Part 3
applications

Chapter 12
bargaining

12.1 The Basic Bargaining Problem: Splitting a Pie

Part 3 of this book is designed to stretch your muscles by providing more applications of the techniques from parts 1 and 2. The next three chapters may be read in any order. They concern three ways that prices might be determined. Chapter 12 is about bargaining – where both sides exercise market power. Chapter 13 is about auctions – where the seller has market power, but sells a limited amount of a good and wants buyers to compete against each other. Chapter 14 is about fixed-price models with a variety of different features such as differentiated or durable goods. One thing all these chapters have in common is that they use new theory to answer old questions.

Bargaining theory attacks a kind of price determination ill described by standard economic theory. In markets with many participants on one side or the other, standard theory does a good job of explaining prices. In competitive markets we find the intersection of the supply and demand curves, while in markets monopolized on one side we find the monopoly or monopsony output. Where theory is less satisfactory is when there are one or few players on both sides of the market. Early in one's study of economics, one learns that under bilateral monopoly (one buyer and one seller), standard economic theory is inapplicable because the traders must bargain. In the chapters on asymmetric information we would have come across this repeatedly except for our assumption that either the principal or the agent faced competition, which we could model as the other side's ability to make a take-it-or-leave-it offer.

Sections 12.1 and 12.2 introduce the archetypal bargaining problem, Splitting a Pie, ever more complicated versions of which make up the rest of the chapter. Section 12.2, where we take the original rules of the game and apply the Nash bargaining solution, is our one dip into cooperative game theory in this book. Section 12.3 looks at bargaining as a finitely repeated process of offers and counteroffers, and section 12.4 views it as an infinitely repeated process, leading up to the Rubinstein model. Section 12.5 returns to a finite number of repetitions (two, in fact), but with incomplete information. Finally, section 12.6 approaches bargaining from the different angle of the Myerson-Satterthwaite model: how people could try to construct a mechanism for bargaining, a prearranged set of rules that would maximize their expected surplus.

Splitting a Pie

PLAYERS
Smith and Jones.

THE ORDER OF PLAY
The players choose shares θ_s and θ_j of the pie simultaneously.

PAYOFFS
If $\theta_s + \theta_j \leq 1$, each player gets the fraction he chose:

$$\begin{cases} \pi_s = \theta_s, \\ \pi_j = \theta_j. \end{cases} \tag{12.1}$$

If $\theta_s + \theta_j > 1$, then $\pi_s = \pi_j = 0$.

Splitting a Pie resembles the game of Chicken except that it has a continuum of Nash equilibria: any strategy profile (θ_s, θ_j) such that $\theta_s + \theta_j = 1$ is Nash. The Nash concept is at its worst here, because the assumption that the equilibrium being played is common knowledge is very strong when there is a continuum of equilibria. The idea of the focal point (section 1.5) might help to choose a single Nash equilibrium. The strategy space of Chicken is discrete and it has no symmetric pure-strategy equilibrium, but the strategy space of Splitting a Pie is continuous, which permits a symmetric pure-strategy equilibrium to exist. That equilibrium is the even split, (0.5, 0.5), which is a focal point.

If the players moved in sequence, Splitting a Pie becomes what is known as the **Ultimatum Game**, which has a tremendous first-mover advantage. If Jones moves first, the unique Nash outcome would be (0, 1), although only weakly, because Smith would be indifferent as to his action. (This is the same open-set problem that was discussed in section 4.3.) In the unique equilibrium, Smith accepts Jones's offer by choosing $\theta_s = 0$ so that $\theta_s + \theta_j = 1$. Of course, if we add to the model even a small amount of ill will by Smith against Jones for making such a selfish offer, Smith would pick $\theta_s > 0$ and reject the offer. That is quite realistic, so depending on the amount of ill will, the equilibrium would have Jones making a more generous offer that depends on Smith's utility tradeoff between getting a share of the pie on the one hand and seeing Jones suffer on the other.

In many applications, this version of Splitting a Pie is unacceptably simple, because if the two players find their fractions add to more than 1 they have a chance to change their minds. In labor negotiations, for example, if manager Jones makes an offer which union Smith rejects, they do not immediately forfeit the gains from combining capital and labor. They lose a week's production and make new offers. We will model just such a sequence of offers, but before we do that let us see how cooperative game theory deals with the original game.

12.2 The Nash Bargaining Solution

A quite different approach to game theory than we have been using in this book is to describe the players and payoff functions for a game, decide upon some characteristics an equilibrium should have based on notions of fairness or efficiency, mathematicize the characteristics, and maybe add a few other axioms to make the equilibrium turn out neatly. This is a reduced-form approach, attractive if the modeller finds it difficult to come up with a convincing order of play but thinks he can say something about what outcome will appear. Nash (1950a) did this for the bargaining problem in what is the best-known application of cooperative game theory. Nash's objective was to pick axioms that would characterize the agreement the two players would anticipate making with each other. He used a game only a little more complicated than Splitting a Pie. In the Nash model, the two players can have different utilities if they do not come to an agreement, and the utility functions can be nonlinear in terms of shares of the pie. Figures 12.1a and b compare the two games.

In figure 12.1, the shaded region denoted by X is the set of feasible payoffs, which we will assume to be convex. The pair of disagreement payoffs or **threat point** is $\overline{U} = (\overline{U}_s, \overline{U}_j)$. The Nash bargaining solution, $U^* = (U_s^*, U_j^*)$, is a function of $\overline{U}$ and X that satisfies the following four axioms.

1 *Invariance*: For any strictly increasing linear function F,

$$U^*[F(\overline{U}), F(X)] = F[U^*(\overline{U}, X)]. \tag{12.2}$$

This says that the solution is independent of the units in which utility is measured.

2 *Efficiency*: The solution is Pareto optimal, so the players cannot both be made better off by any change. In mathematical terms,

$$(U_s, U_j) > U^* \Rightarrow (U_s, U_j) \notin X. \tag{12.3}$$

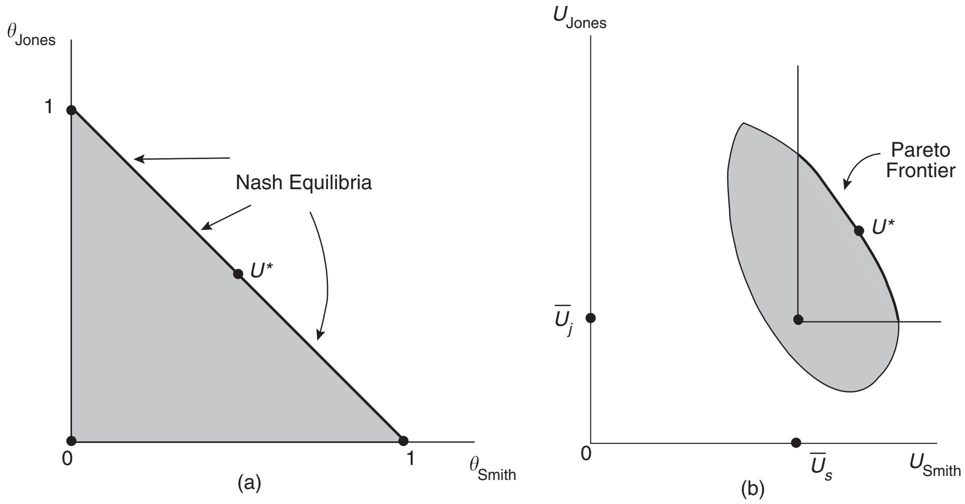

Figure 12.1 (a) Nash Bargaining Game (b) Splitting a Pie.

3 *Independence of irrelevant alternatives*: If we drop some possible utility profiles from X, leaving the smaller set Y, then if U^* was not one of the dropped points, U^* does not change.

$$U^*(\overline{U}, X) \in Y \subseteq X \Rightarrow U^*(\overline{U}, Y) = U^*(\overline{U}, X). \tag{12.4}$$

4 *Anonymity (or symmetry)*: Switching the labels on players Smith and Jones does not affect the solution.

The axiom of Independence of Irrelevant Alternatives is the most debated of the four, but if I were to complain, it would be about the axiomatic approach itself, which depends heavily on the intuition behind the axioms. Everyday intuition says that the outcome should be efficient and symmetric, so that other outcomes can be ruled out a priori. But most of the games in the earlier chapters of this book turn out to have reasonable but inefficient outcomes, and games like Chicken have reasonable asymmetric outcomes.

Whatever their drawbacks, these axioms fully characterize the Nash solution. It can be proven that if U^* satisfies the four axioms above, then it is the unique strategy profile such that

$$U^* = \underset{U \in X,\ U \geq \overline{U}}{argmax}\ (U_s - \overline{U}_s)(U_j - \overline{U}_j). \tag{12.5}$$

Splitting a Pie is a simple enough game that not all the axioms are needed to generate a solution. If we put the game in this context, however, problem (12.5) becomes

$$\underset{\theta_s,\ \theta_j}{Maximize}\ (\theta_s - 0)(\theta_j - 0), \tag{12.6}$$

subject to $\theta_s + \theta_j \leq 1$, which generates the first-order conditions

$$\theta_s - \lambda = 0, \quad \text{and} \quad \theta_j - \lambda = 0, \tag{12.7}$$

where λ is the Lagrange multiplier on the constraint. From (12.6) and the constraint, we obtain $\theta_s = \theta_j = 1/2$, the even split that we found as a focal point of the noncooperative game.

Although Nash's objective was simply to characterize the anticipations of the players, I perceive a heavier note of morality in cooperative than in noncooperative game theory. Cooperative outcomes are neat, fair, beautiful, and efficient. In the next few sections we will look at noncooperative bargaining models that while plausible, lack every one of those features. Cooperative game theory may be useful for ethical decisions, but its attractive features are often inappropriate for economic situations, and the spirit of the axiomatic approach is very different from the utility maximization of economic theory.

It should be kept in mind, however, that the ethical component of cooperative game theory can also be realistic, because people are often ethical, or pretend to be. People very often follow rules that they believe represent virtuous behavior, even at some monetary cost. In bargaining experiments in which one player is given the ability to make a take-it-or-leave it offer (the Ultimatum Game) it is commonly found that he offers a 50–50 split. Presumably this is because either he wishes to be fair or he fears a spiteful response from the other player to a smaller offer. If the subjects are made to feel that they have "earned" the right to be the offering party, they behave much more like the players in noncooperative game theory

(Hoffman & Spitzer [1985]). Frank (1988) and Thaler (1992) describe numerous occasions where simple games fail to describe real-world or experimental results. People's payoffs include more than their monetary rewards, and sometimes knowing the cultural disutility of actions is more important than knowing the dollar rewards. This is one reason why it is helpful to a modeller to keep his games simple: when he actually applies them to the real world, the model must not be so unwieldy that it cannot be combined with details of the particular setting.

12.3 Alternating Offers over Finite Time

In the games of the next two sections, the actions are the same as in Splitting a Pie, but with many periods of offers and counteroffers. This means that strategies are no longer just actions, but rather are rules for choosing actions based on the actions chosen in earlier periods.

Alternating Offers

PLAYERS
Smith and Jones.

THE ORDER OF PLAY

1 Smith makes an offer θ_1.
1* Jones accepts or rejects.
2 Jones makes an offer θ_2.
2* Smith accepts or rejects.
...
T Smith offers θ_T.
T* Jones accepts or rejects.

PAYOFFS
The discount factor is $\delta \leq 1$.
If Smith's offer is accepted by Jones in round m,

$$\pi_s = \delta^m \theta_m,$$
$$\pi_j = \delta^m (1 - \theta_m).$$

If Jones's offer is accepted, reverse the subscripts.
If no offer is ever accepted, both payoffs equal zero.

When a game has many rounds we need to decide whether discounting is appropriate. If the discount rate is r then the discount factor is $\delta = 1/(1+r)$, so, without discounting, $r = 0$ and $\delta = 1$. Whether discounting is appropriate to the situation being modelled depends on whether delay should matter to the payoffs because the bargaining occurs over real time or the game might suddenly end (section 5.2). The game Alternating Offers can be interpreted in either of two ways, depending on whether it occurs over real time or not. If the players made all the offers and counteroffers between dawn and dusk of a single day, discounting would be inconsequential because, essentially, no time has passed. If each offer consumed a week of time, on the other hand, the delay before the pie was finally consumed would be important to the players and their payoffs should be discounted.

Consider first the game without discounting. There is a unique subgame-perfect outcome – Smith gets the entire pie – which is supported by a number of different equilibria. In each equilibrium, Smith offers $\theta_s = 1$ in each period, but each equilibrium is different in terms of when Jones accepts the offer. All of them are weak equilibria because Jones is indifferent between accepting and rejecting, and they differ only in the timing of Jones's final acceptance.

Smith owes his success to his ability to make the last offer. When Smith claims the entire pie in the last period, Jones gains nothing by refusing to accept. What we have here is not really a first-mover advantage, but a last-mover advantage, a difference not apparent in the one-period model.

In the game with discounting, the total value of the pie is 1 in the first period, δ in the second, and so forth. In period T, if it is reached, Smith would offer 0 to Jones, keeping 1 for himself, and Jones would accept under our assumption on indifferent players. In period $(T - 1)$, Jones could offer Smith δ, keeping $(1 - \delta)$ for himself, and Smith would accept, although he could receive a greater share by refusing, because that greater share would arrive later and be discounted.

By the same token, in period $(T - 2)$, Smith would offer Jones $\delta(1 - \delta)$, keeping $(1 - \delta(1 - \delta))$ for himself, and Jones would accept, since with a positive share Jones also prefers the game to end soon. In period $(T - 3)$, Jones would offer Smith $\delta[1 - \delta(1 - \delta)]$, keeping $(1 - \delta[1 - \delta(1 - \delta)])$ for himself, and Smith would accept, again to prevent delay. Table 12.1 shows the progression of Smith's shares when $\delta = 0.9$.

As we work back from the end, Smith always does a little better when he makes the offer than when Jones does, but if we consider just the class of periods in which Smith makes the offer, Smith's share falls. If we were to continue to work back for a large number of periods, Smith's offer in a period in which he makes the offer would approach $1/(1 + \delta)$, which equals about 0.53 if $\delta = 0.9$. The reasoning behind that precise expression is given in the next section. In equilibrium, the very first offer would be accepted, since it is chosen precisely so that the other player can do no better by waiting.

12.4 Alternating Offers over Infinite Time

The Folk Theorem of section 5.2 says that when discounting is low and a game is repeated an infinite number of times, there are many equilibrium outcomes. That does not apply

Table 12.1 Alternating offers over finite time

Round	*Smith's share*	*Jones's share*	*Total value*	*Who offers?*
$T - 3$	0.819	0.181	0.9^{T-4}	Jones
$T - 2$	0.91	0.09	0.9^{T-3}	Smith
$T - 1$	0.9	0.1	0.9^{T-2}	Jones
T	1	0	0.9^{T-1}	Smith

to the bargaining game, however, because it is not a repeated game. It ends when one player accepts an offer, and only the accepted offer is relevant to the payoffs, not the earlier proposals. In particular, there are no out-of-equilibrium punishments such as enforce the Folk Theorem's outcomes.

Let players Smith and Jones have discount factors of δ_s and δ_j which are not necessarily equal but are strictly positive and no greater than one. In the unique subgame-perfect outcome for the infinite-period bargaining game, Smith's share is

$$\theta_s = \frac{1-\delta_j}{1-\delta_s\delta_j}, \tag{12.8}$$

which, if $\delta_s = \delta_j = \delta$, is equivalent to

$$\theta_s = \frac{1}{1+\delta}. \tag{12.9}$$

If the discount rate is high, Smith gets most of the pie: a 1,000 percent discount rate ($r = 10$) makes $\delta = 0.091$ and $\theta_s = 0.92$ (rounded), which makes sense, since under such extreme discounting the second period hardly matters and we are almost back to the simple game of section 12.1. At the other extreme, if r is small, the pie is split almost evenly: if $r = 0.01$, then $\delta \approx 0.99$ and $\theta_s \approx 0.503$.

It is crucial that the discount rate be strictly greater than 0, even if only by a little. Otherwise, the game has the same continuum of perfect equilibria as in section 12.1. Since nothing changes over time, there is no incentive to come to an early agreement. When discount rates are equal, the intuition behind the result is that since a player's cost of delay is proportional to his share of the pie, if Smith were to offer a grossly unequal split, such as (0.7, 0.3), Jones, with less to lose by delay, would reject the offer. Only if the split is close to even would Jones accept, as we will now prove.

Proposition 12.1 (Rubinstein [1982])[1]

In the discounted infinite game, the unique perfect equilibrium outcome is $\theta_s = (1-\delta_j)/(1-\delta_s\delta_j)$, *where Smith is the first mover.*

Proof. We found that in the T-period game Smith gets a larger share in a period in which he makes the offer. Denote by M the maximum nondiscounted share, taken over all the perfect equilibria that might exist, that Smith can obtain in a period in which he makes the offer. Consider the game starting at t. Smith is sure to get no more than M, as noted in table 12.2. (Jones would thus get $(1-M)$, but that is not relevant to the proof.)

The trick is to find a way besides M to represent the maximum Smith can obtain. Consider the offer made by Jones at $(t-1)$. Smith will accept any offer which gives him more than the discounted value of M received one period later, so Jones can make an offer

[1] The proof of proposition 12.1 is not from the original Rubinstein (1982), but is adapted from Shaked & Sutton (1984). The maximum rather than the supremum can be used because of the assumption that indifferent players always accept offers.

Table 12.2 Alternating offers over infinite time

Round	*Smith's share*	*Jones's share*	*Who offers?*
$t-2$	$1-\delta_j(1-\delta_s M)$		Smith
$t-1$		$1-\delta_s M$	Jones
t	M		Smith

of $\delta_s M$ to Smith, retaining $(1-\delta_s M)$ for himself. At $(t-2)$, Smith knows that Jones will turn down any offer less than the discounted value of the minimum Jones can look forward to receiving at $(t-1)$. Smith, therefore, cannot offer any less than $\delta_j(1-\delta_s M)$ at $(t-2)$.

Now we have two expressions for "the maximum which Smith can receive," which we can set equal to each other:

$$M = 1 - \delta_j(1-\delta_s M). \tag{12.10}$$

Solving equation (12.10) for M, we obtain

$$M = \frac{1-\delta_j}{1-\delta_s\delta_j}. \tag{12.11}$$

We can repeat the argument using m, the minimum of Smith's share. If Smith can expect at least m at t, Jones cannot receive more than $(1-\delta_s m)$ at $(t-1)$. At $(t-2)$ Smith knows that if he offers Jones the discounted value of that amount, Jones will accept, so Smith can guarantee himself $(1-\delta_j(1-\delta_s m))$, which is the same as the expression we found for M. The smallest perfect equilibrium share that Smith can receive is the same as the largest, so the equilibrium outcome must be unique. QED.

This model from Rubinstein (1982) is widely used because of the way it explains why two players with the same discount rates tend to split the surplus equally (the limiting case in the model as the discount rate goes to zero and δ goes to one). Unfortunately, as with the Nash bargaining solution, there is no obvious best way to extend the model to three or more players – no best way to specify how they make and accept offers. Haller (1986) shows that for at least one specification, the outcome is not similar to the Rubinstein (1982) outcome, but rather is a return to the indeterminacy of the game without discounting.

No Discounting, but a Fixed Bargaining Cost

There are two ways to model bargaining costs per period: as proportional to the remaining value of the pie (the way used above), or as fixed costs each period, which we analyze next (again following Rubinstein [1982]). To understand the difference, think of labor negotiations during a construction project. If a strike slows down completion, there are two kinds of losses. One is the loss from delay in renting or selling the new building, a loss proportional to its value. The other is the loss from late-completion penalties in the contract, which often take the form of a fixed penalty each week. The two kinds of costs have very different effects on the bargaining process.

Here, let us assume that there is no discounting but whenever a period passes, Smith loses c_s and Jones loses c_j. In every subgame-perfect equilibrium, Smith makes an offer and Jones accepts, but there are three possible cases.

Delay costs are equal

$c_s = c_j = c.$

The Nash indeterminacy of section 12.1 remains almost as bad; any fraction such that each player gets at least c is supported by some perfect equilibrium.

Delay hurts Jones more

$c_s < c_j.$

Smith gets the entire pie. Jones has more to lose than Smith by delaying, and delay does not change the situation except by diminishing the wealth of the players. The game is stationary, because it looks the same to both players no matter how many periods have already elapsed. If in any period t Jones offered Smith x, in period $(t-1)$ Smith could offer Jones $(1-x-c_j)$, keeping $(x+c_j)$ for himself. In period $(t-2)$, Jones would offer Smith $(x+c_j-c_s)$, keeping $(1-x-c_j+c_s)$ for himself, and in periods $(t-4)$ and $(t-6)$ Jones would offer $(1-x-2c_j+2c_s)$ and $(1-x-3c_j+3c_s)$. As we work backwards, Smith's advantage rises to $\gamma(c_j-c_s)$ for an arbitrarily large integer γ. Looking ahead from the start of the game, Jones is willing to give up and accept zero.

Delay hurts Smith more

$c_s > c_j.$

Smith gets a share worth c_j and Jones gets $(1-c_j)$. The cost c_j is a lower bound on the share of Smith, the first mover, because if Smith knows Jones will offer (0, 1) in the second period, Smith can offer $(c_j, 1-c_j)$ in the first period and Jones will accept.

12.5 Incomplete Information

Instant agreement has characterized even the multiperiod games of complete information discussed so far. Under incomplete information, knowledge can change over the course of the game and bargaining can last more than one period in equilibrium, a result that might be called inefficient but is certainly realistic. Models with complete information have difficulty explaining such things as strikes or wars, but if over time an uninformed player can learn the type of the informed player by observing what offers are made or rejected, such unfortunate outcomes can arise. The literature on bargaining under incomplete information is vast. For this section, I have chosen to use a model based on the first part of Fudenberg & Tirole (1983), but it is only a particular example of how one could construct such a model, and not a good indicator of what results are to be expected from bargaining.

Let us start with a one-period game. We will denote the price by p_1 because we will carry the notation over to a two-period version.

One-period Bargaining with Incomplete Information

PLAYERS
A seller, and a buyer called Buyer_{100} or Buyer_{150} depending on his type.

THE ORDER OF PLAY
0 Nature picks the buyer's type, his valuation of the object being sold, which is $b = 100$ with probability γ and $b = 150$ with probability $(1 - \gamma)$.
1 The seller offers price p_1.
2 The buyer accepts or rejects p_1.

PAYOFFS
The seller's payoff is p_1 if the buyer accepts the offer, and otherwise 0.
The buyer's payoff is $(b - p_1)$ if he accepts the offer, and otherwise 0.

Equilibrium:
Buyer_{100}: accept if $p_1 \leq 100$.
Buyer_{150}: accept if $p_1 \leq 150$.
Seller: offer $p_1 = 100$ if $\gamma \geq 1/3$ and $p_1 = 150$ otherwise.

Both types of buyers have a dominant strategy for the last move: accept any offer $p_1 < b$. Accepting any offer $p_1 \leq b$ is a weakly best response to the seller's equilibrium strategy. No equilibrium exists in which a buyer rejects an offer of $p_1 = b$, because we would fall into the open-set problem: there would be no greatest offer in $[0, b)$ that the buyer would accept, and so we could not find a best response for the seller.

The only two strategies that might be optimal for the seller are $p_1 = 100$ or $p_1 = 150$, since prices lower than 100 would lead to a sale with the same probability as $p_1 = 100$, prices in $(100, 150]$ would have the same probability as $p_1 = 150$, and prices greater than 150 would yield zero profits. The seller will choose $p_1 = 150$ if it yields a higher payoff than $p_1 = 100$; that is, if

$$\pi(p_1 = 100) = \gamma(100) + (1 - \gamma)(100) < \pi(p_1 = 150) = \gamma(0) + (1 - \gamma)(150), \tag{12.12}$$

which requires that

$$\gamma < 1/3. \tag{12.13}$$

Thus, if less than a third of buyers have a valuation of 100, the seller will charge 150, gambling that he is not facing such a buyer.

This means, of course, that if $\gamma < 1/3$, sometimes no sale will be made. This is the most interesting feature of the model. By introducing incomplete information into a bargaining model, we have explained why bargaining sometimes breaks down and efficient trades fail to be carried out. This suggests that when wars occur because nations cannot agree, or strikes occur because unions and employers cannot agree, we should look to information asymmetry for an explanation.

This has some similarity to a mechanism design problem. It is crucial that the seller commit to make only one offer. Once the offer $p_1 = 150$ is made and rejected, the seller realizes that $b = 100$. At that point, he would like to make a second offer, of $p_1 = 100$. But of course if he could do that, then rejection of the first offer would not convey the information that $b = 100$.

Now let us move to a two-period version of the same game. This will get quite a bit more complex, so let us restrict ourselves to the case of $\gamma = 1/6$. Also, we will need to make an assumption on discounting – the loss that results from a delay in agreement. Let us assume that each player loses a fixed amount $D = 4$ if there is no agreement in the first period. (As a result, a player can end up with a negative payoff by playing this game, something experienced bargainers will find realistic.)

Two-period Bargaining with Incomplete Information

PLAYERS
A seller, and a buyer called Buyer_{100} or Buyer_{150} depending on his type.

THE ORDER OF PLAY
0 Nature picks the buyer's type, his valuation of the object being sold, which is $b = 100$ with probability $1/6$ and $b = 150$ with probability $5/6$.
1 The seller offers price p_1.
2 The buyer accepts or rejects p_1.
3 The seller offers price p_2.
4 The buyer accepts or rejects p_2.

PAYOFFS
The seller's payoff is p_1 if the buyer accepts the first offer, $(p_2 - 4)$ if he accepts the second offer, and -4 if he accepts no offer.
The buyer's payoff is $(b - p_1)$ if he accepts the first offer, $(b - p_2 - 4)$ if he accepts the second offer, and -4 if he accepts no offer.

Equilibrium Behavior (separating, in mixed strategies)

Buyer_{100}: Accept if $p_1 \leq 104$. Accept if $p_2 \leq 100$.
Buyer_{150}: Accept if $p_1 < 154$. Accept with probabilty $\theta \leq 0.6$ if $p_1 = 154$.
Accept if $p_2 \leq 150$.
Seller: Offer $p_1 = 154$ and $p_2 = 150$.

Buyer_{100}'s strategy

If Buyer_{100} deviates and rejects an offer of p_1 less than 104, his payoff will be -4, which is worse than $(100 - p_1)$. Rejecting p_2 does not result in any extra transactions cost, so he rejects any $p_2 < 100$.

Buyer$_{150}$'s strategy

Buyer$_{150}$'s equilibrium payoff is either

$$\pi_{Buyer_{150}}(Accept\ p_1 = 154) = b - 154 = -4, \tag{12.14}$$

or

$$\pi_{Buyer_{150}}(Reject\ p_1 = 154) = -4 + \theta(b - 150) = -4 \tag{12.15}$$

Thus, Buyer$_{150}$ is indifferent and is willing to mix in the first period. Or, out of equilibrium, if $p_1 < 154$ he can achieve a higher payoff by immediately accepting it. In the second period, the game is just like the one-period game, so he will accept any offer of $p_2 \leq 150$.

The seller's strategy

To check on whether the seller has any incentive to deviate, let us work back from the end. If the game has reached the second period, he knows that the fraction of Buyer$_{100}$'s has increased, since there was some probability that a Buyer$_{150}$ would have accepted $p_1 = 150$. The prior probability was $Prob(Buyer_{100}) = 1/6$, but the posterior is

$$\begin{aligned} &\text{Prob}(\text{Buyer}_{100}|\text{Rejected } p_1 = 154)\\ &\quad = \frac{\text{Prob}(\text{Rejected } p_1 = 154|\text{Buyer}_{100})\text{Prob}(\text{Buyer}_{100})}{\text{Prob}(\text{Rejected } p_1 = 154)}\\ &\quad = \frac{\text{Prob}(\text{Rejected }|100)\text{Prob}(\text{Buyer}_{100})}{\text{Prob}(\text{Rej}|100)\text{Prob}(100) + \text{Prob}(\text{Rej}|150)\text{Prob}(150)}\\ &\quad = \frac{(1)(1/6)}{(1)(\theta) + (1-\theta)(5/6)}\\ &\quad =\leq \frac{1}{3} \text{ if } \theta \leq 0.6. \end{aligned} \tag{12.16}$$

In equilibrium the seller expects the Prob(Buyer$_{100}$) proportion in the second period to be no more than 1/3. If he chooses a price $p_2 > 150$ he will sell to nobody in the second period. If he chooses $p_2 \in (100, 150]$ he will sell only to high-valuing buyers. If he chooses $p_2 \leq 100$ he will sell to both types of buyers. This narrows drawn the possibly optimal prices to $p_2 = 150$ versus $p_2 = 100$. The payoffs from each, as viewed at the start of period 2, are

$$\pi_{seller}/p_2 = 100 - 4 \tag{12.17}$$

and

$$\pi_{seller}(p_2 = 150) = Prob(Buyer_{150}/Rejected\ p_1 = 154) * 150 - 4 \tag{12.18}$$

$$= \left[1 - \left(\frac{(1)(1/6)}{(1)(\theta) + (1-\theta)(5/6)}\right)\right] * 150 - 4. \tag{12.19}$$

If $\theta \leq 0.6$, then $p_2 = 150$ yields the higher payoff.

Thus, neither player has incentive to deviate from the proposed equilibrium.

The most important lesson of this model is that bargaining can lead to inefficiency. Some of the Buyer_{150}s delay their transactions until the second period, which is inefficient since the payoffs are discounted. Moreover, there is a positive probability that the Buyer_{100}s never buy at all, as in the one-period game, and the potential gains from trade are lost.

Note, too, that this is a model in which prices fall over time as bargaining proceeds. The first-period price is definitely $p_1 = 154$, but the second-period price might fall to $p_1 = 150$. This can happen because high-valuation buyers know that though the price might fall if they wait, it might not and they would just incur an extra delay cost. This result has close parallels to the durable-goods monopoly pricing problem that will be discussed in chapter 14.

The price the buyer pays depends heavily on the seller's equilibrium beliefs. If the seller thinks that the buyer has a high valuation with probability 0.5, the price is 100, but if he thinks the probability is 0.05, the price rises to 150. This implies that a buyer is unfortunate if he is part of a group which is believed to have high valuations more often. Even if his own valuation is low, what we might call his bargaining power is low when he is part of a high-valuing group. Ayres (1991) found that when he hired testers to pose as customers at car dealerships, their success depended on their race and gender even though they were given identical predetermined bargaining strategies to follow. Since the testers did as badly even when faced with salesmen of their own race and gender, it seems likely that they were hurt by being members of groups that usually can be induced to pay higher prices rather than out of animus to the group itself.

*12.6 Setting Up a Way to Bargain: The Myerson–Satterthwaite Model

Let us now think about a different way to approach bargaining under incomplete information. We will stay with noncooperative game theory, but now let us ask what would happen under different sets of formalized rules – different mechanisms.

We have seen in section 12.5 that under incomplete information, inefficiency can easily arise in bargaining. This inefficiency varies depending on the rules of the game. Thus, if feasible, the players might like to bind themselves in advance to follow whichever rules are best at avoiding inefficiency – at least as long as they can share the efficiency gains.

Suppose a group of players in a game are interacting in some way. They would like to set up some rules for their interaction in advance that would make the best use of the information they will later have, and this set of rules is what we call a mechanism, the topic of chapter 10. Usually, models analyze different mechanisms without asking how the players would agree upon them, taking that as exogenous to the model. This is reasonable – the mechanism may be assigned by history as an institution of the market. If it is not, then there is bargaining over which mechanism to use, an extra layer of complexity.

Let us consider the situation of two people trying to exchange a good under various mechanisms. The mechanism must do two things:

1 Tell under what circumstances the good should be transferred from seller to buyer; and
2 Tell the price at which the good should be transferred, if it is transferred at all.

Usually these two things are made to depend on **reports** of the two players – that is, on statements they make.

The first mechanisms we will look at are simple.

Bilateral Trading I: Complete Information

PLAYERS
A buyer and a seller.

THE ORDER OF PLAY

0 Nature independently chooses the seller to value the good at v_s and the buyer at v_b using the uniform distribution between 0 and 1. Both players observe these values.
1 The seller reports (v_s^s, v_b^s) and the buyer reports (v_s^b, v_b^b) as their observations of (v_b, v_s), simultaneously.
2 If $(v_s^s, v_b^s) = (v_s^b, v_b^b)$ then the seller keeps the good if $(v_s^s > v_b^s)$ and otherwise the buyer acquires it for a payment to the seller of

$$p = v_s^s + \frac{v_b^s - v_s^s}{2}. \tag{12.20}$$

If $(v_s^s, v_b^s) \neq (v_s^b, v_b^b)$, the good is destroyed and the buyer pays v_b^s to the court.

PAYOFFS
If the seller keeps the good, both players have payoffs of 0. If the buyer acquires the object, the seller's payoff is $(p - v_s)$ and the buyer's is $(v_b - p)$. If the reports disagree, the seller's payoff is $-v_s$ and the buyer's payoff is $-v_b^s$.

I have normalized the payoffs so that each player's payoff is zero if no trade occurs. I could instead have normalized to $\pi_s = v_s$ and to $\pi_b = 0$ if no trade occurred, a common alternative.

This is one of chapter 10's cross checking mechanisms. One equilibrium is for buyer and seller to both tell the truth. That is an equilibrium because if the buyer acquires the object, the payoffs are $\pi_s = (v_b^s - v_s^s)/2$ and $\pi_b = (v_b^s - v_s^s)/2$, both of which are positive if and only if $v_b > v_s$. This will result in the efficient allocation, in the sense that the good ends up with the player who values it most highly. This is, moreover an acceptable mechanism for both players if they expect this equilibrium to be played out, because they share any gains from trade that may exist.[2]

I include Bilateral Trading I to introduce the situation and provide a first-best benchmark, as well as to give another illustration of cross checking. Let us next look at a game of incomplete information and a mechanism which does depend on the players' actions.

Bilateral Trading II: Incomplete Information

PLAYERS
A buyer and a seller.

THE ORDER OF PLAY
0 Nature independently chooses the seller to value the good at v_s and the buyer at v_b using the uniform distribution between 0 and 1. Each player's value is his own private information.
1 The seller reports p_s and the buyer reports p_b.
2 The buyer accepts or rejects the seller's offer. The price at which the trade takes place, if it does, is p_s.

PAYOFFS
If there is no trade, the seller's payoff is 0 and the buyer's is 0.
If there is trade, the seller's payoff is $(p_s - v_s)$ and the buyer's is $(v_b - p_s)$.

This mechanism does not use the buyer's report at all, and so perhaps it is not surprising that the result is inefficient. It is easy to see, working back from the end of the game, that the buyer's equilibrium strategy is to accept the offer if $v_b \geq p_s$ and to reject it otherwise. If the buyer does that, the seller's expected payoff is

$$[p_s - v_s][Prob\{v_b \geq p_s\}] + 0[Prob\{v_b \leq p_s\}] = [p_s - v_s][1 - p_s]. \quad (12.21)$$

Differentiating this with respect to p_s and setting equal to zero yields the seller's equilibrium strategy of

$$p_s = \frac{1 + v_s}{2}. \quad (12.22)$$

This is inefficient because if v_b is just a little bigger than v_s, trade will not occur even though gains from trade do exist that is, even though $v_b > v_s$. In fact, trade will fail to occur whenever $v_b < (1 + v_s)/2$.

Let us try another simple mechanism, which at least uses the reports of both players, replacing move (2) with (2′).

[2] As usual, the efficient equilibrium is not unique. Another equilibrium would be for both players to always report $v_s = 0.5, v_b = 0.4$, which would yield zero payoffs and never result in trade. If either player unilaterally deviated, the punishment would kick in and payoffs would become negative.

(2′) The good is allocated to the seller if $p_s > p_b$ and to the buyer otherwise. The price at which the trade takes place, if it does, is p_s.

Suppose the buyer truthfully reports $p_b = v_b$. What will the seller's best response be? The seller's expected payoff for the p_s he chooses is now

$$[p_s - v_s][Prob\{p_b(v_b) \geq p_s\}] + 0[Prob\{p_b(v_b) \leq p_s\}] = [p_s - v_s][1 - p_s]. \tag{12.23}$$

where the expectation has to be taken over all the possible values of v_b, since p_b will vary with v_b.

Maximizing this, the seller's strategy will solve the first-order condition $1 - 2p_s + v_s = 0$, and so will again be

$$p_s(v_s) = \frac{1 + v_s}{2} = \frac{1}{2} + \frac{v_s}{2}. \tag{12.24}$$

Will the buyer's best response to this strategy be $p_b = v_b$? Yes, because whenever $v_b \geq 1/2 + v_s/2$ the buyer is willing for trade to occur, and the size of p_b does not affect the transactions price, only the occurrence or nonoccurrence of trade. The buyer needs to worry about causing trade to occur when $v_b < 1/2 + v_s/2$, but this can be avoided by using the truthtelling strategy. The buyer also needs to worry about preventing trade from occurring when $v_b > 1/2 + v_s/2$, but choosing $p_b = v_b$ prevents this from happening either.

Thus, it seems that either mechanism (2) or (2′) will fail to be efficient. Often, the seller will value the good less than the buyer, but trade will fail to occur and the seller will end up with the good anyway – whenever $v_b > (1 + v_s)/2$. Figure 12.2 shows when trades will be completed based on the parameter values.

As you might imagine, one reason this is an inefficient mechanism is that it fails to make effective use of the buyer's information. The next mechanism will do better. Its trading rule is called the **double auction mechanism**. The problem is like that of chapter 10's Groves Mechanism, because we are trying to come up with an action rule (allocate the object to the buyer or to the seller) based on the agents' reports (the prices they suggest), under the condition that each player has private information (his value).

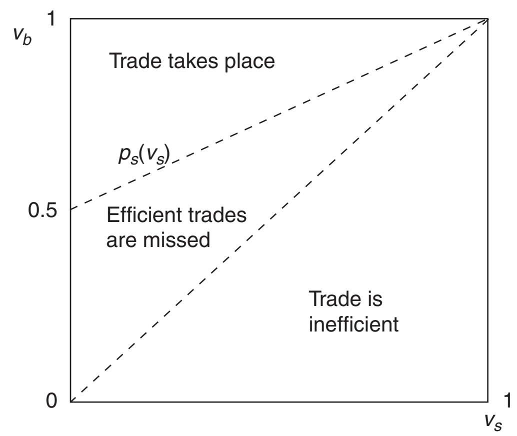

Figure 12.2 Trades in Bilateral Trading II.

Bilateral Trading III: The Double Auction Mechanism

PLAYERS
A buyer and a seller.

THE ORDER OF PLAY
0 Nature independently chooses the seller to value the good at v_s and the buyer at v_b using the uniform distribution between 0 and 1. Each player's value is his own private information.
1 The buyer and the seller simultaneously decide whether to try to trade or not.
2 If both agree to try, the seller reports p_s and the buyer reports p_b simultaneously.
3 The good is allocated to the seller if $p_s \geq p_b$ and to the buyer otherwise. The price at which the trade takes place, if it does, is $p = (p_b + p_s)/2$.

PAYOFFS
If there is no trade, the seller's payoff is 0 and the buyer's is zero. If there is trade, then the seller's payoff is $(p - v_s)$ and the buyer's is $(v_b - p)$.
The buyer's expected payoff for the p_b he chooses is

$$\left[v_b - \frac{p_b + E[p_s|p_b \geq p_s]}{2}\right][Prob\{p_b \geq p_s\}], \tag{12.25}$$

where the expectation has to be taken over all the possible values of v_s, since p_s will vary with v_s.
The seller's expected payoff for the p_s he chooses is

$$\left[\frac{p_s + E(p_b|p_b \geq p_s)}{2} - v_s\right][Prob\{p_b \geq p_s\}], \tag{12.26}$$

where the expectation has to be taken over all the possible values of v_b, since p_b will vary with v_b.

The game has lots of Nash equilibria. Let's focus on two of them, a **one-price equilibrium** and the unique **linear equilibrium**.

In the **one-price equilibrium**, the buyer's strategy is to offer $p_b = x$ if $v_b \geq x$ and $p_b = 0$ otherwise, for some value $x \in [0, 1]$. The seller's strategy is to ask $p_s = x$ if $v_s \leq x$ and $p_s = 1$ otherwise. Figure 12.3 illustrates the one-price equilibrium for a particular value of x. Efficient trade occurs in the shaded region, but is missed in regions A and B. Suppose $x = 0.7$. If the seller were to deviate and ask prices lower than 0.7, he would just reduce the price he receives. If the seller were to deviate and ask prices higher than 0.7, then $p_s > p_b$ and no trade occurs. So the seller will not deviate. Similar reasoning applies to the buyer, and to any value of x, including 0 and 1 (where trade never occurs).

The **linear equilibrium** can be derived very neatly. Suppose the seller uses a linear strategy, so $p_s(v_s) = \alpha_s + c_s v_s$. From the buyer's point of view, p_s will be uniformly distributed from $\alpha_s = 1/4$ to $(\alpha_s + c_s) = 1/12$ with density $1/c_s$, as v_s ranges from 0 to 1. Since $E_b[p_s|p_b \geq p_s] = E_b(p_s|p_s \in [a_s, p_b]) = (a_s + p_b)/2$, the buyer's expected payoff

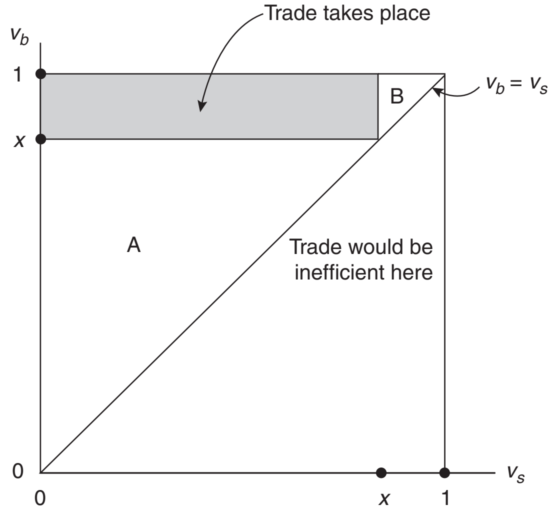

Figure 12.3 Trade in the one-price equilibrium.

(12.25) becomes

$$\left[v_b - \frac{p_b + \frac{\alpha_s + p_b}{2}}{2}\right]\left[\frac{p_b - \alpha_s}{c_s}\right]. \tag{12.27}$$

Maximizing with respect to p_b yields

$$p_b = \left(\frac{2}{3}\right) v_b + \left(\frac{1}{3}\right) \alpha_s. \tag{12.28}$$

Thus, if the seller uses a linear strategy, the buyer's best response is a linear strategy too! We are well on our way to a Nash equilibrium.

If the buyer uses a linear strategy $p_b(v_b) = \alpha_b + c_b v_b$, then from the seller's point of view p_b is uniformly distributed from α_b to $\alpha_b + c_b$ with density $1/c_b$ and the seller's payoff function, expression (12.26), becomes, since $E_s(p_b|p_b \geq p_s) = E_s(p_b|p_b \in [p_s, \alpha_b + c_b] = (p_s + \alpha_b + c_b)/2$,

$$\left[\frac{p_s + \frac{p_s + \alpha_b + c_b}{2}}{2} - v_s\right]\left[\frac{\alpha_b + c_b - p_s}{c_b}\right]. \tag{12.29}$$

Maximizing with respect to p_s yields

$$p_s = \left(\frac{2}{3}\right) v_s + \frac{1}{3}(\alpha_b + c_b). \tag{12.30}$$

Solving equations (12.28) and (12.30) together yields

$$p_b = \left(\frac{2}{3}\right) v_b + \frac{1}{12} \tag{12.31}$$

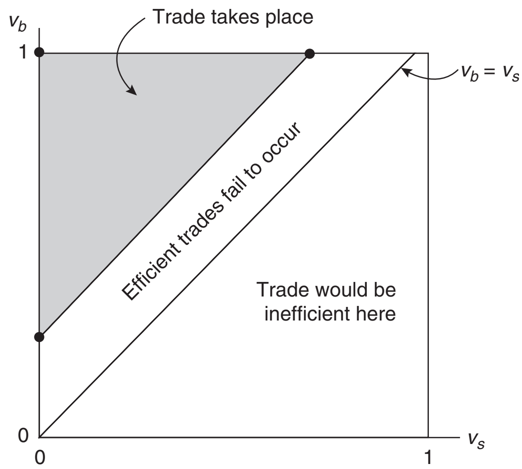

Figure 12.4 Trade in the linear equilibrium.

and

$$p_s = \left(\frac{2}{3}\right) v_s + \frac{1}{4}. \tag{12.32}$$

So we have derived a linear equilibrium. Manipulation of the equilibrium strategies shows that trade occurs if and only if $v_b \geq v_s + (1/4)$, which is to say, trade occurs if the valuations differ enough. The linear equilibrium does not make all efficient trades, because sometimes $v_b > v_s$ and no trade occurs, but it does make all trades with joint surpluses of 1/4 or more. Figure 12.4 illustrates this.

One detail about equation (12.31) should bother you. The equation seems to say that if $v_b = 0$, the buyer chooses $p_b = 1/12$. If that happens, though, the buyer is bidding more than his value! The reason this can be part of the equilibrium is that it is only a weak Nash equilibrium. Since the seller never chooses lower than $p_s = 1/4$, the buyer is safe in choosing $p_b = 1/12$; trade never occurs anyway when he makes that choice. He could just as well bid 0 instead of 1/12, but then he wouldn't have a linear strategy.

The linear equilibrium is not a truth-telling equilibrium. The seller does not report his true value v_s, but rather reports $p_s = (2/3)v_s + 1/4$. But we could replicate the outcome in a truth-telling equilibrium. We could have the buyer and seller agree that they would make reports r_b and r_s to a neutral mediator, who would then choose the trading price p. He would agree in advance to choose the trading price p by (1) mapping r_s onto p_s just as in the equilibrium above, (2) mapping r_b onto p_b just as in the equilibrium above, and (3) using p_b and p_s to set the price just as in the double auction mechanism. Under this mechanism, both players would tell the truth to the mediator. Let us compare the original linear mechanism with a truth-telling mechanism.

The Chatterjee–Samuelson mechanism: *The good is allocated to the seller if* $p_s \geq p_b$ *and to the buyer otherwise. The price at which the trade takes place, if it does, is* $p = (p_b + p_s)/2$.

A direct incentive-compatible mechanism: *The good is allocated to the seller if* $(2/3)p_s + 1/4 \geq (2/3)p_b + 1/12$, *which is to say, if* $p_s \geq p_b - 1/4$, *and to the buyer* *otherwise. The price at which the trade takes place, if it does, is*

$$p = \frac{\left(\left(\frac{2}{3}\right)p_b + \frac{1}{12}\right) + \left(\left(\frac{2}{3}\right)p_s + \frac{1}{4}\right)}{2} = \frac{p_b + p_s}{3} + \frac{1}{6}. \tag{12.33}$$

What I have done is substituted the equilibrium strategies of the two players into the mechanism itself, so now they will have no incentive to set their reports different from the truth. The mechanism itself looks odd, because it says that trade cannot occur unless v_b is more than 1/4 greater than v_s, but we cannot use the rule of trading if $v_b > v_s$ because then the players would start misreporting again. The truth-telling mechanism only works because it does not penalize players for telling the truth, and in order not to penalize them, it cannot make full use of the information to achieve efficiency.

In this game we have imposed a trading rule on the buyer and seller, rather than letting them decide for themselves what is the best trading rule. Myerson & Satterthwaite (1983) prove that of all the equilibria and all the mechanisms that are budget balancing, the linear equilibrium of the double auction mechanism yields the highest expected payoff to the players, the expectation being taken ex ante, before Nature has chosen the types. The mechanism is not optimal when viewed after the players have been assigned their types, and a player might not be happy with the mechanism once he knew his type. He will, however, at least be willing to participate.

What mechanism would players choose, ex ante, if they knew they would be in this game? If they had to choose after they were informed of their type, then their proposals for mechanisms could reveal information about their types, and we would have a model of bargaining under incomplete information that would resemble signalling models. But what if they chose a mechanism before they were informed of their type, and did not have the option to refuse to trade if after learning their type they did not want to use the mechanism?

In general, mechanisms have the following parts.

1 Each agent i simultaneously makes a report p_i.
2 A rule $x(p)$ determines the action (such as who gets the good, whether a bridge is built, etc.) based on the p.
3 Each agent i receives an incentive transfer a_i that in some way depends on his own report.
4 Each agent receives a budget-balancing transfer b_i that does not depend on his own report.

We will denote the agent's total transfer by t_i, so $t_i = a_i + b_i$.

In Bilateral Trading III, the mechanism had the following parts.

1 Each agent i simultaneously made a report p_i.
2 If $p_s \geq p_b$, the good was allocated to the seller, but otherwise to the buyer.
3 If there was no trade, then $a_s = a_b = 0$. If there was trade, then $a_s = (p_b + p_s)/2$ and $a_b = -((p_b + p_s)/2)$.
4 No further transfer b_i was needed, because the incentive transfers balanced the budget by themselves.

It turns out that if the players in Bilateral Trading can settle their mechanism and agree to try to trade in advance of learning their types, an efficient budget-balancing mechanism exists that can be implemented as a Nash equilibrium. The catch will be that after discovering his type, a player will sometimes regret having entered into this mechanism.

This would actually be part of a subgame perfect Nash equilibrium of the game as a whole. The mechanism design literature tends not to look at the entire game, and asks "Is there a mechanism which is efficient when played out as the rules of a game?" rather than "Would the players choose a mechanism that is efficient?"

Bilateral Trading IV: The Expected Externality Mechanism

PLAYERS
A buyer and a seller.

THE ORDER OF PLAY
-1 Buyer and seller agree on a mechanism $(x(p), t(p))$ that makes decisions x based on reports p and pays t to the agents, where p and t are 2-vectors and x allocates the good either to the buyer or the seller.
0 Nature independently chooses the seller to value the good at v_s and the buyer at v_b using the uniform distribution between 0 and 1. Each player's value is his own private information.
1 The seller reports p_s and the buyer reports p_b simultaneously.
2 The mechanism uses $x(p)$ to decide who gets the good, and $t(p)$ to make payments.

PAYOFFS
Player i's payoff is $(v_i + t_i)$ if he is allocated the good, t_i otherwise.

Part (-1) of the order of play is vague on how the two parties agree on a mechanism. The mechanism design literature is also vague, and focuses on efficiency rather than payoff-maximization. To be more rigorous, we should have one player propose the mechanism and the other accept or reject. The proposing player would add an extra transfer to the mechanism to reduce the other player's expected payoff to his reservation utility.

Let me use the term **action surplus** to denote the utility an agent gets from the choice of action.

The **expected externality mechanism** has the following objectives for each of the parts of the mechanism.

1 Induce the agents to make truthful reports.
2 Choose the efficient action.
3 Choose the incentive transfers to make the agents choose truthful reports in equilibrium.
4 Choose the budget-balancing transfers so that the incentive transfers add up to zero.

First I will show you a mechanism that does this. Then I will show you how I came up with that mechanism. Consider the following three-part mechanism:

1 The seller announces p_s. The buyer announces p_b. The good is allocated to the seller if $p_s \geq p_b$, and to the buyer otherwise.
2 The seller gets transfer $t_s = (1 - p_s^2)/2 - (1 - p_b^2)/2$.
3 The buyer gets transfer $t_b = (1 - p_b^2)/2 - (1 - p_s^2)/2$.

This is budget-balancing:

$$\frac{(1 - p_s^2)}{2} - \frac{(1 - p_b^2)}{2} + \frac{(1 - p_b^2)}{2} - \frac{(1 - p_s^2)}{2} = 0. \tag{12.34}$$

The seller's expected payoff as a function of his report p_s is the sum of his expected action surplus and his expected transfer. We have already computed his transfer, which is not conditional on the action taken.

The seller's action surplus is 0 if the good is allocated to the buyer, which happens if $v_b > p_s$, where we use v_b instead of p_b because in equilibrium $p_b = v_b$. This has probability $(1 - p_s)$. The seller's action surplus is v_s if the good is allocated to the seller, which has probability p_s. Thus, the expected action surplus is $p_s v_s$.
The seller's expected payoff is therefore

$$p_s v_s + \frac{(1 - p_s^2)}{2} - \frac{(1 - p_b^2)}{2}. \tag{12.35}$$

Maximizing with respect to his report, p_s, the condition is

$$v_s - p_s = 0, \tag{12.36}$$

so the mechanism is incentive compatible – the seller tells the truth.

The buyer's expected action surplus is v_b if his report is higher, for example, if $p_b > v_s$, and zero otherwise, so his expected payoff is

$$p_b v_b + \frac{(1 - p_b^2)}{2} - \frac{(1 - p_s^2)}{2}. \tag{12.37}$$

Maximizing with respect to his report, p_s, the first-order condition is

$$v_b - p_b = 0, \tag{12.38}$$

so the mechanism is incentive compatible – the buyer tells the truth.

Now let us see how to come up with the transfers. The expected externality mechanism relies on two ideas.

The first idea is that to get the incentives right, each agent's incentive transfer is made equal to the sum of the expected action surpluses of the other agents, where the expectation is calculated conditionally on (1) the other agents reporting truthfully, and (2) our agent's report. This makes the agent internalize the effect of his externalities on the other agents. His expected payoff comes to equal the expected social surplus. Here, this means, for example, that the seller's incentive transfer will equal the buyer's expected action surplus. Thus, denoting the uniform distribution by F,

$$\begin{aligned} a_s &= \int_0^{p_s} (0)\, dF(v_b) + \int_{p_s}^1 v_b dF(v_b), \\ &= 0 + \Big|_{p_s}^1 \frac{v_b^2}{2}, \\ &= \frac{1}{2} - \frac{p_s^2}{2}. \end{aligned} \tag{12.39}$$

The first integral is the expected buyer action surplus if no transfer is made because the buyer's value v_b is less than the seller's report p_s, so the seller keeps the good and the buyer's action surplus is zero. The second integral is the surplus if the buyer gets the good, which occurs whenever the buyer's value, v_b (and hence his report p_b), is greater than the seller's report, p_s.

We can do the same thing for the buyer's incentive, finding the seller's expected surplus.

$$\begin{aligned} a_b &= \int_0^{p_b} 0 dF(v_s) + \int_{p_b}^1 v_s dF(v_s), \\ &= 0 + \Big|_{p_b}^1 \frac{v_s^2}{2}, \\ &= \frac{1}{2} - \frac{p_b^2}{2}. \end{aligned} \tag{12.40}$$

If the seller's value v_s is low, then it is likely that the buyer's report of p_b is higher than v_s, and the seller's action surplus is zero because the trade will take place. If the seller's value v_s is high, then the seller will probably have a positive action surplus.

The second idea is that to get budget balancing, each agent's budget-balancing transfer is chosen to help pay for the other agents' incentive transfers. Here, we just have two agents, so the seller's budget-balancing transfer has to pay for the buyer's incentive transfer. That is very simple: just set the seller's budget-balancing transfer b_s equal to the buyer's incentive transfer a_b (and likewise set b_b equal to a_s).

The intuition and mechanism can be extended to N agents. There are now N reports $p_1, \ldots, p_N$. Let the action chosen be $x(p)$, where p is the N-vector of reports, and the action surplus of agent i is $W_i(x(p), v_i)$. To make each agent's incentive transfer equal to the sum of the expected action surpluses of the other agents, choose it so

$$a_i = E(\Sigma_{j \neq i} W_j(x(p), v_j)). \tag{12.41}$$

The budget balancing transfers can be chosen so that each agent's incentive transfer is paid for by dividing the cost equally among the other $(N-1)$ agents:

$$b_i = \left(\frac{1}{N-1}\right)(\Sigma_{j \neq i} E(\Sigma_{k \neq j} W_k(x(p), v_k))). \tag{12.42}$$

There are other ways to divide the costs that will still allow the mechanism to be incentive compatible, but equal division is the simplest.

The expected externality mechanism does have one problem: the participation constraint. If the seller knows that $v_s = 1$, he will not want to enter into this mechanism. His expected transfer would be $t_s = 0 - (1 - 0.5)^2/2 = -0.125$. Thus, his payoff from the mechanism is $1 - 0.125 = 0.875$, whereas he could get a payoff of 1 if he refused to participate. We say that this mechanism fails to be **interim incentive compatible**, because at the point when the agents discover their own types, but not those of the other agents, the agents might not want to participate in the mechanism or choose the actions we desire.

Ordinarily economists think of bargaining as being less structured than in the Bilateral Trading games, but it should be kept in mind that there are two styles of bargaining: bargaining with loose rules that are "made up as you go along," and bargaining with pre-determined rules to which the players can somehow commit. This second kind of bargaining is more common in markets where many bargains are going to be made and in situations where enough is at stake that the players first negotiate the rules under which the main bargaining will occur. Once bargaining becomes mechanism design, it becomes closer to the idea of simply holding an auction. Bulow & Klemperer (1996) compare the two means of selling an item, observing that a key feature of auctions is involving more traders, an important advantage.

Notes

N12.2 The Nash bargaining solution

- See Binmore, Rubinstein, & Wolinsky (1986) for a comparison of the cooperative and noncooperative approaches to bargaining. For overviews of cooperative game theory see Luce & Raiffa (1957) and Shubik (1982).
- While the Nash bargaining solution can be generalized to n players (see Harsanyi [1977], p. 196), the possibility of interaction between coalitions of players introduces new complexities. Solutions such as the Shapley value (Shapley [1953b]) try to account for these complexities.

  The **Shapley value** satisfies the properties of invariance, anonymity, efficiency, and linearity in the variables from which it is calculated. Let S_i denote a **coalition** containing player i; that is, a group of players including i that makes a sharing agreement. Let $v(S_i)$ denote the sum of the utilities of the players in coalition S_i, and $v(S_i - \{i\})$ denote the sum of the utilities in the coalition created by removing i from S_i. Finally, let $c(s)$ be the number of coalitions of size s containing player i. The Shapley value for player i is then

$$\phi_i = \frac{1}{n}\sum_{s=1}^{n}\frac{1}{c(s)}\sum_{S_i}[v(S_i) - v(S_i - \{i\})]. \tag{12.43}$$

  where the S_i are of size s. The motivation for the Shapley value is that player i receives the average of his marginal contributions to different coalitions that might form. Gul (1989) has provided a noncooperative interpretation.

N12.5 Incomplete information

- Bargaining under asymmetric information has inspired a large literature. In early articles, Fudenberg & Tirole (1983) uses a two-period model with two types of buyers and two types of sellers. Sobel & Takahashi (1983) builds a model with either T or infinite periods, a continuum of types of buyers, and one type of seller. Cramton (1984) uses an infinite number of periods, a continuum of types of buyers, and a continuum of types of sellers. Rubinstein (1985a) uses an infinite number of periods, two types of buyers, and one type of seller, but the types of buyers differ not in their valuations, but in their discount rates. Samuelson (1984) looks at the case where one bargainer knows the size of the pie better than the other bargainer. Perry (1986) uses a model with fixed bargaining costs and asymmetric information in which each bargainer makes an offer in turn, rather than one offering and the other accepting or rejecting. For overviews, see the surveys of Sutton (1986) and Kennan & Wilson (1993).
- The asymmetric information model in section 12.5 has **one-sided** asymmetry in the information: only the buyer's type is private information. Fudenberg & Tirole (1983) and others have also built models with **two-sided** asymmetry, in which buyers' and sellers' types are both private information. In such models a multiplicity of perfect Bayesian equilibria can be supported for a given set of parameter values. Out-of-equilibrium beliefs become quite important, and provided much of the motivation for the exotic refinements mentioned in section 6.2.

N12.6 Setting up a way to bargain: the Myerson–Satterthwaite Model

- The Bilateral Trading model originated in Chatterjee & Samuelson (1983, p. 842), who also analyze the more general mechanism with $p = \theta p_s + (1-\theta)p_b$. I have adapted this description from Gibbons (1992, p. 158).
- Discussions of the general case can be found in Fudenberg & Tirole (1991a, p. 273), and Mas-Colell, Whinston & Green (1994, p. 885). I have taken the term "expected externality mechanism" from MWG. Fudenberg and Tirole use "AGV mechanism" or "AGV-Arrow mechanism" for the same thing, because the idea was first published in and D'Aspremont & Gerard-Varet (1979) and Arrow (1979). It is also possible to add extra costs that depend on the action chosen (for example, a transactions tax if the good is sold from buyer to seller). See Fudenberg & Tirole (1991a, p. 274). Myerson (1991) is also worth looking into.

Problems

12.1: A fixed cost of bargaining and grudges (medium)

Smith and Jones are trying to split 100 dollars. In bargaining round 1, Smith makes an offer at cost 0, proposing to keep S_1 for himself and Jones either accepts (ending the game) or rejects. In round 2, Jones makes an offer at cost 10 of S_2 for Smith and Smith either accepts or rejects. In round 3, Smith makes an offer of S_3 at cost c, and Jones either accepts or rejects. If no offer is ever accepted, the 100 dollars goes to a third player, Dobbs.

(a) If $c = 0$, what is the equilibrium outcome?
(b) If $c = 80$, what is the equilibrium outcome?
(c) If $c = 10$, what is the equilibrium outcome?
(d) What happens if $c = 0$, but Jones is very emotional and would spit in Smith's face and throw the 100 dollars to Dobbs if Smith proposes $S = 100$? Assume that Smith knows Jones's personality perfectly.

12.2: Selling cars (medium)

A car dealer must pay \$10,000 to the manufacturer for each car he adds to his inventory. He faces three buyers. From the point of view of the dealer, Smith's valuation is uniformly distributed between \$12,000 and \$21,000, Jones's is between \$9,000 and \$12,000, and Brown's is between \$4,000 and \$12,000. The dealer's policy is to make a single take-it-or-leave-it offer to each customer, and he knows these three buyers will not be able to resell to each other. Use the notation that the maximum valuation is $\overline{V}$ and the range of valuations is R.

(a) What will the offers be?
(b) Who is most likely to buy a car? How does this compare with the outcome with perfect price discrimination under full information? How does it compare with the outcome when the dealer charges \$10,000 to each customer?
(c) What happens to the equilibrium prices if, with probability 0.25, each buyer has a valuation of \$0, but the probability distribution remains otherwise the same?

12.3: The Nash bargaining solution (medium)

Smith and Jones, shipwrecked on a desert island, are trying to split 100 pounds of cornmeal and 100 pints of molasses, their only supplies. Smith's utility function is $U_s = C + 0.5M$ and Jones's is $U_j = 3.5C + 3.5M$. If they cannot agree, they fight to the death, with $U = 0$ for the loser. Jones wins with probability 0.8.

(a) What is the threat point?
(b) With a 50–50 split of the supplies, what are the utilities if the two players do not recontract? Is this efficient?
(c) Draw the threat point and the Pareto frontier in utility space (put U_s on the horizontal axis).
(d) According to the Nash bargaining solution, what are the utilities? How are the goods split?
(e) Suppose Smith discovers a cookbook full of recipes for a variety of molasses candies and corn muffins, and his utility function becomes $U_s = 10C + 5M$. Show that the split of goods in part (d) remains the same despite his improved utility function.

12.4: Price discrimination and bargaining (easy)

A seller with marginal cost constant at c faces a continuum of consumers represented by the linear demand curve $Q^d = a - bP$, where $a > c$. Demand is at a rate of one or zero units per consumer, so if all consumers between points 1 and 2.5 on the consumer continuum make purchases at a price of 13, we say that a total of 1.5 units are sold at a price of 13 each.

(a) What is the seller's profit if he chooses one take-it-or-leave- it price?
(b) What is the seller's profit if he chooses a continuum of take-it-or-leave-it prices at which to sell, one price for each consumer? (You should think here of a pricing function, since each consumer is infinitesimal.)
(c) What is the seller's profit if he bargains separately with each consumer, resulting in a continuum of prices? You may assume that bargaining costs are zero and that buyer and seller have equal bargaining power.

12.5: A fixed cost of bargaining and incomplete information (medium)

Up to part (c), this problem is identical with problem 12.1. Smith and Jones are trying to split 100 dollars. In bargaining round 1, Smith makes an offer at cost 0, proposing to keep S_1 for himself and Jones either accepts (ending the game) or rejects. In round 2, Jones makes an offer at cost 10 of S_2 for Smith and Smith either accepts or rejects. In round 3, Smith makes an offer of S_3 at cost c, and Jones either accepts or rejects. If no offer is ever accepted, the 100 dollars goes to a third player, Dobbs.

(a) If $c = 0$, what is the equilibrium outcome?
(b) If $c = 80$, what is the equilibrium outcome?
(c) If Jones's priors are that $c = 0$ and $c = 80$ are equally likely, but only Smith knows the true value, what are the players' equilibrium strategies in rounds 2 and 3? (i.e., what are S_2 and S_3, and what acceptance rules will each player use?)
(d) If Jones's priors are that $c = 0$ and $c = 80$ are equally likely, but only Smith knows the true value, what are the equilibrium strategies for round 1? (Hint: the equilibrium uses mixed strategies.)

12.6: A fixed bargaining cost, again (easy)

Apex and Brydox are entering into a joint venture that will yield 500 million dollars, but they must negotiate the split first. In bargaining round 1, Apex makes an offer at cost 0, proposing to keep A_1 for itself. Brydox either accepts (ending the game) or rejects. In round 2, Brydox incurs a cost of 10 million to make an offer that gives A_2 to Apex, and Apex either accepts or rejects. In round 3, Apex incurs a cost of c to make an offer that gives itself A_3, and Brydox either accepts or rejects. If no offer is ever accepted, the joint venture is cancelled.

(a) If $c = 0$, what is the equilibrium? What is the equilibrium outcome?
(b) If $c = 10$, what is the equilibrium? What is the equilibrium outcome?
(c) If $c = 300$, what is the equilibrium? What is the equilibrium outcome?

12.7: Myerson–Satterthwaite (medium)

The owner of a tract of land values his land at v_s and a potential buyer values it at v_b. The buyer and seller do not know each other's valuations, but guess that they are uniformly distributed between 0 and 1. The seller and buyer suggest p_s and p_b simultaneously, and they have agreed that the land will be sold to the buyer at price $p = (p_b + p_s)/2$ if $p_s \leq p_b$.

The actual valuations are $v_s = 0.2$ and $v_b = 0.8$. What is one equilibrium outcome given these valuations and this bargaining procedure? Explain why this can happen.

12.8: Negotiation (Rasmusen [2002]) (hard)

Two parties, the Offeror and the Acceptor, are trying to agree to the clauses in a contract. They have already agreed to a basic contract, splitting a surplus 50–50, for a surplus of Z for each player. The offeror can at cost C offer an additional clause which the acceptor can accept outright, inspect carefully (at cost M), or reject outright. The additional clause is either "genuine," yielding the Offeror X_g and the Acceptor Y_g if accepted, or "misleading," yielding the Offeror X_m (where $X_m > X_g > 0$) and the Acceptor $-Y_m < 0$.

What will happen in equilibrium?

Labor Bargaining: A Classroom Game for Chapter 12[3]

Currently, an employer is paying members of a labor union $46,000 per year, but the union has told its members it thinks $68,000 would be a fairer amount. Every $1,000 increase in salary costs the employer $30 million per year, and benefits the workers in aggregate by $25 million (the missing $5 million going to taxes, which are heavier for the workers).

If the workers go on strike, it will cost the players $25 million per week in foregone earnings, and it will cost the employer $60 million in lost profits. Interest rates are low enough that they can be ignored in this game.

The rules for bargaining are as follows. The union makes the first offer, on May 1 (time 0), and the employer accepts or rejects. If the employer accepts the offer, there is no strike. If the employer rejects it, there is a strike for the next week, but the employer then can make a counteroffer on May 8 (time 1). If it is accepted by the union, the strike has lasted one week. If it is rejected, the union has one week in which to put together its counteroffer for May 15 (time 2).

The workers' morale and bank accounts will run out after 7 weeks of a strike, at time 7. If no other agreement has been reached, the union must then accept an offer as low as $46,000. It will not accept an offer any lower, because the workers angrily refuse to ratify a lower offer.

Students will be put into groups of three that represent either the employer or the union. Employer groups and union groups will then pair up to simultaneously play the game. A group's objective is to maximize its payoff. The instructor will set up place on the blackboard for each group to record its weekly offers. If a group cannot agree on what offer to make and does not write it up on the board in time, then it forfeits its chance to make an offer that week. Each offer must be in thousands of dollars of annual salary – no offers of $52,932 are allowed.

[3] This game is adapted from a classroom game of Vijay Krishna.