September 4, 1999. Overheads created October 5, 1999
Eric Rasmusen, Erasmuse@indiana.edu. Web: Php.indiana.edu/∼erasmuse.

11 Bargaining OVERHEADS

11.1: The Basic Bargaining Problem: Splitting a Pie .

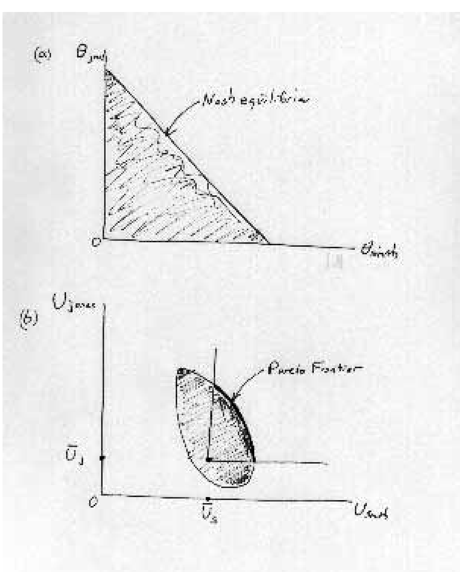
"Splitting a Pie"

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: $\begin{cases} \pi_s = & \theta_s. \\ \pi_j = & \theta_j. \end{cases}$
If $\theta_s + \theta_j > 1$, then $\pi_s = \pi_j = 0$.

Figure 11.1: (a) Splitting a Pie; (b) Nash Bargaining Game



11.2 The Nash Bargaining Solution

The axioms that generate the concept are

(1) Invariance.
For any strictly increasing linear function F,

$$U^*[F(\bar{U}), F(X)] = F[U^*(\bar{U}, X)]. \tag{1}$$

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

(2) Efficiency.
The solution is Pareto optimal, so that not both players can be made better off. In mathematical terms,

$$(U_s, U_j) > U^* \Rightarrow (U_s, U_j) \notin X. \tag{2}$$

(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.

$$U^*(\bar{U}, X) \in Y \subseteq X \Rightarrow U^*(\bar{U}, Y) = U^*(\bar{U}, X). \tag{3}$$

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

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

$$U^* = \underset{U \in X, U \geq \bar{U}}{Argmax} \quad (U_s - \bar{U}_s)(U_j - \bar{U}_j). \tag{4}$$

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 (4) becomes

$$\underset{\theta_s, \theta_j \mid \theta_s + \theta_j \leq 1}{Maximize} \quad (\theta_s - 0)(\theta_j - 0), \tag{5}$$

which generates the first order conditions

$$\theta_s - \lambda = 0, \text{ and } \theta_j - \lambda = 0, \tag{6}$$

where λ is the Lagrange multiplier on the constraint. From (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.

11.3 Alternating Offers over Finite Time

"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,

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

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.

Smith owes his success to his ability to make the last offer.

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 11.1 shows the progression of Smith's shares when $\delta = 0.9$.

Table 11.1 Alternating Offers over Finite Time

Round	bf 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

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 $\frac{1}{1+\delta}$, 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.

11.4 Alternating Offers over Infinite Time

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

$$\theta_s = \frac{1-\delta_j}{1-\delta_s\delta_j}, \tag{7}$$

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

$$\theta_s = \frac{1}{1+\delta}. \tag{8}$$

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 11.1. At the other extreme, if r is small, the pie is split almost evenly: if $r = 0.01$, then $\delta \approx 0.99$ and $\theta_s \approx 0.503$.

Proposition 11.1 (Rubinstein [1982])
In the discounted infinite game, the unique perfect equilibrium outcome is $\theta_s = \frac{1-\delta_j}{(1-\delta_s\delta_j)}$*, 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 11.2. (Jones would thus get 1-M, but that is not relevant to the proof.)

Table 11.2 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

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 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:

$$M = 1-\delta_j(1-\delta_s M). \tag{9}$$

Solving equation (9) for M, we obtain

$$M = \frac{1 - \delta_j}{1 - \delta_s \delta_j}. \tag{10}$$

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.

No Discounting, But a Fixed Bargaining Cost

Assume that there is no discounting, but whenever Smith or Jones makes an offer, he incurs the cost c_s or c_j. In every subgame perfect equilibrium, Smith makes an offer and Jones accepts, but there are three possible cases.

(1) Delay Costs are Equal

$$c_s = c_j = c.$$

The Nash indeterminacy of Section 11.1 remains almost as bad; any fraction such that each player gets at least c is supported by some perfect equilibrium.

(2) 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.

(3) 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.

11.5 Incomplete Information

"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 (acceptance ends the game).
(3) The seller offers a second price p_2.
(4) The buyer accepts or rejects p_2.

Payoffs

$$\pi_{seller} = \begin{cases} p_1 & \text{if } p_1 \text{ is accepted} \\ \delta p_2 & \text{if } p_2 \text{ is accepted} \\ 0 & \text{if no offer is accepted} \end{cases}$$

$$\pi_{buyer} = \begin{cases} b - p_1 & \text{if } p_1 \text{ is accepted} \\ \delta(b - p_2) & \text{if } p_2 \text{ is accepted} \\ 0 & \text{if no offer is accepted} \end{cases}$$

If p_2 is accepted, the buyer's payoff is $\delta(b-p_2)$, rather than $\delta b - p_2$, because the present value of cash paid in the second period is less than that of cash paid in the first period. Consumption in the second period provides less pleasure, but payment provides less pain. Let us set $\delta = 0.9$ for the numerical computations.

A Case of Many Low-Valuation Buyers: $\gamma = 0.5$

We will start by assuming that $\gamma = 0.5$, so the probability is fairly high that the buyer has the valuation 100.

Equilibrium (Pooling). In the first period, $p_1 = 100$, $Buyer_{100}$ accepts $p_1 \leq 100$, and $Buyer_{150}$ accepts $p_1 \leq 105$. In the second period, $p_2 = 100$, $Buyer_{100}$ accepts $p_2 \leq 100$, and $Buyer_{150}$ accepts $p_2 \leq 150$. The seller's beliefs out of equilibrium are that if a buyer rejects $p_1 = 100$, he is $Buyer_{100}$ with probability γ (passive conjectures). The outcome is that $p_1 = 100$ and the buyer accepts.

Let us test that this is a perfect Bayesian equilibrium. As always, work back from the end. Both types of buyers have a dominant strategy for the last move: accept any offer below b. Given the parameters, the seller should not raise p_2 above 100, because with probability $\gamma = 0.5$ he would lose a profit of 100 and his potential revenue is no greater than 150.

$Buyer_{150}$, although looking ahead to $p_2 = 100$, is willing to pay more than 100 in period one because of discounting. His payoff is the same from accepting $p_1 = 105$ as from accepting $p_2 = 100$, because his nominal surplus of 50 from accepting the lower price is discounted to a utility value of 45. $Buyer_{100}$, however, is never willing to pay more than 100, and discounting is irrelevant because he has no surplus to look forward to anyway.

The seller knows that even if $p_1 = 105$ he will still sell to $Buyer_{150}$, but if he tries that and finds no buyer in the first period, he has delayed receipt of his payment, which is discounted. Since $100 > 97.5 (= [1-\gamma] \cdot [105] + \gamma \cdot \delta \cdot [100])$, the seller prefers the safe present price of 100 to the alternative of a gamble between a present 105 and a future 100.

Out-of-equilibrium beliefs are specified for the equilibrium, but they do not really matter. Whatever inference the seller may draw if the buyer refuses $p_1 = 100$, the inference never induces the buyer to change his actions. It might be that the seller believes a refusal indicates that the buyer's value to be 150, so $p_2 = 150$, but that does not change the buyer's incentive to accept $p_1 = 100$.

A Case with Few Low-Valuation Buyers: $\gamma = 0.05$

If the proportion of low-valuation buyers is as small as $\gamma = 0.05$, the equilibrium is separating and in mixed strategies.

Equilibrium (Separating, in Mixed Strategies)
In the first period, $p_1 = 150$, Buyer$_{100}$ accepts $p_1 \leq 100$, and Buyer$_{150}$ accepts p_1 with probability $m(p_1)$, where

$$\begin{cases} m = & 1 \quad \text{if } p_1 \leq 105. \\ m = & \alpha \quad \text{if } 105 < p_1 \leq 150 \\ & (\text{where } 0 \leq \alpha \leq 0.89) \\ m = & 0 \quad \text{if } p_1 > 150. \end{cases}$$

In the second period, $p_2 = 150$ if the seller believes that he faces a Buyer$_{100}$ with probability less than $\frac{1}{3}$, and otherwise $p_2 = 100$. Buyer$_{100}$ accepts $p_2 \leq 100$, and Buyer$_{150}$ accepts $p_2 \leq 150$. The outcome is that $p_1 = 150$, which is sometimes accepted by Buyer$_{150}$; $p_2 = 150$, which is accepted by Buyer$_{150}$; and Buyer$_{100}$ never accepts an offer.

The observed outcome is simple— the price always stays at 150, and some buyers accept in each period while other buyers never accept— but the equilibrium strategies are quite complicated. As we will see, the equilibrium is not even fully determined, because the mixing probability α can take any of a continuum of values.

The strategies in the second-period game are simple enough. In the second period, the buyer accepts if the price is less than his valuation and the seller trades off a safe 100 against a gamble between 0 and 150. He is indifferent between them if

$$100 = 0 \cdot Prob(\text{Buyer}_{100}) + 150 \cdot [1 - Prob(\text{Buyer}_{100})], \qquad (11)$$

which yields a critical value of $Prob(\text{Buyer}_{100}) = \frac{1}{3}$. If neither type of buyer accepted first-period offers, the second period belief would be $Prob(\text{Buyer}_{100}) = \gamma$, which we assumed to be 0.05, so the second-period price would be 150.

The first -period strategies are more complicated. The first period strategy of Buyer_{150} is not the pure strategy of accepting the offer $p_1 = 150$, because if he always accepted in the first period the seller would lower the price in the second period, knowing that a buyer who rejected the first-period offer must be a Buyer_{100}. Anticipating the price drop, Buyer_{150} would refuse $p_1 = 150$, which contradicts the reason for the drop.

In equilibrium, it must be that after a refusal in the first period the seller puts a high enough probability on Buyer_{150} that he decides to keep the price high in the second period. For the seller to want to keep $p_2 = 150$, the probability that a buyer who refused the first-period offer is a Buyer_{150} must be at least $\frac{2}{3}$, from equation (11. 11). If both p_1 and p_2 are equal to 150, the buyer will be indifferent as to when he accepts, so he is willing to follow a mixed strategy. We can calculate the mixing probability $m(150)$ by finding the value that makes the seller willing to keep the price at 150 in the second period. The probability that a buyer is a Buyer_{150} is equal to $1 - \gamma$ in the first period, but a Buyer_{150} only rejects the first-period offer with probability $1 - m(150)$. Therefore, in the second period, using Bayes' Rule,

$$Prob(\text{Buyer}_(100)) = \frac{\gamma}{\gamma + [1 - m(150)][1 - \gamma]}. \tag{12}$$

Plugging in $\gamma = 0.05$ and $Prob(\text{Buyer}_{100}) = \frac{1}{3}$, equation (11. 12) can be solved to yield $m(150) = 0.89$ (rounded). The calculation has ensured that if the Buyer_{150} accepts the first offer with probability 0.89, the probability that a second-period buyer has valuation 100 is $\frac{1}{3}$. The value $\alpha = 0.89$ is the maximum equilibrium probability that a Buyer_{150} refuses to buy in the first period, but a smaller value for α would support an equilibrium *a fortiori* since the probability that a refuser had valuation 150 would be even greater than $\frac{2}{3}$.

There is a continuum of equilibria, which differ in their values for α. Two values are focal points, 0 and 0.89. The value of 0 is a pure strategy, and has the attraction of simplicity. The value of 0.89 is Pareto efficient, because it is the highest equilibrium probability that the buyer accepts immediately, which avoids the lost utility from delay. The qualitative difference between the two equilibria is that in the pure-strategy equilibrium no buyers accept the offer in the first period.

Consider what happens if the seller offers a price of 140 in the first period. For the same reasons as we described for $p_1 = 150$, the equilibrium cannot be in pure strategies. The equilibrium strategies in the out-of-equilibrium subgame are for the Buyer_{150} to mix between accepting and rejecting, and for the seller to mix between $p_2 = 100$ and $p_2 = 150$. Here, unlike on the equilibrium path, the seller must also mix, because otherwise the buyer would strongly prefer to accept 140 rather than wait for 150. The seller is willing to mix only if he believes that there is exactly a one-third probability that the buyer is a Buyer_{100}, so the buyer's strategy is $m(150) = 0.89$, as before. Denote the seller's mixing probability of $p_2 = 100$ by μ. It must take a value that makes the buyer indifferent between accepting and rejecting, so

$$150 - p_1 = .9\mu(150 - 100) + (1 - \mu) \cdot 0, \qquad (13)$$

which solves to $\mu = 3\frac{1}{3} - p_1/45$, or $\mu = 0.22$ for $p_1 = 140$.

*11.6 Setting up a Way to Bargain: The Myerson-Satterthwaite Mechanism

In this section we will 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– 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 p_s.
(2) The buyer chooses "Accept" or "Reject."
(3) The good is allocated to the seller if the buyer accepts and to the buyer otherwise. The price at which the trade takes place, if it does, is $p = p_s$

Payoffs
If there is no trade, both players have payoffs of 0. If there is trade, the seller's payoff is $p - v_s$ and the buyer's is $v_b - p$.

The unique subgame perfect Nash equilibrium of this game is for the seller to report $p_s = v_b$ and for the buyer to accept if $p_s \leq v_b$. This is efficient. But the buyer would be unlikely to agree to this mechanism in the first place, before the game starts, because although it is efficient it always gives all of the social surplus to the seller.

An example of a truthtelling mechanism for this game woudl replace move (3) with

(3′) The good is allocated to the seller if the buyer accepts and to the buyer otherwise. The price at which the trade takes place, if it does, is $p = v_s + \frac{v_b - v_s}{2}$.

"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 endof 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

$$[p_s - v_s]\,[Prob\{v_b \geq p_s\}] + 0\,[Prob\{v_b \leq p_s\}] = [p_s - v_s]\,[1 - p_s]\,. \tag{14}$$

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

$$p_s = \frac{1 + v_s}{2}. \tag{15}$$

This is not efficient because if v_b is just a little bigger than v_s, trade will not occur even though gains from trade do exist. In fact, trade will fail to occur whenever $v_b > \frac{1+v_s}{2}$.

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

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 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

$$[p_s - v_s]\,[Prob\{p_b(v_b) \geq p_s\}] + 0\,[Prob\{p_b(v_b) \leq p_s\}] = [p_s - v_s][1 - p_s]. \tag{16}$$

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

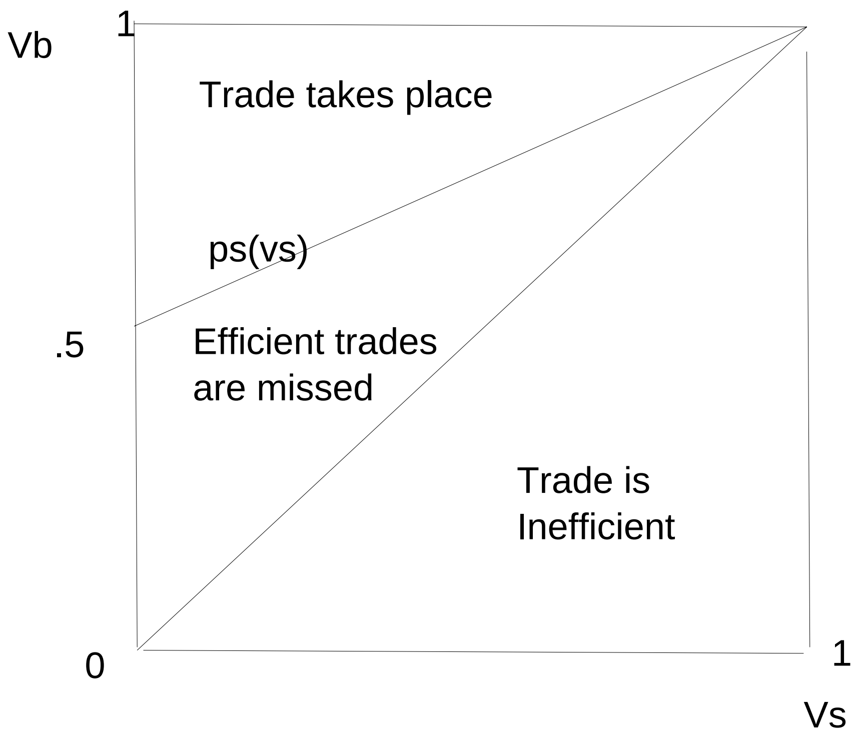
Maximizing this, the seller's strategy will solve the first-order condition $1 - 2p_s + v_s = 0$, and so will again be

$$p_s(v_s) = \frac{1 + v_s}{2} = \frac{1}{2} + \frac{v_s}{2}. \tag{17}$$

Will the buyer's best response to this strategy be $p_b = v_b$? Yes, because whenever $v_b \geq \frac{1}{2} + \frac{v_s}{2}$ 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 $v_b < frac12 + \frac{v_s}{2}$, but this can be avoided by using the truthtelling strategy. The buyer also needs to worry about preventing trade from occurring when $v_b > \frac{1}{2} + \frac{v_s}{2}$, 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 $v_b > \frac{1+v_s}{2}$. Figure 11.2 shows when trades will be completed based on the parameter values.

Figure 11.2 Trades in the Myerson-Satterthwaite Model



"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.
(2) 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 $p = \frac{(p_b+p_s)}{2}$.

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

$$\left[v_b - \frac{p_b + E[p_s|p_b \geq p_s]}{2}\right][Prob\{p_b \geq p_s\}], \tag{18}$$

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

$$\left[-\frac{p_s + E(p_b|p_b \geq p_s)}{2} - v_s\right][Prob\{p_b \geq p_s\}], \tag{19}$$

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

This trading rule is called the **double auction mechanism**. This problem is like that of the 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).

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.

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 $p_s(v_s) = \alpha_s + c_s v_s$. Then from the buyer's point of view, p_s will be uniformly distributed from α_s to $\alpha_s + c_s$ with density $1/c_s$, as v_s ranges from 0 to 1. Since $E_b[p_s|p_b \geq p_s] = E_b(p_s|p_s \in [a_s, p_b]) = \frac{a_s+p_b}{2}$, the buyer's expected payoff (11.18) becomes

$$\left[v_b - \frac{p_b + \frac{\alpha_s+p_b}{2}}{2}\right]\left[\frac{p_b - \alpha_s}{c_s}\right]. \tag{20}$$

Maximizing with respect to p_b yields

$$p_b = \frac{2}{3}v_b + \frac{1}{3}\alpha_s. \tag{21}$$

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 (11.19), becomes, since $E_s(p_b|p_b \geq p_s) = E_s(p_b|p_b \in [p_s, \alpha_b + c_b] = \frac{p_s+\alpha_b+c_b}{2}$,

$$\left[\frac{p_s + \frac{p_s+\alpha_b+c_b}{2}}{2} - v_s\right]\left[\frac{\alpha_b + c_b - p_s}{c_b}\right]. \tag{22}$$

Maximizing with respect to p_s yields

$$p_s = \frac{2}{3}v_s + \frac{1}{3}\left(\alpha_b + c_b\right). \tag{23}$$

Solving equations (11.21) and (11.23) together yields

$$p_b = \frac{2}{3}v_b + \frac{1}{12} \tag{24}$$

and

$$p_s = \frac{2}{3}v_s + \frac{1}{4}. \tag{25}$$

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 by 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.

One detail about equation (11.24) should bother you. The equation seems to say that if $v_b = 0$, the buyer chooses $p_b = 1/12$. If that happens, tho, 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 truthtelling 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 truthtelling equilibrium. We could have the buyer and seller agree that they would make reports r_s and r_b to a neutral mediator, who would then choose the trading price p. He would agree in advance to choose the trading price p by (a) mapping r_s onto p_s just as in the equilibrium above, (b) mapping r_b onto p_b just as in the equilibrium above, and (c) 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 truthtelling 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* $p = \frac{(p_b+p_s)}{2}$

A Direct Incentive-Compatible Mechanism: *The good is allocated to the seller if* $\frac{2}{3}p_s + \frac{1}{4}) \geq \frac{2}{3}p_b + \frac{1}{12}$, *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*

$$p = \frac{(\frac{2}{3}p_b + \frac{1}{12}) + (\frac{2}{3}p_s + \frac{1}{4})}{2} = \frac{p_b + p_s}{3} + \frac{1}{6} \qquad (26)$$

"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 allocated 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.

The **Expected Externality Mechanism** (to use Whinston's term) 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 $t_b = (1-p_b^2)/2 - (1-p_s^2)/2$.

Note that this is budget-balancing:

$$(1-p_s^2)/2 - (1-p_b^2)/2 + (1-p_b^2)/2 - (1-p_s^2)/2 = 0. \qquad (27)$$

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

$$p_s v_s + (1 - p_s^2)/2 - (1 - p_b^2)/2. \tag{28}$$

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

$$v_s - p_s = 0, \tag{29}$$

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, e.g. if $p_b > v_s$, and zero otherwise, so his expected payoff is

$$p_b v_b + (1 - p_b^2)/2 - (1 - p_s^2)/2 \tag{30}$$

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

$$v_b - p_b = 0, \tag{31}$$

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

Now let's 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 (a) the other agents reporting truthfully, and (b) 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,

$$\begin{aligned} a_s &= \int_0^{p_s} 0 dF(v_b) + \int_{p_s}^1 v_b dF(v_b) \\ &= 0 + \frac{v_b^2}{2}\Big|_{p_s}^1 \\ &= \frac{1}{2} - \frac{p_s^2}{2}. \end{aligned} \tag{32}$$

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.

$$\begin{aligned} a_b &= \int_0^{p_b} 0 dF(v_s) + \int_{p_b}^1 v_s dF(v_s) \\ &= 0 + \frac{v_s^2}{2}\Big|_{p_b}^1 \\ &= \frac{1}{2} - \frac{p_b^2}{2}. \end{aligned} \tag{33}$$

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

$$a_i = E_{\theta_{-i}} \left(\Sigma_{j \neq i} W_j(x(p), v_j) \right). \tag{34}$$

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:

$$b_i = \frac{1}{N-1} \left(\Sigma_{j \neq i} E_{\theta_{-j}} \left(\Sigma_{k \neq j} W_k(x(p), v_k) \right) \right). \tag{35}$$

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 $t_s = 0 - (1 - .5)^2/2 = -.125$. Thus, his payoff from the mechanism is $1 - .125 = .875$, 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.