in *Game Theoretic Models of Bargaining*, edited by Alvin Roth, Cambridge University Press, Chapter 8, 149-179, 1985

Sequential bargaining mechanisms

Peter C. Cramton
YALE UNIVERSITY

8.1 Introduction

A fundamental problem in economics is determining how agreements are reached in situations where the parties have some market power. Of particular interest are questions of efficiency and distribution:

- How efficient is the agreement?
- How can efficiency be improved?
- How are the gains from agreement divided among the parties?

Here, I explore these questions in the context of bilateral monopoly, in which a buyer and a seller are bargaining over the price of an object.

Two features of my analysis, which are important in any bargaining setting, are information and impatience. The bargainers typically have private information about their preferences and will suffer some delay costs if agreement is postponed. Information asymmetries between bargainers will often lead to inefficiencies: The bargainers will be forced to delay agreement in order to communicate their preferences. Impatience will tend to encourage an early agreement and will make the parties' communication meaningful. Bargainers with high delay costs will accept inferior terms of trade in order to conclude agreement early, whereas patient bargainers will choose to wait for more appealing terms of trade.

Some authors have examined the bargaining problem in a static context, focusing solely on the role of incomplete information and ignoring the sequential aspects of bargaining. Myerson and Satterthwaite (1983) analyze bargaining as a direct revelation game. In this game, the players agree to a pair of outcome functions: one that maps the players' statements of their types into an expected payment from buyer to seller, and one that maps the players' statements into a probability of trade. These outcome functions are chosen in such a way that truthful reporting is an equilibrium strategy for the players. An important feature of this game is that it is static: Outcome functions are selected, the players report their true types, and then dice are rolled to determine the payment and whether or not trade occurs. To ensure that the players have the proper incentives for truthful reporting, the game will end with positive probability in disagreement even when there are substantial gains from trade. Thus, in the event that the randomization device calls for disagreement, the players may find themselves in a situation in which it is common knowledge that there are gains from trade.

Chatterjee and Samuelson (1983) analyze a strategic game in which both players make offers simultaneously, and trade occurs at a price between the two offers if the seller's offer is less than the buyer's offer. This game is closely related to the direct revelation game, in that it is static. Moreover, it can be shown that for a particular class of examples, the simultaneous-offers game implements the direct revelation game in which the outcome functions are chosen to maximize the players' ex ante utility. As in the direct revelation game, this game ends with positive probability in a state in which both bargainers know that gains are possible (since their respective reservation prices have been revealed), and yet they are forced to walk away from the bargaining table. Thus, the bargaining game assumes implicitly that the players are able to commit to walking away without trading, after it has been revealed that substantial gains from trade exist.

In situations where the bargainers are unable to make binding agreements, it is unrealistic to use a bargaining mechanism that forces them to walk away from known positive gains from trade. Such mechanisms violate a broad interpretation of *sequential rationality* as discussed by Selten (1976) (in terms of subgame perfection), and later by Kreps and Wilson (1982), if one applies sequential rationality not only to the hypothesized game, but to the game form as well. In particular, one should restrict attention to mechanisms that satisfy sequential rationality: It must never be common knowledge that the mechanism induced at any point in time is dominated by an alternative mechanism.

When there is uncertainty about whether or not gains from trade exist, any static game will violate sequential rationality. The players must have time to learn through each other's actions whether gains are possible. In a sequential game, the players communicate their preferences by exhibiting their willingness to delay agreement. Bargainers who anticipate large gains from trade (low-cost sellers and high-valuation buyers) will be unwilling to delay agreement, and so will propose attractive terms of trade that the other is likely to accept early in the bargaining process. On the other hand, high-cost sellers and low-valuation buyers will prefer to wait for better terms of trade. Static games must use a positive probability of disagreement to ensure incentive compatibility, where the probability of disagreement increases as the gains from trade shrink. The advantage of delaying agreement rather than forbidding agreement is that mechanisms can be constructed in which negotiations continue so long as each bargainer expects positive gains. Thus, the bargaining will not end in a state in which it is common knowledge that the players want to renege on their agreed-upon outcome.

Two approaches can be taken in the analysis of perfect bargaining games. The first approach is to examine specific extensive-form games, which determine the set of actions available to the players over time. Intrinsic to any bargaining process is the notion of offers and replies: Bargaining consists of a sequence of offers and decisions to accept or reject these offers. Who makes the offers; the time between offers, responses, and counteroffers; and the possibilities for commitment are determined by the underlying communication technology present in the bargaining setting. This communication technology will imply, in part, a particular bargaining game in extensive form. Cramton (1984), Sobel and Takahashi (1983), and Fudenberg, Levine, and Tirole (Chapter 5 in this volume) illustrate the analysis of particular extensive forms that are perfect bargaining games.

The second approach, and the one adopted in this chapter, is to analyze a general direct revelation game, which maps the players' beliefs into bargaining outcomes. An important distinction between direct revelation games and strategic games is that the direct revelation game does not explicitly model the *process* of bargaining. The sequence of offers and replies that eventually leads to an outcome is not studied in the direct revelation game as it is in strategic games. However, embedded in each sequential bargaining mechanism is a particular form of learning behavior, which can be analyzed. In addition, much can be learned about how information and impatience influence the efficiency of the bargaining outcome and the allocation of gains between players. Thus, even though bargainers will not play direct revelation games in practice, analysis of these games is a useful tool to determine how well the bargainers can hope to do by adopting an appropriate strategic game.

The difference between the static direct revelation game analyzed by Myerson and Satterthwaite (1983) and the sequential direct revelation game considered here is that in the sequential game, the outcome functions not only determine the probability and terms of trade, but also dictate *when* trade is to take place. In the static game trade may occur only at time zero whereas in the sequential game trade may occur at different times depending on the players' reports of their private information. Thus, by analyzing sequential bargaining mechanisms, one is able to infer what the players' learning process is over time. Furthermore, by analyzing mechanisms that are sequentially rational, one can study what bargaining outcomes are possible when the bargainers are unable to make binding agreements.

The introductory discussion presented in this chapter considers the simplest type of sequential bargaining games in which the players' time preferences are described by known and fixed discount rates. I begin by characterizing the class of perfect bargaining mechanisms, which satisfy the desirable properties of incentive compatibility (i.e., each player reports his type truthfully), individual rationality (i.e., every potential player wishes to play the game), and sequential rationality (i.e., it is never common knowledge that the mechanism induced over time is dominated by an alternative mechanism). It is shown that ex post efficiency is unobtainable by any incentive-compatible and individually rational mechanism when the bargainers are uncertain about whether or not they should trade immediately. I conclude by finding those mechanisms that maximize the players' ex ante utility, and show that such mechanisms violate sequential rationality. Thus, the bargainers would be better off ex ante if they could commit to a mechanism before they knew their private information. In terms of their ex ante payoffs, if the seller's delay costs are higher than those of the buyer, then the bargainers are better off adopting a sequential bargaining game rather than a static mechanism; however, when the buyer's delay costs are higher, then a static mechanism is optimal.

The methodology of this paper is based on Myerson and Satterthwaite (1983). I have freely borrowed from their insightful work in much of my analysis. Complete proofs for each proposition, even though many are only slightly different from the proofs found in Myerson and Satterthwaite, are given as an aid to the reader.

8.2 Formulation

Two parties, a buyer and a seller, are bargaining over the price of an object that can be produced by the seller at a cost s and is worth b to the buyer. The seller's cost s and the buyer's valuation b are also called their *reservation prices*, since they represent, respectively, the minimum and maximum price at which each party would agree to trade. Both the buyer and the seller have costs of delaying the bargaining process. Specifically, the value of the object is discounted in the future according to the positive discount rates r for the seller and s for the buyer. Thus, the payoffs, if the bargainers agree to trade at the discounted price x at time t, are $x - se^{-rt}$ for the seller and $be^{-st} - x$ for the buyer. Should the players fail to reach agreement, both of their payoffs are zero. Implicit in this formulation is the assumption that the bargainers discount future money at the same rate, so that at any time t the discounted payment by the buyer equals the discounted revenue to the seller. Without this assumption, it would be possible for the players to achieve an infinite payoff by having the player with the lower discount rate lend an arbitrarily large amount of money to the other player.

The buyer, although aware of his own valuation *b*, does not know the seller's cost of production *s*, but assesses this cost to be distributed according to the distribution $F(s)$, with a positive density $f(s)$ on $[\underline{s},\bar{s}]$. Similarly, the seller knows his cost *s*, but only assess the buyer's valuation to be distributed according to the distribution $G(b)$, with a positive density $g(b)$ on $[\underline{b},\bar{b}]$. Their discount rates and the distributions of the potential buyers and sellers are common knowledge. In addition, it is assumed that both the buyer and the seller are interested solely in maximizing their expected monetary gain.

To summarize, let $\langle F,G,r,s\rangle$ be a sequential direct revelation game, where

F = the distribution of the seller's cost *s* on $[\underline{s},\bar{s}]$,

G = the distribution of the buyer's valuation *b* on $[\underline{b},\bar{b}]$,

r = the seller's discount rate for the object,

s = the buyer's discount rate for the object.

In the revelation game, the players' actions consist of reports of their types, which are mapped into the bargaining outcome by the bargaining mechanism. Thus, the seller *s* reports that his cost is $s' \in [\underline{s},\bar{s}]$, and the buyer *b* reports that his valuation is $b' \in [\underline{b},\bar{b}]$. The revelation game is said to be *direct* if the equilibrium strategies of the players involve truthful reporting, that is, $(s',b') = (s,b)$. The important role of direct revelation games stems from the fact that one can, without loss of generality, restrict attention to direct mechanisms. For any Nash equilibrium of any bargaining game, there is an equivalent direct mechanism that always yields the same outcomes. This well-known result is called the *revelation principle*. Given any mechanism *M* that maps reports into outcomes, and a set of equilibrium strategies *x* that maps true types into reported types, then the composition $\hat{M} = M \circ x$ is a direct mechanism that achieves the same outcomes as the mechanism *M*.

For the revelation game $\langle F,G,r,s\rangle$, a *sequential bargaining mechanism* is the pair of outcome functions $\mathsf{T}(\cdot|\cdot,\cdot)$ and $x(\cdot,\cdot)$, where $\mathsf{T}(t|s,b)$ *is* the probability distribution that the object will be transferred to the buyer at time *t*, and $x(s,b)$ *is* the discounted expected payment from the buyer to the seller, given that the seller and buyer report the reservation prices *s* and *b*, respectively.

Typically, randomization of the outcomes over time is not necessary. Without randomization, the outcome function T can be replaced by the function $t(\cdot,\cdot)$, which determines the time of trade given the players' reports. A sequential bargaining mechanism, then, is the set of outcome functions $\langle t, x\rangle$ where $t(s,b)$ is the time of trade and $x(s,b)$ is the discounted expected payment, given that the seller reports *s* and the buyer reports *b*. Most bargaining mechanisms seen in practice require that the exchange of money and goods take place at the same time. Such a requirement is not restrictive in this model, because there is no benefit to be gained by exchanging money at a different time from the exchange of the good, since both players have identical time preferences for money. For reasons of tractability, I will frequently restrict attention to the simplified mechanism $\langle t, x\rangle$.

8.3 Perfect bargaining mechanisms

The weakest requirements one would wish to impose on the bargaining mechanism $\langle \mathsf{T}, x\rangle$ in the direct revelation game are (1) individual rationality, that is, that everyone wishes to play the game, and (2) incentive compatibility, that is, that the mechanism induces truth telling. In addition, when the bargainers are unable to make binding commitments, one needs the further restriction of sequential rationality: It must never be common knowledge that the mechanism induced over time is dominated by an alternative mechanism. Bargaining schemes that satisfy incentive compatibility, individual rationality, and sequential rationality are called *perfect bargaining mechanisms.* The adjective *perfect* is adopted because of the close relationship between perfect bargaining mechanisms in the direct revelation game and perfect (or sequential) equilibria in an infinite horizon extensive-form game. It remains to be proven that a sequential bargaining mechanism is perfect if and only if it is a perfect equilibrium for some infinite-horizon extensive-form game. This issue will be addressed in future research.

In this section, I derive necessary and sufficient conditions for the sequential bargaining mechanism to be perfect. The incentive-compatibility and individual-rationality conditions were first established in Myerson and Satterthwaite (1983), and later extended to the case of multiple buyers and sellers by Wilson (1982) and Gresik and Satterthwaite (1983). It is important to realize that these properties are actually necessary and sufficient conditions for any Nash equilibrium of any bargaining game, since every Nash equilibrium induces a direct revelation mechanism, as mentioned in Section 8.2.

Incentive compatibility

In order to define and determine the implications of incentive compatibility on the sequential bargaining mechanism $\langle \mathsf{T}, x\rangle$, it is convenient to divide each player's expected payoff into two components as follows. Let

$$S(s) = \int_{\underline{b}}^{\bar{b}} x(s,b)g(b)db, \quad P(s) = \int_{\underline{b}}^{\bar{b}}\int_{0}^{\infty} e^{-rt} d\Im(t|s,b)g(b)db,$$

$$B(b) = \int_{\underline{s}}^{\bar{s}} x(s,b)f(s)ds, \quad Q(b) = \int_{\underline{s}}^{\bar{s}}\int_{0}^{\infty} e^{-st} d\Im(t|s,b)f(s)ds,$$

where $S(s)$ is the discounted expected revenue and $P(s)$ the discounted probability of agreement for seller s, and $B(b)$ *is* the discounted expected payment and $Q(b)$ the discounted probability of agreement for buyer b. Thus, the seller's and buyer's discounted expected payoffs are given by

$$U(s) = S(s) - sP(s) \qquad \text{and} \qquad V(b) = bQ(b) - B(b),$$

respectively.

Formally, the sequential bargaining mechanism $\langle \mathsf{T}, x\rangle$ is *incentive compatible* if every type of player wants to report truthfully his type; that is, for all s and s' in $[\underline{s},\bar{s}]$ and for all b and b' in $[\underline{b},\bar{b}]$,

$$U(s) \geq S(s') - sP(s') \qquad \text{and} \qquad V(b) \geq bQ(b')B(b').$$

Lemma 1. If the sequential bargaining mechanism $\langle \mathsf{T}, x\rangle$ is incentive compatible, then the seller's expected payoff U is convex and decreasing, with derivative $dU/ds = -P$ almost everywhere on $[\underline{s},\bar{s}]$; his discounted probability of agreement P is decreasing; and

$$U(s) - U(\bar{s}) = \int_{\underline{s}}^{\bar{s}} Pdu \qquad \text{and} \qquad S(s) - S(\bar{s}) = \int_{\underline{s}}^{\bar{s}} -udP(u). \tag{S}$$

Similarly, the buyer's expected payoff V is convex and increasing, with derivative $dV/db = Q$ almost everywhere on $[\underline{b},\bar{b}]$; his discounted probability of agreement Q is increasing; and

$$V(b) - V(\underline{b}) = \int_{\underline{b}}^{b} Q(u)du \qquad \text{and} \qquad B(b) - B(\underline{b}) = \int_{\underline{b}}^{b} udQ(u). \tag{B}$$

Proof. By definition, seller s achieves the payoff $U(s) = S(s) - sP(s)$. Alternatively, seller s can pretend to be seller s', in which case his payoff is $S(s') - sP(s')$. In the direct revelation game, the seller s must not want to pretend to be seller s', and so we have $U(s) \geq S(s') - sP(s')$ for all $s,s' \in [\underline{s},\bar{s}]$, or

$$U(s) \geq U(s') - (s - s')P(s'),$$

implying that U has a supporting hyperplane at s' with slope $-P(s') \leq 0$. Thus, U is convex and decreasing with derivative $(dU/ds)(s) = -P(s)$ almost everywhere, and P must be decreasing, which implies the first integral in (S) (I will use the Stieltjes integral throughout, so that any discontinuities in the probability of agreement are accounted for in the integral.) From integration by parts,

$$\int_{s}^{\bar{s}} P(u)du = \bar{s}P(\bar{s}) - sP(s) - \int_{s}^{\bar{s}} udP(u),$$

which, together with the definition of U, yields the second integral in (S). The proof for the buyer is identical.

Lemma 1 indicates the stringent requirements that incentive compatibility imposes on the players′ utilities. In particular, it suggests how one can construct an incentive-compatible payment schedule x, given a probability of agreement distribution Á for which the seller′s discounted probability of agreement $P(s)$ is decreasing in s and the buyer's discounted probability of agreement $Q(b)$ is increasing in b.

Lemma 2. Given the sequential bargaining mechanism $\langle$Á$, x\rangle$ such that P is decreasing, Q *is* increasing, and S and B satisfy (S) and (B) of lemma 1, then $\langle$Á$, x\rangle$ is incentive compatible.

Proof. A mechanism is incentive compatible for the seller if for all s, $s' \in [\underline{s}, \bar{s}]$,

$$S(s) - sP(s) \geq S(s') - sP(s').$$

Rearranging terms yields the following condition for incentive compatibility:

$$s(P(s') - P(s)) + S(s) - S(s') \geq 0. \qquad (S')$$

From (S), we have

$$S(s) = S(s') = \int_{s}^{s'} -udP(u),$$

and from the fundamental theorem of integral calculus,

$$s(P(s') - P(s)) = s\int_{s}^{s'} dP(u).$$

Adding the last two equations results in

$$s(P(s') - P(s)) + S(s) - S(s') = \int_{s}^{s'} (s-u)dP(u) \geq 0,$$

where the inequality follows because the integrand $(s - u)dP(u)$ *is* nonnegative for all s, $u \in [\underline{s}, \bar{s}]$, since P is decreasing. Hence, $\langle \mathsf{T}, x \rangle$ satisfies the incentive-compatibility condition (S'). An identical argument follows for the buyer.

Individual rationality

The sequential bargaining mechanism $\langle \mathsf{T}, x \rangle$ *is individually rational* if every type of player wants to play the game; that is, for all sin $[\underline{s}, \bar{s}]$ and b in $[\underline{b}, \bar{b}]$,

$$U(s) \geq 0 \quad \text{and} \quad V(b) \geq 0.$$

In light of the monotonicity of U and V proven in lemma 1, any incentive-compatible mechanism $\langle \mathsf{T}, x \rangle$ will satisfy individual rationality if the extreme high-cost seller and low-valuation buyer receive a nonnegative payoff; that is, an incentive-compatible mechanism $\langle \mathsf{T}, x \rangle$ is individually rational if and only if $U(\bar{s}) \geq 0$ and $V(\underline{b}) \geq 0$.

The following lemma describes how one can check whether or not a sequential bargaining mechanism is individually rational. It is convenient to state the lemma in terms of the simplified bargaining mechanism $\langle t, x \rangle$ rather than in terms of $\langle \mathsf{T}, x \rangle$. Recall that for the sequential bargaining mechanism $\langle t, x \rangle$, we have

$$S(s) = \int_{\underline{b}}^{\bar{b}} x(s,b)g(b)db, \quad P(s) = \int_{\underline{b}}^{\bar{b}} e^{-rt(s,b)}g(b)db,$$
$$B(b) = \int_{\underline{s}}^{\bar{s}} x(s,b)f(s)ds, \quad Q(b) = \int_{\underline{s}}^{\bar{s}} e^{-st(s,b)}df(s)ds.$$

Lemma 3. If the sequential bargaining mechanism $\langle t, x \rangle$ is incentive compatible and individually rational, then

$$U(\bar{s}) + V(\underline{b}) = \mathsf{E}\left\{\left(b - \frac{1-G(b)}{g(b)}\right)e^{-st(s,b)} - \left(s + \frac{F(s)}{f(s)}\right)e^{-rt(s,b)}\right\} \geq 0, \qquad (IR)$$

where the expectation is taken with respect to s and b.

Proof. First note that from lemma 1, for $\langle t, x \rangle$ to be individually rational, it must be that $U(\bar{s}) \geq 0$ and $V(\underline{b}) \geq 0$. For the seller, we have

$$\int_{\underline{s}}^{\bar{s}} U(s)f(s)ds = U(\bar{s}) + \int_{\underline{s}}^{\bar{s}}\int_{s}^{\bar{s}} P(u)du f(s)$$

$$= U(\bar{s}) + \int_{\underline{s}}^{\bar{s}} F(s)P(s)ds \tag{US}$$

$$= U(\bar{s}) + \int_{\underline{b}}^{\bar{b}}\int_{\underline{s}}^{\bar{s}} F(s)e^{-rt(s,b)}g(b)dsdb,$$

where the first equality follows from lemma 1 and the second equality results from changing the order of integration. Similarly, for the buyer we have

$$\int_{\underline{b}}^{\bar{b}} V(b)g(b)db = V(\underline{b}) + \int_{\underline{b}}^{\bar{b}}\int_{\underline{s}}^{\bar{s}}(1-G(b))e^{-st(s,b)}f(b)dsdb. \tag{UB}$$

Rearranging terms in (*US*) and (*UB*) and substituting the definitions for *U*(*s*) and *V*(*b*), result in the desired expression (*IR*) for $U(\bar{s}) + V(\underline{b})$.

Lemma 4. If the function $t(\cdot,\cdot)$ is such that *P* is decreasing, *Q* is increasing, and (*IR*) is satisfied, then there exists a function $x(\cdot,\cdot)$ such that $\langle t, x\rangle$ is incentive compatible and individually rational.

Proof. The proof is by construction. Let

$$x(s,b) = \int_{\underline{b}}^{b} udQ(u) + \int_{\underline{s}}^{s} udP(u) + c,$$

where *c* is a constant chosen such that $V(\underline{b}) \geq 0.$ To compute *c*, notice that

$$V(\underline{b}) = \underline{b}Q(\underline{b}) - \int_{\underline{s}}^{\bar{s}} x(s,\underline{b})f(s)ds$$

$$= \underline{b}Q(\underline{b}) - c - \int_{\underline{s}}^{\bar{s}}\int_{\underline{s}}^{s} udP(u)f(s)ds$$

$$= \underline{b}Q(\underline{b}) - c + \int_{\underline{s}}^{\bar{s}} s(1-F(s))dP(s) = 0.$$

Thus,

$$c = \underline{b}Q(\underline{b}) + \int_{\underline{s}}^{\bar{s}} s(1-F(s))dP(s).$$

Incentive compatibility for the seller is verified by showing that the seller *s* is better off reporting *s* than s′ ≠ *s*: For all *s*, *s*′ ∈ $[\underline{s},\bar{s}]$,

$$s(P(s') - P(s)) + S(s) - S(s') = s\int_{s}^{s'} dP(u) - \int_{s}^{s'} udP(u)$$

$$= \int_{s}^{s'} (s-u)dP(u) \geq 0,$$

since *P* is decreasing. An identical argument holds for the buyer.

Since $V(\underline{b}) = 0$ and $\langle t, x\rangle$ is incentive compatible and satisfies (*IR*), it follows from lemma 3 that $U(\bar{s}) \geq 0$. Thus, the bargaining mechanism $\langle t, s\rangle$ is incentive compatible and individually rational.

Sequential rationality

To understand how learning takes place in a sequential bargaining mechanism, it is best to interpret the direct revelation game as follows. At time zero (but after the players know their private information), the players agree to adopt a particular sequential bargaining mechanism $\langle t,x\rangle$ that is interim efficient. (Note that any interim-efficient mechanism can be chosen as a Nash equilibrium in an appropriately defined "choice-of-mechanism" game.) The players then report their private information in sealed envelopes to a mediator, who will then implement the mechanism $\langle t,x\rangle$. (Actually, a third party is not necessary, since the role of the mediator can be carried out by a computer programmed by the bargainers to execute the mechanism.) After opening the envelopes, the mediator does not announce the outcome immediately by saying something like, "Trade shall occur two months from now at the price of one thousand dollars," but instead waits until two months have passed and *then* announces, "Trade shall occur now at the price of one thousand dollars." The mediator must wait until the time of trade in order that the mechanism be sequentially rational, since otherwise the bargainers would have an incentive to ignore the mediator's announcement and trade immediately.

As time passes, the players are able to refine their inferences about the other player's private information based on the information that the mediator has not yet made an announcement about. Initially, it is common knowledge that the players' valuations are distributed according to the probability distributions F and G but after τ units of time the common-knowledge belief's become the distributions F and G conditioned on the fact that an announcement has not yet been made; that is,

$$F_t(s) = F(s|t(s,b) > t) \quad \text{and} \quad G_t(b) = G(b|t(s,b) > t).$$

Thus, at any time $t > 0$, the mechanism $\langle t,x\rangle$ induces an outcome function $t(s,b) = t(s,b \mid F_t, G_t)$ for all s and b. A mechanism $\langle t,x\rangle$ is *sequentially rational* if at every time $t \geq 0$, the induced outcome function $t(s,b \mid F_t, G_t)$ *is* interim efficient, that is, there does not exist a mechanism $\langle t', x'\rangle$ preferable to $\langle t,x\rangle$ at some time $t \geq 0$ for all remaining traders and strictly preferred by at least one trader.

The following lemma relates the definition of sequentially rational to common-knowledge dominance.

Lemma 5. A sequential bargaining mechanism $\langle t, x\rangle$ is sequentially rational if and only if it is never common knowledge that the mechanism $t(\cdot,\cdot \mid F_t, G_t)$ that it induces over time is dominated by an alternative mechanism.

Proof. From theorem 1 of Holmström and Myerson (1983), we know that a mechanism is interim efficient if and only if it is not common knowledge dominated by any other incentive-compatible and individually rational mechanism.

A necessary condition for a mechanism to be sequentially rational is that the bargainers continue negotiations so long as each expects positive gains from continuing. For the model here, since there are no transaction costs (only delay costs), this means that negotiations cannot end if there exists a pair of players that have not yet come to an agreement, but *for* which agreement is beneficial at some point in the future. Formally, for the bargaining mechanism $\langle t,x\rangle$ to be sequentially rational, it must be that for all potential players, a failure to reach agreement implies that there is some point beyond which agreement is never beneficial; that is, for all s and b,

$$t(s,b) = \infty \Rightarrow \text{there exists } \hat{t} \geq 0 \text{ such that for every } t > \hat{t}, s \geq be^{(r-s)t}.$$

The condition $s \geq be^{(r-s)t}$ is simply a statement that trade is not beneficial at time t, since

$$x - se^{-rt} + be^{-st} - x \geq 0 \Leftrightarrow se^{-rt} \geq be^{-st} \Leftrightarrow s \geq be^{(r-s)t}.$$

Notice that the strength of this requirement depends on the relative magnitudes of the players' discount rates. When $r > s$, then $e^{(r-s)t} \to \infty$ as $t \to \infty$, and so for all potential pairs of players it is always the case that there exists a time at which trade is beneficial. Thus, when $r > s$, the mechanism $\langle t,x\rangle$ *is* sequentially rational only if trade always occurs; that is, $t(s,b) < \infty$ for all s and b. Likewise, when $r < s$, then $e^{(r-s)t} \to 0$ as $t \to \infty$, and so for every pair of players there is always a point at which trade becomes undesirable for all times in the future. Finally, if $r = s$, then the necessary condition for sequential rationality becomes $t(s,b) = \infty \Rightarrow s \geq b$; that is, trade must occur whenever the gains from trade are initially positive.

To state this necessary condition in a lemma, it will be useful to define B as the set of potential traders for which trade is always beneficial at some time in the future; that is,

$$B = \{(s,b) \mid r > s \text{ or } (r = s \text{ and } s \leq b)\}.$$

Lemma 6. Any mechanism $\langle t,x \rangle$ that excludes trade over a nonempty subset of B violates sequential rationality.

Proof. Let $\mathsf{N} \subset \mathsf{B}$ be the set for which trade never occurs. Then, at some point in time t, the induced mechanism has $t(s,b \mid F_t, G_t) = \infty$ for all remaining traders, which includes N. However, this mechanism is not interim efficient, since it is dominated by a mechanism that results in a positive probability of trade for some traders in N (a partially pooling equilibrium with this property will always exist).

I claim that sequential rationality is a necessary condition for rationality in games with incomplete information in which commitment is not possible. If a mechanism is not sequentially rational, then at some point in time it is common knowledge that *all* potential agents would prefer an alternative mechanism and hence this alternative mechanism will be adopted by the agents at that point in time. Thus, it would be inconsistent for the players to believe that the original mechanism would be carried out faithfully.

Necessary, and sufficient conditions for perfection

Lemmas I through 5 are summarized in the following theorem, which gives necessary and sufficient conditions for the sequential bargaining mechanism $\langle t,x \rangle$ to be perfect.

Theorem 1. A sequential bargaining mechanism $\langle t,x \rangle$ is incentive compatible if and only if the functions

$$S(s) = \int_{\underline{b}}^{\bar{b}} x(s,b)g(b)db, \quad P(s) = \int_{\underline{b}}^{\bar{b}} e^{-rt(s,b)}g(b)db,$$

$$B(b) = \int_{\underline{s}}^{\bar{s}} x(s,b)f(s)ds, \quad Q(b) = \int_{\underline{s}}^{\bar{s}} e^{-st(s,b)}df(s)ds$$

are such that P is decreasing, Q is increasing, and

$$S(s) - S(s') = \int_{s}^{\bar{s}} -udP(u) \quad \text{and} \quad B(b) - B(\underline{b}) = \int_{\underline{b}}^{b} -udQ(u). \tag{IC}$$

Furthermore, for t such that P is decreasing and Q is increasing, there exists an x such that $\langle t,x \rangle$ is incentive compatible and individually rational if and only if

$$U(\bar{s}) + V(\underline{b}) = \mathsf{E}\left\{\left(b - \frac{1-G(b)}{g(b)}\right)e^{-st(s,b)} - \left(s + \frac{F(s)}{f(s)}\right)e^{-rt(s,b)}\right\} \geq 0. \tag{IR}$$

Finally, the mechanism $\langle t,x \rangle$ is sequentially rational if and only if it is never common knowledge that the mechanism it induces over time is dominated by an alternative mechanism.

8.4 Efficiency

The set of perfect bargaining mechanisms is typically quite large, which means that there are many extensive-form games with equilibria satisfying incentive compatibility, individual rationality, and sequential rationality. To narrow down this set, it is natural to assume additional efficiency properties. Three notions of efficiency, described at length by Holmström and Myerson (1983), are ex post, interim, and ex ante efficiency. The difference between these concepts centers on what information is available at the time of evaluation: Ex ante efficiency assumes that comparisons are made before the players know their private information, interim efficiency assumes that the players know only their private information, and ex post efficiency assumes that all information is known.

Ex post efficiency

Ideally, one would like to find perfect bargaining mechanisms that are ex post efficient. The mechanism $\langle t,x \rangle$ is *ex post efficient* if there does not exist an alternative mechanism that can make both players better off in terms of their ex post utilities (after all of the information is revealed). (This is often called *full-information efficiency* in the literature. Holmström and Myerson (1983) term this "ex post classical efficiency" to distinguish it from their concept of ex post incentive-efficiency, in which incentive constraints are recognized.) Equivalently, for a mechanism to be ex post efficient, it must maximize a weighted sum $a_1(s, b)u(s) + a_2(s, b)v(b)$ of the players' ex post utilities for all s and b, where $a_1(\cdot,\cdot)$, $a_2(\cdot,\cdot) \geq 0$ and the ex post utilities of seller s and buyer b are

$$u(s,b) = x(s,b) - se^{-rt(s,b)} \quad \text{and} \quad v(s,b) = be^{-st(s,b)} - x(s,b).$$

Since the payoff functions are additively separable in money and goods, and thus utility is transferable between players, we can assume equal weights (i.e., $a_1(s, b) = a_2(s, b) = 1$ for every s,b) without loss of generality. To simplify notation, define $p(s,b)= e^{-t(s,b)}$, so that $p(s,b)^r$ is the discounted probability of agreement for seller s given that the buyer has valuation b, and $p(s,b)^s$ *is* the discounted probability of agreement for buyer b given that the seller has cost s. With this change, a sequential bargaining mechanism becomes the pair of functions $\langle p,x\rangle$ where p: $[\underline{s},\bar{s}]$ X $[\underline{b},\bar{b}] \rightarrow [0,1]$. The bargaining mechanism $\langle p,x\rangle$, then, is ex post efficient if for all $s \in [\underline{s},\bar{s}]$ and $b \in [\underline{b},\bar{b}]$, the function $p(s,b)$ is chosen to solve the program

$$\max_{p\in[0,1]} p(p) = bp^s - sp^r.$$

The first-order condition is

$$\frac{dp}{dp} = sbp^{s-1} - rsp^{r-1} = 0$$

or

$$p = \left(\frac{sb}{rs}\right)^{1/(r-s)}.$$

Checking the boundary conditions and assuming that $\underline{s}$, $\underline{b} \geq 0$ yields

$$p^*(s,b) = \begin{cases} 1 & \text{if } s < b, rs \leq sb, \\ \left(\dfrac{sb}{rs}\right)^{1/(r-s)} & \text{if } r > s, rs \geq sb, \\ 0 & \text{if } r \leq s, s \geq b. \end{cases} \qquad (EP)$$

The following theorem demonstrates that it is impossible to find ex post-efficient mechanisms if the bargainers are uncertain whether or not trade should occur immediately. This result is shown in an example in Cramton (1984).

Theorem 2. There exists an incentive-compatible, individually rational bargaining mechanism that is ex post efficient if it is common knowledge that trade should occur immediately. However, an ex post-efficient mechanism does not exist if the buyer's delay cost is at least as great as the seller's and it is not common knowledge that gains from trade exist.

Proof. Suppose that it is common knowledge that trade should occur immediately. Then, three cases are possible: (1) $r \leq s$ and $\bar{s} \leq \underline{b}$, (2) $r > s$ and $r\bar{s} \leq s\underline{b}$, and (3) $r = \infty$ and $s < \infty$. What needs to be shown is that $p^*(s,b) = 1$ for all s,b satisfies (IR). For cases (1) and (2),

$$\begin{aligned} U(\bar{s}) + V(\underline{b}) &= \mathrm{E}\left\{\left(b - \frac{1-G(b)}{g(b)}\right) - \left(s + \frac{F(s)}{f(s)}\right)\right\} \\ &= \mathrm{E}\left\{b - \frac{1-G(b)}{g(b)}\right\} - \mathrm{E}\left\{s + \frac{F(s)}{f(s)}\right\} \\ &= \int_{\underline{b}}^{\bar{b}} (bg(b) - 1 + G(b))db - \int_{\underline{s}}^{\bar{s}} (sf(s) + F(s))ds \\ &= bG(b)\Big|_{\underline{b}}^{\bar{b}} - \bar{b} + \underline{b} - sF(s)\Big|_{\underline{s}}^{\bar{s}} \\ &= \underline{b} - \bar{s} \geq 0, \end{aligned}$$

where the integration is done by parts. In case (3),

$$U(\bar{s}) + V(\underline{b}) = \mathrm{E}\left\{b - \frac{1-G(b)}{g(b)}\right\} = \underline{b} \geq 0.$$

Then, by lemma 4, there exists an x such that $\langle p,x\rangle$ is incentive compatible and individually rational.

Now, assume that it is not common knowledge that gains from trade exist and the buyer’s delay cost is at least as great as the seller’s (i.e. $r \le s$). Notice that when $r \le s$, we find that $\langle p,x \rangle$ is ex post efficient if trade occurs without delay whenever there are positive gains from trade:

$$p^*(s,b) = \begin{cases} 1 \text{ if } s < b, \\ 0 \text{ if } s \ge b. \end{cases}$$

Substituting this function for p into (IR) yields

$$\begin{aligned} U(\bar{s}) + V(\underline{b}) &= \int_{\underline{b}}^{\bar{b}} \int_{\underline{s}}^{\min(b,\bar{s})} (bg(b) + G(b) - 1) f(s) ds db - \int_{\underline{b}}^{\bar{b}} \int_{\underline{s}}^{\min(b,\bar{s})} (sf(s) + F(s)) ds g(b) db \\ &= \int_{\underline{b}}^{\bar{b}} (bg(b) + G(b) - 1) f(b) db - \int_{\underline{b}}^{\bar{b}} \min\{bF(b), \bar{s}\} g(b) db \\ &= -\int_{\underline{b}}^{\bar{b}} (1 - G(b)) F(b) db + \int_{\underline{s}}^{\bar{b}} (b - \bar{s}) g(b) db \\ &= -\int_{\underline{b}}^{\bar{b}} (1 - G(b)) F(b) db + \int_{\bar{s}}^{\bar{b}} (1 - G(b)) db \\ &= -\int_{\underline{b}}^{\bar{s}} (1 - G(u)) F(u) du. \end{aligned}$$

Thus any incentive-compatible mechanism that is ex post efficient must have

$$U(\bar{s}) + V(\underline{b}) = -\int_{\underline{b}}^{\bar{b}} (1 - G(u)) F(u) du < 0,$$

and so it cannot be individually rational.

When the seller’s delay cost is greater than the buyer’s and it is not common knowledge that trade should occur immediately, a general proof that ex post efficiency is not achievable cannot be given due to the complicated expression for $p^*(s,b)$ in this case. However, analysis of examples (see Section 8.5) suggests that ex post efficiency is typically unobtainable.

Ex ante efficiency

The strongest concept of efficiency, other than ex post efficiency (which is generally unobtainable), that can be applied to games of incomplete information is ex ante efficiency. A player’s ex ante utility is his expected utility *before* he knows his type. Thus, given the sequential bargaining mechanism $\langle p,x \rangle$, the seller’s and buyer’s ex ante utilities are

$$\begin{aligned} \mathsf{U} &= \int_{\underline{s}}^{\bar{s}} U(s) f(s) ds = \int_{\underline{s}}^{\bar{s}} \int_{\underline{b}}^{\bar{b}} (x(s,b) - sp(s,b)^r) g(b) db f(s) ds, \\ \mathsf{V} &= \int_{\underline{b}}^{\bar{b}} V(b) g(b) db = \int_{\underline{b}}^{\bar{b}} \int_{\underline{s}}^{\bar{s}} (bp(s,b)^s - x(s,b)) f(s) ds g(b) db. \end{aligned}$$

The mechanism $\langle p,x \rangle$ is *ex ante efficient* if there does not exist an alternative mechanism that can make both players better off in terms of their ex ante utilities. Thus, for a mechanism to be ex ante efficient, it must maximize a weighted sum $a_1\mathsf{U} + a_2\mathsf{V}$ of the players’ ex ante utilities, where $a_1, \ a_2 \ge 0$. For tractability and reasons of equity, I will assume equal weights (i.e., $a_1 = a_2 = 1$). (One might think that the assumption of equal weights is made without loss of generality, because the payoff functions here are additively separable in money and goods, and thus utility is transferable between players. Although this intuition is correct in a setting of complete information, it is false when there is incomplete information, because an ex ante transfer of utility will violate individual rationality for some players.) The use of unequal weights would not significantly change the results but would greatly complicate the analysis.

If the bargainers were to choose a bargaining mechanism *before* they knew their types, it would seem reasonable that they would agree to a scheme that was ex ante efficient. It is generally the case, however, that the players know their private information before they begin negotiations, and therefore would be unable to agree on an ex ante-efficient mechanism, since the players are concerned with their *interim* utilities $U(s)$ and $V(b)$ rather than their ex ante utilities U and V. Nevertheless, it may be that the sequential bargaining mechanism is chosen by an uninformed social planner or arbitrator, in which case the selection of an ex ante-efficient mechanism would be justified. Alternatively, one might suppose that the choice of a bargaining mechanism is based on established norms of behavior and that these norms have evolved over time in such a way as to produce ex ante-efficient mechanisms. In situations where the choice of a bargaining mechanism does not occur before the players know their types or is not handled by an uninformed third party, ex ante efficiency is too strong a requirement. The weaker requirement of *interim efficiency*-that there does not exist a dominating mechanism in terms of the players' interim utilities $U(s)$ and $V(b)$-is more appropriate.

The sum of the players' ex ante utilities for the bargaining mechanism $\langle p,x\rangle$ is given by

$$\mathsf{U} + \mathsf{V} = \int_{\underline{b}}^{\bar{b}}\int_{\underline{s}}^{\bar{s}} (bp(s,b)^s - sp(s,b)^r)\, f(s)g(b)\,ds\,db.$$

A bargaining mechanism, then, is ex ante efficient if it maximizes this sum subject to incentive compatibility and individual rationality:

$$\max_{p(\cdot,\cdot)} \mathsf{E}\{bp(s,b)^s - sp(s,b)^r\}$$

such that

$$\mathsf{E}\left\{\left(b - \frac{1-G(b)}{g(b)}\right)p(s,b)^s - \left(s + \frac{F(s)}{f(s)}\right)p(s,b)^r\right\} \geq 0, \qquad (P)$$

where p is chosen such that P is decreasing and Q is increasing. Multiplying the constraint by $l \geq 0$ and adding it to the objective function yields the Lagrangian

$$\mathsf{L}\ (p,l) = \mathsf{E}\left\{\left((1+l)b - l\,\frac{1-G(b)}{g(b)}\right)p(s,b)^s - \left((1+l)s + l\,\frac{F(s)}{f(s)}\right)p(s,b)^r\right\}$$

$$= (1+l)\mathsf{E}\left\{\left[b - \left(\frac{l}{1+l}\right)\left(\frac{1-G(b)}{g(b)}\right)\right]p(s,b)^s - \left[s + \left(\frac{l}{1+l}\right)\left(\frac{F(s)}{f(s)}\right)\right]p(s,b)^r\right\}.$$

For any $a \geq 0$, define the functions

$$c(s,a) = s + a\,\frac{F(s)}{f(s)} \qquad \text{and} \qquad d(b,a) = b - a\,\frac{1-G(b)}{g(b)}.$$

Then, the Lagrangian (ignoring the constant $(1 + l)$) becomes

$$\mathsf{L}\ (p,l) = \mathsf{E}\{d(b,a)p(s,b)^s - s(s,a)p(s,b)^r\},$$

which is easily maximized by pointwise optimization. The first-order condition is

$$\frac{d\mathsf{L}}{dp} = s\,dp^{s-1} - r\,cp^{r-1}$$

or

$$p = \left(\frac{sd}{rc}\right)^{1/(r-s)}.$$

Establishing the boundary conditions and noticing that $c(\cdot,\cdot) \geq 0$ yields the optimal solution

$$p_a(s,b) = \begin{cases} 1 & \text{if } c(s,a) < d(b,a), rc(s,a) \leq sd(b,a), \\ \left(\dfrac{sd(b,a)}{rc(s,a)}\right)^{1/(r-s)} & \text{if } r > s, rc(s,a) > sd(b,a) > 0, \\ 0 & \text{if } (r \leq s, c(s,a) \geq d(b,a)) \text{ or } d(b,a) \leq 0. \end{cases}$$

The following theorem determines how to find an ex ante-efficient mechanism for any sequential bargaining game.

Theorem 3. If there exists an incentive-compatible mechanism $\langle p,x \rangle$ such that $p = p_a$ for some a in [0,1] and $U(\bar{s}) = V(\underline{b}) = 0$, then this mechanism is ex ante efficient. Moreover, if $c(\cdot,1)$ and $d(\cdot,1)$ are increasing functions on $[\underline{s},\bar{s}]$ and $[\underline{b},\bar{b}]$, respectively, and ex post efficiency is unobtainable, then such a mechanism must exist.

Proof. The first statement in this theorem follows from the fact that the Lagrangian $\mathsf{L}(p,l)$ is maximized by the function p_a with $a = l/(1+l)$. Hence, p_a yields an ex ante-efficient mechanism provided that the individual-rationality constraint is binding.

To prove the existence part of the theorem, suppose that $c(\cdot, 1)$ and $d(\cdot, 1)$ are increasing, and that the players are uncertain whether or not trade should occur immediately. Then, for every $a \in [0,1]$, $c(\cdot, a)$ and $d(\cdot, a)$ are increasing, which implies that $p_a(s,b)$ is increasing in s and decreasing in b. Thus, P is decreasing and Q is increasing, as required by incentive compatibility.

It remains to be shown that there is a unique $a \in [0,1]$, for which the individual-rationality constraint is binding. Define

$$R(a) = \mathsf{E}\{d(b,1)(p_a(s,b))^s - c(s,1)(p_a(s,b))^r\}$$

so that $R(a)$ is the value of the integral in the individual-rationality constraint as a function of a. First, notice that $R(1) \geq 0$, since the term in the expectation is nonnegative for all s and b. Furthermore, $R(0) < 0$, since there does not exist an ex post-efficient mechanism. Therefore, if $R(a)$ is continuous and strictly increasing in a, then there is a unique $a \in [0,1]$ for which $R(a) = 0$.

The continuity and monotonicity of $R(\cdot)$ are most easily verified by considering two cases.

CASE 1 ($r \leq s$). When $r \leq s$, then

$$p_a(s,b) = \begin{cases} 1 & \text{if } c(s,a) < d(b,a), \\ 0 & \text{if } c(s,a) \geq d(b,a). \end{cases}$$

Thus, $p_a(s,b)$ is decreasing in a, since

$$d(b,a) - c(s,a) = (b-s) - a\left(\frac{1-G(b)}{g(b)} + \frac{F(s)}{f(s)}\right)$$

is decreasing in a. Thus, for or $a < b$. $R(b)$ differs from $R(a)$ only because $0 = p_b(s,b) < p_a(s,b) = 1$ for some (s,b) where $d(b,b) < c(s,b)$, and so $d(b,1) < c(s,1)$. Therefore, $R(\cdot)$ is strictly increasing.

To prove that $R(\cdot)$ is continuous, observe that if $c(s,1)$ and $d(b,1)$ are increasing in s and b, then $c(\cdot,a)$ and $d(\cdot,a)$ are strictly increasing for any $a < 1$. So, given b and a, the equation $c(s,a) = d(b,a)$ has at most one solution in s, and this solution varies continuously in b and a. Hence, we may write

$$R(a) = \int_{\underline{b}}^{\bar{b}} \int_{\underline{s}}^{r(b,a)} (d(b,1) - c(s,1)) f(s) g(b) ds db,$$

where $r(b,a)$ is continuous in b and a. Thus, $R(a)$ is continuous in a.

CASE 2 ($r > s$). When $r > s$, then

$$p_a(s,b) = \begin{cases} 1 & \text{if } sc(s,a) \le sd(b,a), \\ \left(\dfrac{sd(b,a)}{rc(s,a)}\right)^{1/(r-s)} & \text{if } rc(s,a) > sd(b,a) > 0, \\ 0 & \text{if } d(b,a) \le 0. \end{cases}$$

Since

$$sd(b,a) - rc(s,a) = sb - rs - a\left(s\,\frac{1-G(b)}{g(b)} + r\,\frac{F(s)}{f(s)}\right)$$

$$\frac{sd(b,a)}{rc(s,a)} = \frac{s}{r}\left(\frac{b - a\{(1-G(b))/g(b)]}{S + a[F(s)/f(s)]}\right),$$

and $d(b,a)$ are decreasing in a, $p_a(s,b)$ is decreasing in a. Thus, for $a < b$, $R(b)$ differs from $R(a)$ only because $p_b(s,b) < p_a(s,b)$ for some (s,b) where $sd(b,a) < rc(s,a)$. Therefore, $R(\cdot)$ is strictly increasing.

Since $c(\cdot,a)$ and $d(\cdot,a)$ are strictly increasing for any $a < 1$, the equation $d(b,a) = 0$ has at most one solution in b and the equation $rc(s,a) = sd(b,a)$ has at most one solution in s, and the solutions vary continuously in b and a. Hence, we may write

$$R(a) = \int_{q(a)}^{\bar{b}} \Bigg(\int_{\underline{s}}^{r(b,a)} (d(b,1) - c(b,1)) f(s) ds + \int_{r(b,a)}^{\bar{s}} [d(b,1)(p_a(s,b))^s - c(s,1)(p_a(s,b))^s] f(s) ds \Bigg) g(b) db,$$

where $q(a)$ and $r(b, a)$ are continuous in b and a. Therefore, $R(a)$ is continuous in a.

Since $R(\cdot)$ is continuous and strictly increasing, with $R(0) < 0$ and $R(\mathrm{l}) \geq 0$, there must be a unique $a \in [0,\mathrm{l}]$ such that $R(a) = 0$ and $p_a(s,b)$ is ex ante efficient.

It is worthwhile to point out that the requirement in the existence part of theorem 3 that $c(\cdot,1)$ and $d(\cdot,1)$ be increasing functions is satisfied by a large range of distribution functions. A sufficient condition for $c(\cdot,1)$ and $d(\cdot,1)$ to be increasing is for the ratio of the distribution and the density to be increasing. This is a local characterization of the monotone likelihood ratio property and is satisfied by many distributions, such as the uniform, exponential, normal, chi-square, and Poisson distributions.

I now prove that the ex ante-efficient mechanism typically violates sequential rationality, and hence show that bargainers who are unable to make binding commitments are worse off (in an ex ante sense) than bargainers who are able to commit to particular strategies.

Corollary 1. If ex post efficiency is unobtainable, $c(\cdot,1)$ and $d(\cdot,1)$ are increasing functions, and $d(\underline{b},1) < 0$ if $r > s$, then the ex ante-efficient mechanism violates sequential rationality.

Proof. By theorem 3, the ex ante-efficient mechanism exists and is given by p_a for some $a > 0$. Consider the set of traders who never trade under p_a, but for whom trade is always beneficial at some point in the future:

$$\mathsf{N} = \{(s,b) \mid p_a(s,b) = 0 \text{ and } [r > s \text{ or } (r = s \text{ and } s \le b]\}.$$

By our hypothesis, this set is nonempty. Thus, from lemma 6, the mechanism p_a violates sequential rationality.

8.5 The case of uniform symmetric exchange: An example

To illustrate the theory presented in the earlier sections, it will be useful to look at an example. In particular, consider the case of uniform symmetric exchange in which both the seller's cost and the buyer's valuation are uniformly distributed on [0,1]. Then, $c(s,a) = (1 + a)s$ and $d(b, a) = (1 + a)b - a$, which are strictly increasing when $a = 1$, and so by theorem 3 we know that, for some $a \in [0,1]$, the mechanism $p = p_a$ is ex ante efficient. The desired a is found by setting $R(a)$ to zero, so that $U(\bar{s}) = V(\underline{b}) = 0$. Again, it will be useful to consider two cases depending on whether $r \le s$ or $r > s$.

CASE 1 ($r \le s$). When $r \le s$, then

$$p_a(s,b) = \begin{cases} 1 \text{ if } s < b - \dfrac{a}{1+a}, \\ 0 \text{ if } s \geq b - \dfrac{a}{1+a}. \end{cases}$$

Define $m = a/(1 + a)$. Then, we wish to find $m \in [0,\frac{1}{2}]$ such that

$$R(a) = \int_m^1 \int_0^{b-m} (2(b-s)-1)\,ds\,db = 0$$

Performing the integration yields

$$\left(m - \frac{1}{4}\right)(m+1)^2 = 0,$$

which has a root in $[0,\frac{1}{2}]$ at $m = \frac{1}{4}$. Thus, $a = \frac{1}{3}$ and

$$p_a(s,b) = \begin{cases} 1 \text{ if } s < b - \dfrac{1}{4}, \\ 0 \text{ if } s \geq b - \dfrac{1}{4}. \end{cases}$$

When $r \leq s$, ex ante efficiency is obtained by a mechanism that transfers the object without delay if and only if the buyer's valuation exceeds the seller's by at least $\frac{1}{4}$. Perhaps somewhat surprisingly, the ex ante-efficient mechanism in this case does not depend on r or a. Since the value of the object is declining more rapidly for the buyer than for the seller, it is always better to transfer the item immediately, if at all. Hence, even though the players can reveal information by delaying agreement, in the ex ante-efficient mechanism they choose to trade immediately or not at all, so that a static mechanism ex ante dominates any sequential bargaining mechanism. This static mechanism, however, is not sequentially rational, which illustrates corollary 1.

An extensive-form game that implements the ex ante-efficient mechanism when $r \leq s$ has been studied by Chatterjee and Samuelson (1983). They consider the simultaneous-offers game, in which the players simultaneously announce prices and the object is traded if the buyer's bid exceeds the seller's offer. For this example, the seller's optimal strategy is to offer the price $\frac{2}{3}s + \frac{1}{4}$, and the buyer's best response is to bid $\frac{2}{3}b + \frac{1}{12}$, which implies that trade occurs provided that $\frac{2}{3}s + \frac{1}{4} < \frac{2}{3}b + \frac{1}{12}$ or $s < b - \frac{1}{4}$, as in the ex ante-efficient mechanism. For this equilibrium, the price at which the object is sold is

$$x(s,b) = \begin{cases} \dfrac{1}{3}(b+s) + \dfrac{1}{6} & \text{if } s < b - \dfrac{1}{4}, \\ 0 & \text{if } s \geq b - \dfrac{1}{4}. \end{cases}$$

The sum of the players' ex ante utilities is

$$U + V = \int_{1/4}^1 \int_0^{b-1/4} (b-s)\,ds\,db = \frac{9}{64},$$

whereas the total utility from the ex post-efficient mechanism is

$$\int_0^1 \int_0^b (b-s)\,ds\,db = \frac{1}{6},$$

Thus, 15.6 percent of the gains from trade are lost when $r \leq s$, due to delays in agreement.

CASE 2 ($r > s$). When $r > s$, then

$$p_a(s,b) = \begin{cases} 1 & \text{if } s \le \dfrac{s}{r}\left(b - \dfrac{a}{1+a}\right), \\ \left(\dfrac{s\left(b - \dfrac{a}{1+a}\right)}{rs}\right)^{1/(r-a)} & \text{if } s > \dfrac{s}{r}\left(b - \dfrac{a}{1+a}\right), \\ 0 & \text{if } b \le \dfrac{a}{1+a}. \end{cases}$$

Making the substitution, $m = a/(1+a)$, we wish to find $m \in [0,\frac{1}{2}]$ such that

$$\int_m^1 \left[\int_0^{(s/r)(b-m)} (2b-1-2s)ds + \int_{(s/r)(b-m)}^1 \left[\left[(2b-1)\left(\frac{s\,(b-m)}{rs}\right)^{s/(r-s)} - 2s\left(\frac{s\,(b-m)}{rs}\right)^{s/(r-s)} \right] ds \right] db = 0.$$

Let $d = s/r$ and $g = s/(r-s)$, so that $1 + g = r/(r-s)$. After this substitution, we have

$$\int_m^1 \int_0^{d(b-m)} (2b-1-2s)ds + \int_{d(b-m)}^1 [(2b-1)[d(b-m)]^g - 2[d(b-m)]^{1+g}]s^{-g}\,ds]db = 0.$$

Performing the inner integration (assuming $g \neq 1$) yields

$$\int_m^1 \Big[d(b-m)[2-d)b + dm - 1]$$

$$+ \frac{1}{1-g}[2(1-d)b + 2dm - 1][d(b-m)]^g \{1 - [d(b-m)]^{1-g}\} \Big] db$$

$$= \int_m^1 \Big[d\{(2-d)b^2 - (2m(1-d)+1]b + m(1-dm)\}$$

$$+ \frac{1}{1-g}[2(1-d)b + 2dm - 1]\{[d(b-m)]^g\, d(b-m)\} \Big] db$$

$$= \int_m^1 \frac{d}{1-g}\Big[[d - g(2-d)]b^2 - \{g[1+2m(1-d)] - 2dm\}b$$

$$+ dm^2 - gm(1-dm) + [2(1-d)b + 2dm - 1]d^{g-1}(b-m)^g \Big] db = 0.$$

Since

$$\int_m^1 (b-m)^g\, db = \frac{(1-m)^{1+g}}{1+g} \qquad \text{and}$$

$$\int_m^1 b(b-m)^g\, db = \frac{(1-m)^{1+g}}{1+g}\left(1 - \frac{1-m}{2+g}\right),$$

after integration we have

$$\frac{d}{1-g}\Big[\frac{1}{3}(1-m^3)[d - g(2-d)] + \frac{1}{2}(1-m^2)\{g[1+2m(1-d)] - 2dm\}$$

$$+ (1-m)[dm^2 - gm(1-dm)] + \frac{d^{g-1}(1-m)^{1+g}}{1+g}$$

$$\left[2dm - 1 + 2(1-d)\left(1 - \frac{1-m}{2+g}\right) \right] \Big] = 0.$$

Dividing by $d(1-m)/(1-g)$, yields

$$\frac{1}{3}(1+m+m^2)[d-g(2-d)]+\frac{1}{2}(1+m)\{g[1+2m(1-d)]-2dm\}$$
$$+dm^2-gm(1-dm)+\frac{d^{g-1}(1-m)^g}{1+g}\left[2dm-1+2(1-d)\left(1-\frac{1-m}{2+g}\right)\right]\Big]=0. \quad (R)$$

Given $d = s/r$, a root $m \in [0,\frac{1}{2}]$ to (R) is easily found numerically.

The sum of the players' ex ante utilities is computed as follows:

$$U+V=\int_m^1\left[\int_0^{d(b-m)}(b-s)ds+\int_{d(b-m)}^0\left[b\left(\frac{d(b-m}{s}\right)^g-s\left(\frac{d(b-m}{s}\right)^{1+g}\right]ds\right]db$$
$$=\int_m^1\left[d(b-s)\left(\left(1-\frac{1}{2}d\right)b+\frac{1}{2}dm\right)+\frac{1}{1-g}[(1-d)b+dm][d(b-m)]^g\{1-[d(b-m)]^{1-g}\}\right]db$$
$$=\int_m^1\frac{d}{1-g}\left[\left(\frac{1}{2}d(1+g)-g\right)b^2+[mg-dm(1+g)]b+\frac{1}{2}dm^2(1+g)+d^{g-1}[(1-d)b+dm](b-m)^g\right]db$$
$$=\frac{d(1-m)}{1-g}\left[\frac{1}{3}\left(1+m+m^2\right)\left(\frac{1}{2}d(1+g)-g\right)+\frac{1}{2}(1+m)[mg-dm(1+g)]\right.$$
$$\left.+\frac{1}{2}dm^2(1+g)+\frac{d^{g-1}(1-m)^g}{1+g}\left(dm+(1-d)\frac{1+m+g}{2+g}\right)\right].$$

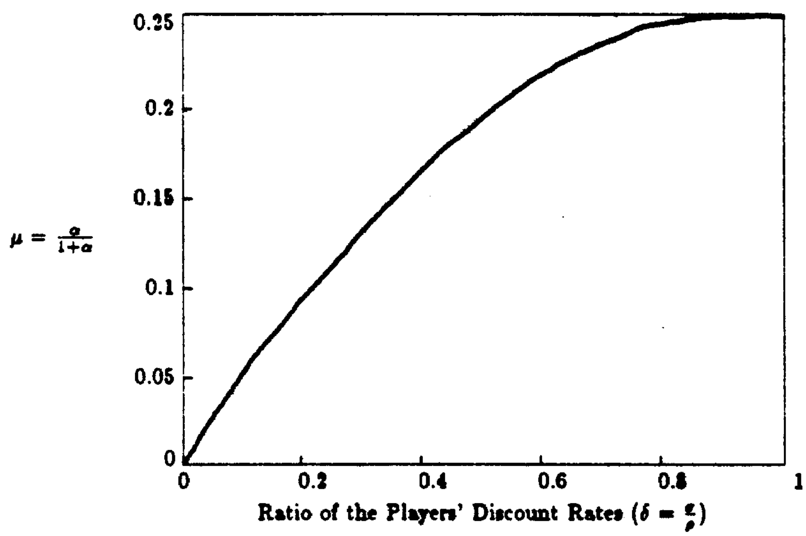

Figure 8.1 Value of μ as a function of the ratio of the players' discount rates.

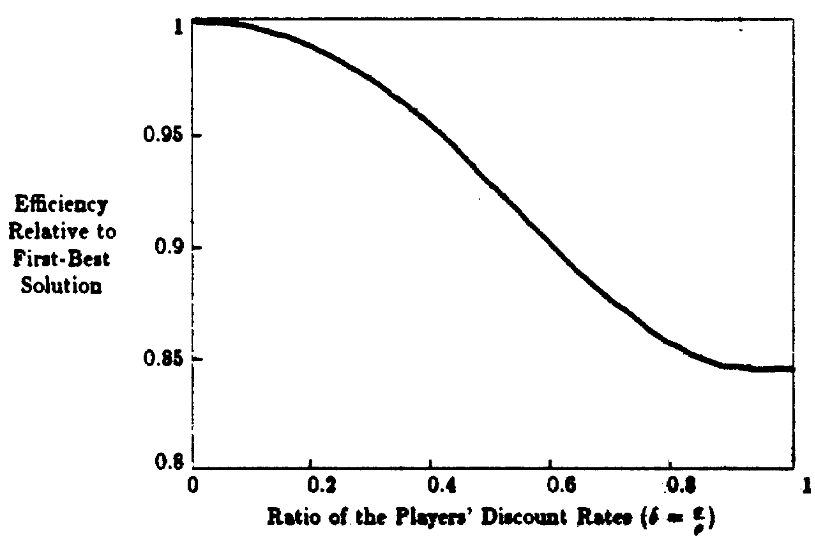

Figure 8.2 Efficiency as a function of the ratio of the players' discount rates.

The value of m and the efficiency of the ex ante-efficient mechanism relative to the first-best (full-information) solution are shown in Figure 8.1 and Figure 8.2, respectively, as the ratio of the players' discount rates is varied from 0 to 1. Bargaining efficiency improves as the seller's discount rate is increased relative to the buyer's. When the players' discount rates are equal, 15.6 percent of the gains from trade are lost due to delays in agreement. This inefficiency decreases to zero as $r \to \infty$, illustrating theorem 2.

8.6 Conclusion

Two important features of any bargaining setting are information and time. Bargainers typically have incomplete information about each other's preferences, and therefore must communicate some of their private information in order to determine whether or not gains from trade exist. One means of communication is for the agents to signal their private information through their willingness to delay agreement: Bargainers who anticipate large gains from trade will be unwilling to delay agreement and so will propose attractive terms of trade that the other is likely to accept early in the bargaining process, whereas bargainers expecting small gains will prefer to wait for better offers from their opponent. In this chapter, I have described the properties of such a bargaining model, by analyzing a sequential direct revelation game.

Modeling the bargaining process as a sequential game, where the agents communicate their private information over time, has two main advantages. First, from the point of view of realism, one commonly observes bargaining taking place over time. Second, any static bargaining mechanism, because it does not permit the agents to learn about their opponent's preferences, must end with positive probability in a situation where gains from trade are possible and yet no agreement is reached. If both bargainers know that gains from trade exist, what prevents them from continuing negotiations until an agreement is reached? By introducing the time dimension, and hence allowing the bargainers to communicate through their actions over time, one is able to construct perfect bargaining mechanisms, in which the bargainers continue to negotiate so long as they expect positive gains from continuing.

When the bargainers discount future gains according to known and fixed discount rates, it was found that the bargainers may be better off (in terms of their ex ante utilities) using a sequential bargaining mechanism than a static scheme. This is the result of the time dimension introducing an additional asymmetry into the problem, which may be exploited to construct sequential bargaining mechanisms that ex ante dominate the most efficient static mechanisms. Even in situations where a static mechanism is ex ante efficient, it is unlikely that such a mechanism would be adopted by the bargainers, since it necessarily would violate sequential rationality.

The analysis presented here represents an early step toward understanding how agreements are reached in conflict situations under uncertainty. Several simplifying assumptions have been made in order to keep the analysis manageable. First, modeling the agents' time preferences with constant discount rates is an appealing example, but not an accurate description of all bargaining settings. (Fishburn and Rubinstein (1982) derive under which circumstances the discounting assumption is valid. In particular, they prove that any preferences over bargaining outcomes that are monotonic, continuous, and stationary can be represented by discounting provided the bargainers exhibit impatience over all outcomes except that of no agreement.) Second, the agents have been assumed to be risk neutral, but in many bargaining situations the agents' willingness to take risks is an important bargaining factor. Third, I have restricted attention to rational agents who can calculate (at no cost) their optimal strategies. Certainly, few agents are so consistent and calculating. With less-than-rational agents, an agent's capacity to mislead his opponent becomes an important variable in determining how the gains from trade are divided. Finally, I have assumed that the players' valuations are independent. However, in many settings the bargainers' valuations will be correlated, and so, for example, the seller's willingness to trade may be a signal of the valuation of the object to the buyer.

Although it would be useful in future research to weaken the simplifying assumptions made here, perhaps the most fruitful avenue for further study is the analysis of specific extensive-form bargaining games. The advantage of looking at specific extensive-form games is that the bargaining rules are independent of the probabilistic beliefs that the players have about each other's preferences. In a direct revelation game, on the other hand, the bargaining rule depends in a complicated way on these probabilistic beliefs. Because of this dependence, direct revelation games are not played in practice.

Can one find a strategic game that comes close to implementing the ex ante-efficient bargaining mechanism over a wide range of bargaining situations? Initial studies along these lines have been conducted by Cramton (1984), Fudenberg and Tirole (1983), and Sobel and Takahashi (1983). All three papers consider a model in which only one of the bargainers makes offers. When the players' reservation prices are uniformly distributed on [0,1] and their discount rates are equal, it was found that this model results in 32 percent of the gains from trade being lost, as opposed to 16 percent being lost when the ex ante-efficient bargaining mechanism is adopted

(Cramton (1984)). Thus, the players' inability to commit to ending negotiations results in a bargaining outcome that is significantly less efficient than if commitment were possible.

Perhaps a better candidate for a strategic bargaining game that is nearly ex ante efficient is the game in which the bargainers alternate offers. This game was analyzed by Rubinstein (1982) in a setting of complete information, but an analysis with incomplete information has yet to be done. Of particular interest is the alternating-offers game as the time between offers goes to zero, because this strategic game represents a very general bargaining rule: At any time, a bargainer may make a new offer or accept the most recent offer of his opponent. It would be a pleasant surprise if such a reasonable bargaining game was nearly ex ante efficient over a variety of circumstances.

A second promising area for research is further study on the implications of sequential rationality to bargaining and to more general games of incomplete information. I intend to address this issue in depth in future research.

REFERENCES

Chatterjee, Kalyan, and William Samuelson (1983): Bargaining under Incomplete Information. *Operations Research*, Vol. 31, pp. 835 -51.

Cramton. Peter C. (1984): Bargaining with Incomplete Information: An Infinite Horizon Model with Continuous Uncertainty. *Review of Economic Studies*, Vol. 51, pp. 579-93.

Fishburn, Peter C., and Ariel Rubinstein (1982): Time Preference. *International Economic Review* Vol. 23, pp. 719 - 36.

Fudenberg, Drew, and Jean Tirole (1983): Sequential Bargaining with Incomplete Information. *Review of Economic Studies*, Vol. 50, pp. 221 -47.

Gresik, Thomas A., and Mark A. Satterthwaite (1983): *The Number of Traders Required To Make a Market Competitive. The Beginnings of a Theory*. Research Paper No.551, Graduate School of Management, Northwestern University.

Holmström, Bengt, and Roger B. Myerson (1983): Efficient and Durable Decision Rules with Incomplete Information. *Econometrica*, Vol. 51, 1799 1820.

Kreps, David M., and Robert Wilson (1982): Sequential Equilibria. *Econometrica*, Vol. 50, pp. 863-94.

Myerson, Roger B.. and Mark A. Satterthwaite (1983): Efficient Mechanism for Bilateral Trading. *Journal of Economic Theory* Vol. 28, pp. 265-81.

Selten, R. (1976): Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games. *International Journal of Game Theory*, Vol.4. pp. 25-53.

Sobel. Joel. and Ichiro Takahashi (1983): A Multi-Stage Model of Bargaining. *Review of Economic Studies*, Vol. 50. pp. 411 - 26.

Wilson, Robert (1982): *Double Auctions.* Technical Report No. 391, Stanford Institute for Mathematical Studies in the Social Sciences.