- a single seller wants to sell a single indivisible unit of some object to one of n buyers. Refer to the set of buyers as N
- each buyer has a type s_i drawn from a set M_i - each buyer's value for the object that the seller has to offer is $w^i : M_i \to \mathbb{R}^+$.
- let $M = M_1 \times \cdots \times M_n$
- a buyer of type s_i who purchases the good from the seller at price x enjoys utility $w^i(s_i) - x$. If the buyer expects to buy with probability p at expected price x then the buyer's expected payoff is $w^i(s_i)p - x$, so buyers are risk neutral
- types of all the buyers are jointly drawn from a distribution π on $M = M_1 \times \cdots \times M_n$
- the conditional distribution is $\pi(s_{-i}|s_i)$
- an *auction* is a collection $\{p_i, x_i\}_{i \in N}$ with $p_i : M \to \mathbb{R}^+$, and $x_i : M \to \mathbb{R}$ such that $\sum_{i \in N} p_i(s) \leq 1$ for all $s \in M$

- the seller's expected payoff in the auction is

$$\sum_{s \in M} \pi(s) \left[\sum_{i \in N} x_i(s) \right]$$

- an auction *extracts the full surplus* if the seller's expected payoff is equal to

$$\sum_{s \in M} \pi(s) \left[\max_{i \in N} w^i(s_i) \right]$$

- an auction is *individually rational* if

$$\sum_{s_{-i} \in M_{-i}} \pi(s_{-i}|s_i) \left[p_i(s_i, s_{-i}) w^i(s_i) - x_i(s_i, s_{-i}) \right] \geq 0$$

for every i and s_i.

- an auction is *Bayesian incentive compatible* if

$$\sum_{s_{-i} \in M_{-i}} \pi(s_{-i}|s_i) \left[p_i(s_i, s_{-i}) w^i(s_i) - x_i(s_i, s_{-i}) \right] \geq$$

$$\sum_{s_{-i} \in M_{-i}} \pi(s_{-i}|s_i) \left[p_i(t, s_{-i}) w^i(s_i) - x_i(t, s_{-i})\right]$$

for every $i \in N$, $s_i \in M_i$ and $t \in M_i$.

- A Bayesian incentive compatible, individually rational auction that extracts the full surplus exists if and only if there is no type for any buyer whose beliefs about the other buyers' types can be written as a positive linear combination of his own beliefs when he has other types.

- intuition (due to Zvika Neeman) - start with a second price auction in which the highest bidder gets the good and pays the second highest bid. All bidders bid their true values in such an auction, so ex post, the good is always given to the bidder with the highest value. Let $v_i(s_i)$ be the expected surplus of a buyer of type s_i in such an auction (usually called the *information rent*). Since the high bidder only pays the second highest bid, $v_i(s_i) > 0$ and such an auction will fail to extract all the surplus for the seller who will only get the expectation of the second highest bid conditional on

s_i.

- Now imagine requiring that each bidder pay a fee that is conditional on the bids of the other buyers (the fee only depends on the other buyers bids so that no buyer has an incentive to misreport his type to get a lower fee). To extract the surplus, we want to know whether it might be possible to construct a fee $f_i(m_{-i})$ whose expectation conditional on s_i is equal to $v_i(s_i)$ which the information rent earned by a bidder of type s_i.

- Now choose two different values for s_i, say s_i and s'_I. Generally $v_i(s_i) \neq v_i(s'_i)$. The fee we charge has to be independent of the buyer's own type because we don't know it. The cost of the fee to the buyer will only depend on the buyer's beliefs about the types of the other bidders. So if two different bidder types have the same beliefs about the others (the case where the s_i are independent), then the expected cost of the fee to both of them will be the same and we won't be able to extract two different surplus levels.

- more broadly, suppose some type s_i has beliefs that are just an unweighted average of the beliefs of two other types $s'$and s''. Once we design fees that extract the surplus from these other two types, then the expected fee for the third type must be an unweighted average of the expected fees for the other two types regardless of the information rent that type s_i really earns.

- Formal condition - we can't find a type s_i for any buyer i and a vector of non-negative weights $\rho_i(t_i)$ such that

$$\pi(s_{-i}|s_i) = \sum_{t_i \neq s_i} \rho_i(t_i)\pi(s_{-i}|t_i)$$

for every s_{-i}.

- Farkas Lemma: exactly one of the following two systems has a solution

$$Ax = b$$

$$x \geq 0$$

or

$$A^T y \geq 0$$

$$b^T y < 0$$

- Proof that the formal condition implies the existence of a surplus extracting mechanism.

- Oddly enough, construct the fees first: Think of $A_i(s_i)$ as a matrix whose rows are indexed by the various arrays of types s_{-i} for the other bidders, and whose columns are indexed by the types other than s_i. Each cell in the matrix is then the conditional probability that the types of the others take on the array of values corresponding to the row index, where the condition is that i has the type given by the column index.

- Now $\pi(\cdot|s_i)$ can be written as a column vector, and the condition in the theorem is that there does not exist a non-negative column vector $\rho_i \in \mathbb{R}^{n-1}$ such that $A\rho_i = \pi(\cdot|s_i)$. To make the analogy with Farkas Lemma, the vector ρ_i corresponds with the vector x, while $\pi(\cdot|s_i)$ coincides with the vector b. So by Farkas Lemma, there is a vector $y_i(s_i)$ such that $A^T y_i(s_i) \geq 0$ and $\pi(\cdot|s_i)^T y_i(s_i) < 0$.

Define the fee $f_i(s_i)_{s_{-i}} = \gamma_i(s_i)\left[y_i(s_i)_{s_{-i}} - \pi(\cdot|s_i)^T y_i(s_i)\right]$ where each $\gamma_i(s_i)$ is potentially a large positive number.

- By construction

$$\sum_{s_{-i}\in M_{-i}} \pi(s_{-i}|s_i) f_i(s_i)_{s_{-i}} =$$

$$\sum_{s_{-i}\in M_{-i}} \pi(s_{-i}|s_i)\gamma_i(s_i)\left[y_i(s_i)_{s_{-i}} - \pi(\cdot|s_i)^T y_i(s_i)\right] =$$

$$\gamma_i(s_i)\left[\sum_{s_{-i}\in M_{-i}} \pi(s_{-i}|s_i) y_i(s_i)_{s_{-i}} - \pi(\cdot|s_i)^T y_i(s_i)\right]$$

$$= 0$$

while

$$\sum_{s_{-i}\in M_{-i}} \pi(s_{-i}|t) f_i(s_i)_{s_{-i}} =$$

$$\sum_{s_{-i}\in M_{-i}} \pi(s_{-i}|t)\gamma_i(s_i)\left[y_i(s_i)_{s_{-i}} - \pi(\cdot|s_i)^T y_i(s_i)\right] =$$

$$\gamma_i(s_i)\left[\sum_{s_{-i}\in M_{-i}} \pi(s_{-i}|t) y_i(s_i)_{s_{-i}} - \pi(\cdot|s_i)^T y_i(s_i)\right] > 0$$

by the Farkas Lemma result.

- Now we have a collection of fee schedules $f_i(s_i)$ one for each buyer and type. The fee designed for a buyer of type s_i has zero expected cost if the buyer truly has type s_i, but if he has any other type, the fee has a strictly positive, and potentially very large cost, depending on $\gamma_i(s_i)$.

- Let p be the allocation rule associated with a second price auction and define

$$x_i(s_i, s_{-i}) = p_i(s_i, s_{-i})w^i(s_i) - \gamma_i(s_i)f_i(s_i)_{s_{-i}}$$

- this auction is incentive compatible, individually rational and extracts the full surplus.