c + c_W) \geq \pi_W \geq (c + c_S)/(c + c_W)$, then, for any $a^* \geq c_W$ there exists a ("separating") sequential equilibrium in which a weak player 2 accepts the opening demand of a^* by player 1, but a strong player 2 rejects this demand and makes a counter-offer of $a^* - c_W$, which player 1 accepts.

The multiplicity of equilibria arises because of the freedom permitted by the concept of sequential equilibrium in attributing beliefs to players after they have observed a deviation from equilibrium. Such deviations are zero probability events and so cannot be dealt with by Bayesian updating.

We illustrate the ideas underlying these results by considering case (2). Let $1 - c - c_S \geq a^* \geq c$. We construct a sequential equilibrium in terms of two commonly held states-of-mind labeled I (for initial) and O (for optimistic). In state I, it is common knowledge that 1 believes that 2 is weak with probability π_W. In state O, it is common knowledge that 1 believes that 2 is weak for sure. The players continue in state I until there is a deviation by 2 after which they switch to state O. The transition occurs immediately after player 2 makes an offer, before player 1 responds. Once in state O, they remain there no matter what. (These conjectures are called "optimistic". They are useful in rendering deviations unattractive and hence in constructing multiple equilibria.) Player 1 and the weak type of player 2 behave is state O precisely as in the complete information case when it is certain that 2 has the high cost rate c_W; the strong type of player 2 uses a best response against player 1's strategy. In state I:

1. Player 1 demands a^* and accepts an offer of a if and only if $a \geq a^* - c$.
2. A strong player 2 offers $a^* - c$ and accepts only a demand of $a \leq a^*$.
3. A weak player 2 offers $a^* - c$ and accepts only a demand of $a < a^* - c + c_W$.

Some comments on why the parameters need to be restricted in order to sustain the equilibrium may be helpful. Notice that, if 1 demands more than a^* and less than $a^* - c + c_W$ at time 0, then he remains in state I and expect less than $\pi_W(a^* - c + c_W) + (1 - \pi_W)(a^* - c - c)$. The condition that this quantity not exceed a^* is that $2c/(c + c_W) \geq \pi_W$. The requirement that $a^* \geq c$ is simply to ensure that the offer $a^* - c$ be feasible. Finally, observe that the reason that it is optimal for a strong 2 to accept an opening demand of a^* is that $1 - a^* \geq c - c_S$. The latter quantity is what 2 gets by refusing. (The state shifts to O, in which player 1 believes he faces a weak opponent. The strong type of player 2 therefore makes, at time $\tau = 1$, an offer of c which 1 accepts. His expected payoff from the refusal is therefore $c - c_S$.)

Prolonged disagreement

Case 2 from the preceding subsection is now used to construct a sequential equilibrium in which the bargaining may be prolonged for many periods before agreement is achieved.

Choose three numbers x^*, y^* and z^* which satisfy $c \leq x^* < y^* < z^* \leq 1 - c + c_S$. The time that elapses in equilibrium before agreement is reached is denoted by N, where N is chosen to be the largest even integer smaller than

$$\min\{(y^* - x^*)/c,\ (z^* - y^* + c_W - c)/c_W,\ (z^* - y^* + c_S - c)/c_S\}.$$

Until the N^{th}th period, player 1 and both types of player 2 are required to hold out for the entire surplus. The players switch to a sequential equilibrium as described in case 2 of the previous subsection as follows:

(1) $a^* = x^*$ if 1 deviates from the prescribed behavior during the first N periods;

(2) $a^* = z^*$ if 2 deviates during the first N periods;

(3) $a^* = y^*$ if the N^{th} period is reached without a deviation.

A switch occurs immediately after a deviation. The bound on N ensures that 1 does not deviate at time 0. The prescribed play yields a payoff of $y^* - Nc$ as opposed to his best alternative, which is to demand x^*. It also ensures that a weak 2 does not deviate at the second period. The prescribed play yields a payoff of $1 - y^* - Nc_W$ as opposed to his best alternative, which is to offer $z^* - c$, whose acceptance yields a payoff of $1 - z^* + c - c_W$.

When the length of a period is τ, the parameters c, c_S and c_W in the above must be multiplied by τ and the delay time to agreement becomes $T = N\tau$. The limit of the latter as $\tau \rightarrow 0+$ is positive. Thus, there may be significant delay in reaching agreement, even when τ is small, although no information is revealed along the equilibrium path after a deviation occurs. Any deviation is interpreted as signalling weakness and leads to an equilibrium that favors the non-deviant.

Gul and Sonnenschein [1988] do not accept that such non-stationary equilibria are reasonable. In this context, stationarity refers to the assumption that players do not change their behavior so long as 1 does not change his belief about 2's type. A version of their result for the fixed costs model that we have been using as an example is that any sequential equilibrium in which 2's strategies are stationary must lead to an agreement no later than the second period.

In their paper, Gul and Sonnenschein analyze a more complex bargaining model between a seller and a buyer in which the seller's reservation value is 0 and the buyer's reservation value has a continuous distribution F with support [1, h]. They impose two properties in addition to stationarity on sequential equilibrium. The Monotonicity property requires that, for histories after which the seller's posterior distribution for the buyer's reservation value is the conditional distribution of F given [1, x], the seller's offer must be increasing in x. The No Free Screening property requires that the buyer's offer can influence the seller's beliefs only after histories in which at least one of the buyer's equilibrium offers is supposed to be accepted by the seller.

Result 10 (Gul/Sonnenschein, 1988) For all $\epsilon > 0$ there is $\tau^* > 0$ such that for all positive $\tau < \tau^*$, in every sequential equilibrium which satisfies stationarity, monotonicity and no free screening, the probability that bargaining continues after time ϵ is at most ϵ.

Gul and Sonnenschein conclude from Result 10 that bargaining with one-sided uncertainty leads to vanishingly small delays when the interval between successive proposals becomes sufficiently small. We are not convinced that such a sweeping conclusion is legitimate, although we do not deny that actual delays in real-life bargaining must often be caused by factors that are more complex than the uncertainties about the tastes or beliefs of a player as we have modeled them. Uncertainties about how rational or irrational an opponent is are probably at least as important. The reason for our skepticism lies in the fact that, as is shown by Ausubel and Deneckere [1989] and others, the result relies heavily on the stationarity assumption. As explained in Section 2, stationarity assumptions do more than attribute simplicity of behavior to the players: they also make players' beliefs insensitive to past events.

Note that Result 10 and that of Gul/Sonnenschein/Wilson [1986] have an importance beyond bargaining theory because of their significance for the "Coase conjecture". Note also that Vincent [1987] demonstrates that, if the seller and the buyer have correlated valuations for the traded item, then delay is possible when the time between offers goes to zero even under stationarity assumptions.

Refinements of sequential equilibrium in bargaining models [Rubinstein 1985a,b]

Our study of the fixed costs model shows that the concept of sequential equilibrium needs to be refined if unique equilibrium outcomes are to be obtained. To motivate the refinement that we propose, consider the following situation. Player 1 makes a demand of a which is rejected by 2 who makes a counter-offer of b, where $a - c_W < b < a - c_S$. If the rejection and the counter-offer are out of equilibrium, then the sequential equilibrium concept does not preclude 1 from assigning probability one to the event that 2 is weak. Is this reasonable? Observe that 2 rejects the demand a in favor an offer of b which, if accepted, leads to a payoff of $1 - b - c_S > 1 - a$ for the strong 2, but only $1 - b - c_W < 1 - a$ for the weak 2. One can therefore "rationalize" the offer of b on the part of the strong player but not on the part of the weak player. Should not this offer therefore convince 1 that his opponent is strong?

The next result, which is a version of Rubinstein [1985a,b], explores the hypothesis that players' beliefs incorporate such "rationalizations" about their opponents. The precise requirements for rationalizing conjectures are that, in any history after which player 1 is not certain that he faces the weak type of player 2:

1. If 2 rejects the offer a and makes a counter-offer b satisfying $a - c_W < b < a - c_S$, then 1 assigns probability one to the event that his opponent is strong.
2. If 2 rejects the offer a and makes a counter-offer b satisfying $a - c_W \geq b$, then 1 does not increase the probability he attaches to 2's being strong.

Result 11 (Rubinstein, 1985a,b) For any sequential equilibrium with rationalizing conjectures:

1. If $\pi_W > 2c/(c + c_W)$ then if 2 is weak there is an immediate agreement in which 1 gets the entire surplus, while if 2 is strong the agreement is delayed by one period, at which time 1 gets $1 - c_W$.
2. If $(c + c_S)/(c + c_W) < \pi_W < 2c/(c + c_W)$ then if 2 is weak there is an immediate agreement in which 1 gets c_W, while if 2 is strong the agreement is delayed by one period, at which time 1 gets nothing at all.
3. If $\pi_W < (c + c_S)/(c + c_W)$ then there is an immediate agreement in which 1 gets c_S whatever 2's type.

Rubinstein [1985a] provides a more general result applied to the family of time preferences explored in section 2. Various refinements of a similar nature have been proposed by numerous authors. In particular, Grossman and Perry [1986] propose a refinement they call "perfect sequential equilibrium" which seems to lead to plausible outcomes in bargaining models for which it exists.

Strategic delay [Admati/Perry, 1987]

One may modify the previous model by allowing a responding player to choose how much time may pass before he makes his counter-offer. He may either immediately accept the proposal with which he is currently faced or he may refuse the demand and choose a pair (a,Δ), where $a \in A$ is his counter-proposal and $\Delta \geq \tau$ is the length of the delay during which no player may make a new offer. (Without incomplete information, this modification has no bite. In equilibrium, each player minimizes the delay and chooses $\Delta = \tau$.)

The refinement of sequential equilibrium described here (which is somewhat stronger than that offered by Admati/Perry [1987]) is similar to that of the preceding section:

1. After any history that does not convince 1 that 2 is weak for sure, suppose that 1 demands a and that this demand is rejected by 2 who then counters with an offer of b after a delay of $\Delta \geq \tau$. If $1 - b - c_W\Delta < 1 - a \leq 1 - b - c_S\Delta$, then 1 concludes that 2 is strong for sure.
2. Suppose that 1 is planning to accept an offer a if this is delayed by Δ but that 2 delays a further $d > 0$ before making an offer b satisfying $1 - b - c_Wd < 1 - a \leq 1 - b - c_Sd$. Then, whatever the previous history, 1 concludes that 2 is strong for sure.

For $2c/(c + c_W) < \pi_W$, Admati/Perry [1987] show that any sequential equilibrium satisfying these additional assumptions has player 1 demanding the entire surplus at time 0. A weak player 2 accepts, but a strong player 2 rejects and makes a counter-offer of 0 after a delay of $1/c_W$ which player 1 accepts.

The result is to be compared with case 1 of Result 11 in which agreement is delayed by a vanishingly small amount when $\tau \to 0+$. Here, the delay in reaching agreement when 2 is strong does not depend on τ, and hence the delay persists in the limiting case as $\tau \to 0+$. (The constraint on π_W is necessary. See Admati/Perry [1987] for details.)

Other results

Strategic sequential bargaining models with incomplete information are surveyed by Cramton [1984], by Wilson [1987c], and by several contributors to Roth [1985]. We have dealt only with one-sided uncertainty. Cramton [1988] constructs a sequential equilibrium for the alternating offers model with two-sided uncertainty.

Bikhchandani [1985] points out that the sensitivity of the results on prolonged disagreement to certain changes in the bargaining procedure and in the solution concept employed. Perry [1986] seeks to endogenize the choice of the initial proposer.

The complexity of the analysis is reduced substantially if only two possible agreements are available; sharp results can then be obtained. See Chatterjee/Samuelson [1987]. Notice that this case is strongly related to games of attrition as studied in other game-theoretic contexts.

Finally, it has to be emphasized that many of the issues in bargaining with incomplete information that we have studied in this paper also arise in models in which only the uninformed party is allowed to make offers. For a presentation of such a model see, for example, Fudenberg/Levine/Tirole [1985].

8. Bargaining and Mechanism Design

The mechanism design literature regards a theory of bargaining as providing a mapping from the space of problem parameters to a solution to the bargaining problem. Attention is focused on the mappings or mechanisms that satisfy certain interesting properties, the aim being to study simultaneously the Nash equilibria for a large class of bargaining games of incomplete information without the need to specify each of the bargaining games in detail.

The rest of the section follows ideas appearing in the path-breaking paper of Myerson and Satterthwaite [1983]. The idea is explained in the context of a particularly simple case analyzed by Matsuo [1988].

A seller and a buyer of a single indivisible good have to negotiate a price. Both buyer and seller may be strong or weak, it being common knowledge that the prior probability of each possible pairing of types is the same. A player's strength or weakness depends on his reservation value which may be s_1, s_2, b_1 or b_2, where $0 = s_1 < b_1 < s_2 < b_2$. We let $s_2 = b_1 + \delta$ and assume that $b_1 = s_1 + \epsilon$, and $b_2 = s_2 + \epsilon$.

A mechanism M in this context is a mapping that assigns an outcome to each realization of (s,b). An outcome is a pair consisting of a price and a probability. Thus a mechanism is a pair of functions (p,π). The interpretation is that when the realization ia (s,b) then with probability $\pi(s,b)$ agreement is reached on the price $p(s,b)$, and with probability $1 - \pi(s,b)$ there is disagreement. The expected utility gain to a seller with reservation value s from the use of the mechanism M is $U(s) = E_b\pi(s,b)(p(s,b) - s)$. The expected utility gain to a buyer with reservation utility b is $V(b) = E_s\pi(s,b)(b - p(s,b))$.

Suppose that the buyer and the seller negotiate by choosing strategies in a non-cooperative bargaining game. A mechanism M can then be constructed by making a selection from the Nash equilibria of this game. It should be noted that the restriction of the set of outcomes to consist of a price and a probability significantly limits the scope of bargaining games to which the current investigation is applicable.

If the bargaining game has the properties that each player's security level is at least as large as his reservation value and that the action spaces are independent of the type of a buyer or a seller, then the mechanism must satisfy the following constraints:

Individual Rationality For all s and b we have $U(s) \geq 0$ and $V(b) \geq 0$.
Incentive Compatibility For all s, s', b and b' we have

$$U(s) \geq E_b\pi(s',b)(p(s',b) - s) \quad \text{and} \quad V(b) \geq E_s\pi(s,b')(b - p(s,b')).$$

If the mechanism represents the outcome of a game, the second condition asserts that no player prefers to use the strategy employed by another player. (Note that we are not necessarily discussing a direct mechanism and so the strategies need not consist simply of an announcement of a player's type. However, one could, of course, apply the "revelation principle" (Chapter ???) and thereby study only direct mechanisms without loss of generality.)

An efficient mechanism is a mechanism which induces an agreement whenever a surplus exists (i.e. b > s). In our example, a surplus exists except when a low reservation value buyer confronts a high reservation value seller (i.e.

$s = s_2$ and $b = b_1$).

We now explain why an efficient mechanism satisfying Individual Rationality and Incentive Compatibility exists if and only if $2\epsilon \geq \delta$.

Assume first that $2\epsilon < \delta$. Let $\sigma(s)$ denote the probability with which a seller with reservation value s reaches agreement. Let $\beta(b)$ be similarly defined for buyers. The incentive compatibility constraints can then be rewritten as $(s_2 - s_1)\sigma(s_2) \leq U(s_1) - U(s_2) \leq (s_2 - s_1)\sigma(s_1)$ and $(b_2 - b_1)\beta(b_1) \leq V(b_2) - V(b_1) \leq (b_2 - b_1)\beta(b_2)$.

If an efficient mechanism exists, then $\sigma(s_2) = \beta(b_1) = 1/2$. It follows that $U(s_1) \geq U(s_1) - U(s_2) \geq (s_2 - s_1)/2$ and $V(b_2) \geq V(b_2) - V(b_1) \geq (b_2 - b_1)/2$.

The sum of the expected gains to a strong buyer and a strong seller is then $U(s_1)/2 + V(b_2)/2 \geq (s_2 - s_1 + b_2 - b_1)/4 = (\epsilon + \delta)/2$, but the total expected surplus is only $[(b_2 - s_1) + (b_2 - s_2) + (b_1 - s_1) + 0]/4 = \epsilon + \delta/4 < (\epsilon + \delta)/2$. No efficient mechanism can therefore exist.

Next, assume that $\delta \leq 2\epsilon$. We now construct a game in which there is a Nash equilibrium that induces an efficient mechanism. In the game, the seller announces either s_1 or s_2 and the buyer announces either b_1 or b_2. The following table indicates the prices (<u>not</u> payoffs) that are then enforced (D means disagreement).

	b_1	b_2
s_1	b_1	$(s_1 + b_2)/2$
s_2	D	s_2

This game has a Nash equilibrium in which all types tell the truth and in which an efficient outcome is achieved. Notice in particular that if a weak seller is honest and reports s_1, he gets a payoff of $(s_1 + b_2 + 2b_1)/4$, while if he is dishonest and reports s_2, he gets $(s_1 + s_2)/2$. But $(s_1 + b_2 + 2b_1)/4 - (s_1 + s_2)/2 = (2\epsilon - \delta)/4 \geq 0$.

■

The above example illustrates some of the ideas of Myerson/Satterthwaite [1983]. They offer some elegant characterization results for incentive-compatible mechanisms from which they are able to deduce a number of interesting conclusions. In particular:

Result 12 (Myerson/Satterthwaite, 1983) Let $\bar{s} \leq \underline{b} < \bar{s} \leq \underline{b}$. If s is distributed with positive density over the interval $[\underline{s}, \bar{s}]$ and b is independently distributed with positive density over the interval $[\underline{b}, \bar{b}]$, then <u>no</u> incentive-compatible, individually rational mechanism is efficient.

Given this result, it is then natural to ask what can be said about the mechanisms that maximize expected total gains from trade. The conclusion of Myerson and Satterthwaite in the case when both s and b are uniformly distributed on [0,1] is a neat one: the expected gains from trade are maximized by a mechanism that transfers the object if and only if $b \geq s + 1/4$. Chatterjee and Samuelson [1983] had previously shown that the sealed-bid, double auction, in which the object is sold to the buyer at the average of the two bid prices whenever the buyer's bid exceeds the seller's, admits an equilibrium in which this maximal gain from trade is achieved. (The seller proposes the price $2s/3 + 1/4$ and the buyer proposes $2b/3 + 1/12$.)

The mechanism design approach is more general than that of noncooperative bargaining theory with which this chapter has been mostly concerned. However, the above mechanism design results, although wide in their scope of the situations to which they apply, do no more than to classify scenarios in which efficient outcomes are or are not achievable in equilibrium. Even when an efficient outcome is achievable, it need not be the realized outcome in the class of noncooperative games that is actually relevant in a particular applied context. This trade-off between generality and immediate applicability is one which we noted before in comparing cooperative and noncooperative game theory. As in that case, the two approaches should be seen as complementary, each providing insights where the other is silent.

9. Final Comments

In the past decade John Nash's [1950, 1953] pioneering work on noncooperative bargaining theory has been taken up again and developed by numerous authors. We see three directions in which progress has been particularly fruitful:

1. Sequential models have been introduced in studying specific bargaining procedures.
2. Refinements of Nash equilibrium have been applied.
3. Bargaining models have been embedded in market situations to provide insights into markets with decentralized trading.

In spite of this progress, important challenges are still ahead. The most pressing is that of establishing a properly founded theory of bargaining under incomplete information. A resolution of this difficulty must presumably await a major breakthrough in the general theory of games of incomplete information. From the perspective of economic theory in general, the main challenge remains the modeling of trading institutions (with the nature of "money" the most obvious target).

Because many of the results of noncooperative bargaining theory are relatively recent, there are few sources of a general nature that can be recommended for further reading. Harsanyi [1977] provides an interesting early analysis of some of the topics covered in the chapter. Roth [1985] and Binmore/Dasgupta [1987] are collections of papers the scope of which coincides with that of this chapter. Sutton [1986] and Rubinstein [1987] are survey papers. Osborne and Rubinstein [1990] contains a more detailed presentation of the material in this chapter.

References

Admati, A. R. and Perry, M. (1987), "Strategic Delay in Bargaining", Review of Economic Studies **54**, 345–364.

Ausubel, L. M. and R. J. Deneckere (1989), "Reputation in Bargaining and Durable Goods Monopoly", Econometrica, (to appear).

Bernheim, D. (1984), "Rationalizable Strategic Behavior", Econometrica **52**, 1007–1028.

Bester, H. (1988), "Bargaining, Search Costs and Equilibrium Price Distributions", Review of Economic Studies **55**, 201–214.

Bikhchandani, S. (1985), "A Bargaining Model with Incomplete Information", unpublished paper, Graduate School of Business, Stanford.

Binmore, K. G. (1985), "Bargaining and Coalitions" pp. 269–304 in Roth (1985).

Binmore, K. G. (1986), "Modelling Rational Players", ST/ICERD discussion paper 86/133, London School of Economics.

Binmore, K. G. (1987a), "Nash Bargaining Theory I, II, III", pp. 27–46, 61–76, and 239–256 in Binmore and Dasgupta (1987).

Binmore, K. G. (1987b), "Perfect Equilibria in Bargaining Models", pp. 77–105 in Binmore and Dasgupta (1987).

Binmore, K. G. (1987c), "Nash Bargaining and Incomplete Information", pp. 155–192 in Binmore and Dasgupta (1987).

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

Binmore, K. G. and M. J. Herrero (1988a), "Matching and Bargaining in Dynamic Markets", Review of Economic Studies **55**, 17–31.

Binmore, K. G. and M. J. Herrero (1988b), "Security Equilibrium," Review of Economic Studies **55**, 33–48.

Binmore, K. G., A. Rubinstein and A. Wolinsky (1986), "The Nash Bargaining Solution in Economic Modeling", Rand Journal of Economics **17**, 176–188.

Binmore, K. G., A. Shaked and J. Sutton (1988), "An Outside Option Experiment", unpublished paper, London School of Economics.

Butters, G. (1977), "Equilibrium Price Distributions in a Random Meetings Market", unpublished paper, Princeton University.

Chae, S. and J. Yang (1988), "The Unique Perfect Equilibrium of an N-person Bargaining Game", Economics Letters **28**, 221–223.

Chatterjee, K. and W. Samuelson (1983), "Bargaining under Incomplete

Information", Operations Research 31, 835-851.

Chatterjee, K. and L. Samuelson (1987), "Bargaining with Two-Sided Incomplete Information: An Infinite Horizon Model with Alternating Offers", Review of Economic Studies, 54, 175-192.

Chikte, S.P. and S. P. Deshmukh (1985), "The Role of External Search in Bilateral Bargaining", Discussion Paper 659, Center for Mathematical Studies in Economics and Management Science, Northwestern University.

Cramton, P. (1984), "The Role of Time and Information in Bargaining", Ph.D. thesis, Graduate School of Business, Stanford University.

Cramton, P. (1988), "Strategic Delay in Bargaining with Two-Sided Uncertainty", Working Paper 36, Yale School of Organization and Management.

Crawford, V. P. (1982), "A Theory of Disagreement in Bargaining", Econometrica 50, 607-638.

Cross, J. (1969), The Economics of Bargaining, New York: Basic Books.

Diamond, P.A. (1981), "Mobility Costs, Frictional Unemployment and Efficiency", Journal of Political Economy 89, 798-811.

Diamond, P. A. and E. Maskin (1979), "An Equilibrium Analysis of Search and Breach of Contract, I: Steady States", Bell Journal of Economics 10, 282-316.

Dow, G. K. (1989), "Knowledge is Power: Informational Precommitment in the Capitalist Firm", European Journal of Political Economy (to appear).

Fishburn, P. C. and A. Rubinstein (1982), "Time Preference", International Economic Review 23, 677-694.

Fudenberg, D. D. Levine, and J. Tirole (1985), "Infinite Horizon Models of Bargaining with One-Sided Incomplete Information", pp. 73-98 in Roth (1985).

Gale, D. (1986a), "Bargaining and Competition Part I: Characterization", Econometrica 54, 785-806.

Gale, D. (1986b), "Bargaining and Competition Part 2: Existence" Econometrica 54, 807-818.

Gale, D. (1986c), "A Simple Characterization of Bargaining Equilibrium in a Large Market Without the Assumption of Dispersed Characteristics", Working Paper 86-05, Center for Analytical Research in Economics and the Social Sciences, University of Pennsylvania.

Gale, D. (1987), "Limit Theorems for Markets with Sequential Bargaining", Journal of Economic Theory 43, 20-54.

Grossman, S. and M. Perry (1986), "Sequential Bargaining under Asymmetric Information", Journal of Economic Theory **39**, 120–154.

Gul, F. (1989), "Bargaining Foundations of Shapley Value", Econometrica (to appear).

Gul, F. and H. Sonnenschein (1988), "Bargaining with One-Sided Uncertainty", Econometrica **56**, 601–612.

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

Harsanyi, J. (1967/68), "Games with Incomplete Information Played by Bayesian Players", Parts I, II, III, Management Science **14**, 159–182, 320–334, 486–502.

Harsanyi, J. (1977), Rational Behavior and Bargaining Equilibrium in Games and Social Situations, Cambridge University Press.

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

Harsanyi, J. and R. Selten (1988), A General Theory of Equilibrium Selection in Games, Cambridge: MIT Press.

Horn, H. and A. Wolinsky (1986), "Worker Substitubality and Patterns of Unionization", unpublished paper.

Horn, H. and A. Wolinsky (1988), "Worker Substitutability and Patterns of Unionization", Economic Journal **98**, 484–497.

Jun, B. (1987), "A Structural Consideration on 3-Person Bargaining", Ph.D. thesis, University of Pennsylvania

Jun, B. (1989), "Noncooperative Bargaining and Union Formation", Review of Economic Studies **56**, 59–76.

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

Kreps, D. M. and Wilson, R. (1982), "Sequential Equilibrium" Econometrica **50**, 863–894.

Leininger, W., P. Linhart, and R. Radner (1986), "The Sealed Bid Mechanism for Bargaining with Incomplete Information", Journal of Economic Theory, (to appear).

Matsuo, T. (1988), "On Incentive Compatible, Individually Rational and ex Post Efficient Mechanisms for Bilateral Trading", unpublished paper, University of Pennsylvania.

Matthews, S. and A. Postlewaite (1987), "Pre-Play Communication in Two-Person, Sealed-Bid, Double Auctions", Discussion Paper 87-12, Center for Analytic Research in Economics and the Social Sciences, University of Pennsylvania.

McLennan, A. (1982), "A Noncooperative Definition of Two-Person Bargaining", Working Paper 8303, University of Toronto.

McLennan, A. and H. Sonnenschein (1989), "Sequential Bargaining as a Noncooperative Foundation for Walrasian Equilibrium", Econometrica (to appear).

Mortensen, D. (1982a), "Property Rights and Efficiency in Mating, Racing, and Related Games", American Economic Review **72**, 968–979.

Mortensen, D. (1982b), "The Matching Process as a Noncooperative Bargaining Game", pp. 233–258 in J. J. Mccall (ed.), The Economics of Information and Uncertainty, Chicago: University of Chicago Press.

Moulin, H. (1982), "Bargaining and Noncooperative Implementation", Discussion Paper A239 0282, Laboratorie d'Econometrie, Paris.

Moulin, H. (1984), "Implementing the Kalai-Smoridinsky Bargaining Solution" Journal of Economic Theory **33**, 32–45.

Muthoo, A. (1988), Bargaining Theory, Ph.D. thesis, Cambridge University.

Myerson, R. and M. Satterthwaite, (1983), "Efficient Mechanisms for Bilateral Trading", Journal of Economic Theory **29**, 265–281.

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

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

Osborne, M. J. and A. Rubinstein (1990), Bargaining and Markets, Academic Press.

Owen, G. (1982), Game Theory, Academic Press, second edition.

Pearce, D. (1984), "Rationalizable Strategic Behavior and the Problem of Perfection", Econometrica **52**, 1029–1050.

Perry, M. (1986), "An Example of Price Formation in Bilateral Situations: a Bargaining Model with Incomplete Information", Econometrica **54**, 313–321.

Rosenthal, R. and H. Landau (1981), "Repeated Bargaining with Opportunities for Learning", Journal of Mathematical Sociology **8**, 61–74.

Roth, A. E. (1978), "The Nash Solution and the Utility of Bargaining", Econometrica **45**, 657–664.

Roth, A. E. (1985) (ed.), Game Theoretic Models of Bargaining, Cambridge University Press.

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

Rubinstein, A. (1985a), "A Bargaining Model with Incomplete Information about Time Preferences", Econometrica **53**, 1151–1172.

Rubinstein, A. (1985b), "The Choice of Conjectures in a Bargaining Game with Incomplete Information", pp. 99–114 in Roth (1985).

Rubinstein, A. (1987), "A Sequential Strategic Theory of Bargaining, pp. 197–224 in T. Bewley (ed.), Advances in Economic Theory, Cambridge University Press.

Rubinstein, A. (1988), "Comments on the Interpretation of Game Theory", unpublished paper, London School of Economics.

Rubinstein, A. (1989), "Competitive Equilibrium in a Market with Decentralized Trade and Strategic Behavior: An Introduction" pp. 243-259 in G. Feiwel (ed.), The Economics of Imperfect Competition and Employment, London: Macmillan.

Rubinstein, A. and A. Wolinsky (1985), "Equilibrium in a Market with Sequential Bargaining", Econometrica **53**, 1133–1150.

Rubinstein, A. and A. Wolinsky (1986), "Decentralized Trading, Strategic Behavior and the Walrasian Outcome", Working Paper 86–12, Center for Analytic Research in Economics and the Social Sciences, University of Pennsylvania.

Samuelson, L. (1987), "Disagreement in Markets with Matching and Bargaining", unpublished paper, Pennsylvania State University.

Schelling, T. (1960), The Strategy of Conflict, Cambridge: Harvard University Press.

Shaked, A. (1987), "Opting Out: Bazaars versus 'Hi Tech' Markets", Discussion Paper 87/159, ST/ICERD, London School of Economics.

Shaked, A. and J. Sutton (1984), "Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model", Econometrica **52**, 1351–1364.

Ståhl, I. (1967), "Studier i bilaterala monolets teori". Licentiat thesis, HHS, Stockholm.

Ståhl, I. (1972), Bargaining Theory, Stockholm: Economics Research Institute.

Ståhl, I. (1988), "A Comparison Between the Rubinstein and Ståhl Bargaining Models", Research Report 6437, Stockholm School of Economics.

Sutton, J. (1986), "Non-Cooperative Bargaining Theory: An Introduction", Review of Economic Studies **53**, 709–724.

Vincent, D. (1987), "Bargaining with Common Values", Ph.D. thesis, Princeton

University.

von Neumann, J. and O.Morgenstern (1944), Theory of Games and Economic Behavior, Princeton University Press.

Wilson, R. (1985), "Notes on Market Games with Complete Information", Graduate School of Business, Stanford University.

Wilson, R. (1987a), "Game Theoretic Analysis of Trading", pp. 33-70 in T. Bewley (ed.), Advances in Economic Theory: The Fifth World Congress, Cambridge University Press.

Wilson, R. (1987b), "Efficient Performance in Two Agent Bargaining", Journal of Economic Theory **41**, 154-172.

Wilson, R. (1987c), "Bilateral Bargaining", Technical Notes, Graduate School of Business, Stanford University.

Wolinsky, A. (1987a), "Matching, Search, and Bargaining", Journal of Economic Theory **42**, 311-333.

Wolinsky, A. (1987b), "Information Revelation in a Market with Pairwise Meetings", unpublished paper, University of Pennsylvania.

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

Zeuthen, F. (1930), Problems of Monopoly and Economic Warfare, London: Routledge.

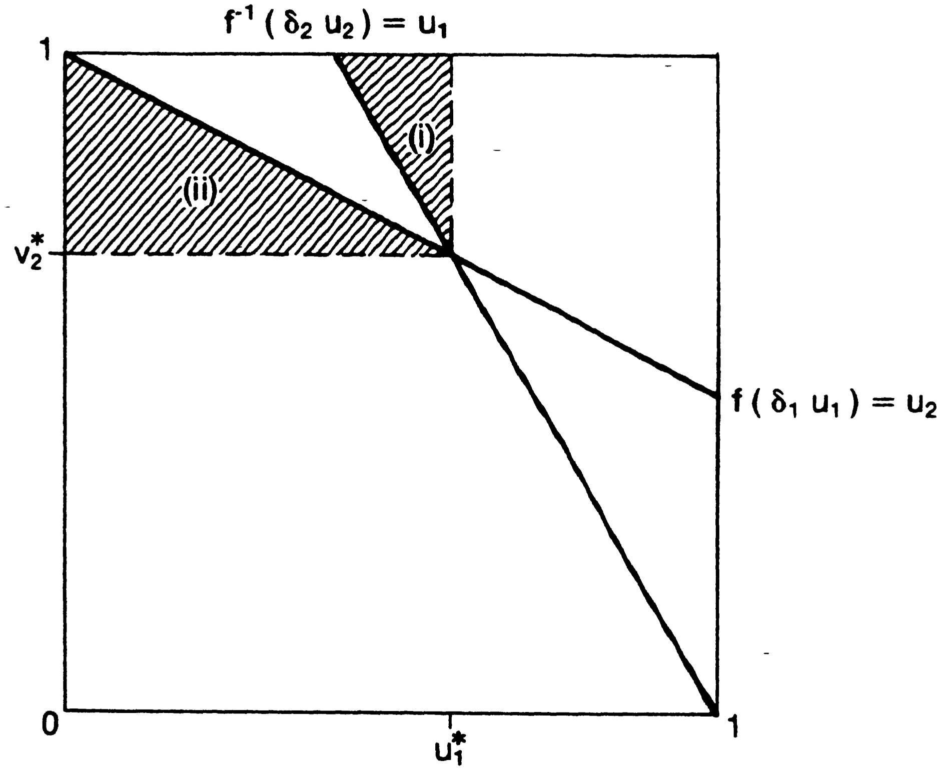

Figure 1

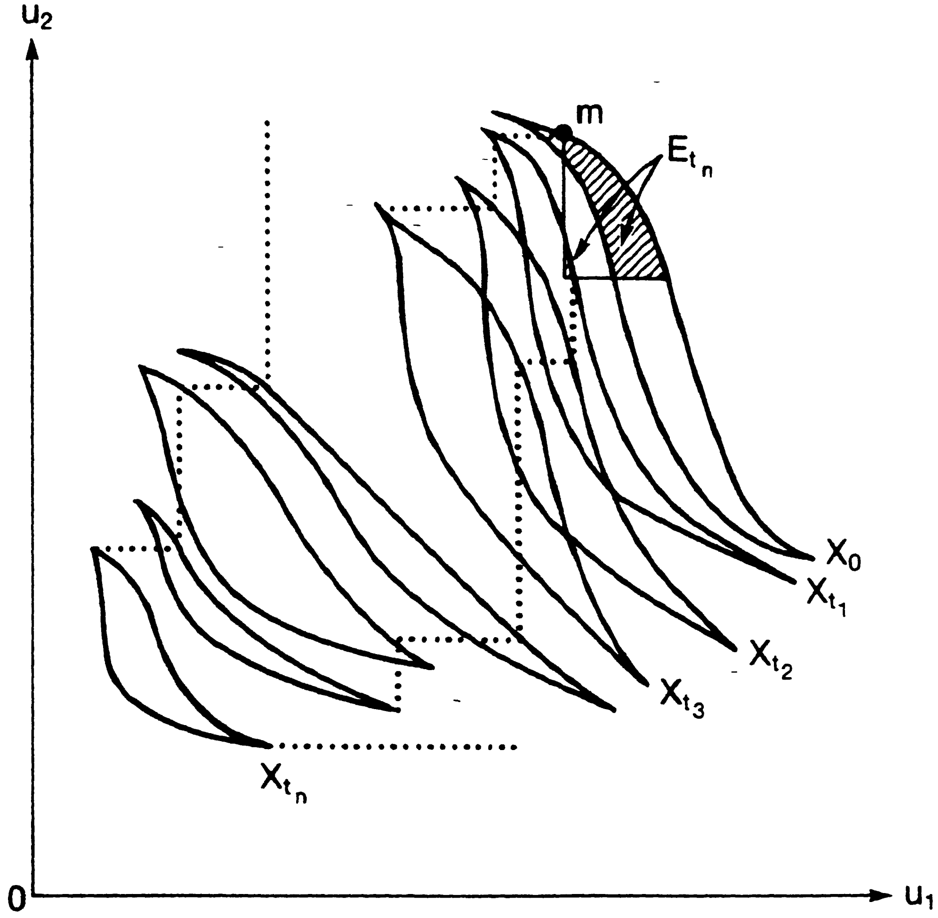

Figure 2

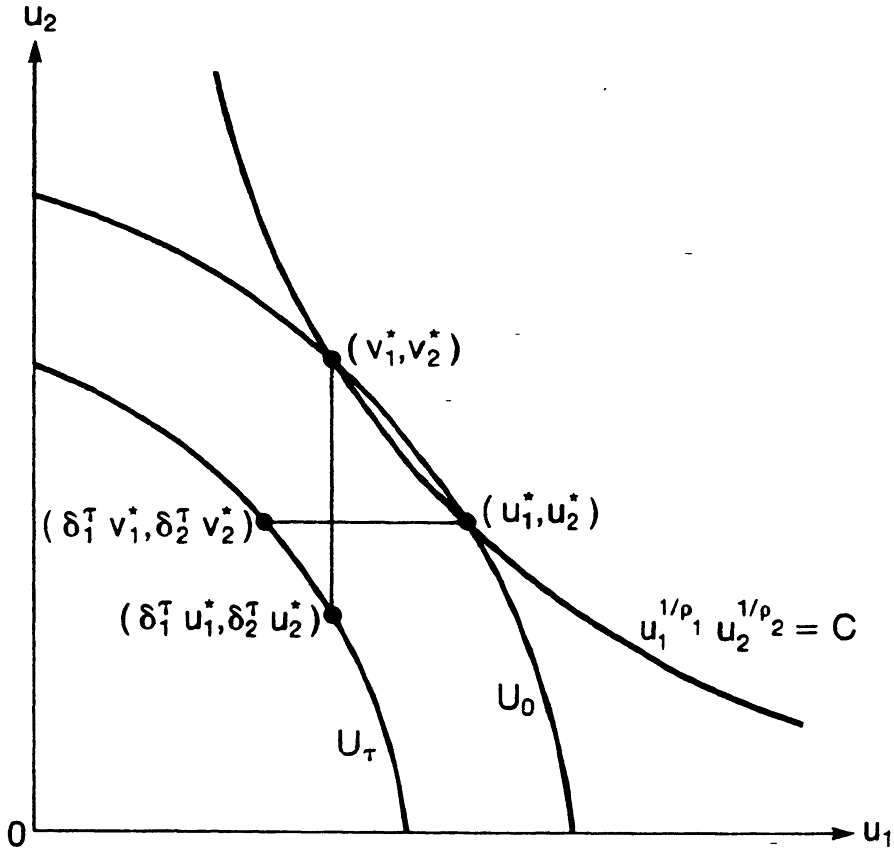
u_2
(v_1^*, v_2^*)
$(\delta_1^\tau v_1^*, \delta_2^\tau v_2^*)$
(u_1^*, u_2^*)
$(\delta_1^\tau u_1^*, \delta_2^\tau u_2^*)$
$u_1^{1/\rho_1} u_2^{1/\rho_2} = C$
U_0
U_τ
0
u_1

Figure 3

Recent CREST Working Papers

89–01: Mark Bagnoli, Severin Borenstein, "Carrot and Yardstick Regulation: Enhancing Market Performance with Output Prizes," October, 1988.

89–02: Ted Bergstrom, Jeffrey K. MacKie–Mason, "Some Simple Analytics of Peak-Load Pricing," October, 1988.

89–03: Ken Binmore, "Social Contract I: Harsanyi and Rawls," June, 1988.

89–04: Ken Binmore, "Social Contract II: Gauthier and Nash," June, 1988.

89–05: Ken Binmore, "Social Contract III: Evolution and Utilitarianism," June, 1988.

89–06: Ken Binmore, Adam Brandenburger, "Common Knowledge and Game Theory," July, 1989.

89–07: Jeffrey A. Miron, "A Cross Country Comparison of Seasonal Cycles and Business Cycles," November, 1988.

89–08: Jeffrey A. Miron, "The Founding of the Fed and the Destabilization of the Post-1914 Economy," August, 1988.

89–09: Gérard Gaudet, Stephen W. Salant, "The Profitability of Exogenous Output Contractions: A Comparative-Static Analysis with Application to Strikes, Mergers and Export Subsidies," July, 1988.

89–10: Gérard Gaudet, Stephen W. Salant, "Uniqueness of Cournot Equilibrium: New Results from Old Methods," August, 1988.

89–11: Hal R. Varian, "Goodness-of-fit in Demand Analysis," September, 1988.

89–12: Michelle J. White, "Legal Complexity," October, 1988.

89–13: Michelle J. White, "An Empirical Test of the Efficiency of Liability Rules in Accident Law," November, 1988.

89–14: Carl P. Simon, "Some Fine-Tuning for Dominant Diagonal Matrices," July, 1988.

89–15: Ken Binmore, Peter Morgan, "Do People Exploit Their Bargaining Power? An Experimental Study," January, 1989.

89–16: James A. Levinsohn, Jeffrey K. MacKie–Mason, "A Simple, Consistent Estimator for Disturbance Components in Financial Models," April 25, 1989.

89–17: Hal R. Varian, "Sequential Provision of Public Goods," July, 1989.

89–18: Hal R. Varian, "Monitoring Agents with Other Agents," June, 1989.

89–19: Robert C. Feenstra, James A. Levinsohn, "Distance, Demand, and Oligopoly Pricing," July 17, 1989.

89–20: Mark Bagnoli, Shaul Ben-David, Michael McKee, "Voluntary Provision of Public Goods," August, 1989.

89–21: N. Gregory Mankiw, David Romer, Matthew D. Shapiro, "Stock Market Forecastability and Volatility: A Statistical Appraisal," August, 1989.

89–22: Arthur J. Robson, "Efficiency in Evolutionary Games: Darwin, Nash and the Secret Handshake," 1989.

89–23: Mark Bagnoli, Ted Bergstrom, "Log-Concave Probability and Its Applications," September 7, 1989.

89–24: Gérard Gaudet, Stephen W. Salant, "Towards a Theory of Horizontal Mergers," July, 1989.

89–25 (evolved from 87-35): Stephen W. Salant, Eban Goodstein, "Predicting Committee Behavior in Majority-Rule Voting Experiments," July, 1989.

89–26: Ken Binmore, Martin J. Osborne, Ariel Rubinstein, "Noncooperative Models of Bargaining," 1989.

89–27: Avery Katz, "Your Terms or Mine? The Duty to Read the Fine Print in Contracts," February 19, 1989.