Economics 703
Advanced Microeconomics **Prof. Peter Cramton**

Lecture Note 3: Mechanism Design[1]

Outline

A. Introduction to Games with Incomplete Information
   1. Imperfect Information vs. Incomplete Information
   2. Bayesian Games
   3. The Revelation Principle

B. Bilateral Trading Mechanisms
   1. War of Attrition
   2. Simultaneous Offers
   3. The Public Choice Problem
   4. A General Model
   5. Efficiency in Games with Incomplete Information
   6. Durability

C. Multilateral Trading Mechanisms
   1. Dissolving a Partnership
   2. Optimal Auctions

A. Introduction to Games with Incomplete Information

1. Imperfect Information vs. Incomplete Information

Game of *imperfect information*: one or more players do not know the full history of the game.
Game of *incomplete information*: the players have different private information about their preferences and abilities.
Example: Porter's model on cartel maintenance is an example of a game of imperfect information: each firm is well aware of the other firms' preferences and abilities, but does not know the production quantities of the other firms. It would be a game of incomplete information if a firm's cost function were known privately (unknown preferences) or a firm's productive capacity were known privately (unknown ability).
Note that uncertainty about strategy spaces can be modeled as uncertainty about preferences, by simply associating a sufficiently negative payoff to unavailable strategies.

The key to analyzing games of incomplete information is to transform them into games of imperfect

[1] These notes are based without restraint on notes by Robert Gibbons, MIT.

information by letting nature move first, randomly selecting each player's payoff function.

Example 1: Symmetric oligopoly model with unknown costs. Firm i's marginal cost c_i may be either low or high {l,h}. Firm i observes its cost (nature's choice of l or h) but not the costs of the other firms. In the game of imperfect information, j does not observe natures choice of c_i, but j holds probabilistic beliefs about the likelihoods of nature's choice, summarized by a probability p that nature chose l.

Example 2: An auction is a another example of a game with incomplete information. Suppose a seller decides to use a first-price, sealed-bid auction to allocate a good to one of two buyers. Let nature's choice of the buyers' valuations for the good, v_1 and v_2, be independently and uniformly distributed on [0,1]. Based on v_i, player i submits a bid $b_i(v_i)$. The player with the highest bid gets the good and pays her bid (a coin flip can break ties).

2. Bayesian Games (Harsanyi, *Management Science* 1967-8)

normal form game $G = \{A_1,...,A_n; u_1,...,u_n\}$

Bayesian game $\Gamma = \{A_1,...,A_n; T_1,...,T_n; p_1,...,p_n; u_1,...,u_n\}$ where

A_i = strategy set for i, actions: $a = (a_1,...,a_n) \in A = A_1 \times ... \times A_n$.

T_i = type space for i, types: $t = (t_1,...,t_n) \in T = T_1 \times ... \times T_n$

p_i = beliefs for i, $p_i(t_{-i} \mid t_i)$ = i's belief about types t_{-i} given type t_i.

u_i = utility function for i, $u_i(a,t)$ depends on both actions a and types t.

Beliefs $\{p_1,...,p_n\}$ are *consistent* if they can be derived from Bayes' rule from a common joint distribution p(t) on T; i.e., there exists p(t) such that

$$p_i(t_{-i}|t_i) = \frac{p(t)}{p(t_i)} \quad \text{where} \quad p(t_i) = \sum_{t_{-i} \in T_{-i}} p(t_{-i}, t_i) \quad \text{for all i and } t_i.$$

Beliefs are consistent if nature moves first and types are determined according to p(t) and each i is informed only of t_i.

Auction example: types t_i are the valuations v_i, actions a_i are the bids b_i, $A_i = [0,\infty)$, $T_i = [0,1]$, $p_i(t_j|t_i) = 1$ for all t_i and t_j, and

$$u_i(a,t) = \begin{cases} t_i - a_i & \text{if } a_i > a_j \\ (t_i - a_i)/2 & \text{if } a_i = a_j \\ 0 & \text{if } a_i < a_j. \end{cases}$$

The unique symmetric equilibrium bidding strategy here is $b(v_i) = v_i/2$. That is, if player i conjectures that j is bidding one-half of his valuation, then player i's best response is to bid one-half of her valuation. This strategy reflects the optimal tradeoff between bidding low to increase the payoff from winning and bidding high to increase the probability of winning.

In determining an optimal strategy, player i must consider what every possible type of player j (each v_j) will do, but since player j's strategy depends on what each possible type of i will do, player i must specify a strategy for each of her possible types, even though she knows which type she is. Player i must not only recognize her own ignorance by conjecturing what each possible j type will do, but recognize (and hopefully exploit) j's ignorance by conjecturing what she would do if she was someone else. Formally, a *strategy* for i is a plan of action for each of i's possible types $s_i\!:\!T_i \rightarrow A_i$. That is, what to do in every possible contingency (each of the possible types).

A strategy profile $s = (s_1,\ldots,s_n)$ is a *Bayesian equilibrium* of Γ if

$$\sum_{t_{-i}\in T_{-i}} p_i(t_{-i}|t_i)u_i[s(t),t] \geq \sum_{t_{-i}\in T_{-i}} p_i(t_{-i}|t_i)u_i[(s_{-i}(t_{-i}),a_{-i}),t] \;\; \forall\, i,\ a_i \in A_i.$$

Existence of a Bayesian equilibrium when the type sets and pure-strategy spaces are finite follows from the standard existence theorem for finite games. Indeed, given consistent beliefs, a Bayesian equilibrium of Γ is simply a Nash equilibrium of the game with imperfect information in which nature moves first. Any game of incomplete information *with consistent beliefs* can be transformed into a standard normal form game.

3. Revelation Principle (Myerson, *Econometrica* 1979 and others)

An equilibrium of a Bayesian game Γ can be represented by a simple equilibrium of a modified Bayesian game Γ' as follows:

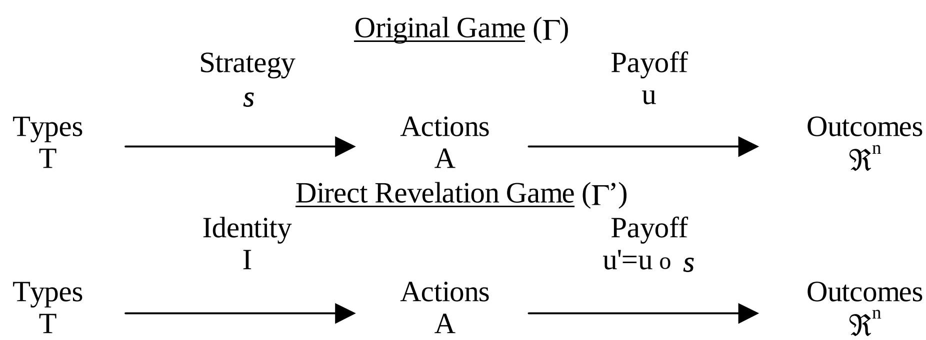
$\Gamma = \{A_1,\ldots,A_n; T_1,\ldots,T_n; p_1,\ldots,p_n; u_1,\ldots,u_n\}$

$\Gamma' = \{A_1',\ldots,A_n'; T_1,\ldots,T_n; p_1,\ldots,p_n; u_1',\ldots,u_n'\}$ where

$A_i' = T_i$ (each player reports her private information (possibly dishonestly))

$u_i'(a',t) = u_i[s(a'),t]$ (by reporting type t_i you get the payoff that t_i gets by playing the equilibrium strategy $s_i(t_i)$ in Γ)

For any Bayesian equilibrium s of Γ, reporting your true type is a Bayesian equilibrium of Γ'. In the game Γ, types are mapped into actions via strategies and then these actions are mapped into outcomes via the utility functions. In the game Γ', types are mapped directly into outcomes, by the composition of the utility and strategy functions. This relationship is shown below:



In the direct revelation game, the players' equilibrium strategy profile is simply the identity map I(t) = t. Consider the auction example in which the equilibrium bid is $b(v_i) = v_i/2$ where v_i is i's valuation. In the direct mechanism Γ', reporting a value (type) of v_i is the same as bidding $v_i/2$ in the original game, so truthful reporting is a best response.

A direct mechanism Γ' in which truthful reporting is a Bayesian equilibrium is call *incentive compatible*. The *revelation principle* states that without loss of generality, the analysis of Bayesian equilibria can be restricted to incentive compatible direct mechanisms.

Typically, the original game involves complex strategies depending on the players' subjective probability beliefs and a simple map of the resulting actions into outcomes; whereas, the direct revelation game involves trivial strategies (the identity map), but complex outcome functions which map the players' reports into outcomes. In some situations, however, there exists a truth-inducing mechanism that has extremely simple outcome functions mapping reports into payoffs. The key to these mechanisms is their clever use of the reported information. Two examples, a public choice problem and an auction problem, are considered.

Example 1: Provision of a Public Good (Groves, *Econometrica* 1973)

A public good is an asset that is enjoyed by many (the public), but for which it is difficult to charge for the use of the asset. Examples include: a road, a bridge, the judicial system, defense, health-care, a library, and a computer center. There, of course, is a free-rider problem here: everyone would like the public good to be provided, but would rather not pay for it. The important questions are: Should a public good be provided? If so, who should pay for it?

The difficulty in answering these questions stems from the fact that often the individuals' valuations of the public good are known privately. How should the decision maker elicit information about valuations? As an example, consider a dean who must decide how much to spend on computers based on the preferences of the faculty. Suppose the dean based the budget on the *mean response* of the faculty. Then individual faculty members would have an incentive to misrepresent their responses by overstating if they wanted a large budget or understating if they wanted a small budget. As a result, the dean might find it difficult to determine the appropriate budget size. Alternatively, the dean might base the budget on the *median response*. Then each member would have an incentive to tell the truth, since a member's information is only used when the dean does what the individual wants. Indeed, truth-telling is a dominant strategy: regardless of what strategy the others' adopt the best response is to tell the truth. (Can you prove this?)

As another example suppose you are one of three households who live on an unpaved private road. You and your neighbors have often discussed the road's shortcomings: the rocks and dips are a nuisance to drivers and cause additional wear and tear on tires and shock absorbers, the mud holes after a rainstorm make walking unpleasant, and cycling is virtually impossible.

The cost of paving the road is c. Requests for county funds to pay for all or part of the road have been repeatedly dismissed, and you do not foresee any change in that position.

You know that the other two households on the road have at least some interest in seeing the road paved, and after some discussion with your neighbors, the three of you mutually agree to give some focused attention to the question of whether the road should be paved.

After some hard thinking you conclude that you value a paved road at an amount v_i; that is, you are willing to pay at most v_i to see the road paved. You also assess that your neighbors' valuations are drawn from some distribution.

You have mutually agreed to meet together to discuss whether the three households should build the road. If the road is built, the cost c must be paid from some combination of funds from the three households; no subsidies from outside sources are available. What procedure should you use in determining whether the road should be built and who should pay how much?

A natural procedure would be for each of you to simultaneously announce your values, pave the road if the announced values exceed the cost, and pay in proportion to the announced values (bids). Unfortunately, this bidding procedure is not incentive compatible. Each household has an incentive to understate its true valuation. In deciding how much to shade your valuation, you will trade off the benefit of understating (pay less if the road is paved) with the cost (a lower probability of paving the road). Of course, this tradeoff depends on what the others are doing. Although it is not obvious, your best response will have the property that the more honest the others are the more you should lie. This misrepresentation leads to inefficiency.

But what if a benefactor suggests the following procedure. Each household simultaneously reports its valuation by making the bid $b_i(v_i)$. If the sum of the bids exceed c, then the road is paved, and each contributes the amount the other players' bids fall short of c (or zero if the others' bids exceed c). That is, letting $b = b_1 + b_2 + b_3$,

$$u_i(b_1, b_2, b_3, v_i) = \begin{cases} 0 & \text{if } b < c \\ v_i - (c - b_j - b_k) & \text{if } b \geq c \text{ and } b_j + b_k < c, \\ v_i & \text{if } b \geq c \text{ and } b_j + b_k \geq c. \end{cases}$$

This procedure, called a *Groves Mechanism*, is important because truth-telling is a dominant strategy. To see why notice that your bid *does not influence how much you pay*, only whether the road is paved. If the others' bids sum to $x \geq c$, then your payoff is v_i regardless of what you bid, so bidding v_i is a best response.

If $x < c$, then your payoff is $v_i - (x - c)$ if $b_i \geq x - c$ and 0 otherwise. Your payoff is maximized by bidding so that the road is paved whenever $v_i - (x - c) \geq 0$, which is accomplished by bidding $b_i = v_i$. Hence, we have a truth-dominant mechanism, which is ex post efficient—the road is paved whenever the value exceeds the cost.

There are, however, two problems with this procedure. First, the sum of the payments $3c - 2b \leq c$, since $b \geq c$, so having a benefactor to make up the deficit is essential. Worse yet, the *more* the households value the road, the *more* the benefactor must contribute. Hence, the players have a strong incentive to collude and *overstate* their valuations, so that the benefactor pays a larger share of the road.

Example 2: Second-Price Auction

Consider again the private-value auction model in which each of n bidders has a valuation v_i for the good being auctioned, where each bidders' valuation is private information. Suppose the seller uses a second-price auction to allocate the good: each bidder simultaneously submits a bid and the good goes to the highest bidder, who pays the seller a price equal to the *second highest* bid. Notice that, as in a Groves mechanism, a player's bid does not influence the terms of trade (if the player wins), but does affect whether the player wins.

Truth (bidding your valuation) is a dominant strategy here as well. Since player i's bid does not influence the price paid if i wins, the optimal bid is such that player i wins whenever the price is less than v_i. But this is accomplished by bidding v_i: by bidding $b_i < v_i$, i stands to lose some profitable opportunities (whenever there is a bid b between b_i and v_i), and by bidding more than v_i, i may lose by winning (whenever there is a bid between v_i and b_i).

B. Bilateral Trading

1. War of Attrition (Bishop, Cannings, and Smith, *J. of Theor. Biology* 1978)

The war of attrition is a concession game in which each player selects an optimal time to concede to the other. It is useful in modeling animal conflict, arms races, strikes, exit in oligopoly, and other disputes. The both-pay auction is an example of a war of attrition with complete information. Here we look at a symmetric model with incomplete information.

Two animals are fighting for a prize (a piece of meat, a mate, etc.). Each knows the value of the prize to himself, but not to the other. The valuations are independent and identically distributed random variables with distribution F and density f on [0,1]. Fighting is costly: each incurs a cost of c for each minute the fight continues. How long should an animal i with valuation v_i wait before conceding?

We wish to determine the symmetric Bayesian equilibrium for this game. Let $t_i(v_i)$ be the stopping time for animal i with valuation v_i. Suppose that the optimal stopping time is strictly increasing in one's valuation. Then let $x_i(t_i) = t_i^{-1}(t_i)$ be the valuation of animal i if it concedes at time t_i. Animal i's payoff is

$$u_i(v_1, v_2, t_1, t_2) = \begin{cases} v_i - ct_j & \text{if } t_j \le t_i \\ -ct_i & \text{if } t_j > t_i. \end{cases}$$

Animal i seeks to maximize her expected utility given animal j's strategy $t_j(\cdot)$; that is for each v_i, t_i is chosen to

$$\max_{t_i} \int_0^{x_{j(t_i)}} [v_i - ct_j] f(v_j) dv_j \quad - ct_i[1 - F(x_j(t_i))].$$

The first-order condition is

$$x_j'(t_i) v_i f(x_j(t_i)) - c[1 - F(x_j(t_i))] = 0.$$

By symmetry, we have $x_j(\cdot) = x_i(\cdot) = x(\cdot)$ and $t_j(\cdot) = t_i(\cdot) = t(\cdot)$, hence the first-order condition can be rewritten as

$$x'(t) = \frac{c[1 - F(x(t))]}{vf(x(t))}$$

But in equilibrium, $x(t) = v$, and $x'(t) = 1/t'(v)$, so

$$t'(v) = \frac{vf(v)}{c[1 - F(v)]}.$$

Hence, the equilibrium strategy is given by

$$t(v) = \int_0^v \frac{zf(z)}{c[1 - F(z)]} dz .$$

2 *Simultaneous Offers* (Chatterjee and Samuelson, *Operations Research* 1983)

A seller and a buyer are engaged in the trade of a single object worth s to the seller and b to the buyer. These valuations are known privately, as summarized below.

Traders	Value	Distributed	Payoff	Private Info	Common Knowledge	Strategy (Offer)
Seller	s	$s \sim F$ on $[\underline{s}, \overline{s}]$	$u = P - s$	s	F, G	p(s)
Buyer	b	$b \sim G$ on $[\underline{b}, \overline{b}]$	$v = b - P$	b	F, G	q(b)

Independent private value model: s and b are independent random variables.

Ex post efficiency: trade if and only if $s < b$. Does private information about valuations lead to inefficiencies?

Game: Each player simultaneously names a price; if $p \leq q$ then trade occurs at the price $P = (p + q)/2$; if $p > q$ then no trade (each player gets zero).

Payoffs: Seller Buyer

$$u(p,q,s,b) = \begin{cases} P - s & \text{if } p \leq q \\ 0 & \text{if } p > q \end{cases} \qquad u(p,q,s,b) = \begin{cases} P - s & \text{if } p \leq q \\ 0 & \text{if } p > q \end{cases}$$

where the trading price is $P = (p + q)/2$.

Example: Let F and G be independent uniform distributions on [0,1]. What are the equilibrium offer functions p(s) and q(b)? Equilibrium conditions:

$$(1)\ \forall\ s \in [\underline{s},\bar{s}],\ p(s) \in \operatorname*{argmax}_p E_b\{u(p,q,s,b) \mid s,q(\cdot)\}$$

$$(2)\ \forall\ b \in [\underline{b},\bar{b}],\ q(b) \in \operatorname*{argmax}_q E_s\{v(p,q,s,b) \mid b,p(\cdot)\}$$

Assume the players' strategies are strictly increasing in their private information, so they are invertible, and let $x(\cdot) = p^{-1}(\cdot)$ and $y(\cdot) = q^{-1}(\cdot)$.

Optimization in (1) can be stated as

$$\max_p \int_{y(p)}^{1} [(p + q(b))/2 - s]db$$

which yields the first-order condition

$$-y'(p)[p - s] + [1 - y(p)]/2 = 0,$$

since $q(y(p)) = p$. Similarly, optimization in (2) can be stated as

$$\max_q \int_0^{x(q)} [b - (p(s) + q)/2]ds$$

which yields the first-order condition

$$x'(q)[b - q] - x(q)/2 = 0,$$

since $p(x(q)) = q$.

The first-order conditions determine each trader's best response to the other's offer function. In equilibrium, one's conjecture about the other's strategy is confirmed. Hence, we require that $s = x(p)$ and $b = y(q)$. The equilibrium first-order conditions then are

(1') $-2y'(p)[p - x(p)] + [1 - y(p)] = 0,$

(2') $2x'(q)[y(q) - q] - x(q) = 0.$

Solving (2') for y(q) and replacing q with p yields

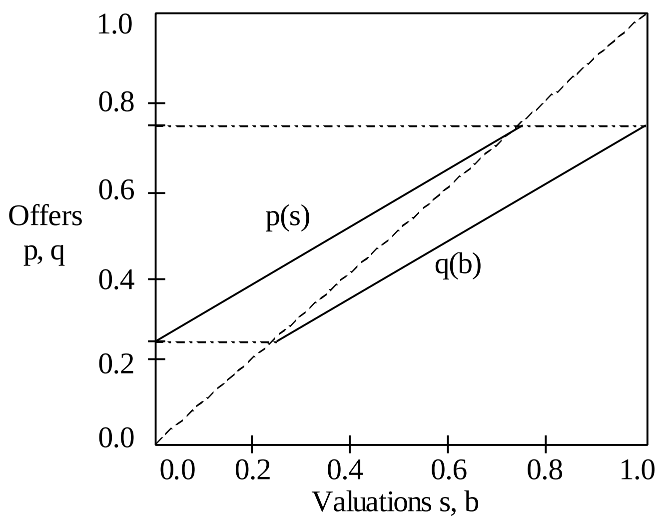
$$(2'') \qquad y(p) = p + \frac{1}{2}\frac{x(p)}{x'(p)}, \quad \text{so} \quad y'(p) = \frac{3}{2} - \frac{1}{2}\frac{x(p)x''(p)}{[x'(p)]^2}.$$

Substituting into (1') then yields

$$(1'') \qquad [x(p) - p]\left[3 - \frac{x(p)x''(p)}{[x'(p)]^2}\right] + \left[1 - p - \frac{1}{2}\frac{x(p)}{x'(p)}\right] = 0.$$

This second-order differential equation in $x(\cdot)$ has a two-parameter family of solutions, including the linear solution $x(p) = \alpha p + \beta$. It is easy to show that $\alpha = 3/2$ and $\beta = -3/8$. Using (2") then yields $y(q) = 3/2\ q - 1/8$. Inverting these functions results in $p(s) = 2/3\ s + 1/4$ and $q(b) = 2/3\ b + 1/12$, as shown in Figure 1.

Figure 1



Trade occurs if and only if $p(s) \leq q(b)$, or $b - s \geq 1/4$. Hence, the gains from trade must be at least 1/4 or no trade takes place, so the outcome is inefficient. Is this inefficiency inevitable? The answer is yes as we will see in a moment.

3. The Public Choice Problem

Returning to the public choice problem, suppose there are two members of society $i \in \{1,2\}$. They must decide whether to undertake a project for which there is no choice of scale $d \in \{0,1\}$. The net benefits of the project to player 1 have a monetary value of $v_i \in (-\infty,\infty)$, where v_i is known privately to i. Ex post efficiency requires that the project be undertaken whenever the net benefits are positive:

$$d^*(v_1,v_2)=\begin{cases}1 & \text{if } v_1+v_2\geq 0,\\ 0 & \text{otherwise.}\end{cases}$$

The players want to design a game to implement this efficient choice rule. They allow the possibility that one should pay the other a transfer as compensation. By the Revelation Principle, they can restrict attention to incentive-compatible direct mechanisms $\{d(v),t(v)\}$, where $v = \{v_1,v_2\}$, $d:\Re^2\to\{0,1\}$ determines the decision as a function of the reports, and $t:\Re^2\to\Re^2$ determines the transfers between the players where $t(v) = \{t_1(v),t_2(v)\}$ and $t_i(v)$ is the transfer that player i receives.

We would like to find a mechanism that satisfies:

(1) efficient social choice: $d(v)\equiv d^*(v)$, and

(2) dominant-strategy incentive compatibility: for all $\hat{v}_j$,

$$v_i \in \underset{\hat{v}_i}{\operatorname{argmax}}\, v_i d(\hat{v}_i,\hat{v}_j)+t_i(\hat{v}_i,\hat{v}_j).$$

The Groves mechanism is such a procedure: if the reported types are $\hat{v}=(\hat{v}_1,\hat{v}_2)$, then

(D) $$d(\hat{v})=\begin{cases}1 & \text{if } \hat{v}_1+\hat{v}_2\geq 0,\\ 0 & \text{otherwise,}\end{cases}$$

and $t_i(\hat{v})=d(\hat{v})\hat{v}_j+h_i(\hat{v}_j)$, for some function $h_i:\Re\to\Re$. Clearly, incentive compatibility implies (1). To see that truth is a dominant strategy, note that v_i solves

$$\max_{\hat{v}_i}\hat{v}_i d(\hat{v}_i,\hat{v}_j)+t_i d(\hat{v}_i,\hat{v}_j)$$

if and only if v_i solves

$$\max_{\hat{v}_i} d(\hat{v}_i,\hat{v}_j)[v_i+\hat{v}_j].$$

But this is the case, since reporting $\hat{v}_i=v_i$, makes $d(\hat{v}_i,\hat{v}_j)$ equal to 1 if and only if $v_i+\hat{v}_j\geq 0$.

The transfer $t_i(\hat{v})$ is constructed so that, given honest reporting by player j, player i faces the social optimization problem. This is what leads to an efficient social choice. Green and Laffont (1977), Holmstrom (1979), and Laffont and Maskin (1980) show that the Groves scheme is the *only* direct mechanism satisfying (1) and (2).

As mentioned before, the problem with this scheme is that the transfers do not satisfy budget balance. There may be a net surplus or deficit. Budget balance requires

$$t_1(v)+t_2(v)=(v_1+v_2)d(v_1,v_2)+h_1(v_2)+h_2(v_1)=0, \text{ or}$$

$$h_1(v_2)+h_2(v_1)=\begin{cases}-(v_1+v_2) & \text{if } v_1+v_2\geq 0,\\ 0 & \text{otherwise.}\end{cases}$$

But this cannot happen if h_i is independent of v_i.

One way to guarantee budget balance is to weaken the incentive compatibility criterion, so that truth is merely a Bayesian equilibrium rather than a dominant strategy:

(2') Bayesian incentive compatibility:

$$v_i \in \underset{\hat{v}_i}{\operatorname{argmax}} E_{\hat{v}_j}[v_i d(\hat{v}_i, v_j) + t_i d(\hat{v}_i, v_j) \mid v_i]$$

Arrow (1979) and d'Aspremont and Gerard-Varet (1979) show that replacing (2) with (2') allows use to satisfy

(3) budget balance: $t_1(v) + t_2(v) = 0$ for all v.

Consider the Bayesian game $\Gamma = \{A_1, A_2; V_1, V_2; p_1, p_2; u_1, u_2\}$, where $A_i = V_i = \Re$, $u_i(a,v) = v_i d(a) + t_i(a)$, and $p_i(v_j \mid v_i) = f_j(v_j)$, so types are independent. We wish to construct a mechanism satisfying (1), (2'), and (3). Our decision rule must be as in (D) to satisfy (1). Consider the transfers $t_i(v) = g_i(v_i) - g_j(v_j)$, where

$$g_i(v_i) = \int_{-\infty}^{\infty} v_j d(v_i, v_j) f_j(v_j) dv_j \quad .$$

Clearly, the transfers balance, but is truth telling a Bayesian equilibrium? Player i chooses the report $\hat{v}_i$ to solve

$$\max_{\hat{v}_i} \int_{-\infty}^{\infty} [v_i d(\hat{v}_i, v_j) + t_i(\hat{v}_i, v_j)] f_j(v_j) dv_j \quad .$$

Substituting the definition of $t_i(\cdot)$ yields

$$\max_{\hat{v}_i} \int_{-\infty}^{\infty} (v_i + v_j) d(\hat{v}_i, v_j) f_j(v_j) dv_j \quad ,$$

which is equivalent to

$$\max_{\hat{v}_i} \int_{-\hat{v}_i}^{\infty} (v_i + v_j) f_j(v_j) dv_j \quad ,$$

when the decision rule in (D) is used. The first-order condition for this last optimization is $(v_i - \hat{v}_i) f_j(-\hat{v}_i) = 0$, so $\hat{v}_i = v_i$.

Unfortunately, there is a problem with this mechanism, too: the players knowing their types may not wish to participate in the mechanism; that is, they may be better off with the status quo (zero) than their expected payoff from participation. Formally, we wish to require

(4) interim individual rationality:

$$U_i(v_i) = \int_{-\infty}^{\infty} [v_i d(v_i, v_j) + t_i(v_i, v_j)] f_j(v_j) dv_j \quad \geq 0 \text{ for all } v_i \in V_i.$$

Substituting for $d(\cdot)$ and $t(\cdot)$ yields:

$$U_i(v_i) = \int_{-\hat{v}_i}^{\infty} (v_i + v_j) f_j(v_j) dv_j \quad - \int_{-\infty}^{\infty} v_i f_i(v_i)[1 - F_i(-v_i)] dv_i$$

$$= v_i[1 - F_j(-v_i)] + \int_{-\hat{v}_i}^{\infty} v_j f_j(v_j) dv_j \quad - \int_{-\infty}^{\infty} v_i f_i(v_i)[1 - F_j(-v_i)] dv_i \quad .$$

Note that $U_i'(v_i) = 1 - F_j(-v_i) \geq 0$, so if interim individual rationality fails, it fails for the lowest values of v_i. Assume that the means and variances of v_i and v_j are finite. Then as $v_i \rightarrow -\infty$ the first and second terms approach zero. So if the integral in the third term is positive (for which it is sufficient that the mean of v_i be positive) then for sufficiently low values of v_i, $U_i(v_i) < 0$.

Much of the mechanism design literature adopts the perspective that (2') incentive compatibility, (3) budget balance, and (4) individual rationality are necessary conditions and then optimize over this feasible set in an attempt to get close to (1).

4. A General Model (Myerson and Satterthwaite, *JET* 1983)

We now turn to a direct revelation game analysis of the Chatterjee and Samuelson model of bilateral exchange. The seller's valuation s (known privately) is distributed according to the distribution F with positive density f on $[\underline{s}, \bar{s}]$; similarly, the buyer's private valuation b is drawn (independently from F) from the distribution G with positive density g on $[\underline{b}, \bar{b}]$. F and G are common knowledge. In the direct revelation game, the traders report their valuations and then an outcome is selected. Given the reports (s,b), an outcome specifies a probability of trade and the terms of trade. Specifically, a direct mechanism is a pair of outcome functions $\langle p,x \rangle$, where p(s,b) is the probability of trade given the reports (s,b), and x(s,b) is the expected payment from the buyer to the seller. Given reports (s,b), the seller's ex post utility is u(s,b) = x(s,b) - sp(s,b) and the buyer's ex post utility is v(s,b) = bp(s,b) - x(s,b); i.e., both traders are risk neutral and there are no income effects. The traders' reports are chosen to maximize their interim utility from the mechanism. It is helpful to break the interim utilities into two components, depending on the expected price and the expected probability of trade. In particular, for the mechanism $\langle p,x \rangle$ define:

$$X(s) = \int_{\underline{b}}^{\overline{b}} x(s,b)g(b)db \qquad Y(b) = \int_{\underline{s}}^{\overline{s}} x(s,b)f(s)ds$$

$$P(s) = \int_{b}^{\overline{b}} p(s,b)g(b)db \qquad Q(b) = \int_{\underline{s}}^{\overline{s}} p(s,b)f(s)ds.$$

where X(s) is the seller's expected revenue given the report s, Y(b) is the buyer's expected payment given the report b, P(s) is the seller's probability of trade, and Q(b) is the buyer's probability of trade. The seller's and buyer's interim utilities (U(s) and V(b)) are then given by

$$U(s) = X(s) - sP(s) \qquad V(b) = bQ(b) - Y(b).$$

The mechanism $\langle p,x \rangle$ is *incentive compatible* if for all $s,\ s' \in [\underline{s}, \overline{s}]$ and $b,\ b' \in [\underline{b}, \overline{b}]$

(IC) $\quad U(s) \geq X(s') - sP(s') \qquad V(b) \geq bQ(b') - Y(b').$

The mechanism $\langle p,x \rangle$ is *individually rational* if for all $s \in [\underline{s}, \overline{s}]$ and $b \in [\underline{b}, \overline{b}]$

(IR) $\quad U(s) \geq 0 \qquad V(b) \geq 0.$

Notice that this is *interim* individual rationality, not *ex post* IR: the traders are allowed to refuse to participate after they know their type but before the game begins. Once they agree to the mechanism, they are *committed* to follow it even if the outcome leads to an ex post loss for one or both of the players. There is no guarantee that $u(s,b) \geq 0$ and $v(s,b) \geq 0$ for all (s,b).

The following lemma is fundamental throughout the self-selection literature. It depends, however, on two limiting assumptions: independent valuations and utility that is linear in money and the good.

Lemma 1 (Mirrlees, Myerson): The mechanism $\langle p,x \rangle$ is IC if and only if $P(\cdot)$ is decreasing, $Q(\cdot)$ is increasing, and

(IC') $\quad U(s) = U(\overline{s}) + \int_{s}^{\overline{s}} P(t)dt \qquad V(b) = V(\underline{b}) + \int_{\underline{b}}^{b} Q(t)dt \quad .$

Proof: *Only if.*

By definition, U(s) = X(s) - sP(s) and U(s') = X(s') - s'P(s'). This and (IC) imply

$U(s) \geq X(s') - sP(s') = U(s') + (s' - s)P(s')$, and

$U(s') \geq X(s) - s'P(s) = U(s) + (s - s')P(s).$

Putting these inequalities together yields

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

Taking s' > s implies that $P(\cdot)$ is decreasing. Dividing by (s' - s) and letting $s' \rightarrow s$, then yields dU(s)/ds = -P(s) a.e., and integrating produces (IC'). The same is true for the buyer.

If. To prove (IC) for the seller, note that it suffices to show that

$$s[P(s) - P(s')] + [X(s') - X(s)] \le 0 \text{ for all } s,\ s' \in [\underline{s}, \bar{s}].$$

Substituting for X(s') and X(s) using (IC') and the definition of U(s) yields

$$X(s) = sP(s) + U(\bar{s}) + \int_s^{\bar{s}} P(t)dt \quad .$$

Then it suffices to show for every $s,\ s' \in [\underline{s}, \bar{s}]$ that

$$0 \ge s[P(s) - P(s')] + s'P(s') + \int_{s'}^{\bar{s}} P(t)dt \quad - sP(s) - \int_s^{\bar{s}} P(t)dt$$

$$= (s' - s)P(s') + \int_{s'}^{s} P(t)dt \quad = \int_{s'}^{s} [P(t) - P(s')]dt \quad ,$$

which holds because $P(\cdot)$ is decreasing. The proof for the buyer is similar.¦

The next lemma states that given an incentive compatible mechanism, we need only check IR for the highest seller type and lowest buyer type.

Lemma 2. An incentive compatible mechanism $\langle p,x \rangle$ is individually rational if and only if

(IR') $\quad U(\bar{s}) \ge 0 \quad$ and $\quad V(\underline{b}) \ge 0.$

Proof. Clearly, (IR') is necessary for $\langle p,x \rangle$ to be IR. By Lemma 1, $U(\cdot)$ is decreasing; hence, (IR') is sufficient as well. ¦

Corollary. An incentive-compatible, individually rational mechanism $\langle p,x \rangle$ satisfies

(*) $$U(\bar{s}) + V(\underline{b}) = \int_{\underline{b}}^{\bar{b}} \int_{\underline{s}}^{\bar{s}} \left[b - \frac{1 - G(b)}{g(b)} - s - \frac{F(s)}{f(s)} \right] p(s,b)f(s)g(b)dsdb \ge 0.$$

Proof. Using (IC') and the definition of U(s) yields

$$X(s) = sP(s) + U(\bar{s}) + \int_s^{\bar{s}} P(t)dt.$$

Taking the expectation with respect to s (and substituting in the definitions of X(s) and P(s)) shows that

$$\int_{\underline{b}}^{\bar{b}} \int_{\underline{s}}^{\bar{s}} x(s,b)f(s)g(b)dsdb = U(\bar{s}) + \int_{\underline{b}}^{\bar{b}} \int_{\underline{s}}^{\bar{s}} sp(s,b)f(s)g(b)dsdb$$
$$+ \int_{\underline{b}}^{\bar{b}} \int_{\underline{s}}^{\bar{s}} p(s,b)F(s)g(b)dsdb \quad .$$

The third term in the right hand side follows, since

$$\int_{\underline{s}}^{\bar{s}} \int_{s}^{\bar{s}} p(t,b)f(s)dtds \quad = \int_{\underline{s}}^{\bar{s}} \int_{\underline{s}}^{t} p(t,b)f(s)dsdt \quad = \int_{\underline{s}}^{\bar{s}} p(t,b)F(s)ds.$$

Preceding analogously for the buyer yields

$$\int_{\underline{b}}^{\bar{b}}\int_{\underline{s}}^{\bar{s}} x(s,b)f(s)g(b)dsdb \quad = \quad -V(\underline{b}) + \int_{\underline{b}}^{\bar{b}}\int_{\underline{s}}^{\bar{s}} bp(s,b)f(s)g(b)dsdb$$

$$- \int_{\underline{b}}^{\bar{b}}\int_{\underline{s}}^{\bar{s}} p(s,b)f(s)[1-G(b)]dsdb.$$

Equating the righthand sides of the last two equations and applying (IR') completes the proof.¦

Theorem. If it is not common knowledge that gains exist (the supports of the traders' valuations have non-empty intersection), then no incentive-compatible, individually rational trading mechanism can be ex-post efficient.

Proof. A mechanism is ex-post efficient if and only if trade occurs whenever $s \le b$:

$$p(s,b) = \begin{cases} 1 & \text{if } s \le b \\ 0 & \text{if } s > b. \end{cases}$$

To prove that ex-post efficiency cannot be attained, it suffices to show that the inequality (*) in the Corollary fails when evaluated at this p(s,b). Hence,

$$\int_{\underline{b}}^{\bar{b}}\int_{\underline{s}}^{\min\{b,\bar{s}\}} \left[b - \frac{1-G(b)}{g(b)} - s - \frac{F(s)}{f(s)} \right] f(s)g(b)dsdb$$

$$= \int_{\underline{b}}^{\bar{b}}\int_{s}^{\min\{b,\bar{s}\}} [bg(b) + G(b) - 1]f(s)dsdb - \int_{\underline{b}}^{\bar{b}}\int_{s}^{\min\{b,\bar{s}\}} [sf(s) + F(s)]dsg(b)db$$

$$= \int_{\underline{b}}^{\bar{b}} [bg(b) + G(b) - 1]F(b)db - \int_{\underline{b}}^{\bar{b}} \min\{bF(b), \bar{s}\}g(b)db$$

$$= -\int_{\underline{b}}^{\bar{b}} [1-G(b)]F(b)db + \int_{\bar{s}}^{\bar{b}} (b-\bar{s})g(b)db$$

$$= -\int_{\underline{b}}^{\bar{b}} [1-G(b)]F(b)db + \int_{\bar{s}}^{\bar{b}} [G(b)-1)db$$

$$= -\int_{\underline{s}}^{\bar{s}} [1-G(t)]F(t)dt < 0, \text{ since } \underline{b} < \bar{s}.$$

The second term in the third line follows, since by integrating by parts

$$\int_{\underline{s}}^{x} [sf(s) + F(s)]ds = xF(x).$$

Since ex-post efficiency is unattainable, we need a weaker efficiency criterion with which to measure a mechanism's performance.

5. Efficiency in Games with Incomplete Information

(Holmstrom and Myerson, *Econometrica* 1983)

The fundamental notion of efficiency is *Pareto optimality*. An allocation is Pareto optimal if there does not exist an alternative allocation that makes no parties worse off and at least one party strictly better off. A natural extension of Pareto efficiency to games of incomplete information is something like: "a decision rule is efficient if and only if no other *feasible* decision rule can be *found* that may make some individuals *better off* without ever making any other individual worse off." There are three ambiguities in this seemingly straight-forward definition:

(1) What is meant by a feasible decision rule? Are we to recognize incentive constraints?

(2) What is meant by better off or worse off? Since there is uncertainty, expected utility is the relevant criterion, but on what information should the expectation be conditioned? Three alternatives are:

  (a) *Ex ante information*: a planner's information at the beginning of the game (no knowledge of types).

  (b) *Interim information*: a player's private information at the beginning of the game (each player knows her own type t_i).

  (c) *Ex post information*: all the private information (the vector of types t is known).

(3) Who is to "find" the potentially better decision rule, and at what information stage? If a player proposes a particular decision rule after learning her private information, the other players may infer something about the player's type from the information.

Consider a Bayesian game $\Gamma = \{A_1,...,A_n;T_1,...,T_n;p_1,...,p_n;u_1,...u_n\}$, where each action set A_i and type set T_i is finite, and the beliefs p_i are consistent. Let D be the set of probability distributions over $A = A_1 \times ... \times A_n$. A decision rule (or direct mechanism) $\delta:T \to D$ maps reports into a randomization over feasible actions. The utility function $u_i(d,t):D \times T \to \Re$ maps the decision and types into payoffs. A decision rule $\delta \in \Delta \equiv \{\delta:T \to D\}$ is incentive compatible if for all i and $t_i \in T_i$

(IC) $$\sum_{t_{-i} \in T_{-i}} p_i(t_{-i}|t_i)u_i(d(t),t) \geq \sum_{t_{-i} \in T_{-i}} p_i(t_{-i}|t_i)u_i(d(t_{-i},\hat{t}_i),t)$$

for all $\hat{t}_i \in T_i$. Let $\Delta^* = \{\delta:T \times D \to \Re\}$ be the set of all incentive-compatible decision rules. By the revelation principle, we can restrict attention to $\delta \in \Delta^*$.

Our interpretation is that given a decision rule δ, the players report their types to a central planner, who then requires them to carry out the actions specified by the decision rule given the reports. If the central planner was not able to enforce the decision rule by making sure that the actions specified by the decision rule are carried out, say because actions are unobservable, then we would need to include *obedience* constraints in addition to the *honesty* constraints given by (IC). In the remainder of this section, we will assume that obedience is not a problem.

For a decision rule $\delta(\cdot)$, the expected utility at each of the three information stages are

(1) *Ex Ante Utility*: $U_i(\delta) = \sum_{t \in T} p(t) u_i(\delta(t), t)$,

(2) *Interim Utility*: $U_i(\delta | t_i) = \sum_{t_{-i} \in T_{-i}} p_i(t_{-i} | t_i) u_i(\delta(t), t)$, and

(3) *Ex Post Utility*: $U_i(\delta | t) = u_i(\delta(t), t)$.

These three evaluation measures lead to three different notions of domination:

(1) γ *ex ante dominates* δ iff $U_i(\gamma) \geq U_i(\delta)\ \forall\ i$

with at least one strict inequality; (Before you know your type, do you prefer γ to δ?)

(2) γ *interim dominates* δ iff $U_i(\gamma|t_i) \geq U_i(\delta|t_i)\ \forall$ i and $t_i \in T_i$

with at least one strict inequality; (After you know your type, do you prefer γ to δ?)

(3) γ *ex post dominates* δ iff $U_i(\gamma|t) \geq U_i(\delta|t)\ \forall$ i and $t \in T$

with at least one strict inequality. (After all types are known, do you prefer γ to δ?)

There are six concepts of efficiency depending on the information stage (ex ante, interim, or ex post) and on whether the incentive constraints are recognized (whether the feasible set is Δ or Δ^*). Of these, three are the most important (in order from strongest to weakest):

(1) δ is *ex post (classically) efficient* iff there does not exist $\gamma \in \Delta$ that ex post dominates δ.

(2) δ is *ex ante (incentive) efficient* iff there does not exist $\gamma \in \Delta^*$ that ex ante dominates δ.

(3) δ is *interim (incentive) efficient* iff there does not exist $\gamma \in \Delta^*$ that interim dominates δ.

Ideally, we would like our decision rule to be ex post efficient: for all realizations of the state of nature, no player can be make better off without making another worse off (ignoring incentive constraints); however, when incentive constraints are recognized, ex post efficiency is often unattainable. Ex ante efficiency, on the other hand, is a sensible goal for a social planner that does not have any private information. If the players are selecting a decision rule after they have their private information, then interim efficiency is a sensible requirement.

Since ex post efficiency is not possible in the Myerson and Satterthwaite trading game, it is natural to determine the ex ante efficient mechanism that maximizes the expected gains from trade:

$$\int_{\underline{s}}^{\bar{s}} U(s)f(s)ds + \int_{\underline{b}}^{\bar{b}} V(b)g(b)db$$

over all incentive compatible decision rules. Myerson and Satterthwaite show that the ex ante efficient decision rule (probability of trade) is

$$p^{\alpha}(s,b) = \begin{cases} 1 & \text{if } c(s,\alpha) \le d(b,\alpha) \\ 0 & \text{if } c(s,\alpha) > d(b,\alpha), \end{cases}$$

where

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

and α is chosen so that U(s) = V($\underline{b}$) = 0. Notice that if $\alpha = 0$, then p^{α} is ex post efficient and that if $\alpha = 1$, p^{α} maximizes the expression in (*).

Also, notice that the ex ante efficient trading rule has the property that, given the reports, trade either occurs with probability one or not at all. This has an interesting interpretation if we think of a dynamic trading mechanism $\langle t,x\rangle$, rather than static trading rule $\langle p,x\rangle$, where t(s,b) is the time of trade as a function of the reports and x(s,b) is the payment from buyer to seller.[2] If the traders discount future trades at the discount rate r, then the dynamic mechanism $\langle t,x\rangle$ is equivalent to the static mechanism $\langle p,x\rangle$ where $p(s,b) = e^{-rt(s,b)}$. Trading with probability one is the same as trading without delay (at time t = 0). Hence, ex ante efficiency requires that the traders trade immediately or not at all. But such a trading mechanism violates sequential rationality, since after the first instant if the traders did not trade, they still have an incentive to continue bargaining because not all the gains from trade were realized in the first instant. A natural extension of perfection to the dynamic mechanism $\langle t,x\rangle$ is that it is never common knowledge that the mechanism induced by $\langle t,x\rangle$ at any time is dominated by an alternative mechanism. Unfortunately, this characterization of perfection has proven to be intractable.

When the traders' valuations are uniformly distributed on [0,1], then the ex ante efficient mechanism is the linear equilibrium in the simultaneous offer model of Chatterjee and Samuelson, in which trade occurs if and only if the gains from trade are at least 1/4. If the traders cannot commit to walking away from gains from trade, then they would be unable to implement this mechanism. So long as it is not common knowledge that gains exist, the traders will, with positive probability, make incompatible demands in situations where gains from trade exist. What is keeping the traders from continuing negotiations in this

[2]See Cramton [1985], "Sequential Bargaining Mechanisms," in *Game Theoretic Models of Bargaining*, edited by Alvin Roth.

case?

It is worth summarizing the accomplishments of the direct revelation analysis in this model: (1) a characterization of the set of all Bayesian equilibria of all bargaining games in which the players' strategies map their private valuations into a probability of trade and a payment from buyer to seller, (2) a proof that ex post efficiency is unattainable if it is uncertain that gains from trade exist, (3) a determination of the set of ex ante efficient mechanisms, (4) a proof that ex ante efficiency is incompatible with sequential rationality: the players must be able to commit to walk away from gains from trade in order to achieve ex ante efficiency.

6. Durability (Holmstrom and Myerson, 1983)

We have argued that a decision rule is feasible only if it is compatible with the players incentives ($\delta \in \Delta^*$). Are there further restrictions on the feasible set of decision rules δ that should be made? One such restriction, mentioned above, is that if the players have limited abilities to make binding commitments, then this limitation poses a further restriction on the set of feasible decision rules.

A second restriction on the set of feasible decision rules can come from the process of deciding on which decision rule to implement. Suppose that an ex ante efficient decision rule δ is proposed. Is it ever the case that a player, knowing her private information, could suggest an alternative decision rule γ that the others would surely prefer? The answer is yes, as can be seen from the following example. Each of two players, 1 and 2, is of one of two types, a or b, with all four possible combinations of types equally likely. The players' utilities as a function of their types and which of three possible decisions {A,B,C} are shown below.

	1a	1b	2a	2b
d = A	2	0	2	2
d = B	1	4	1	1
d = C	0	9	0	-8

The ex ante efficient decision rule that maximizes the sum of the players' payoffs is:

$\delta(1a,2a) = A$ $\delta(1a,2b) = B$

$\delta(1b,2a) = C$ $\delta(1b,2b) = B$.

No outsider could suggest an alternative decision rule that would make some type better without making another type worse off. But if player 1's type is 1a, then player 1 can suggest to 2 that decision A be adopted, and 2 would surely accept such a proposal. In the words of Holmstrom and Myerson, the decision rule δ is not durable. A decision rule is *durable* iff the players would never unanimously approve a change to any other decision rule.

C. Multilateral Trading Mechanisms

1. Dissolving a Partnership (Cramton, Gibbons, and Klemperer, 1987)

Consider generalizing the Myerson and Satterthwaite (MS) problem to the case of n traders who share in the ownership of a single asset. Specifically, each trader $i \in \{1,...,n\}$ owns a share $r_i \geq 0$ of the asset, where $r_1 + ... + r_n = 1$. As in MS, player i's valuation for the entire good is v_i, and the utility from owning a share r_i is $r_i v_i$, measured in monetary terms. The v_i's are independent and identically distributed according to $F(\cdot)$ on $[\underline{v}, \overline{v}]$. A *partnership* (r,F) is fully described by the vector of ownership rights $r = \{r_1,...,r_n\}$ and the traders' beliefs F about valuations.

MS consider the case n = 2 and r = {1,0}. They show that there does not exist a Bayesian equilibrium s of the trading game such that (1) s is (interim) individually rational and (2) s is ex post efficient. In contrast, we show that if the ownership shares are not too unequally distributed, then it is possible to satisfy both (1) and (2). As in MS, the definition of a Bayesian trading game includes budget balance, so this possibility result satisfies all of the criteria described in the public-choice problem.

In addition to exploring the MS impossibility result, this paper considers the dissolution of partnerships, broadly construed. In a situation of joint ownership, who should buy out whom and at what price? Applications include divorce and estate fair-division problems, and also public choice. For example, when several towns jointly need a hazardous-waste dump, which town should provide the site and how should it be compensated by the others?

In this context, ex post efficiency means giving the entire good to the partner with the highest valuation. We will say that a partnership (r,F) can be *dissolved efficiently* if there exists a Bayesian equilibrium s of a Bayesian trading game such that s is interim individually rational and ex post efficient.

Theorem: The partnership (r,F) can be dissolved efficiently if and only if

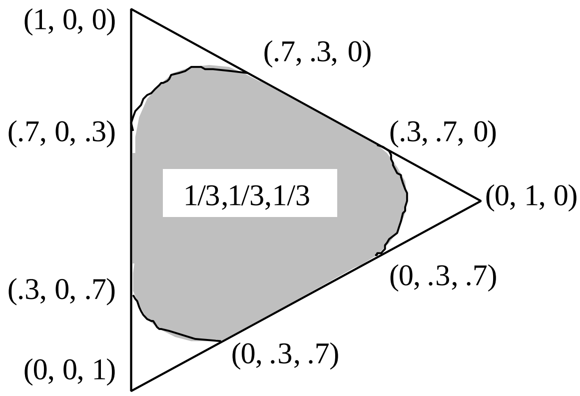
$$(*) \qquad \sum_{i=1}^{n}\left[\int_{v_i^*}^{\overline{v}}[1-F(u)]u\,dG(u) - \int_{\underline{v}}^{v_i^*} F(u)u\,dG(u)\right] \geq 0$$

where v_i^* solves $F(v_i)^{n-1} = r_i$ and $G(u) = F(u)^{n-1}$.

The proof is an application of the tools used in the MS analysis and so is omitted. As an example, consider the case of n=3, $F(v_i) = v_i$. Then (*) becomes

$$\sum_{i=1}^{3} r_i^{3/2} \leq 3/4,$$

which is illustrated in the figure below.



Proposition: For any distribution F, the one-owner partnership r = {1,0,0,...,0} cannot be dissolved efficiently.

The second proposition generalizes the MS impossibility result to the case of many buyers. The one-owner partnership can be interpreted as an auction. Ex post efficiency is unattainable because the seller's reservation value v_1 is private information: the seller finds it in her best interest to set a reservation price above her value v_1. An optimal auction maximizes the seller's expected revenue over the set of feasible (ex post inefficient) mechanisms.

The proposition also speaks to the time-honored tradition of solving complex allocation problems by resorting to lotteries: even if the winner is allowed to resell the object, such a scheme is inefficient because the one-owner partnership that results from the lottery cannot be dissolved efficiently.

One problem with the direct mechanisms is that most of the computational burden is on the mechanism designer, who must consider the incentive-compatibility constraints, which depend in a complicated way on the players' probability beliefs. The players, in contrast, simply tell the truth. Most observed (indirect) mechanisms allocate the burden differently: the rules of the game are simple, and the players solve complex problems to compute equilibrium strategies. In this spirit, consider the following equivalent indirect mechanism.

Theorem: If a partnership (r,F) can be dissolved efficiently, then the unique symmetric equilibrium of the following bidding game is interim individually rational and achieves ex-post efficiency: given an arbitrary minimum bid $\underline{b}$,

(a) the players choose bids $b_i \in [\underline{b},\infty)$;

(b) the good goes to the highest bidder;

(c) *each* bidder i pays

$$p_i(b_1,\ldots,b_n) = b_i - \frac{1}{n-1}\sum_{j\neq i}^{n} b_j \text{ ; and}$$

(d) each player receives a side-payment, independent of the bidding,

$$c_i(r_1,\ldots,r_n) = \int_{\underline{v}}^{v_i^*} u dG(u) - \tfrac{1}{n}\sum_{j=1}^{n}\int_{\underline{v}}^{v_j^*} u dG(u).$$

Note that the side-payments balance, as do the payments made as a function of the bids. The side-payments are used to compensate large shareholders, who are effectively dispossessed in the bidding game since the prices p_i are independent of the shares r_i and so treat shareholders alike. They can be thought of as entry fees for those bringing a small share to the auction, and bonuses for those bringing a large share. The side-payments are zero if each of the n partners owns share 1/n.

Also note that several bidders may pay positive prices, but not win anything. This feature of the auction makes it less attractive if the bidders are risk averse. As in a standard auction, a higher bid buys the bidder a larger probability of winning. Here, however, making a higher bid is like buying more lottery tickets in that the purchase price of losing tickets is not refunded.

2. Optimal Auctions (Myerson, 1981)

This paper analyzes the problem faced by a seller who owns one (indivisible) object and would like to sell it to one of n possible buyers. Each buyer's willingness to pay for the object is private information. The seller wishes to induce the buyers to participate in the auction that maximizes the expected revenue the seller collects.

Let the n buyers be indexed by $i \in \{1,...,n\}$. Let each buyer i's willingness to pay for the object be $t_i \in [\underline{a}_i, \bar{a}_i]$, and let i's type be distributed independently according to the (strictly positive) density $f_i(\cdot)$. The seller's type, t_0, is *common knowledge* (unlike in the partnership analysis above).

A *Bayesian auction* consists of bids spaces $\{B_1,...,B_n\}$ and outcome functions $\tilde{p}_i : B \rightarrow [0,1]$ and $\tilde{x}_i B \rightarrow \Re$, where $B = B_1 \times ... \times B_n$, $\tilde{p}_i$ is the probability that player i gets the object when the bids are $b = \{b_1,...,b_n\} \in B$, and $\tilde{x}_i$ is the payment from i to the seller when the bids are $b \in B$. For each b,

$$\sum_{i=1}^{n} \tilde{p}_i \leq 1,$$

which allows for the possibility that the seller may keep the object. Also, note that the $\tilde{x}_i$ functions are allowed to take on both positive and negative values, and they are *not* constrained to be zero when i is not the winning bidder.

Given these outcome functions, the utility functions $\tilde{u}_i(b,t)$ are

$$\tilde{u}_i(b,t) \equiv t_i \tilde{p}_i(b) - \tilde{x}_i(b),$$

where $t = \{t_1,...,t_n\}$. These utility functions complete the description of the Bayesian game $\Gamma = \{B_1,...,B_n; T_1,...,T_n; f_1,...,f_n; \tilde{u}_1,..., \tilde{u}_n\}$. A *strategy* for bidder i in this game is $b_i:T_i \rightarrow B_i$. A strategy profile $b=\{b_1,...,b_n\}$ is a *Bayesian equilibrium* if for each $t_i \in T_i$, the prescribed bid $b_i(t_i)$ is a best response to the $n - 1$ other strategies b_{-i}.

The seller wishes to design a Bayesian auction (i.e., choose $\{B_i, \tilde{p}_i, \tilde{x}_i\}$ to maximize the expected revenue

$$E_b\left\{\left[1-\sum_{i=1}^{n}\tilde{p}_i(b)\right]t_0 + \sum_{i=1}^{n}\tilde{x}_i(b)\right\}$$

with respect to the distribution of bids implied by an equilibrium strategy profile b(t) and the joint density $f(t) \equiv f_1(t_1)\cdot\ldots\cdot f_n(t_n)$, subject to the constraint that each bidder receive non-negative expected utility from participating in the auction: for each i,

$$E_{b_{-i}}\{t_i\tilde{p}_i(b_i,b_{-i}) - \tilde{x}_i(b_i,b_{-i}) | b_i = b_i(t_i)\} \geq 0.$$

Rather than optimize over the complex space of Bayesian auctions, one can appeal to the Revelation Principle, which states that any Bayesian equilibrium in any Bayesian game can be represented as a truth-telling equilibrium in a direct revelation game, in which the players submit claims about their types rather than bids. The seller's problem then is to choose outcome functions $p_i(t)$ and $x_i(t)$, which determine who gets the good and the payments between players as a function of the reports t, to

$$\text{(ER)} \qquad \max\int_T\left\{t_0\left(1-\sum_{i=1}^{n}p_i(t)\right)+\sum_{i=1}^{n}x_i(t)\right\}f(t)dt$$

subject to the constraints that

1. for all i, for all t, $p_i(t) \geq 0$ and $p_1(t) + ... + p_n(t) \leq 1$,
2. truth-telling is a Bayesian equilibrium, and
3. each player's expected utility is non-negative.

To express the last two constraints precisely, define

$$v_i(\tau_i,t_i) \equiv \int_{T_{-i}} u_i(\tau_i,t_{-i},t_i,t_{-i})f_{-i}(t_{-i})dt_{-i},$$

which is bidder i's expected utility from reporting τ_i when i's true type is t_i and the other players are conjectured to report their true types. Then we have:

(IC) $V_i(t_i) \equiv v_i(t_i,t_i) \geq v_i(\tau_i,t_i)$ for all i and for all $\tau_i, t_i \in T_i$, and

(IR) $V_i(t_i) \geq 0$ for all i.

The seller's problem of maximizing (ER) subject to (IC), (IR), and feasibility is considerably simplified by the following two lemmas. The first lemma involves one additional definition:

$$P_i(t_i) \equiv \int_{T_{-i}} p_i(t) f_{-i}(t_{-i}) dt_{-i}$$

is the conditional probability that player i gets the object when i's type is t_i.

Lemma 1: $\{p_i(\cdot),x_i(\cdot)\}$ satisfies (IC) and (IR) iff for all i

(i) $P_i(t_i)$ is weakly increasing,

(ii) $V_i(t_i) = V_i(\underline{a}_i) + \int_{\underline{a}_i}^{t_i} P_i(\mathit{t}_i) d\mathit{t}_i$ for all $t_i \in T_i$, and

(iii) $V_i(\underline{a}_i) \geq 0$.

The proof is as in the last section. Substituting (ii) and the definition of $V_i(t_i)$ into (ER) and changing the order of integration yields Lemma 2.

Lemma 2: If $\{p_i(\cdot),x_i(\cdot)\}$ satisfies (IC) and (IR), then (ER) becomes

$$\text{(ER')} \qquad t_0 - \sum_{i=1}^{n} V_i(\underline{a}_i) + \int_T \left[\sum_{i=1}^{n} \left\{ t_i - \frac{1-F_i(t_i)}{f_i(t_i)} - t_0 \right\} p_i(t) \right] f(t)dt.$$

Proof: By definition,

$$V_i(t_i) = \int_{T_{-i}} [t_i p(t_i, t_{-i}) - x_i(t_i, t_{-i})] f_{-i}(t_{-i}) dt_{-i},$$

and by (ii) and the definition of $P_i(t_i)$,

$$V_i(t_i) = V_i(\underline{a}_i) + \int_{\underline{a}_i}^{t_i} P_i(\mathit{t}_i) d\mathit{t}_i$$

$$= V_i(\underline{a}_i) + \int_{\underline{a}_i}^{t_i} P_i(\mathit{t}_i, t_{-i}) f_{-i}(t_{-i}) dt_{-i} d\mathit{t}_i.$$

Rearranging terms yields

$$\int_{T_{-i}} x_i(t) f_{-i}(t_{-i}) dt_{-i} = -V_i(\underline{a}_i) + \int_{T_{-i}} \left[t_i p_i(t) - \int_{\underline{a}_i}^{t_i} p_i(\mathit{t}_i, t_{-i}) d\mathit{t}_i \right] f_{-i}(t_{-i}) dt_{-i}$$

and integrating with respect to $f_i(t_i)$ produces

$$\int_T x_i(t)f(t)dt = -V_i(\underline{a}_i) + \int_T \left[t_i p_i(t) - \int_{\underline{a}_i}^{t_i} p_i(\mathit{t}_i, t_{-i})d\mathit{t}_i \right] f(t)dt$$

$$= -V_i(\underline{a}_i) + \int_T \left[t_i - \frac{1-F_i(t_i)}{f_i(t_i)} \right] p_i(t)f(t)dt$$

after changing the order of integration. Substituting into (ER) then completes the proof.¦

In some case, Lemma 2 allows the problem of optimal auction design to be solved almost by inspection, as follows. Note that the choice variables $\{x_i(\cdot)\}$ do not appear in (ER'). But for fixed $\{p_i(\cdot)\}$ we can choose

(EP) $$x_i(t) = t_i p_i(t) - \int_{\underline{a}_i}^{t_i} p_i(t_{-i}, \mathit{t}_i)d\mathit{t}_i,$$

thereby satisfying (ii). Suppose the seller sets $V_i(\underline{a}_i) = 0$ for each i. Then the problem has simplified to choosing $\{p_i(\cdot)\}$ to maximize (ER') subject to (i) and the feasibility constraint. In "regular" cases, the solution to this problem *ignoring* (i) happens to satisfy (i), and so solves the general problem. In "irregular" cases, a subtle argument is needed which is only sketched below.

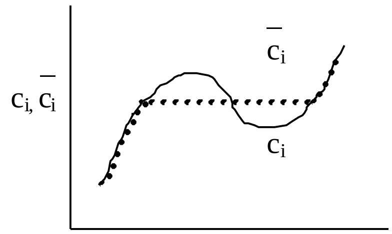
Consider choosing $\{p_i(\cdot)\}$ to maximize (ER') subject only to the feasibility constraint. To ease the notation, define

$$c_i(t_i) \equiv t_i - \frac{1-F_i(t_i)}{f_i(t_i)}$$

and for fixed t let j maximize $c_i(t_i)$ over $i \in \{1,...,n\}$ such that $c_j(t_j) \geq c_i(t_i)$ for all i. (If more than one j achieves this maximum, choose one arbitrarily. Note that such ties are immaterial in expected-value calculations because they happen with probability zero.) Now we can choose $\{p_i(\cdot)\}$ to maximize (ER') pointwise in t: for each fixed t, if $c_j(t_j) - t_0 > 0$ then set $p_j(t) = 1$ and $p_i(t) = 0$ for all $i \neq j$, and if $c_j(t_j) - t_0 \leq 0$ then set $p_i(t) = 0$ for all i. This defines an optimal auction if (i) holds.

A sufficient condition for (i) to hold when $\{p_i(\cdot)\}$ is defined as above is that for each i, $c_i(t_i)$ is weakly increasing in t_i, in which case we call the problem "regular." (Many familiar distributions are regular. For a partial list see Baron and Besanko [1984].)

Two papers that treat the irregular case are Baron and Myerson [1982] and the Myerson paper under discussion. The results for the irregular case are phrased in terms of new functions $\bar{c}_i(t_i)$ that are constructed from the functions $c_i(t_i)$ and are guaranteed to be increasing. The basic idea is summarized in the picture below. When $c_i(t_i)$ is increasing, the construction of $\bar{c}_i(t_i)$ simply mimics $c_i(t_i)$.



Returning to the regular case, it remains to show that when $c_i(t_i)$ is increasing then (i) holds. Consider $\tau_i < t_i$. Then $c_i(\tau_i) \leq c_i(t_i)$, and so $p_i(\tau_i,t_{-i}) \leq p_i(t_i,t_{-i})$ for any t_{-i}. Therefore $P_i(t_i)$ is weakly increasing, completing the optimization.

An important idea in the literature on self-selection is that ex ante (or more precisely here, interim) efficiency may require ex post inefficiency. In other words, to maximize expected revenue when the buyers know their types but the seller does not, the seller may need to design an auction that sometimes *fails* to award the object to the player with the highest willingness to pay.

For example, let $t_0 = 0$, and $F_i(t_i) = t_i$ for t_i on [0,1]. Then $c(t_i) = t_i - (1 - t_i) = 2t_i - 1$ and $c_i(t_i) - t_0 > 0$ if and only if $t_i > 1/2$. In effect, the seller sets a *reservation price* that deters half the types from collecting the object, even though this would be efficient since $t_0 = 0$, because this reservation price increases the bids of the other half of the types. (Note that this is identical, when n = 1, to the problem of a monopolist facing demand Q(p) = 1 - p and having zero costs. See the papers by Gul, Sonnenschein, and Wilson (1986) and Gul and Sonnenschein (1986) for more on this parallel.) One interesting feature of this example is that the reservation price, which is essential when n = 1 to get any profits, does not go to zero as $n \to \infty$. Indeed, the reservation price of 1/2 does not depend on n.

As a second example, consider the case of asymmetric bidders. Again, ex post inefficiency can result. Let n=2, $f_1(t_1) = 1/((\bar{a}_1 - \underline{a}_1)$ on $[\underline{a}_1 - \bar{a}_1]$, and $f_2(t_2) = 1/(\bar{a}_2 - \underline{a}_2)$ on $[\underline{a}_2 - \bar{a}_2]$. As before, let $t_0 = 0$. Then $c_i(t_i) = 2t_i - \bar{a}_i$, and it could happen that $2t_1 - \bar{a}_1 > 2t_2 - \bar{a}_2 > 0$ and $t_2 > t_1$, so that 1 gets the object even though 2 values it more. (Note that this requires $\bar{a}_2 > \bar{a}_1$, so the intuition is analogous to that behind the previous example: the seller withholds the object from low types of bidder 2 in order to extract more from the high types.

It would be nice to be able to interpret the optimal auction in terms of auctions observed in practice. Unfortunately, this is not so easy, but there are two partial results in this direction.

First, consider the symmetric, regular case: for each i, $T_i = T_1$, $f_i(t_i) = f_1(t_1)$, and

$$c_1(t_1) = t_1 - \frac{1 - F_1(t_1)}{f_1(t_1)}$$

is weakly increasing. Suppose that $c_1(t_1)$ is strictly increasing. Then there is an inverse function $c^{-1}(\cdot)$, where we ignore the subscript for convenience. For fixed t, let j denote the bidder with highest type: $t_j > t_i$ for all $i \neq j$. In the optimal auction, bidder j gets the object if $c(t_j) - t_0 > 0$, or $t_j > c^{-1}(t_0)$, and pays

(EP) $$x_j(t) = t_j p_j(t) - \int_{\underline{a}_j}^{t_j} p_j(t_{-j}, \mathit{t}_j) d\mathit{t}_j,$$

which is simply

$$\max\{c^{-1}(t_0), \max_{i \neq j} t_i\},$$

the second-highest type (including the seller). In other words, in the symmetric, (strictly) regular case, a *second-price auction* in which the seller bids (or sets the reservation price) $c^{-1}(t_0)$ is optimal.

Note that no claim about uniqueness is being made here. Indeed, the second result concerning familiar auctions suggests quite the reverse.

Recall that $\{x_i(\cdot)\}$ disappeared from (ER'). Therefore, $\{p_i(\cdot)\}$ and $\{V_i(\underline{a}_i)\}$ completely determine the seller's expected revenue. In particular, one form of the *Revenue Equivalence Theorem* states that if

(1) $V(\underline{a}_i) = 0$ for all i, and

(2) for each t, $p_j(t) = 1$ if $t_j > \max_{j \neq i} t_i$,

then the seller's expected revenue is the expected value of the second-highest type. This can be proven either by manipulating (ER') or by appealing to (EP) as above.

Note that, in the absence of reservation prices, symmetric equilibria of the English (ascending oral bids), Dutch (descending prices until one bidder accepts), first-price (sealed bids), and second-price (sealed bids) auctions satisfying these conditions. Something else then must explain the predominance (75% of all auctions) of the English auction in practice. Milgrom and Weber [1982] show that an explanation emerges from models that violate the *private values* assumption made here. An assumption at the opposite extreme is that the bidders observe noisy signals about a *common value*; this model is often used analyze bidding for oil, gas, and mineral rights. In this setting, the highest bidder is probably the one with the most over-optimistic estimate of the common value. This "winner's curse" is alleviated in auctions that release information about other bidders' signals, as the English auction does as the bid ascends, and this increases the expected revenue.

A limitation of this analysis is that it takes the number of bidders and the distributions of valuations as given exogenously. But surely in a real situation, where the preparation of bids is costly, the set of bidders (both the number and the distribution of valuations) will depend on the particular auction design chosen by the seller. When the set of bidders is endogenous, the optimal auction design might be quite different from what is optimal when the set of bidders is exogenous.