11

Strategic Bargaining

Of the many possible applications of strategic interaction among a small number of players, one ubiquitous situation surely comes to mind: bargaining. Examples are plentiful: a firm and a union bargaining over wages and benefits; a local municipality bargaining with a private provider over the terms of service; the head of a political party bargaining with other party members over campaign issues; and even the mundane bargaining that occurs between a buyer and a merchant at a street bazaar. In all of these examples, deals are struck by parties bargaining over terms, payments, and other aspects that all boil down to one basic fact: each party wants to get the best deal it can.

How can we model and analyze bargaining through the lens of game theory? As demonstrated by the examples just mentioned, the issue is often about a surplus that has to be split among the parties, and it involves the parties making proposals, responding to them, and trying to settle on an agreement. Following some early work by Ståhl (1972, 1977), Rubinstein (1982) offered a particular stylized model of bargaining that has become the standard in most applications. The simple framework considers two players that need to split a "pie" (representing the surplus from an agreement or the gains from trade). The pie is assumed to have a total value that is normalized to equal 1, and then the parties bargain on how to split the pie.

In order to analyze this process as a strategic game, the bargaining process itself needs to be structured to include actions, outcomes that result from actions, and preferences over the outcomes, as with any other game we can imagine. Ståhl suggested a prespecified procedure in which the game starts with player 1 making an offer to split the pie and player 2 either accepting or rejecting the offer. If the offer is accepted, the split proposed by player 1 is implemented, and the game ends. If player 2 rejects player 1's offer, the players then switch roles: player 2 makes the offer and player 1 responds by accepting or rejecting the offer. The game can either continue this way until some exogenous deadline arrives or simply go on forever.

Yet, as the saying goes, "time is money," and delay ought to impose some loss on the players. There are two ways to impose this loss. First, there may be a fixed cost of progressing from one round to another, so that a constant piece of the pie is removed every time a rejection occurs. Alternatively we can impose discounting in the same way we did for repeated games: after every period of rejection, the total size of the pie is only a fraction of what it was before. This second approach is the one we will be following here, and we can summarize the game as follows.

In the first round:

- Player 1 offers shares $(x, 1-x)$, where player 1 receives x and player 2 receives $1-x$.
- Player 2 then chooses to accept, causing the game to end with payoffs $v_1 = x$ and $v_2 = 1-x$, or reject, causing the game to move to the second round.

In the second round:

- A share $1-\delta$ of the pie is dissipated (the pie is discounted with discount factor $0 < \delta < 1$).
- Player 2 offers shares $(x, 1-x)$, where player 1 receives x and player 2 receives $1-x$.
- Player 1 chooses to accept, causing the game to end with payoffs $v_1 = \delta x$ and $v_2 = \delta(1-x)$, or reject, causing the game to move to the third round.

In the third round and beyond:

- The game continues in the same way, where following rejection in an odd period player 2 gets to offer in the next even period, and vice versa.
- Each period has further discounting of the pie, so that in period t the total pie is worth δ^{t-1}.

From the specification of the bargaining procedure we see that once a player agrees to the offer made by his opponent, the game will end with an agreement. If instead the players keep rejecting offers then the game may continue either indefinitely or until some prespecified deadline is reached. Assume for now that the game ends in some prespecified final period T, which will be reached if the players did not reach an agreement in the $T-1$ periods that preceded T. Furthermore assume that period T is a "hard deadline," so that both players receive a payoff of zero if they fail to reach an agreement in this last period. This, for example, can be the case with a fisherman bargaining with a restaurateur over a fresh catch. If they wait more than a few hours, the fish will go bad and there will no longer be gains from trade. The heuristic extensive form of this game with an odd number of periods T (so that player 1 is first and last) is depicted in Figure 11.1.

Notice that this game has an interesting structure that is different from what we have seen so far. On the one hand, it has some features of a finitely repeated game as follows: If we think of an "odd" round as one in which player 1 proposes and player 2 responds, and an "even" round as the converse, then we can treat each *pair of rounds* as *one stage* that repeats itself as long as an agreement is not reached. Furthermore as rounds proceed the value of the pie that these players need to split is shrinking according to the discount factor δ. On the other hand, there are two features that are not part of the repeated-game structure we developed in Chapter 10. First, the game can end during any round if the proposal is accepted, and second, payoffs are obtained only when the game actually ends and not as a flow of stage-game payoffs.

Before proceeding with the analysis, an obvious question is whether this stylized bargaining game is a reasonable caricature of reality. We need to convince ourselves that this game is, at least at some level, representative of what we believe bargaining is about. To a large extent, sequences of offers and counteroffers are natural components of any type of bargaining, and clearly once agreements are reached the bargaining stage is over. However, the fixed end date T, at which an all-or-nothing agreement

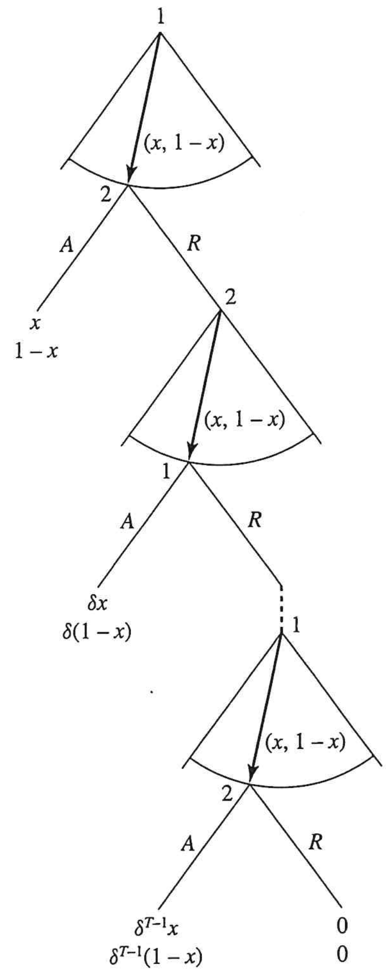

FIGURE 11.1 An odd T-period alternating-offer bargaining game.

is reached, is not extremely appealing. For simplicity, we will first analyze the game with a finite end date; later we will extend the analysis to eliminate this artificial termination stage.

11.1 One Round of Bargaining: The Ultimatum Game

We begin the analysis with an easy case in which there is only one round of bargaining, so that $T = 1$. This is often referred to in the literature as an **ultimatum game:** player 1 makes an ultimatum take-it-or-leave-it offer to player 2, and player 2 either accepts it, in which case the pie is split between the players, or rejects it, in which case both players receive a payoff of zero. Because this is a game of perfect information (player 2 sees the offer of player 1 before he needs to accept or reject it), we can apply backward induction to find which paths of play can be supported by a subgame-perfect equilibrium.

As a benchmark, we begin the analysis by finding which paths of play can be supported by a Nash equilibrium without requiring sequential rationality. The result is quite striking:

Proposition 11.1 *In the bargaining game if $T = 1$ then any division of surplus $x^* \in [0, 1]$, $(v_1, v_2) = (x^*, 1 - x^*)$, can be supported as a Nash equilibrium.*

Proof We will construct a pair of strategies that are mutual best responses and that lead to $(x^*, 1 - x^*)$ as the division of surplus. Let player 1's strategy be "I propose x^*," and let player 2's strategy be "I accept any offer $x \leq x^*$ and reject any offer $x > x^*$." It is easy to see that these two strategies are mutual best responses independent of the value of $x^* \in [0, 1]$. ■

This proposition tells us that the concept of Nash equilibrium has no bite for this simple one-stage bargaining game because it can rationalize *any division of surplus.* In other words, we cannot predict in any meaningful way what the outcome of such a game will be if we require only that people choose strategies that are mutual best responses.

That said, a close observation of the strategies constructed to support an arbitrary division of surplus will immediately reveal that sequential rationality is almost always violated: Player 2 is saying "there is a minimum that I am willing to accept." But what if player 1 offers him less? In particular what if he offers him $(1 - x) = \varepsilon > 0$ with ε being a very small number? If player 2 rejects this offer he will get 0, while if he accepts it he will get a payoff of $\varepsilon > 0$, implying that his best response is to accept *any strictly positive payoff.* Anticipating this, player 1 should offer player 2 the smallest possible amount. This logic leads to the following result:

Proposition 11.2 *The bargaining game with $T = 1$ admits a unique subgame-perfect equilibrium in which player 1 offers $x = 1$ and player 2 accepts any offer $x \leq 1$.*

Proof We have established that player 2 must accept any positive share $(x < 1)$. Player 2 is indifferent between accepting or rejecting $x = 1$, so the proposed strategy is sequentially optimal, and the unique best response of player 1 to player 2's strategy is to offer $x = 1$. The only other sequentially rational strategy for player 2 is to accept any strictly positive share $(x < 1)$ and reject getting 0 $(x = 1)$. But player 1 does not have a best response to this strategy,[1] and therefore it cannot be part of a subgame-perfect equilibrium. ■

This result is very stark, especially when compared with the analysis of Nash equilibrium. Requiring only mutual best responses yields no meaningful prediction, but sequential rationality predicts a unique and extreme outcome: player 2 should accept anything, and thus player 1 has an extreme form of take-it-or-leave-it first-mover advantage, giving him the whole pie of surplus.

Remark Many experiments have been performed by researchers to see what behavior actually prevails for the ultimatum game, in which player 1 can offer to split a sum of money with player 2. If player 2 accepts then the split is realized and both players receive the proposal of player 1. If player 2 rejects then the players receive

1. This is similar to what happens in the asymmetric-cost Bertrand competition model presented in Section 5.2.4, in which player 1's payoff function is discontinuous, and hence he does not have a well-defined best response. To see it in this context, imagine that player 2's strategy is to accept any offer $x < 1$ and reject $x = 1$. Clearly $x = 1$ is not a best response because any offer $x < 1$ is better. The problem is that no offer $x < 1$ is a best response either. If player 1 offers some $x' < 1$, then the offer $x = x' + \frac{1-x'}{2}$ is better than x'. The conclusion is that player 1's payoff function is discontinuous at $x = 1$, and he does not have a best response to player 2's strategy.

nothing. Contrary to the theory, and maybe not too surprisingly, the experiments show that those players in the role of player 1 offer significant shares to those in the role of player 2, typically between 25% and 50% of the total value. Furthermore, when player 2 is offered a small amount, typically less than 20–25% of the total value, then he will reject these offers. This suggests that subjects in the role of player 2 act with some "spite" when they are offered low shares because they prefer getting nothing to an outcome in which player 1 gets a large fraction of the total value. See Camerer (2003) for an excellent summary of the experimental results and some theories that predict this behavior.

11.2 Finitely Many Rounds of Bargaining

When we consider bargaining games that extend beyond one period but that will still end at some specified date $T < \infty$, the Nash equilibrium concept continues to have no predictive power. In particular we can easily construct the kind of strategies used in the proof of proposition 11.1 to *any* horizon, *including an infinite horizon.* To construct a Nash equilibrium that results in an arbitrary division of surplus $x^* \in [0, 1]$ we can use the following time-independent strategies: In every odd round in which player 1 proposes, player 1 will propose x^* and player 2 will accept proposals of $x \leq x^*$. In every even round in which player 2 proposes, player 2 will propose x^* and player 1 will accept proposals of $x \geq x^*$. It is easy to see that these strategies are mutual best responses, and an agreement is reached in the first round with the division of surplus $(x^*, 1 - x^*)$.

What is even more disappointing is that, using the Nash concept, we cannot even pin down when an agreement will be reached. For example, we can adopt the strategies that result in any $(x^*, 1 - x^*)$ with the following modification: in the first period player 1 offers $x = 1$ and player 2 rejects *any* offer. It is easy to see that these form a pair of best responses, resulting in a Nash equilibrium in which an agreement is reached in the second period. In exercise 11.1, you are asked to show that similar strategies can be constructed to support an agreement of *any* division of surplus in *any* period.

Imposing sequential rationality will, once again, offer a stark contrast. Indeed it should not be surprising that applying sequential rationality and subgame perfection to any finite-length bargaining game will result in a unique outcome, because backward induction will apply in a similar way that it did for the case of one round of bargaining. In the last round the proposer will offer to keep the entire pie to himself, the responder will agree, and the solution will proceed sequentially using backward induction.

In the one-round ultimatum game analyzed previously, player 1 got all the surplus in the unique subgame-perfect equilibrium because by construction the only round is the last round. Sequential rationality implies that in the last round the player making the offer has all the "bargaining power," which in turn implies that he will get the whole surplus. We can ask then, what will happen in a game with two rounds of bargaining?

As Figure 11.2 demonstrates, the second round is a one-round bargaining game, which implies that *if it is reached* then player 2 gets the whole pie in any subgame-perfect equilibrium. Continuing with backward induction, and assuming that $\delta \leq 1$, we can find the subgame-perfect equilibrium for the two-round game as follows. In the first round player 2 knows that in the next (and last) round he will receive the whole pie, which will be worth only δ because of discounting. Thus if he is offered less than δ for himself (player 1 offers $x > 1 - \delta$) then player 2 should reject the offer and get the whole pie in the next period. Hence the unique sequentially rational strategy for

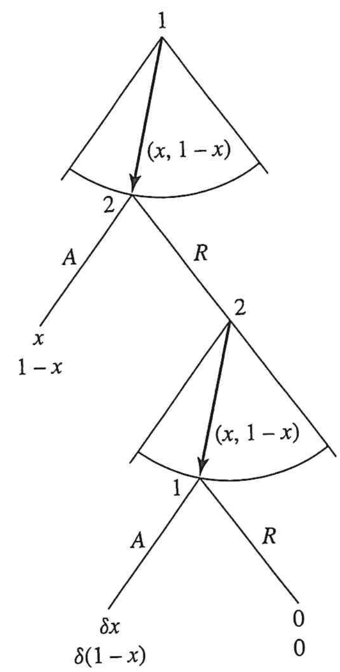

FIGURE 11.2 An alternating-offer bargaining game with two periods.

player 2 at the accept/reject stage of the first round is to accept any offer of $x \le 1 - \delta$. Moving up the tree to the node at which player 1 makes his first offer, he realizes that any offer $x > 1 - \delta$ will be rejected, while any offer $x \le 1 - \delta$ will be accepted. Thus player 1's best response is to offer $x = 1 - \delta$, leading to a division of surplus equal to $(v_1, v_2) = (1 - \delta, \delta)$.

Note that if $\delta = 1$ then player 2 gets the whole pie because he then has the ultimate form of a take-it-or-leave-it advantage, and with $\delta < 1$ player 1 has a bit of a first-mover advantage owing to the discounting. This observation is useful because it suggests an interesting pair of forces. Essentially the backward induction logic that we have used for the two-round bargaining model will generalize to any finite-round game. Therefore a tension will emerge between the ability to make a last take-it-or-leave-it offer and the first-mover advantage that is a result of discounting given $\delta < 1$. Whether player 1 or player 2 is the last mover will artificially depend on whether the game has an odd or even number of periods.

Consider the case with an odd number of rounds $T < \infty$, implying that player 1 has both the first-mover and last-mover advantages. The following backward induction argument applies:

- In period T player 2 accepts any offer, so player 1 offers $x = 1$ and payoffs are $v_1 = \delta^{T-1}$ and $v_2 = 0$.
- In period $T - 1$ (even period, player 2 offers), by backward induction player 1 should accept anything resulting in a payoff of $v_1 \ge \delta^{T-1}$. If player 1 is offered x in period $T - 1$ then $v_1 = \delta^{T-2}x$. This implies that in period $T - 1$ player 1 will accept any $x \ge \delta$ and by backward induction player 2 should offer $x = \delta$, which yields player 1 a payoff of $v_1 = \delta \times \delta^{T-2} = \delta^{T-1}$. Payoffs are then $v_1 = \delta^{T-1}$ and $v_2 = \delta^{T-2}(1 - \delta)$.

- In period $T-2$ (odd period), conditional on the analysis for $T-1$, player 2's best response is to accept any x that gives him $\delta^{T-3}(1-x) \geq \delta^{T-2}(1-\delta)$. Player 1's best response to this is to offer the largest x that satisfies this inequality, and solving it with equality yields player 1's best response: $x = 1-\delta+\delta^2$. This offer followed by player 2's acceptance yields $v_1 = \delta^{T-3}x = \delta^{T-3} - \delta^{T-2} + \delta^{T-1}$ and $v_2 = \delta^{T-2} - \delta^{T-1}$.
- In period $T-3$ (even period), using the same backward induction argument, player 1's best response is to accept any x that satisfies $\delta^{T-4}x \geq \delta^{T-3} - \delta^{T-2} + \delta^{T-1}$, and player 2 then makes an offer that satisfies this constraint with equality, $x = \delta - \delta^2 + \delta^3$, yielding $v_1 = \delta^{T-3} - \delta^{T-2} + \delta^{T-1}$ and $v_2 = \delta^{T-4}(1-x) = \delta^{T-4} - \delta^{T-3} + \delta^{T-2} - \delta^{T-1}$.

We can continue with this tedious exercise only to realize that a simple pattern emerges. If we consider the solution for an odd period $T-s$ (s being even because T is assumed to be odd) then the backward induction argument leads to the sequentially rational offer

$$x_{T-s} = 1 - \delta + \delta^2 - \delta^3 + \cdots + \delta^s,$$

while for an even period $T-s$ (s being odd) then the backward induction argument leads to the sequentially rational offer

$$x_{T-s} = \delta - \delta^2 + \delta^3 - \delta^4 + \cdots + \delta^s.$$

We can use this pattern to solve for the subgame-perfect equilibrium offer in the first period, x_1, which by backward induction must be accepted by player 2 and it is equal to

$$\begin{aligned} x_1 &= 1 - \delta + \delta^2 - \delta^3 + \delta^4 + \cdots + \delta^{T-1} \\ &= (1 + \delta^2 + \delta^4 + \cdots + \delta^{T-1}) - (\delta + \delta^3 + \delta^5 + \cdots + \delta^{T-2}) \\ &= \frac{1-\delta^{T+1}}{1-\delta^2} - \frac{\delta - \delta^T}{1-\delta^2} \\ &= \frac{1+\delta^T}{1+\delta}, \end{aligned}$$

and this in turn implies that

$$v_1^* = x_1 = \frac{1+\delta^T}{1+\delta} \quad \text{and} \quad v_2^* = (1-x_1) = \frac{\delta - \delta^T}{1+\delta}. \tag{11.1}$$

We can now offer some insights into this solution, which has some very appealing properties. The first observation is important:

Proposition 11.3 *Any subgame-perfect equilibrium must have the players reach an agreement in the first round.*

The proof follows a simple logic. If an agreement is reached in a later round with payoffs (v_1', v_2') then discounting implies that part of the surplus is wasted and that $v_1' + v_2' < 1$. But then player 1 could deviate and offer $x = 1 - v_2' - \varepsilon$ for some small $\varepsilon > 0$, which guarantees player 2 the payoff $v_2' + \varepsilon$ immediately in the first round. Sequential rationality implies that player 2 should accept this immediately, and for ε small enough this gives player 1 a payoff greater than v_1'.

A second observation follows from the fact that player 1 has both the last-mover take-it-or-leave-it advantage and the first-mover discounting advantage. This implies that $v_1^* > v_2^*$ for any discount factor $\delta \in [0, 1]$, as can be seen from the equilibrium values in (11.1). Because T is fixed, however, it is hard to isolate these two effects, especially if we wish to measure the way in which patience, represented by δ, affects the equilibrium payoffs. For example, if players are very impatient and $\delta = 0$ then the first-mover advantage kicks in with full force. Regardless of T, this is basically equivalent to a one-round game in which player 1 gets the whole surplus. On the other hand, if we consider the limit of $\delta = 1$ then $v_1^* = 1$ and $v_2^* = 0$, because no discounting implies that moving last guarantees the whole surplus, regardless of the length of the game T.

This last observation is somewhat disturbing: if players are very patient then the last mover's take-it-or-leave-it advantage just flows up the game tree no matter how long the game is! This rather counterintuitive result is precisely due to the artificial stopping period T. If $\delta = 1$ then the only thing that matters is who has the last word—there is no loss from waiting, and hence the forward-looking players know that the last person to make an offer holds all the bargaining power.

To overcome this artificially dictated outcome, we will first address the source of the problem by fixing the discount factor δ and observing what happens to the equilibrium payoffs as the game gets longer and longer. More precisely,

$$\lim_{T\to\infty} v_1^* = \lim_{T\to\infty} \frac{1+\delta^T}{1+\delta} = \frac{1}{1+\delta} \quad \text{and} \quad \lim_{T\to\infty} v_2^* = \lim_{T\to\infty} \frac{\delta-\delta^T}{1+\delta} = \frac{\delta}{1+\delta}. \tag{11.2}$$

As we can see from the limit payoffs in (11.2), the artificial end-period effect disappears. We can clearly observe the first-mover discounting advantage in that for any $0 < \delta < 1$ it must be that $\lim_{T\to\infty} v_1^* > \lim_{T\to\infty} v_2^*$. We can also use the limiting equilibrium payoffs to ask what happens as the patience of the players changes without the problem posed by an artificial end period T. For the extreme case of impatience with $\delta = 0$ we get the same result: the game is equivalent to a one-round ultimatum game. When the players become very patient, we now obtain a more sensible result,

$$\lim_{\delta\to 1}\lim_{T\to\infty} v_1^* = \lim_{\delta\to 1} \frac{1}{1+\delta} = \frac{1}{2} \quad \text{and} \quad \lim_{\delta\to 1}\lim_{T\to\infty} v_2^* = \lim_{\delta\to 1} \frac{\delta}{1+\delta} = \frac{1}{2}. \tag{11.3}$$

The limit payoffs in (11.3) suggest a rather appealing and intuitive result for very long horizons with very patient players: the very long horizon eliminates the last mover's take-it-or-leave-it advantage for any given discount factor δ, leaving only the first-mover advantage due to discounting. If we take the level of patience to its limit of $\delta = 1$ then this eliminates the first-mover advantage as well, and we have the two players splitting the pie equally in the unique subgame-perfect equilibrium.[2]

As exercise 11.3, you can solve for the subgame-perfect equilibrium when the number of periods is even, and see that at the limit we get the same results.

2. Note that to reach this appealing conclusion the order of limits matters. If we reverse the order of limits and take $\delta \to 1$ first, followed by $T \to \infty$ second, then $\lim v_1^* = 1$ and $\lim v_2^* = 0$, preserving the artificial take-it-or-leave-it advantage.

11.3 The Infinite-Horizon Game

We now analyze the infinite-horizon bargaining game in which, if an agreement is never reached, the players continue to alternate roles indefinitely. In this case the assumption that disagreement leads to zero payoffs is natural because of discounting: if the two players disagree forever then there is nothing left on which to agree. However, there is a crucial difference between analyzing the infinite-horizon game and considering the infinite limit of the finite-horizon game: the path of perpetual disagreement is of infinite length, and we cannot therefore apply a backward induction argument to find the subgame-perfect equilibrium.

There is, however, an interesting feature of the infinite-horizon bargaining game: the *stationary structure of the game* following a disagreement at any stage. Every odd period is the same, with player 1 making an offer, and the continuation game has a potentially infinite horizon. Similarly every even period is the same, with player 2 making the offer. This allows us to apply a rather appealing, and not too difficult, logic to solve for what turns out to be the unique subgame-perfect equilibrium of the infinite-horizon bargaining game.[3]

The first important observation is that, following the same logic we introduced earlier in proposition 11.3, sequential rationality implies that an agreement must be reached in the first period. The reason is that wasteful discounting will not be tolerated because it can be avoided in the first stage. Hence any subgame-perfect equilibrium must have agreement reached in the first period.

Now imagine that there is more than one subgame-perfect equilibrium. If this were the case then when player 1 makes an offer he must have a *best* subgame-perfect equilibrium, yielding him a value of $\overline{v}_1$, and a *worst* subgame-perfect equilibrium, yielding him a value of $\underline{v}_1$. Similarly when player 2 makes an offer after rejecting an offer made by player 1 then he too has a *best* subgame-perfect equilibrium, yielding him a value of $\overline{v}_2$, and a *worst* subgame-perfect equilibrium, yielding him a value of $\underline{v}_2$.

This is where the stationary structure of the game is useful, because it implies that $\overline{v}_1 = \overline{v}_2 = \overline{v}$ and $\underline{v}_1 = \underline{v}_2 = \underline{v}$. The reason is that in any odd period the game looks exactly the same as in any even period, just with the identities of the players switched. Hence if the structure of the game results in the existence of different subgame-perfect equilibria that support certain payoffs for player 1 starting in period 1, the same must be true for player 2 in period 2.

Now we take advantage of the fact that what one player gives up, the other player must receive (and recall that agreement must be reached in the first period), which results in a straightforward relationship between the best- and worst-payoff subgame-perfect equilibria of the game. What is worst for player 1 must be best for player 2 and vice versa. Hence in the subgame-perfect equilibrium that supports the payoff $\underline{v}$ for player 1 in period 1, it must be true that if player 2 rejects player 1's offer, he can secure himself a payoff of $\overline{v}$ when it is his turn to make the offer.

Immediate agreement implies that player 1 must offer player 2 a share of the pie that will deter player 2 from rejecting player 1's offer. If player 2 can get $\overline{v}$ after rejection, this implies that player 1 must offer player 2 a payoff of at least $\delta\overline{v}$ in the first period. Of course player 1 will offer the least he can to secure an agreement,

3. The method used here was developed by Shaked and Sutton (1984), but the infinite model was proposed, and the unique solution was identified, in the seminal paper by Rubinstein (1982).

which implies that he will offer player 2 a share exactly equal to $\delta\overline{v}$, which in turn implies that player 1's payoff from this subgame-perfect equilibrium must be

$$\underline{v} = 1 - \delta\overline{v}. \tag{11.4}$$

A symmetric argument applies for the subgame-perfect equilibrium that results in player 1 obtaining a payoff of $\overline{v}$ and player 2 obtaining a payoff of $\underline{v}$, which results in the equation

$$\overline{v} = 1 - \delta\underline{v}. \tag{11.5}$$

Taking (11.4) and (11.5) we obtain that

$$\overline{v} = \underline{v} = \frac{1}{1+\delta}.$$

This result characterizes the unique subgame-perfect equilibrium of the infinite-horizon bargaining game. Notice that it coincides with the limit of the unique subgame-perfect equilibrium of the finite-horizon bargaining game as shown in Section 11.2. Furthermore because we know that agreement is reached in the first round we can use this payoff characterization to spell out the strategies that together support this outcome. In the unique subgame-perfect equilibrium the strategies are as follows: In each odd round player 1 offers to keep $x = \frac{1}{1+\delta}$ to himself and player 2 accepts the offer if $x \leq \frac{1}{1+\delta}$. In each even round player 2 offers to give player 1 $x = \frac{\delta}{1+\delta}$ and player 1 accepts any $x \geq \frac{\delta}{1+\delta}$.

Remark As you might have imagined, bargaining games have received considerable attention in the literature, in both economics and political science, because bargaining situations are so ubiquitous. Many variants of the bargaining game described here have been analyzed, in which the relationship might break down, the players each have an "outside option" that is greater than zero, there are more than two players that can be matched randomly, . . . the list goes on and on. Nash (1950b) offered an analysis of the bargaining problem, but interestingly he did not follow the framework of noncooperative game theory that he himself had founded. Instead he offered an axiomatic approach to suggest a sensible solution to the two-person bargaining problem. The resulting "Nash bargaining solution" is widely used in economics. For a broad treatment of bargaining games and solutions see Osborne and Rubinstein (1990) and Muthoo (1999).

11.4 Application: Legislative Bargaining

The bargaining model analyzed here offers a reasonable depiction of many two-person bargaining situations. In particular there is the potential for give-and-take, and at the end an agreement has to be reached by both parties for the deal to be sealed. Another setting in which bargaining is ubiquitous is in legislatures, where lawmakers bargain over the allocation of surplus through either bills, budget agreements, or regulations. In most political settings the convention is that once a majority of players concur on an offer, the allocation is agreed upon despite the dismay of those who may still oppose it.

In an important contribution, Baron and Ferejohn (1989) build on the Rubinstein-Ståhl model of bargaining and offer a framework for analyzing multilateral bargaining with a simple majority rule. The two main departures from the two-player model are as follows: First, there is an odd number N of players, where $\frac{N+1}{2}$ votes are needed to secure an agreement. Second, instead of players alternating their roles in a prespecified order, the bargaining protocol follows a random-assignment rule: in every period each player has an equal probability of being the proposer.

Baron and Ferejohn consider two bargaining rules that to a large extent mimic common practices: the closed rule, under which proposals cannot be amended, and the open rule, under which they can. In what follows, we adhere to the approach used in Section 11.3 to analyze the outcomes and expected payoffs from a potentially infinite-horizon model. The technique is to use the payoffs that some players expect to get if they reject an offer to determine what the equilibrium payoffs must be to support the proposed equilibrium behavior.

11.4.1 Closed-Rule Bargaining

The closed-rule setting is very much an N-player version of the Rubinstein infinite-horizon model. When called upon to propose, the proposer makes an offer to the legislative body of the form $x = (x_1, x_2, \ldots, x_N)$, where x_i is the share of surplus offered to player i, with $\sum_{i=1}^{n} x_i \leq 1$. If a majority accepts the proposal then the game ends, and if not then the current "pie" is discounted according to the discount factor δ.

It turns out that when there are three or more players in an N-player version of the Rubinstein infinite-horizon bargaining game then there are many possible subgame-perfect equilibria.[4] What Baron and Ferejohn suggest is a simple way around this indeterminacy. They focus attention on *symmetric stationary equilibria* in which first, each player proposes the same split of the pie every time he is called to propose, regardless of the history of the game, and second, the respondents vote based only on the current proposal as it compares to their expectations about future proposals. What is clever about these assumptions is that together they imply that, like the two-player version of Rubinstein's model, the game effectively "starts over" every time a proposal is rejected, and because the proposer is chosen randomly at the beginning of every stage, the continuation value of each player is the expected utility of playing the game.[5]

Let v be the expected payoff for any player i to play the subgame-perfect equilibrium of this game at the beginning of any stage,[6] and consider a player i who is choosing to respond to a proposal that awards him x_i. If he rejects the proposal then

4. This was noted first by Herrero (1985). For a large number of players and a discount factor close to 1, any split of the pie can be supported by a subgame-perfect equilibrium. Showing this is beyond the scope of this text, but the interested reader can consult Baron and Ferejohn (1989) for a proof.

5. More precisely, Baron and Ferejohn (1989) define two subgames as *structurally equivalent* if (1) the agenda at the initial nodes of the subgames is identical; (2) the set of players who can be recognized is the same; and (3) the strategy sets of players are identical. This means that two subgames in which a previous offer was rejected and someone is chosen to propose a bill are structurally equivalent. An equilibrium is then defined to be *stationary* if the continuation values for each structurally equivalent subgame are the same. Symmetry then implies that each player behaves in the same way. As a consequence of these restrictions, before any subgame of proposal starts, all players have the same expected payoff from playing the game.

6. Given the assumption that players are playing a symmetric stationary equilibrium, and because each player is equally likely to be chosen to be a proposer at the beginning of the game and at any stage, then this value v must be the same for every player.

he expects to get v in the next period, and hence he will accept the proposal if and only if $x_i \geq \delta v$. As a consequence, any player who plays the role of a proposer anticipates that he must propose at least δv to at least $\frac{N+1}{2}$ other players.

Given this observation, the proposer's best response is to choose exactly $\frac{N-1}{2}$ other players and offer them exactly δv each, keeping the rest of the pie for himself. Notice that $\frac{N-1}{2}$ is half of the group of responders, which together with the proposer makes up a majority in favor of the proposed split. Hence the portion of the pie that the proposer keeps for himself in the symmetric subgame-perfect equilibrium is

$$k = 1 - \frac{N-1}{2}\delta v. \tag{11.6}$$

To maintain symmetry, each proposer chooses his $\frac{N-1}{2}$ coalition partners randomly from the remaining $N-1$ players, implying that each of the $N-1$ responders gets chosen to be part of the coalition with probability $\frac{1}{2}$.

The equilibrium analysis leads to the computation of v. At the beginning of the game (and of every stage after rejection), each player expects to play in one of three roles. He can either be the proposer and receive k, which occurs with probability $\frac{1}{N}$, or be a responder, which occurs with probability $\frac{N-1}{N}$. If he is a responder, then with probability $\frac{1}{2}$ he will be in the winning coalition and receive δv, or he can be outside the coalition and receive zero. This implies that

$$v = \frac{k}{N} + \frac{N-1}{2N}\delta v. \tag{11.7}$$

Combining (11.6) and (11.7) we obtain that $v = \frac{1}{N}$. This should not be surprising because, after all, we focused on a symmetric equilibrium. As a consequence, we force all the players to get an equal share in expectations. In addition, because any agreement that could be reached at a later period can be reached at the beginning of the game, there is no need to have a delay.

It is more interesting to focus attention on the value of k, the share obtained by the proposer, which is

$$k(N) = 1 - \delta\left(\frac{N-1}{2N}\right). \tag{11.8}$$

Two implications of (11.8) are noteworthy. The first, which is congruent with Rubinstein's two-player model, is that as the discount factor δ increases, the share obtained by the first-mover proposer decreases. The intuition is that the proposer has to pay more to the responders in order to discourage them from waiting for their turns because the future becomes more valuable, and in the future they have a chance of being proposers themselves.

The second is the relationship between the share that the proposer has to offer his coalition partners and the size of the coalition needed. On the one hand, as the number of players grows the per-partner offer goes down, because, as we have seen, $v = \frac{1}{N}$. On the other hand, the proposer needs to pay more partners to form a winning coalition. It turns out that the second force overcomes the first, and as N grows the share of the proposer decreases. More concretely if $N = 3$ then the share obtained by the proposer is $k(3) = 1 - \delta\frac{1}{3} > \frac{2}{3}$, while as N increases this share drops and at the limit is $\lim_{N\to\infty} k(N) = \frac{2-\delta}{2}$. It is interesting to try to interpret this result as bargaining power. As N increases, it is true that in *absolute* terms the share of the proposer drops. However, in *relative* terms, compared to the share of other players (that is, in terms of the ratio of the proposer's share to the share of a coalition member), it increases and approaches infinity. Intuitively, as there are more and more players, the proposer has more options to form a coalition, and hence has more relative to the others in his coalition. Nonetheless to form that coalition he has to give up more resources.

11.4.2 Open-Rule Bargaining

Now let's introduce the option of amending a proposal, as is allowed in the U.S. Congress. Of course there are many ways in which bills and proposals can be amended, and just as with the closed-rule model we have to make some assumptions on the protocol to make the model tractable. As such, assume that after the proposer makes his proposal $x = (x_1, x_2, \ldots, x_N)$, one of the remaining $N - 1$ players is selected to offer an amendment; we refer to him as the amender. The amender can second the proposer's original proposal and then bring it to a yes-or-no vote. In this case a simple majority of $\frac{N+1}{2}$ is needed to pass the proposal x.

Alternatively the amender can offer an amendment, $x' = (x'_1, x'_2, \ldots, x'_N)$, in which case the game proceeds in two steps: First, the players vote between the original proposal x and the amendment x'. The winner of this vote becomes the baseline proposal for the next period of play, in which a new amender is chosen randomly among the N players to either second the baseline proposal or offer a new amendment. The game proceeds in this way until a proposal both passes the stage at which an amender seconds it and is accepted by a simple majority. As you can imagine, this model is more complex than the closed-rule model, given its protocol of allowing for amendments. As such, we take the simplest case of $N = 3$.

In the open-rule model not only is the original proposer constrained to offer a proposal that might be accepted by a simple majority, but he must also offer enough to deter a random amender from proposing his own amendment that he hopes to pass instead of the original proposal. Hence, whatever the original proposer offers, it must be enough to discourage any potential amender from offering an alternative amendment.

We proceed therefore with two possible approaches. In the first, called "guaranteed success," the original proposal will offer the same share to each of the two potential amenders, and in equilibrium each amender will second the proposal, and it will pass. In the second approach, called "risky success," the original proposer will offer a share to only one of the two potential amenders, taking the risk that the other player will be selected as an amender and in turn will "punish" the original proposer by offering him nothing.

11.4.2.1 Guaranteed Success We will construct a symmetric stationary subgame-perfect equilibrium in which the first proposal passes, regardless of who the amender is. Let the proposer keep k for himself and offer each of the other two players $\frac{1-k}{2}$, with the expectation that either of the two potential amenders will second the proposal and it will pass. As before, we will use the equilibrium conditions to compute the equilibrium level of k. In addition, symmetry implies that in expectations all the players get the same expected payoff from playing the game.

Just as we defined v in the closed-rule model to be the expected payoff for any player i to play the subgame-perfect equilibrium of this game at the beginning of any stage, let $v(k)$ be the equilibrium expected payoff of a player beginning an amendment stage with a previous proposal that offers k to one of the players, say i, and $\frac{1-k}{2}$ to the other two players.

Hence in an equilibrium in which the original proposal is adopted, a proposer must offer each of the other two players at least $\delta v(k)$ to induce each to second the proposal, implying that $\frac{1-k}{2} \geq \delta v(k)$. The proposer will seek to maximize his share given this constraint, so his best response is to choose $\frac{1-k}{2} = \delta v(k)$. But notice that owing to symmetry each of the potential amenders could follow the same strategy and get the equilibrium proposal adopted, with k for himself. As such, it must be the case that $v(k) = k$, and this observation leads to the equation that characterizes this symmetric subgame-perfect equilibrium:

$$\frac{1-k}{2} = \delta k \qquad \text{or} \qquad k = \frac{1}{1+2\delta}. \tag{11.9}$$

Comparing this outcome to that of the closed-rule model, there is one predictable similarity, together with a perhaps more surprising difference. As in the closed-rule case, and as in any dynamic bargaining model with discounting, the original proposer has a first-mover advantage, which is decreasing in the discount factor: the more patient the players are, the more the proposer has to offer the others.

What is more interesting is the effect of amendments. By comparing the solution in (11.9) with the solution of the closed-rule case (11.8), we can show that the original proposer is better off in the closed-rule case.[7] The reason is that in the closed-rule case the proposer can target a subset of the players and play them off against the others, thus gaining extra bargaining power. With an open rule, such targeting is impossible if the proposer wants to guarantee that his proposal will pass, because if he caters to some players at the expense of others then those players who were shortchanged will have an incentive to amend the proposal at the expense of the proposer. As (11.9) shows, when the discount factor approaches 1 then the proposer has no bargaining power at all and the pie is equally split among the players, unlike the closed-rule case, in which his share approaches $\frac{2}{3}$ of the pie.

11.4.2.2 Risky Success We concluded the previous analysis with the observation that the proposer can guarantee himself a share of the pie equal to $k = \frac{1}{1+2\delta}$. An interesting question is the following: can it be beneficial for the proposer to target only one of the other two lawmakers and offer him a larger share at the expense of the remaining one, while taking the risk that the proposal might fail?

To analyze this situation assume again that the proposer keeps k for himself, but instead of splitting the remaining surplus among the two other players he targets only one and offers him $1-k$, while the remaining player is offered nothing. The equilibrium for which we are aiming is one in which the player who was offered a share of $1-k$ moves to second the offer, and hence have it accepted by a majority of two players. However, the strategy is risky because the player who was offered nothing by the proposer will certainly propose an amendment. It is easy to see that any amendment he offers that would give the other player at least $1-k$ will beat the original proposal.

7. Subtracting the proposer's guaranteed-success equilibrium payoff in the open-rule model from his equilibrium payoff in the closed-rule model with three players yields, for all $1 > \delta > 0$,

$$\left(1 - \frac{\delta}{3}\right) - \left(\frac{1}{1+2\delta}\right) = \frac{\delta(5-2\delta)}{3(1+2\delta)} > 0.$$

Of course we have to determine what the player who was offered nothing will himself propose as an amendment. To make things simpler, while focusing as before on a symmetric subgame-perfect equilibrium, we assume that players follow a simple *selective inclusion strategy*. To illustrate this strategy, imagine that the proposer, i, offers k to himself, $1-k$ to player j, and 0 to player h. If j is the randomly chosen amender then he will second the proposal and have it accepted. If instead h is the randomly chosen amender then he will make an amendment in which he proposes k to himself, $1-k$ to player j, and 0 to player i. Because player j is indifferent, we can assume that he will vote for the amendment and cause it to be accepted.

Now we can calculate the symmetric subgame-perfect equilibrium value of k given the equilibrium restrictions, which in this case are a bit trickier than in the guaranteed-success case because there is asymmetry between the two responders at the proposal stage. Define $v(k)$ to be the equilibrium expected payoff of a proposer who offers k to himself, $1-k$ to his coalition partner, and 0 to the third player. Let $v(0)$ be the equilibrium expected payoff of the player who will be offered 0.

Given the selective inclusion strategies, the proposer's offer will be accepted with probability $\frac{1}{2}$, in which case he will get k, and will be rejected with probability $\frac{1}{2}$, in which case there will be a delay of one period and then the proposer will find himself in the shoes of the player who gets offered 0. Hence the equilibrium expected payoff of the proposer is

$$v(k) = \tfrac{1}{2}k + \tfrac{1}{2}\delta v(0). \tag{11.10}$$

Similarly consider the situation of the player who was offered 0. With probability $\frac{1}{2}$ he will not be selected as the amender, in which case he will receive the proposed 0 because the other responder will second the original proposal and it will pass. However, with probability $\frac{1}{2}$ he will be selected, in which case he will find himself in the position of the proposer after the pie is discounted. Hence the equilibrium expected payoff of the player who is offered 0 is

$$v(0) = \tfrac{1}{2}\delta v(k). \tag{11.11}$$

These two restrictions offer us two equations with three unknowns, $v(0)$, $v(k)$, and k. The final restriction is the equilibrium requirement that the player who is proposed the share of $1-k$ will indeed accept it. Since he can put himself immediately in the shoes of the proposer if he is willing to delay the deal by one period, his equilibrium payoff must be at least as large as what he can get from becoming a proposer, hence $1-k \geq \delta v(k)$, and because the proposer will offer as little as possible to get this player to respond, we get the third equilibrium restriction, which is

$$1-k = \delta v(k). \tag{11.12}$$

Solving equations (11.10)–(11.12) yields the solution, which is

$$k = \frac{4-\delta^2}{4+2\delta-\delta^2}$$

and

$$v(k) = \frac{2}{4+2\delta-\delta^2}. \tag{11.13}$$

Recall that because there is a probability of $\frac{1}{2}$ that the proposer's offer is rejected, his equilibrium expected utility is given by $v(k)$ in equation (11.13). By comparing the solution in (11.13) with the solution of the closed-rule case in (11.8), we can show that the original proposer is better off in the closed-rule case.[8] As before, the closed rule offers the proposer more bargaining power.

It is most interesting to compare the risky-success symmetric equilibrium payoff to the guaranteed-success symmetric equilibrium payoff. The risky-success outcome will yield a higher expected payoff than the guaranteed-success outcome if and only if the value in (11.13) is greater than the value in (11.9), that is,

$$\frac{2}{4+2\delta-\delta^2} = \frac{1}{1+2\delta}$$

or

$$\delta > \sqrt{3} - 1.$$

That is, when the discount factor is above the value $\sqrt{3} - 1$, the expected payoff from the risky-success equilibrium is greater than that obtained from the guaranteed-success equilibrium.

This result not only has interesting implications but also has some intuition behind it. When the discount factor is high then players are patient, and it becomes quite costly for the proposer to prevent amendments from both potential amenders. As a consequence the proposer would rather opt to cater to only one of the two potential amenders, at the risk of having an amendment on the table that hurts him.

As for the results implication, if we are willing to take seriously the idea that the equilibrium played by the players will be the one that yields them the highest expected payoff, then the result suggests two interesting features of open-rule legislative bargaining. The first is that sometimes proposals will pass with a large (in the guaranteed-success case, unanimous) coalition, while at other times there will be a minimum coalition. Second, in the event that the minimum coalition is expected, there is some probability that the proposal fails and amendments impose costly delay in equilibrium, because every time a proposal fails the surplus is discounted.

11.5 Summary

- Bargaining situations are common across many social settings. Stylized models of bargaining games can help shed light on the outcomes that are more likely to occur as a consequence of the bargaining protocol.
- A key determinant of how players will share the surplus from an agreement is the discount factor. The higher the discount factor the more patient the players, and the surplus will be shared more equally among them.

8. Subtracting the proposer's risky-success expected equilibrium payoff in (11.13) from his equilibrium payoff in the closed-rule model with three players yields, for all $1 > \delta > 0$,

$$\left(1-\frac{\delta}{3}\right) - \left(\frac{2}{4+2\delta-\delta^2}\right) = \frac{6+2\delta-5\delta^2+\delta^3}{12+6\delta-3\delta^2} > 0.$$

- The bargaining protocol and agreement rules will have a significant impact on how the surplus is shared. In finite games with prespecified alternating moves, the last player to move has a last-mover strategic advantage, while the first player has an advantage owing to discounting.
- When the game is not finite or when the role of proposer is determined randomly with equal probability, there is no last-mover advantage and only the first-mover advantage remains.

11.6 Exercises

11.1 **Disagreement:** Construct a pair of strategies for the ultimatum game ($T = 1$ bargaining game) that constitute a Nash equilibrium and together support the outcome that there is no agreement reached by the two players and the payoffs are zero to each. Show that this disagreement outcome can be supported by a Nash equilibrium regardless of the number of bargaining periods.

11.2 **Holdup:** Consider an ultimatum game ($T = 1$ bargaining game) in which before player 1 makes his offer to player 2, player 2 can invest in the size of the pie. If player 2 chooses a low level of investment (L) then the size of the pie is small, equal to v_L, while if player 2 chooses a high level of investment (H) then the size of the pie is large, equal to v_H. The cost to player 2 of choosing L is c_L, while the cost of choosing H is c_H. Assume that $v_H > v_L > 0$, $c_H > c_L > 0$, and $v_H - c_H > v_L - c_L$.

   a. What is the unique subgame-perfect equilibrium of this game? Is it Pareto optimal?
   b. Can you find a Nash equilibrium of the game that results in an outcome that is better for both players as compared to the unique subgame-perfect equilibrium?

11.3 **Even/Odd Symmetry:** In Section 11.2 we analyzed the alternating-offer bargaining game for a finite number of periods when T was odd. Repeat the analysis for T even.

11.4 **Constant Delay Cost:** Consider a two-player alternating-offer bargaining game in which instead of the pie shrinking by a discount factor $\delta < 1$, the players each pay a cost $c_i > 0$, $i \in \{1, 2\}$, to advance from one period to the next. So if player i receives a share of the pie that gives him a value of x_i in period t then his payoff is $v_i = x_i - (t - 1)c_i$. If the game has T periods then a sequence of rejections results in each player receiving $v_i = -(T - 1)c_i$.

   a. Assume that $T = 2$. Find the subgame-perfect equilibrium of the game and show in which way it depends on the values of c_1 and c_2.
   b. Are there Nash equilibria in the two-period game that are not subgame perfect?
   c. Assume that $T = 3$. Find the subgame-perfect equilibrium of the game and show in which way it depends on the values of c_1 and c_2.

11.5 **Asymmetric Patience 1:** Consider a three-period sequential (alternating-offer) bargaining model in which two players have to split a pie worth 1 (starting with player 1 making the offer). Now the players have different discount factors, δ_1 and δ_2.

   a. Compute the outcome of the unique subgame-perfect equilibrium.

b. Show that when $\delta_1 = \delta_2$ player 1 has an advantage.

c. What conditions on δ_1 and δ_2 give player 2 an advantage? Why?

11.6 **Asymmetric Patience 2:** Consider the analysis of the infinite-horizon bargaining model in Section 11.3 and assume that the players have different discount factors δ_1 and δ_2. Find the unique subgame-perfect equilibrium using the same techniques, and show that as δ_1 and δ_2 become closer in value, the solution you found converges to the solution derived in Section 11.3.

11.7 **Legislative Bargaining Revisited:** Consider a finite T-period version of the Baron and Ferejohn legislative bargaining game with an odd number N of players and with a closed rule as described in Section 11.4.1.

a. Find the unique subgame-perfect equilibrium for $T = 1$. Also find a Nash equilibrium that is not subgame perfect.

b. Find the unique subgame-perfect equilibrium for $T = 2$ with a discount factor $0 < \delta \leq 1$. Also find a Nash equilibrium that is not subgame perfect.

c. Compare what the first period's proposer receives in the subgame-perfect equilibrium you found in (b) to what a first-period proposer receives in the two-period, two-person Rubinstein-Ståhl bargaining game. What intuitively accounts for the difference?

d. Compare the subgame-perfect equilibrium you found in (b) to the solution of the infinite-horizon model in Section 11.4.1. What intuitively accounts for the similarity?