Mechanism Design

Vanderbilt Economics. Greg Leo

December 3, 2021

Introduction

Birds

Doesn't bird coloration seem inefficient and dangerous for a tasty bird? Maybe, but it helps solve an asymmetric information problem. Female birds want to mate with high-quality mates. However, if you ask any bird his quality, he will lie- it's cheap talk. Color is not cheap talk. A colorful bird is more likely to be spotted and eaten by predators. The only birds who can afford to be brightly colored are the ones that are not affected by this; the ones who can escape anyway- the high-quality birds. In this way, the cost of predation leads to a situation where bird coloring carries real information.

Common Kingfisher

Denote Bird Color $C \in \left[\frac{1}{2}, 1\right]$ and bird quality: $Q \in [0, 1]$. Let the predation probability be: $P(Q, C) = \left(1 - \frac{Q}{2}\right) C$ and the mating probability: $M(C)$. A bird's fitness is the probability that he successfully mates. That is, he is not eaten and mates:

Birds (Fitness) Utility: $U = (1 - P(Q, C)) M(C)$

Suppose C is a choice for the birds (in reality it is not, but this is a nice story anyway). Let us start with a situation where the female birds ignore the information that would be carried by coloring. What if $M(C) = m$? Of course, there is no reason to take a risk of being brightly colored:

$U = \left(1 - C + C\frac{Q}{2}\right) m$

$\frac{\delta U}{\delta c} = m\left(-1 + \frac{Q}{2}\right) < 0$

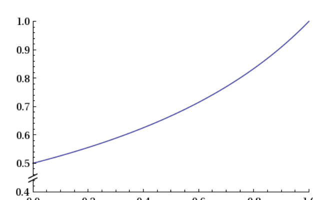
$C = \frac{1}{2}$.

But, if female birds pay attention to the signal, things are different. What if $M(C) = C$?

$U = \left(C - C^2 + C^2\frac{Q}{2}\right)$

$\frac{\delta U}{\delta c} = \left(1 - 2C\left(1 - \frac{Q}{2}\right)\right)$

$C = \frac{1}{2\left(1 - \frac{Q}{2}\right)} = \frac{1}{2 - Q}$



Color as a Function of Quality

is this in the interest of female birds? Maybe, if their fitness is increasing in the quality of their mate, and reducing uncertainty is important enough. It comes at a cost, though, the probability of bird survival is constant under this policy $P = \frac{1}{2}$ while under a constant coloration of $C = \frac{1}{2}$, $P = \left(\frac{1}{2} - \frac{Q}{4}\right)$. We would need more detail to determine whether the mating policy of female birds: $M(C) = C$ is in some way optimal. I hope it's clear that it *could* better than ignoring color.

Brains

Education can help solve an asymmetric information problem which is similar to the one Birds face. Employers want high-quality employees. Attaining a high level of education is easier for high quality types. Because of this, education has information content.

Common Graduation

Denote education: $e \in [1, 10]$ and worker quality: $\theta \in [1, 2]$. Suppose the cost of education to a person with quality θ is: $\frac{e}{\theta}$.

The ex-post Productivity (and also the wage in a competitive environment) isθ, but a company can't observe θ. The only observe e. Thus, the wage they pay is the ex-ante productivity: $E(\theta|e)$.

Let $f(e) = E(\theta|e)$. A worker is paid $f(e)$ and has a cost of $\frac{e}{\theta}$. Thus, worker utility is: $U(e, \theta) = f(e) - \frac{e}{\theta}$. This is concave in e when f is concave. The maximum occurs where $f'(e) = \frac{1}{\theta}$. Thus, if f is concave, the optimal e must be increasing in θ. *What a lovely signal.*

This story is most famously told in Michael Spence (1973). "Job Market Signaling". **Quarterly Journal of Economics**. 87 (3): 355–374

Screening

Designing incentives to elicit information from a single individual.

Pricing a Single Indivisible Good.

Buyer:

Type: θ

Known only to the buyer, not the seller.

Transfers: t

$u(1|\theta) = \theta - t$

Utility of consuming the good.

$u(0|\theta) = 0$

No consumption.

Seller:

Maximizes revenue.

Believes $\theta \sim F(\theta)$, $\theta \in \Theta = [\underline{\theta}, \bar{\theta}]$

We assume $f(\theta) > 0$ everywhere.

Definition. *Mechanism* $\Gamma(\Sigma, t, q)$**:**

Strategies $\sigma \in \Sigma$

Transfers: $t(\sigma)$

In this environment, this is the price of the good.

Allocations: $q(\sigma) : A \to [0, 1]$

Probability of sale.

Definition. *Direct Mechanism* $\Gamma(\Theta, t, q)$**:**

In a direct mechanism, strategies are "claims" about type.

Transfers: $t(\theta)$

In this environment, this is the price of the good.

Allocation: $q(\theta) : \Theta \to [0, 1]$

Probability of sale.

Why focus on direct mechanisms?

Definition. *Incentive Compatible*

a direct mechanism is incentive compatible if the strategy θ is optimal for a player of type θ.

Proposition. Revelation Principle.

For every mechanism there is an equivalent incentive compatible direct mechanism. That is, for every $\Gamma(\Sigma, t, q)$ with optimal strategies $\sigma(\theta)$ there is a direct $\Gamma'(t', q')$ with optimal strategies $\sigma'(\theta) = \theta$ such that $t(\sigma(\theta)) = t'(\theta)$ and $q(\sigma(\theta)) = q'(\theta)$.

Determine the types who play each σ in the indirect mechanism. Construct the direct mechanism by assigning the same allocation q and payment t to anyone who claims to be one of these types in the direct mechanism. Telling the truth ($\sigma'(\theta) = \theta$) is optimal in the direct mechanism by construction since each type is being assigned an outcome that that type had no incentive to move away from in the indirect mechanism.

Notation. Utility Under Direct Mechanism

$u(\tilde{\theta}|\theta) = \theta q(\tilde{\theta}) - t(\tilde{\theta})$.

The first term in $u(\tilde{\theta}|\theta)$ is the "claimed" type, the second term is the actual type.

Notation. Incentive Compatible [I.C.]:

$u(\theta|\theta) \geq u(\tilde{\theta}|\theta)\ \forall \theta, \tilde{\theta} \in \Theta$

Notation. Utility Under I.C. Direct Mechanism

We can suppress the "claimed type."

$u(\theta) = \theta q(\theta) - t(\theta)$.

Notation. Individually Rational [I.R.]

$u(\theta) > 0\, \forall \theta \in \Theta$

Technically, Interim Individual Rationality

Posted Price

Example. Our First Mechanism.

Exercise. Is posted price I.C.? Is it I.R.? What is the seller's expected revenue for uniform F when $p = \frac{1}{2}$.

$$q(\theta) = \begin{cases} 1 & \theta \geq p \\ 0 & \theta < p \end{cases}$$

$$t(\theta) = \begin{cases} p & \theta \geq p \\ 0 & \theta < p \end{cases}$$

Lottery Selling

Example. Our Second Mechanism.

Exercise. Is lottery selling I.C.? **Is it I.R.?** What is the seller's expected revenue? Is there a p and $\tilde{p}$ such that expected revenue is larger than the posted-price mechanism?

$$q(\theta) = \begin{cases} 1 & \theta \geq p \\ \frac{1}{2} & \theta \in (\tilde{p}, p) \\ 0 & \theta < p \end{cases}$$

$$t(\theta) = \begin{cases} p & \theta \geq p \\ \tilde{p} & \theta \in (2\tilde{p}, p) \\ 0 & \theta < p \end{cases}$$

Characterizing I.C. Mechanisms.

Lemma. *Monotonicity of q.*

If Γ is I.C., q increases in θ.

This is an intuitive result and comes from subtracting these two inequalities from each other $\theta q(\theta) - t(\theta) \geq \theta q(\tilde{\theta}) - t(\tilde{\theta})$ and $\tilde{\theta} q(\tilde{\theta}) - t(\tilde{\theta}) \geq \tilde{\theta} q(\theta) - t(\theta)$.

Lemma. *Monotonicity, convexity and smoothness of $u(\theta)$.*

If Γ is I.C., $u(\theta)$ increases and is convex in θ and is almost-everywhere smooth. Where it is smooth, $u'(\theta) = q(\theta)$.

This is perhaps less intuitive. $\theta q(\tilde{\theta}) - t(\tilde{\theta})$ is increasing and convex in θ. Its max is as well. Smoothness result follows from this and the derivative is a result of the envelope theorem.

Lemma. *Payoff Equivalence*

$u(\theta) = u(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} q(x)\,dx$

The utility is the integral of its derivative with $u(\underline{\theta})$ added as a constant of integration. Some additional technical arguments are needed to guarantee this. See text.

Lemma. *Revenue Equivalence*

$t(\theta) = t(\underline{\theta}) + [\theta q(\theta) - \underline{\theta} q(\underline{\theta})] - \int_{\underline{\theta}}^{\theta} q(x)\,dx$

From payoff equivalence: $u(\theta) = u(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} q(x)\,dx$

$\theta q(\theta) - t(\theta) = \underline{\theta} q(\underline{\theta}) - t(\underline{\theta}) + \int_{\underline{\theta}}^{\theta} q(x)\,dx$

Proposition. *A mechanism is I.C. if and only if.*

1. *q is increasing in θ.*

and

2. *Revenue equivalence holds.*

Necessity has been established above. Sufficiency is proven in the text on page 14. I.C. can be established from these two conditions.

Lottery Selling

Our Second Mechanism.

Problem. Confirm Revenue Equivalence

$$q(\theta) = \begin{cases} 1 & \theta \geq p \\ \frac{1}{2} & \theta \in (\tilde{p}, p) \\ 0 & \theta < p \end{cases}$$

$$t(\theta) = \begin{cases} p & \theta \geq p \\ \tilde{p} & \theta \in (2\tilde{p}, p) \\ 0 & \theta < p \end{cases}$$

Constructing Prices for Some Mechanisms

Problem. *Everyone Plays.*

Construct *I.C.* t:

In this mechanism, everyone wins.

$\Theta = [0, 1]$

$q(\theta) = \frac{1}{2}$

$t(0) = 0$

Problem. *Linear Lottery #1*

Construct *I.C.* t:

In this mechanism, your probability of winning is your valuation. $t(\theta) = \frac{1}{2}\theta^2$

$\Theta = [0, 1]$

$q(\theta) = \theta$

$t(0) = 0$

Problem. *Linear Lottery #2*

In this everyone has at least $\frac{1}{2}$ probability of getting the item.

Construct *I.C.* t:

$\Theta = [0, 1]$

$q(\theta) = \frac{\theta+1}{2}$

$t(0) = 0$

Problem. Which yields a higher expected revenue if $\theta \sim U(0,1)$?

Characterizing I.R. Mechanisms.

Proposition. *Characterization of I.R. and I.C. Mechanisms.*

An $I.C.$ mechanism is $I.R.$ iff $u(\underline{\theta}) \geq 0$.

This is implied by $I.R.$ since I.R. requires this for every type. It implies $I.R.$ since u is monotonic (established above).

Optimal Mechanisms

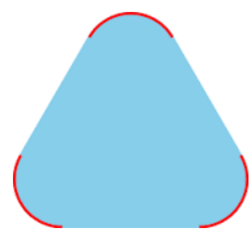
Proposition. *A Revenue Maximizing Mechanism is a Posted Price Mechanism with $p^* = Arg.Max_{p\in[\underline{\theta},\bar{\theta}]} p(1 - F(p))$.*

This follows from the results below.

Note that $p(1 - F(p))$ is the expected revenue since $1 - F(p)$ is the probability the buyer is willing to pay p.

Theorem. *Bauer Maximum Principle.*

A continuous convex function attains its maximum on a compact convex set at an extreme point.



An extreme point is the generalization of a corner. It is any point such that if that point is removed, the resulting set remains convex.

Lemma. *Extreme Points of the Set of I.C. and I.R. q are steps.*

The extreme points of q are all q such that $q(\theta) = 0$ for $\theta < \theta^*$ and $q(\theta) = 1$ for $\theta > \theta^*$.

Lemma. *Expected Revenue is a Linear Function of q.*

This requires $f(\alpha q) = \alpha f(q)$ and $f(q + q') = f(q) + f(q')$. Proving this would be a valuable exercise.

Pricing a Single Divisible Good

Buyer:

We now assume buyer gets $\theta v(q)$ for consuming q of the good.

$u(q,t) = \theta v(q) - t$.

v is strictly increasing and strictly concave$v' > 0, v'' < 0$.

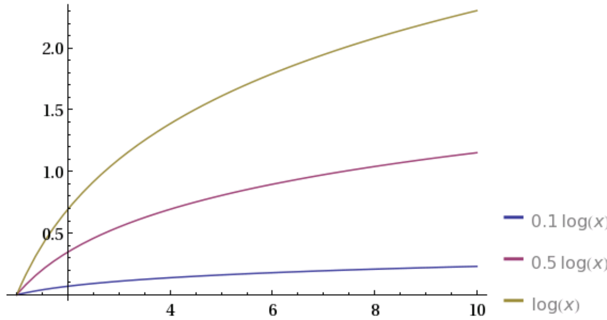
This form of the utility function is limiting. Each type has a utility function that is simply a vertically stretched version of the others.

Seller:

Knows $v(q)$ but not θ.

Variable cost function is qc.

$c < \bar{\theta}v(0)$ and there is some $q* < \infty$ such that $c = \bar{\theta}v(q^*)$. These ensure that something can be sold and that there is never an infinite amount sold.



$\theta v(q)$ might look something like this.

Characterizations

Proposition. *Characterization of I.C. Mechanisms*

$\Gamma(q,t)$ is I.C. iff

1) q is increasing

2) $t(\theta) = t(\underline{\theta}) + [\theta v(q(\theta)) - \underline{\theta} v(q(\underline{\theta}))] - \int_{\underline{\theta}}^{\theta} v(q(x))\,dx$

Note the similarity in the transfer characterization. It can be derived in the same way as in the previous section. $[\theta v(q(\theta)) - \underline{\theta} v(q(\underline{\theta}))]$ is the difference in utility to the lowest type.

Proposition. *Characterization of I.R. Mechanisms*

$\Gamma(q,t)$ that is I.C. is I.R. iff

1) $t(\underline{\theta}) \leq \underline{\theta} v(q(\underline{\theta}))$

Proposition. *Optimal Transfers in I.C. and I.R. Mechanisms*

In an optimal mechanism, $t(\theta) = \theta v(q(\theta)) - \int_{\underline{\theta}}^{\theta} v(q(x))\,dx$

Notice that $t(\theta) = t(\underline{\theta}) + [\theta v(q(\theta)) - \underline{\theta} v(q(\underline{\theta}))] - \int_{\underline{\theta}}^{\theta} v(q(x))\,dx$ is increasing in $t(\underline{\theta})$ which can be at most $\underline{\theta} v(q(\underline{\theta}))$. Plugging this in provides the expression.

The transfers of an optimal mechanism are the difference between $(\theta v(q(\theta))$ the utility to a buyer of type θ of getting $q(\theta))$ and $\int_{\underline{\theta}}^{\theta} v(q(x))\,dx$ an "information rent."

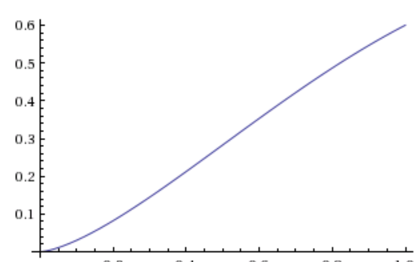
Notice $\int_{\underline{\theta}}^{\theta} v(q(x))\,dx = \frac{1}{\theta-\underline{\theta}} \int_{\underline{\theta}}^{\theta} (\theta - \underline{\theta})\, v(q(x))\,dx$. Thus, the information rent is the average utility difference between buyers of types θ and $\underline{\theta}$ of claiming any type between these two.

If, for instance, $\underline{\theta} = 0$ then this is just the average utility a type θ buyer would get by claiming some type below θ.

Example. *A Linear Mechanism*

What are the optimal transfer functions for a mechanism with $\Theta = [0,1]$, $v = \sqrt{\theta}$ and $q = \theta$?

$\theta\sqrt{\theta} - \int_0^{\theta} x\sqrt{x}dx = \theta^{\frac{3}{2}} - \frac{2}{5}\theta^{\frac{5}{2}} = \theta^{\frac{3}{2}}\left[1 - \frac{2}{5}\theta\right]$



Optimization

A seller's expected profit is:

$$\Pi(q,t)=\int_{\underline{\theta}}^{\bar{\theta}}\left[\theta v(q(\theta))-\int_{\underline{\theta}}^{\theta}v(q(x))\,dx\right]f(\theta)\,d\theta-\int_{\underline{\theta}}^{\bar{\theta}}cq(\theta)f(\theta)\,d\theta$$

$$\Pi(q,t)=\int_{\underline{\theta}}^{\bar{\theta}}\theta v(q(\theta))f(\theta)\,d\theta-\int_{\underline{\theta}}^{\bar{\theta}}\int_{\underline{\theta}}^{\theta}v(q(x))\,dxf(\theta)\,d\theta-\int_{\underline{\theta}}^{\bar{\theta}}cq(\theta)f(\theta)\,d\theta$$

What is $\int_{\underline{\theta}}^{\bar{\theta}}\int_{\underline{\theta}}^{\theta}v(q(x))\,dxf(\theta)\,d\theta$. It is the expected value of the function $g(\theta)=\int_{\underline{\theta}}^{\theta}v(q(x))\,dx$. We can use the property:

$$E(g(\theta))=g(\underline{\theta})+\int_{\underline{\theta}}^{\bar{\theta}}g'(t)(1-F(t))\,dt$$

$$g(\theta)=g(\underline{\theta})+\int_{\underline{\theta}}^{\theta}g'(t)\,dt$$

$$E(g(\theta))=\int_{\underline{\theta}}^{\bar{\theta}}g(\underline{\theta})f(\theta)\,d\theta+\int_{\underline{\theta}}^{\bar{\theta}}\left(\int_{\underline{\theta}}^{\theta}g'(t)\,dt\right)f(\theta)\,d\theta$$

Switch order:

$$E(g(\theta))=g(\underline{\theta})+\int_{\underline{\theta}}^{\bar{\theta}}\left(g'(t)\right)\left(\int_{t}^{\bar{\theta}}f(\theta)\,d\theta\right)dt$$

$$g(\theta)=g(\underline{\theta})+\int_{\underline{\theta}}^{\bar{\theta}}\left(g'(t)\right)(1-F(t))\,dt$$

This gives:

$$\int_{\underline{\theta}}^{\bar{\theta}}\int_{\underline{\theta}}^{\theta}v(q(x))\,dxf(\theta)\,d\theta=\int_{\underline{\theta}}^{\bar{\theta}}v(q(\theta))(1-F(\theta))\,d\theta$$

We now have:

$$\Pi(q,t)=\int_{\underline{\theta}}^{\bar{\theta}}\theta v(q(\theta))f(\theta)\,d\theta-\int_{\underline{\theta}}^{\bar{\theta}}v(q(\theta))(1-F(\theta))\,d\theta-\int_{\underline{\theta}}^{\bar{\theta}}cq(\theta)f(\theta)\,d\theta$$

$$\Pi(q,t)=\int_{\underline{\theta}}^{\bar{\theta}}\left[v(q(\theta))\left(\theta-\frac{1-F(\theta)}{f(\theta)}\right)-cq(\theta)\right]f(\theta)\,d\theta$$

One way to maximize such a thing is to pick $v\left(q\left(\theta\right)\right)\left(\theta-\frac{1-F(\theta)}{\theta}\right)-cq\left(\theta\right)$ that is maximum pointwise, initially ignoring the monotonicity demand on q.

This is strictly concave by the concavity of v.

$$v'\left(q_\theta\right)\left(\theta-\frac{1-F\left(\theta\right)}{f\left(\theta\right)}\right)=c$$

Since $v'\left(q_\theta\right)$ is decreasing by the concavity of v, this has a (unique) solution iff:

$$v'\left(q_\theta\right)\left(\theta-\frac{1-F\left(\theta\right)}{f\left(\theta\right)}\right)>c$$

Otherwise, the optimal choice is $q\left(\theta\right)=0$. Suppose $v'\left(q_\theta\right)\left(\theta-\frac{1-F(\theta)}{\theta}\right)>c$ is true, then when is the solution monotonic?

$$v'\left(q\left(\theta\right)\right)=\frac{c}{\theta-\frac{1-F(\theta)}{f(\theta)}}$$

$q\left(\theta\right)$ must be increasing, which implies by concavity that $v'\left(q\left(\theta\right)\right)$ must be decreasing. This will be true **if and only if** $\theta-\frac{1-F(\theta)}{\theta}$ **is increasing** or **if** $\frac{f(\theta)}{1-F(\theta)}$ **is increasing.**

The second, sufficient condition, is known as increasing hazard rate condition.

Logarithmic Example

Example. $v\left(q\right)=ln\left(q+1\right)$. $\theta\sim U\left(0,1\right)$. $c=\frac{1}{2}$.

Note that $\frac{f(\theta)}{1-F(\theta)}=\frac{1}{1-\theta}$ is increasing.

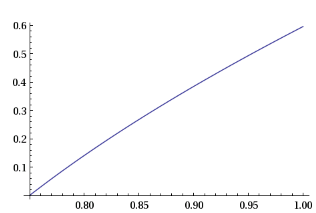
$$\frac{1}{q\left(\theta\right)+1}\left(2\theta-1\right)=\frac{1}{2}$$

As a more general result, when utility is logarithmic, and the distribution of types is uniform, the allocation function is linear at any point where $q\left(\theta\right)>0$.

$$q\left(\theta\right)=\left(4\theta-3\right)$$

$$q\left(\theta\right)=\begin{cases}4\theta-3 & \theta\geq\frac{3}{4}\\ 0 & \theta<\frac{3}{4}\end{cases}$$

$$t\left(\theta\right)=\begin{cases}\theta ln\left(4\theta-2\right)-\int_{\frac{3}{4}}^{\theta}ln\left(4x-2\right)dx & \theta\geq\frac{3}{4}\\ 0 & \theta<\frac{1}{2}\end{cases}$$



Optimal $t\left(\theta\right)$ for $\theta\geq\frac{3}{4}$

Generalized Screening

Environment

Notation. Types $\theta \in \Theta$

Outcomes $a \in A$

Utility of agent: $u(a,\theta) - t$

Note that utility is still quasi-linear. This is one aspect where our analysis is not as general as it could be.

Notation. Direct Mechanisms

$\Gamma(q,t)$ is made up of:

$q : \Theta \to A$

$t : \Theta \to \mathbb{R}$

q is called the "decision rule"

Implementability

Definition. *Incentive Compatibility*

In this setting *I.C. means* $u(q(\theta),\theta) - t(\theta) \geq u(q(\theta'),\theta) - t(\theta')$ for all $\theta, \theta' \in \Theta$

Definition. *Implementable q*

A decision rule q is implementable if there is $\Gamma(q,t)$ which is *I.C.*

Proposition. *Necessary Condition for Implementability*

q is implementable if for $\theta^1, \theta^2 \in \Theta$: $u(q(\theta^1),\theta^1) - u(q(\theta^2),\theta^1) \geq u(q(\theta^1),\theta^2) - u(q(\theta^2),\theta^2)$

This property is called "weak monotonicity." Two types θ^1 and θ^2 might both like what θ^1gets over what θ^2 gets, but the strength of that preference must be weakly stronger for θ^1. This ensures that for any particular pair of types, we can find transfers that gets the types to sort appropriately.

Examples

Example. *Is* $q(\theta^1) = a$, $q(\theta^2) = b$, $q(\theta^3) = c$ *weakly monotone?*

	θ^1	θ^2	θ^3
a	4	2	0
b	1	1	1
c	0	-1	3

Is it implementable?

Yes. $t(\theta^1) = 3$, $t(\theta^2) = 1$, $t(\theta^3) = 2$

Example. *Is $q\left(\theta^1\right) = a$, $q\left(\theta^2\right) = b$, $q\left(\theta^3\right) = c$ weakly monotone?*

	θ^1	θ^2	θ^3
a	0	-1	1
b	1	0	-1
c	-1	1	0

Is it implementable?

No. What happens here is that, while we can find transfers to satisfy truth-telling for any single pair, there is no set of transfers that simultaneously satisfy truth-telling for every pair.

Characterization

Proposition. *Necessary and Sufficient Condition for Implementability*

q is implementable if and only if for every cyclic, finite sequence of types $\left(\theta^1, \theta^2, \ldots, \theta^k\right)$ that begins and ends with the same $\theta^1 = \theta^k$, $0 \leq \sum_{i=1}^{k-1}\left[u\left(q\left(\theta^i\right), \theta^i\right) - u\left(q\left(\theta^i\right), \theta^{i+1}\right)\right]$.

This condition is called cyclic monotonicity. As we tour around the type-space, we keep track of the differences in utilities that θ^i and θ^{i+1} get for θ^i's allocation. This condition says that for any tour, this sum must be positive, or equivalently, the average $\frac{\sum_{i=1}^{k-1}\left[u\left(q\left(\theta^i\right),\theta^i\right)-u\left(q\left(\theta^i\right),\theta^{i+1}\right)\right]}{k-1}$ must be positive.

Example. *Is $q\left(\theta^1\right) = a$, $q\left(\theta^2\right) = b$, $q\left(\theta^3\right) = c$ monotone?*

	θ^1	θ^2	θ^3
a	0	-1	1
b	1	0	-1
c	-1	1	0

No, consider the cyclic sequence $\left(\theta^1, \theta^3, \theta^2, \theta^1\right) = -1 - 1 - 1 = -3 < 0$

Corollary. *A Simple Sufficient Condition for Implementability*

For every pair θ^1, θ^2, $u\left(q\left(\theta^1\right), \theta^1\right) - u\left(q\left(\theta^1\right), \theta^2\right) \geq 0$.

Every type likes what that he gets better than anyone else likes it.

Exercise. *Check this sufficient condition for $q\left(\theta^1\right) = a$, $q\left(\theta^2\right) = b$, $q\left(\theta^3\right) = c$.*

	θ^1	θ^2	θ^3
a	4	2	0
b	1	1	1
c	0	-1	3

Environments where Weak Monotonicity is Sufficient for Implementability

One-Dimensional Environments.

In these environments, types in Θ can be sorted onto the real line based on their preferences over some ordering of the outcomes A.

Definition. *Type Ordering* $\succ_R$

$\succ_R$ may not be complete even if R is.

$u(a,\theta) - u(a',\theta) > u(a,\theta') - u(a',\theta')$ for all a,$a' \in A$ where $a\, R\, a'$ and not $a'\, R\, a$

That is, type θ is higher than θ' when he cares more about the differences between the outcomes in the ordering induced by R.

$u(a,\theta) - u(a',\theta) = 0$ for all a,$a' \in A$ where $a\, R\, a'$ and $a'\, R\, a$

That is, all types care equally about outcomes which are not strictly ordered.

Definition. *Monotone Decision Rules*

A decision rule is "monotone with respect to R when $\theta \succ_R \theta'$ implies $q(\theta)\; R\, q(\theta')$.

Higher types get higher outcomes.

Lemma. *Weakly Monotone Decision Rules are Monotone*

Weakly Monotone q are monotone with respect to R. '

In fact, they are equivalent when $\succ_R$ is complete.

Definition. *One Dimensional Type-Space*

Θ is said to be one-dimensional with respect to R when $\succ_R$ is complete.

Proposition. *Monotonicity is Sufficient for Implementability in One-Dimensional Spaces*

ForΘ is bounded[1] *and one-dimensional with respect to ordering R on A. A decision rule that is monotonic with respect to R is implementable.*

Imagine taking a cyclic tour of the type-space, keeping track of the average differences. By time you return to the starting point, the sum must be positive.

[1] Bounded means $u(a',\theta) - u(a,\theta)$ has a finite upper-bound.

Rich Environments.

Definition. *Rich Type-Space*

A type space if rich if for any $a, b \in A$ (finite A) where aRb then every possible $u(a)$ such that $u(a) > u(b)$ is contained in the set of possibility utilities $u(a) \in \{u(a,\theta) : \theta \in \Theta\}$.

Proposition. *Sufficiency of Weak Monotonicity in Rich Type-Spaces*

Θ is rich, then weak monotonicity of q is sufficient for implementability.

Revenue Equivalence and Individual Rationality.

Assumption. *Θ is convex and $u(a,\theta)$ is a convex function.*

Proposition. *Revenue Equivalence in General Settings*

If $\Gamma(q,t)$ is I.C. then $\Gamma(q,t')$ is I.C. if and only if $t' = t + a$ for $a \in \mathbb{R}$.

Definition. *Individual Rationality with Outside Options*

A mechanism is individually rational with respect to outside option a if for all θ, $u(q(\theta),\theta) - t(\theta) \geq u(a,\theta)$

Lemma. *Individual Rationality in One-Dimensional Settings*

A mechanism is individually rational with respect to outside option a if and only if for all $\underline{\theta}$, $u(q(\underline{\theta}),\underline{\theta}) - t(\underline{\theta}) \geq u(a,\underline{\theta})$

Here, $\underline{\theta}$ is the "smallest" element with respect to the order $\succ_R$.

Bayesian Mechanism Design

Selling a Single Indivisible Good Among Multiple Buyers

Buyer:

Buyers: $I = \{1, 2, .., N\}$

Types: θ_i

Known only to the buyer, not the seller.

Transfers: t_i

$u(1|\theta_i) = \theta_i - t$

Utility of consuming the good.

$u(0|\theta_i) = 0$

No consumption.

Seller:

Maximizes revenue.

Believes $\theta_i \sim F(\theta)$ independently, $\theta \in \Theta = \left[\underline{\theta}, \bar{\theta}\right]^N$

We assume $f(\theta) > 0$ everywhere.

Definition. *Direct Mechanism $\Gamma(\Theta, t, q)$*:

Transfers: $t(\theta)$

Allocation: $q(\theta) : \Theta \to \Delta$

In a direct mechanism, strategies are "claims" about type.

This is now a vector.

Δ is the set of probability vectors of length N adding up to at most 1. These are the probabilities of sale.

Notation. Expected Utilities

Expected allocation: $Q_i(\theta_i) = \int_{\Theta_{-i}} q_i(\theta_i, \theta_{-i}) f_{-i}(\theta_{-i}) d\theta_{-i}$

Expected transfer $T_i(\theta_i) = \int_{\Theta_{-i}} t_i(\theta_i, \theta_{-i}) f_{-i}(\theta_{-i}) d\theta_{-i}$

Expected Utility of Generic Claim: $U_i\left(\tilde{\theta}_i|\theta_i\right) = \theta_i Q_i\left(\tilde{\theta}_i\right) + T_i\left(\tilde{\theta}_i\right)$

Expected Utility of Truth: $U_i(\theta_i) = U_i(\theta_i|\theta_i) = \theta_i Q_i(\theta_i) + T_i(\theta_i)$

We take these expectations against "truthful" play of other types.

Definition. *Incentive Compatible*

A direct mechanism is incentive compatible if truth-telling is a Bayesian Nash Equilibrium of Γ.

$U_i\left(\theta_i|\theta_i\right) \geq U_i\left(\tilde{\theta}_i|\theta_i\right)$ for all i and $\tilde{\theta}_i \in \left[\underline{\theta}, \bar{\theta}\right]$.

Truth-telling is optimal when others are assumed to tell the truth.

Proposition. *Revelation Principle*

For any mechanism Γ with equilibrium strategies σ. There is a direct mechanism Γ' and a truthful equilibrium such that these equilibria in Γ and Γ' have the same outcome distributions and transfers for each type vector θ.

The proof is a straightforward outcome. Construct q and t for Γ'. If there is no incentive to deviate in Γ, then there will be no incentive to deviate in Γ'.

Definition. *Individual Rationality*

A direct mechanism that is I.C. is individually rational if

$U\left(\theta_i\right) \geq 0$ for all i and $\tilde{\theta}_i \in \left[\underline{\theta}, \bar{\theta}\right]$.

Characterizations

Lemma. *Monotonicity of Q_i*

For I.C. mechanisms, $Q_i\left(\theta_i\right)$ is weakly increasing in θ_i for all i.

The proof is very similar to the one used in the screening section.

Lemma. *Monotonicity, Convexity of U_i*

For I.C. mechanisms, U_i is increasing and convex. Further, $U_i'\left(\theta_i\right) = Q_i\left(\theta_i\right)$.

Lemma. *Payoff Equivalence*

$U_i\left(\theta_i\right) = U_i\left(\underline{\theta}\right) + \int_{\underline{\theta}}^{\theta_i} Q_i\left(x\right) dx$

Again, this is the integration of the derivative of U_i found above with $U_i\left(\underline{\theta}\right)$ as constant of integration.

Proposition. *Revenue Equivalence*

$T_i\left(\theta_i\right) = T_i\left(\underline{\theta}\right) + \left[\theta_i Q_i\left(\theta_i\right) - \underline{\theta} Q_i\left(\underline{\theta}\right)\right] - \int_{\underline{\theta}}^{\theta_i} Q_i\left(x\right) dx$

Proposition. *Characterization of I.C. Mechanisms*

Γ is I.C. iff,

1) Q_i is increasing.

2) $T_i(\theta_i) = T_i(\underline{\theta}) + [\theta_i Q_i(\theta_i) - \underline{\theta} Q_i(\underline{\theta})] - \int_{\underline{\theta}}^{\theta_i} Q_i(x)\, dx$

Proposition. *Characterization of I.C. & I.R. Mechanisms*

Γ is I.C. iff,

1) Q_i is increasing.

2) $T_i(\theta_i) = T_i(\underline{\theta}) + [\theta_i Q_i(\theta_i) - \underline{\theta} Q_i(\underline{\theta})] - \int_{\underline{\theta}}^{\theta_i} Q_i(x)\, dx$

3) $T_i(\underline{\theta}) \leq \underline{\theta} Q_i(\underline{\theta})$

Examples

Example. *Auction with Uniform Distribution and Two Bidders*

$Q_i(\theta_i) = \theta_i$, $\theta_i \in [0, 1]$ and $T_i(0) = 0$.

$T_i(\theta_i) = \theta_i^2 - \int_0^{\theta_i} x dx = \theta_i^2 - \int_0^{\theta_i} x dx = \frac{1}{2}\theta_i^2$

Why $Q_i(\theta_i) = \theta_i$?

Is this accurate for first-price auction? What about second?

Example. *Auction with Uniform Distribution and Two Bidders and a Reserve*

$Q_i(\theta_i) = 0$ when $\theta_i \leq \frac{1}{2}$ and $Q_i(\theta_i) = \theta_i$ when $\theta_i \geq \frac{1}{2}$

$T_i(\theta) = [\theta_i^2] - \int_{\frac{1}{2}}^{\theta_i} x dx = \frac{1}{2}\theta_i^2 + \frac{1}{8}$

Which do you think is more profitable for the seller?

Maximizing Revenue

In a revenue maximizing mechanism, it must be: $T_i(\underline{\theta}) = \underline{\theta} Q_i(\underline{\theta})$.

All transfers are increasing in $T_i(\underline{\theta})$

Lemma. *Transfer Functions Under Revenue Maximization*

$T_i(\theta_i) = \theta_i Q_i(\theta_i) - \int_{\underline{\theta}}^{\theta_i} Q_i(x)\, dx$

Note. Rewriting the Revenue Function

$\sum_{i=1}^{N}\int_{\underline{\theta}}^{\bar{\theta}}\left(\theta_i Q_i(\theta_i)-\int_{\underline{\theta}}^{\theta_i} Q_i(x)\,dx\right) f(\theta_i)\,d\theta_i$

$=\sum_{i=1}^{N}\int_{\underline{\theta}}^{\bar{\theta}} Q_i(\theta_i)\left(\theta_i-\frac{1-F(\theta_i)}{f(\theta_i)}\right) f(\theta_i)\,d\theta_i$

$=\sum_{i=1}^{N}\int_{\Theta} q_i(\theta_i)\left(\theta_i-\frac{1-F(\theta_i)}{f(\theta_i)}\right) f(\theta_i)\,d\theta_i$

The first equality follows in the same way as in the screening section. The second comes from the fact that $Q_i(\theta_i)=\int_{\Theta_{-i}} q_i(\theta_i)\,d\Theta_{-i}$.

As before, we can try to maximize this pointwise. To do this, simply pick $q_i(\theta_i)=1$ whenever $\left(\theta_i-\frac{1-F(\theta_i)}{f(\theta_i)}\right)$ is maximal within the N values (unless that maximum is also negative). We get:

Lemma. *Optimal q_i Ignoring Monotonicity*

For any θ, $q_i(\theta_i)=1$ when $\psi(\theta_i)=max\{\psi(\theta_1),\psi(\theta_2),...,\psi(\theta_n)\}>0$ and $q_i(\theta_i)=0$ otherwise.

Let $\theta_i-\frac{1-F(\theta_i)}{f(\theta_i)}=\psi(\theta_i)$.

Suppose $\psi(\theta_i)$ is strictly increasing. Then, $q(\theta_i)$ is increasing which implies $Q_i(\theta_i)$ is increasing.

Proposition. *Optimal q_i when $\psi_i(\theta_i)$ is Strictly Increasing*

For any θ, $q_i(\theta_i)=1$ when $\psi(\theta_i)=max\{\psi_1(\theta_1),\psi_2(\theta_2),...,\psi_n(\theta_n)\}>0$ and $q_i(\theta_i)=0$ otherwise.

However, notice that if all ψ_i are identical and increasing, $\psi(\theta_i)=max\{\psi(\theta_1),\psi(\theta_2),...,\psi(\theta_n)\}>0$ if and only if $\theta_i=max(\theta)$. Thus we have:

Where $\theta_i=max(\theta)$ means θ_i is the maximal element of the vector θ.

Proposition. *Optimal q_i when F are identical and $\psi(\theta_i)$ is Strictly Increasing*

For any θ, $q_i(\theta_i)=1$ when $\theta_i=max(\theta)$, and $\psi(\theta_i)\geq 0$, $q_i(\theta_i)=0$ otherwise.

One might think that any auction which is incentive compatible and awards the item to the highest bidder would be optimal. This, however, does not incorporate the optimality condition $\psi(\theta_i)\geq 0$. Under what conditions would this be a problem?

Example. *Auction with Uniform Bidders Revisited*

Suppose:$\theta_i \sim U[0,1]$

The optimal auction format for two bidders with uniform value is not simply second price, but second price with a reserve!

$\psi(\theta_i) = 2\theta_i - 1 \geq 0$ when $\theta_i \geq \frac{1}{2}$.

Optimal Rule:

The optimal rule is i wins when $\theta_i > \theta_j \geq \frac{1}{2}$. Thus $Q_i(\theta_i) = 0$ when $\theta_i \leq \frac{1}{2}$ and $Q_i(\theta_i) = \theta_i$ when $\theta_i \geq \frac{1}{2}$

Transfers:

$T_i(\theta) = \left[\theta_i^2\right] - \int_{\frac{1}{2}}^{\theta_i} x dx = \frac{1}{2}\theta_i^2 + \frac{1}{8}$

Revenue:

$2\int_0^1 \left(\frac{1}{2}\theta_i^2 + \frac{1}{8}\right) d\theta_i = \frac{5}{12}$

What about in the case of no reserve? $2\int_0^1 \left(\frac{1}{2}x^2\right) dx = \frac{1}{3}$

Maximizing Welfare

Designer wants to maximize: $\sum_I \theta_i q_i(\theta)$

The optimal choice is $q_i(\theta) = 1$ if $\theta_i = \max\{\theta_1, ..., \theta_n\}$ with transfer functions: $T_i(\theta_i) = \theta_i Q_i(\theta_i) - \int_{\underline{\theta}}^{\theta_i} Q_i(x)\, dx$

Public Goods

Notation. Public good: $g = \{0, 1\}$

Utility: $\theta_i g - t_i$

Decision Rule: $q : \theta \to \{0, 1\}$

Cost of producing good: $c > 0$

Budget Balance

Definition. *Ex-Post Budget Balance*

$\sum t_i(\theta) \geq cq(\theta)$

Definition. *Ex-Ante Budget Balance*

$\int_\Theta \sum t_i(\theta) f(\theta)\, d\theta \geq c \int_\Theta \sum q(\theta) f(\theta)\, d\theta q(\theta)$

Proposition. *Equivalence of Ex-Ante and Ex-Post BB*

Ex-post implies ex-ante. Furthermore, for any mechanism that is ex-ente BB, there is a mechanism with the same decision rule and equivalent expected transfers for every type which is ex-post BB.

Pick two agents. The first one covers the ex-post deficit but is given a bonus equal to the expected ex-post deficit given type. Clearly, on average, these balance. A second agent is then asked to cover the expected ex-post deficit given the type of the first. But, this expected deficit, averaged over all types of the first must also be zero! If there is an ex-ante surplus, subtract that ex-ante surplus as a constant from some random agent, and continue.

	0	3
0	0,0	0,1
3	1,0	1,1

c is 4.

Value is 3. For each person.

Ex-ante cost is 1.

Let Player 1 cover the deficit.

	0	3		0	3
0	0,0	0-1,1	0	0,0	-1,1
3	1-1,0	1+2,1	3	0,0	3,1

Now notice that 1 has different incentives, and does not want to tell the truth!

If $\theta_1 = 3$, expected value is 0 vs $\frac{1}{2}$ for saying $\theta_1 = 0$.

Player 2 now compensates (or is compensated by) 1 for the *expected amount of the original deficit* based on 1's type. If $\theta_1 = 0$, player 2 takes $\frac{1}{2}$ from the government, and the government takes $\frac{1}{2}$ from player 1. If $\theta_1 = 3$, player 2 gives the government $\frac{1}{2}$ and the government gives $\frac{1}{2}$ to player 1. These are ex-post neutral transfers!

	0	3			
0	0+$\frac{1}{2}$,0-$\frac{1}{2}$	-1+$\frac{1}{2}$,1-$\frac{1}{2}$		$\frac{1}{2}$,$-\frac{1}{2}$	-$\frac{1}{2}$,$\frac{1}{2}$
3	0-$\frac{1}{2}$,0+$\frac{1}{2}$	3-$\frac{1}{2}$,1+$\frac{1}{2}$		-$\frac{1}{2}$,$\frac{1}{2}$	$\frac{5}{2}$,$\frac{3}{2}$

Notice that these bring the ex-ante incentives for player 1 back to their original values. $\theta_1 = 0$ expects utility 1 for telling the truth and 0 for saying $\theta_1 = 0$ and again has incentive to tell the truth.

Lemma. *Ex-Ante Budget Characterization*

$\sum_{i \in I} -U_i\left(\underline{\theta}\right) + \int_{\Theta} q\left(\theta\right) \left[\sum_{i \in I} \left(\theta_i - \frac{1-F_i(\theta_i)}{f_i(\theta_i)}\right)\right] f\left(\theta\right) d\theta$

Previously, we've been looking at revenue maximizing mechanisms in which case it is always optimal$U_i\left(\underline{\theta}\right) = 0$.

Definition. *Pivot Mechanism*

$t_i\left(\theta\right) = \underline{\theta} q^*\left(\underline{\theta}, \theta_{-i}\right) + \left(q^*\left(\theta\right) - q^*\left(\underline{\theta}, \theta_{-i}\right)\right)\left(c - \sum_{i \neq j} \theta_j\right)$

If $\underline{\theta} = 0$, $q^*\left(\theta\right) = q^*\left(\underline{\theta}, \theta_{-i}\right)$ then $t_i = 0$. Otherwise, pays $c - \sum_{i \neq j} \theta_j$

Lemma. *Pivot is Incentive Compatible and Individually Rational.*

Proof. Consider a particular value of θ_{-i}. Player is pivotal if $\theta_i \geq \left(c - \sum_{i \neq j} \theta_j\right)$. Player wants to be pivotal if $\theta_i \geq \left(c - \sum_{i \neq j} \theta_j\right)$. Thus, if $\theta_i \geq \left(c - \sum_{i \neq j} \theta_j\right)$, the player has incentive to report θ_i (or anything higher), if $\theta_i < \left(c - \sum_{i \neq j} \theta_j\right)$, the player has incentive to report θ_i or anything lower. This is true for *any* θ_{-i}, it must be true on average as well. Further, if $\theta_i = 0$, the player is not pivotal and $t_i = 0$. □

Example. ***Pivot Mechanism***

$\theta_i \in [0, 30]$

$\theta_1 = 10$

$\theta_2 = 20$

$\theta_3 = 30$

$c = 40$

Suppose q is the efficient rule.

1 pays 0

2 pays 0

3 pays $40 - (10 + 20) = 10$

Notice that this is not ex-post budget balanced!

Impossibility

Proposition. *Impossibility of Budget Balanced I.C. I.R. Mechanisms*

A first best, incentive compatible and individually rational mechanism does not exist except in trivial *cases.*

Proof. (*Sketch*) Suppose we could find a mechanism where total revenue is $\int_\Theta q^*(\theta) \left[\sum_{i \in I} \left(\theta_i - \frac{1-F_i(\theta_i)}{f_i(\theta_i)}\right)\right] f(\theta)\, d\theta$.

Notice that this has nothing to do with any particulars of the mechanism.

We can, but no such mechanism will produce revenue at least $\int_\Theta q^*(\theta)\, c$.

Maximum Revenue of the Pivot

The Pivot Achieves Highest Possible Ex-Ante Revenue. If it isn't budget-balanced, no mechanism is!

$c - \sum_{i \neq j} \theta_j$ for any pivotal type and $\underline{\theta}$ for any non-pivotal agent.

$\sum_{i \in P} \left(c - \sum_{i \neq j} \theta_j\right) + \sum_{i \in NP} \underline{\theta}$

$\left(\sum_{i \in P} c - \sum_{i \in P} \sum_{i \neq j} \theta_j\right) + \sum_{i \in NP} \underline{\theta}$

$\left(\sum_{i \in P} c - \left[(P-1) \sum_{i \in P} \theta_i + P \sum_{i \in NP} \theta_i\right]\right) + \sum_{i \in NP} \underline{\theta}$

$\sum_{i \in P} c - (P-1) \sum_I \theta_i - \sum_{i \in NP} \theta_i + \sum_{i \in NP} \underline{\theta}$

$Pc - (P-1) \sum_I \theta_j - \sum_{i \in NP} \left(\theta_j - \underline{\theta}\right)$ □

What is possible?

A budget balanced, mechanism has:

$\int_{\Theta} q(\theta)\left[\sum_{i\in I}\left(\theta_i - \frac{1-F_i(\theta_i)}{f_i(\theta_i)}\right) - c\right] f(\theta)\, d\theta \geq 0$

Ex-ante welfare:

$\int_{\Theta} q(\theta)\left[\sum \theta_i - c\right] f(\theta)\, d\theta$

Use Lagrange Method:

$\int_{\Theta} q(\theta)\left[\sum \theta_i - c\right] f(\theta)\, d\theta - \lambda \int_{\Theta} q(\theta)\left[\sum_{i\in I}\left(\theta_i - \frac{1-F_i(\theta_i)}{f_i(\theta_i)}\right) - c\right] f(\theta)\, d\theta$

$\int_{\Theta} q(\theta)(1+\lambda)\left[\sum\left(\theta_i - \frac{\lambda}{1+\lambda}\frac{1-F_i(\theta_i)}{f_i(\theta_i)}\right) - c\right] f(\theta)\, d\theta$

Set $q = 1$ when $\sum_I \theta_i > c + \sum_I \frac{\lambda}{1+\lambda}\frac{1-F_i(\theta_i)}{f_i(\theta_i)}$

What λ allows a balanced budget, though? That's a difficult problem.

Suppose θ_1 and θ_2 are uniformly distributed from 0 to 1.

$\sum_I \theta_i > c + \sum_I \frac{\lambda}{1+\lambda}\frac{1-F_i(\theta_i)}{f_i(\theta_i)}$

$\theta_1 + \theta_2 > c + \frac{\lambda}{1+\lambda}(2 - \theta_1 - \theta_2)$

$\frac{1+2\lambda}{1+\lambda}(\theta_1 + \theta_2) > c + \frac{2\lambda}{1+\lambda}$

$(\theta_1 + \theta_2) > \frac{1+\lambda}{1+2\lambda}c + \frac{2\lambda}{1+2\lambda}$

Let $\frac{1+\lambda}{1+2\lambda}c + \frac{2\lambda}{1+2\lambda} = s(c)$

$q = 1$ when $\theta_1 + \theta_2 \geq s(c)$

What we need is,

$\int_{\Theta} q(\theta)\left[\sum_{i\in I}\left(\theta_i - \frac{1-F_i(\theta_i)}{f_i(\theta_i)}\right) - c\right] f(\theta)\, d\theta = 0$

$\int_{\Theta} q(\theta)\left[2(\theta_1 + \theta_2) - 2 - c\right] f(\theta)\, d\theta = 0$

$\int_{\theta_1+\theta_2 \geq s(c)} \left[2(\theta_1 + \theta_2) - 2 - c\right] f(\theta)\, d\theta = 0$

Suppose $c = 1$.

$\int_{s-\theta_1}^{1}\int_{s-1}^{1}\left[2(\theta_1 + \theta_2) - 3\right] d\theta_1 d\theta_2$

$\int_{s-1}^{1}\int_{s-x}^{1}\left[2(x+y) - 2\right] dy dx = \frac{1}{6}(s-2)^2(4s-5)$

$\frac{1}{6}(s-2)^2(4s-5) = 0$

$s = \frac{5}{4}$

What is the efficiency lost?

$\int_{s-1}^{1}\int_{s-x}^{1}\left[x + y - 1\right] dy dx = \frac{1}{6}(-2+s)^2(-1+2s)$

For $s = 1$, $\frac{1}{6}$

For $s = \frac{5}{4}$, $\frac{9}{64}$

84.38% of first-best efficiency is possible.

What would a monopolist do?

$Max \int_{\Theta} q(\theta) \left[\sum_{i \in I} \left(\theta_i - \frac{1-F_i(\theta_i)}{f_i(\theta_i)}\right) - c\right] f(\theta)\, d\theta$

$q = 1$ if $\sum_{i \in I} \left(\theta_i - \frac{1-F_i(\theta_i)}{f_i(\theta_i)}\right) \geq c$

$2(\theta_1 + \theta_2) - 2 \geq c$

$\theta_1 + \theta_2 \geq \frac{c+2}{2}$

If $c = 1$, $\theta_1 + \theta_2 \geq \frac{3}{2}$

Compare these:

First Best: $\theta_1 + \theta_2 \geq 1$

Second Best $\theta_1 + \theta_2 \geq \frac{5}{4}$

Monopolist $\theta_1 + \theta_2 \geq \frac{3}{2}$

Bilateral Trade

The usual utility functions apply. However, t_s will be assumed to be the transfer that the seller receives.

Seller eats: $\theta_s + t_s$ otherwise t_s

Buyer eats: $\theta_b - t_b$ otherwise $-t_b$

Notation. Type of the seller $\theta_s \in \Theta_s = \left[\underline{\theta}_s, \bar{\theta}_s\right]$

Type of the buyer $\theta_b \in \Theta_b = \left[\underline{\theta}_b, \bar{\theta}_b\right]$

Definition. A direct mechanism is q, t_s, t_b where $q : \Theta_s \times \Theta_b \to [0,1]$ (probability of sale).

Notation. Ex-Ante Utility

$Q_s(\theta_s) = E_{\theta_b}(q(\theta_b, \theta_s))$ and $T_s(\theta_s) = E_{\theta_b}(t(\theta_b, \theta_s))$

$Q_b(\theta_b) = E_{\theta_s}(q(\theta_b, \theta_s))$ and $T_b(\theta_b) = E_{\theta_s}(t(\theta_b, \theta_s))$

$U_S(\theta_s) = (1 - Q_s(\theta_s))\theta_s + T_b(\theta_b)$

$U_b(\theta_b) = (Q(\theta_b))\theta_b - T_b(\theta_b)$

Almost nothing is new here.

Definition. *Individual Rationality*

$U_s(\theta_s) \geq \theta_s \ \forall \theta_s$

$U_b(\theta_b) \geq 0 \ \forall \theta_b$

Lemma. *Monotonic Utility (the sellers in the reverse of the usual way).*

Proof. By the envelope condition, the derivatives are:

$U_S(\theta_s) = (1 - Q_s(\theta_s))$ which is decreasing in θ_s.

$U_b(\theta_b) = (Q(\theta_b))$ which is increasing in θ_b. □

Corollary. *Characterization for IR*

$U_s\left(\bar{\theta}_s\right) \geq \bar{\theta}_s$

$U_b\left(\underline{\theta}_b\right) \geq 0$

Proposition. Characterization of I.C. Mechanisms

For the buyer, we've already done this.

Γ is I.C. for the buyer iff,

1) Q_b *is increasing.*

2) $T_b\left(\theta_b\right) = T_b\left(\underline{\theta}_b\right) + \left[\theta_b Q_b\left(\theta_b\right) - \underline{\theta}_b Q_b\left(\underline{\theta}_b\right)\right] - \int_{\underline{\theta}_b}^{\theta_b} Q_b\left(x\right) dx$

The seller has a sign or two reversed, we've already done this.

Γ is I.C. for the buyer iff,

1) Q_s *is decreasing.*

2) $T_s\left(\theta_s\right) = T_s\left(\underline{\theta}_s\right) + \left[\bar{\theta}_s\left[1 - Q_s\left(\bar{\theta}_s\right)\right] - \underline{\theta_s}\left[1 - Q_s\left(\theta_s\right)\right]\right] - \int_{\theta_s}^{\bar{\theta}_s}\left(1 - Q_s\left(x\right)\right) dx$

Budget

We might like $t_s = t_b$. But in general, we will say:

Definition. *Ex-post Budget Balance*

$t_s - t_b \leq 0.$

The remainder is "burned" money. It is an interesting open question as to where and what extent money burning is useful in economic mechanisms.

Definition. *Ex-ante Budget Balance*

$T_s - T_b \leq 0.$

Proposition. *Ex-post and Ex-ante Budget Balance Are Equivalent*

As in the public goods environment. If one is satisfied, the other can be satisfied as well, and for the same reason (risk neutrality).

Welfare Maximization

We seek to maximize the expected value of $q\left(\theta_s, \theta_b\right)\theta_b - t_b + \left(1 - q\left(\theta_s, \theta_b\right)\right)\theta_s + t_s$ **.**

It is also an interesting open question as to where and what extent withholding is useful in economic mechanisms.

First best is $q = 1$ if and only if $\theta_b \geq \theta_s$.

Theorem. *First Best is impossible with and I.C. I.R. Mechanism*

This is the celebrated Myerson, Satterthwaite (1983) result. One of the most important impossibilities in mechanism design, and an important negative reflection of the welfare theorems of economics. ***Efficiency is not possible when there is asymmetric information****.*

Proof. The pivot mechanism is I.C. and I.R. and minimizes the ex-ante deficit. And yet, under the pivot mechanism, there is always an ex-post deficit. Thus, there is always an ex-ante deficit. □

Definition. *The Pivot Mechanism*

$t_s(\theta) = q^*(\bar{\theta}_s, \theta_b)\,\bar{\theta}_s + (q^*(\theta) - q^*(\bar{\theta}_s, \theta_b))\,\theta_b$

$t_b(\theta) = q^*(\theta_s, \underline{\theta}_b)\,\underline{\theta}_b + (q^*(\theta) - q^*(\theta_s, \underline{\theta}_b))\,\theta_s$

Lemma. *Pivot is I.C. and I.R.*

Proof. (Sketch) Intuitively, if I am pivotal, I pay the other person's valuation. But, if I am pivotal, the other's valuation is below mine and so I am always willing to tell the truth about my valuation. Furthermore:

$U_s(\bar{\theta}_s) = \bar{\theta}_s + 0$

$U_b(\bar{\theta}_b) = 0$ □

Lemma. *Pivot Minimizes Ex-Ante Revenue Deficit*

Proof. $T_s(\theta_s) = T_s(\bar{\theta}_s) + \bar{\theta}_s[1 - Q_s(\bar{\theta}_s)] - \underline{\theta_s}[1 - Q_s(\theta_s)] - \int_{\theta_s}^{\bar{\theta}_s}(1 - Q_s(x))\,dx$

$T_s(\theta_s) = U_s(\bar{\theta}_s) - \underline{\theta_s}[1 - Q_s(\theta_s)] - \int_{\theta_s}^{\bar{\theta}_s}(1 - Q_s(x))\,dx$

But, $U_s(\bar{\theta}_s)$ is maximized and so T_S is minimized.

Similarly,

$T_b(\theta_b) = T_b(\underline{\theta}_b) + [\theta_b Q_b(\theta_b) - \underline{\theta_b} Q_b(\underline{\theta}_b)] - \int_{\underline{\theta}_b}^{\theta_b} Q_b(x)\,dx$

$T_b(\theta_b) = -U_b(\underline{\theta}_b) + [\theta_b Q_b(\theta_b)] - \int_{\underline{\theta}_b}^{\theta_b} Q_b(x)\,dx$

U_b is minimized, and so T_b is maximized.

Together $T_s - T_b$ is minimized. □

Lemma. *Pivot Maintains Ex-Post Revenue Deficit*

Proof. $t_s(\theta) - t_b(\theta) = q^*(\bar{\theta}_s, \theta_b)\,\bar{\theta}_s + (q^*(\theta) - q^*(\bar{\theta}_s, \theta_b))\,\theta_b - q^*(\theta_s, \underline{\theta}_b)\,\underline{\theta}_b - (q^*(\theta) - q^*(\theta_s, \underline{\theta}_b))\,\theta_s$

As long as q is not constant:

$= q^*(\bar{\theta}_s, \theta_b)[\bar{\theta}_s - \theta_b] + q^*(\theta_s, \underline{\theta}_b)[\theta_s - \underline{\theta}_b]$

Note that if $q^*(\bar{\theta}_s, \theta_b) = 1$ then $\theta_b \geq \bar{\theta}_s$ and so $q^*(\bar{\theta}_s, \theta_b)[\bar{\theta}_s - \theta_b] \leq 0$

Similarly, if $q^*(\theta_s, \underline{\theta}_b) = 1$ then $\underline{\theta}_b \geq \theta_s$ and so $q^*(\theta_s, \underline{\theta}_b)[\theta_s - \underline{\theta}_b] \leq 0$

Furthermore, these are strictly negative for some values whenever trade is strictly efficient sometimes. □

Welfare Maximization-Second Best

The designer seeks to maximize $E[q(\theta)\theta_b + (1 - q(\theta))\theta_s] = E(\theta_s) + E[q(\theta)[\theta_b - \theta_s]]$

We can focus on $E[q(\theta)[\theta_b - \theta_s]]$.

Buyers transfer $-U_b(\underline{\theta}_b) + E\left(q(\theta)\left[\theta_b - \frac{1-F_b(\theta_b)}{f_b(\theta_b)}\right]\right)$

Seller's transfer $U_s(\bar{\theta}_s) - E\left((1 - q(\theta))\left[\theta_s + \frac{F_s(\theta_s)}{f_s(\theta_s)}\right]\right)$

Exact-ex-ante Budget Balance:

$-U_b(\underline{\theta}_b) + E\left(q(\theta)\left[\theta_b - \frac{1-F_b(\theta_b)}{f_b(\theta_b)}\right]\right) = U_s(\bar{\theta}_s) - E\left((1 - q(\theta))\left[\theta_s + \frac{F_s(\theta_s)}{f_s(\theta_s)}\right]\right)$

$E\left(q(\theta)\left[\theta_b - \frac{1-F_b(\theta_b)}{f_b(\theta_b)}\right]\right) = U_s(\bar{\theta}_s) + U_b(\underline{\theta}_b) - E\left((1 - q(\theta))\left[\theta_s + \frac{F_s(\theta_s)}{f_s(\theta_s)}\right]\right)$

Since $U_s(\bar{\theta}_s) + U_b(\underline{\theta}_b) \geq \bar{\theta}_s$

$E\left(q(\theta)\left[\theta_b - \frac{1-F_b(\theta_b)}{f_b(\theta_b)}\right]\right) \geq \bar{\theta}_s - E\left((1 - q(\theta))\left[\theta_s + \frac{F_s(\theta_s)}{f_s(\theta_s)}\right]\right)$

Modified:

$E\left(q(\theta)\left[\theta_b - \frac{1-F_b(\theta_b)}{f_b(\theta_b)}\right]\right) \geq \bar{\theta}_s - E\left(\left[\theta_s + \frac{F_s(\theta_s)}{f_s(\theta_s)}\right]\right) + E\left(q(\theta)\left[\theta_s + \frac{F_s(\theta_s)}{f_s(\theta_s)}\right]\right)$

$E\left(q(\theta)\left[\theta_b - \frac{1-F_b(\theta_b)}{f_b(\theta_b)} - \theta_s - \frac{F_s(\theta_s)}{f_s(\theta_s)}\right]\right) \geq \bar{\theta}_s - E\left(\left[\theta_s + \frac{F_s(\theta_s)}{f_s(\theta_s)}\right]\right)$

Notice the right side doesn't depend on q. Call it K. The budget constraint is:

$E\left(q(\theta)\left[\theta_b - \frac{1-F_b(\theta_b)}{f_b(\theta_b)} - \theta_s - \frac{F_s(\theta_s)}{f_s(\theta_s)}\right]\right) - K \geq 0$

Now maximize:

$E[q(\theta)[\theta_b - \theta_s]] + \lambda\left[E\left(q(\theta)\left[\theta_b - \frac{1-F_b(\theta_b)}{f_b(\theta_b)} - \theta_s - \frac{F_s(\theta_s)}{f_s(\theta_s)}\right]\right) - K\right]$

$E[q(\theta)[\theta_b - \theta_s]] + E\left(q(\theta)\left[\lambda\theta_b - \lambda\frac{1-F_b(\theta_b)}{f_b(\theta_b)} - \lambda\theta_s - \lambda\frac{F_s(\theta_s)}{f_s(\theta_s)}\right]\right) - \lambda K$

Combine:

$E\left[q(\theta)\left[(1+\lambda)\theta_b - \lambda\frac{1-F_b(\theta_b)}{f_b(\theta_b)} - (1+\lambda)\theta_s - \lambda\frac{F_s(\theta_s)}{f_s(\theta_s)}\right]\right] - \lambda K$

In each case, we are maximizing the expected value of:

$q(\theta)\left[(1+\lambda)\theta_b - \lambda\frac{1-F_b(\theta_b)}{f_b(\theta_b)} - (1+\lambda)\theta_s - \lambda\frac{F_s(\theta_s)}{f_s(\theta_s)}\right]$

Set $q = 1$ when the inside is positive:

$\theta_b - \frac{\lambda}{1+\lambda}\frac{1-F_b(\theta_b)}{f_b(\theta_b)} \geq \theta_s + \frac{\lambda}{1+\lambda}\frac{F_s(\theta_s)}{f_s(\theta_s)}$

And find the λ that balances the budget.

Welfare Maximization-Second Best Numerical:

Suppose both are uniform $0,1$.

$\theta_b - \frac{\lambda}{1+\lambda}(1-\theta_b) \geq \theta_s + \frac{\lambda}{1+\lambda}\theta_s$

$\theta_b - \frac{\lambda}{1+\lambda}(1-\theta_b) \geq \theta_s + \frac{\lambda}{1+\lambda}\theta_s$

$\theta_b - \theta_s \geq \frac{\lambda}{1+2\lambda}$

Thus, the optimal rule has a "gap" structure. Let $s = \frac{\lambda}{1+2\lambda}$. Budget balance is:

$E\left(\tilde{q}(\theta)\left[\theta_b - \frac{1-F_b(\theta_b)}{f_b(\theta_b)} - \theta_s - \frac{F_s(\theta_s)}{f_s(\theta_s)}\right]\right) = \bar{\theta}_s - E\left(\left[\theta_s + \frac{F_s(\theta_s)}{f_s(\theta_s)}\right]\right)$

$E(\tilde{q}(\theta)[2\theta_b - 1 - 2\theta_s]) = 1 - E(2\theta_s)$

$E(\tilde{q}(\theta)[2\theta_b - 1 - 2\theta_s]) = 1 - E(2\theta_s)$

$E(\tilde{q}(\theta)[2\theta_b - 1 - 2\theta_s]) = 0$

$\int_0^{1-s}\int_{\theta_s+s}^{1}[2\theta_b - 1 - 2\theta_s]\,d(\theta_b)\,d(\theta_s) = 0$

$\frac{1}{6}(s-1)^2(4s-1) = 0$

$s = \left\{\frac{1}{4}, 1\right\}$

The only budget balanced gaps are 1 and $\frac{1}{4}$, the welfare maximizing of these is $\frac{1}{4}$.

Dominant Strategy Mechanism Design

Dominant Strategy Auctions

Back to the auction setting. $\theta_i \in [\underline{\theta}, \bar{\theta}]$. $I \in \{1, 2, ..., N\}$.

Buyer's Utility: $\theta_i - t_i$ if good is awarded, $-t_i$ otherwise.

Seller's Utility: $\sum_{i=1}^{N} t_i$.

Remark. In dominant strategy mechanism design, ex-ante calculations ARE NOT RELEVANT for individuals. We should not calculate ex-ante utility nor ex-ante budget balance.

Proposition. *Revelation Principle*

Suppose we have a generic game. Γ which has terminal nodes with outcomes that specify a probability distribution over who gets the good and *transfers. Suppose $\sigma_i(\theta_i)$ is a dominant strategy for i and type θ_i, then there is a direct mechanism Γ' which has action space Θ_i for each player, the outcome of any $\sigma(\theta)$ is the same and θ_i is a dominant strategy for i with θ_i.*

Definition. *Dominant Strategy Incentive Compatible [DSIC]*

A direct mechanism $\Gamma(q, t)$ is DSIC if for each type and each buyer:

$\theta_i q_i(\theta_i, \theta_{-i}) - t_i(\theta_i, \theta_{-i}) \geq \theta_i q_i\left(\tilde{\theta}_i, \theta_{-i}\right) - t_i\left(\tilde{\theta}_i, \theta_{-i}\right)$

For all θ_i, $\tilde{\theta}_i$ and $\theta_{-i} \in \Theta_{-i}$.

Example. $\theta_i \in [0, 1]$.

$\theta_i \geq 0.5, \theta_j \geq 0.5$, $q_i = 0.5$, $t_i = 0.25$.

$\theta_i \geq 0.5, \theta_j \leq 0.5$, $q_i = 1$, $t_i = 0.25$.

$\theta_i \leq 0.5, \theta_j \leq 0.5$, $q_i = 0.5$, $t_i = 0.$.

Exercise. Is it DSIC?

Suppose $\theta_j \leq 0.5$. Does i want to tell the truth?

$\theta_i \geq 0.5$. $\theta_i - 0.25 \geq \frac{\theta_i}{2}$? Yes!

$\theta_i \leq 0.5$ $\frac{\theta_i}{2} \geq \theta_i - 0.25$? Yes!

Suppose $\theta_j \geq 0.5$. Does i want to tell the truth?

$\theta_i \geq 0.5$. $\frac{\theta_i}{2} - 0.25 \geq 0$? Yes!

$\theta_i \leq 0.5$ $0 \geq \frac{\theta_i}{2} - 0.25$? Yes!

Definition. *Ex-Post Individually Rational*

A direct mechanism is *Ex-post* individually rational [EPIR] if:

$\theta_i q_i(\theta_i, \theta_{-i}) - t_i(\theta_i, \theta_{-i}) \geq 0$

For all θ_i and $\theta_{-i} \in \Theta_{-i}$.

Exercise. Is the above mechanism *EPIR?*

Note. A dominant strategy mechanism can be thought of as a series of screening mechanisms.

DSIC requires $\theta_i q_i(\theta_i, \theta_{-i}) - t_i(\theta_i, \theta_{-i}) \geq \theta_i q_i\left(\tilde{\theta}_i, \theta_{-i}\right) - t_i\left(\tilde{\theta}_i, \theta_{-i}\right)$ for all θ_i, $\tilde{\theta}_i$ and $\theta_{-i} \in \Theta_{-i}$.

Fix a θ_{-i}. Our goal is to get i to tell the truth. That is a screening mechanism. To construct a DSIC mechanism, we need to construct a series of incentive compatible screening mechanisms for every θ_{-i}!

Because of this, we can characterize dominant strategy mechanisms using what we learned from screening!

Proposition. *Characterization of DSIC Mechanisms*

A direct mechanism is DSIC if and only if for every i and θ_{-i}

1. $q_i(\theta_i, \theta_{-i})$ *is increasing in* θ_i.
2. $t_i(\theta_i, \theta_{-i}) = t_i(\underline{\theta}, \theta_{-i}) + (\theta_i q_i(\theta_i, \theta_{-i}) - \underline{\theta} q_i(\underline{\theta}, \theta_{-i})) - \int_{\underline{\theta}}^{\theta_i} q_i(x, \theta_{-i})\, dc$

Lemma. *Characterization of EPIR Mechanisms*

A direct mechanism is EPIR if and only if for every i and θ_{-i}:

1. $\underline{\theta} q(\underline{\theta}, \theta_{-i}) - t_i(\underline{\theta}, \theta_{-i}) \geq 0$

Proposition. *A class of DSIC and EPIR mechanisms.*

$q_i(\theta) = \frac{1}{n}$ *if* $\psi_i(\theta_i) \geq 0$ *and* $\theta_i \geq \theta_i$. *Where n is number of* $\theta_j = max\{\theta_1, \theta_2, ..., \theta_N\}$

$q_i(\theta) = 0$ *otherwise.*

$t_i = q_i\left[min\left\{\tilde{\theta}_i | q_i\left(\tilde{\theta}_i, \theta_{-i}\right) > 0\right\}\right]$

Exercise. Is it individually rational? Check for $\underline{\theta}$. Is it incentive compatible? Suppose $q = 0$. To get $q \geq 0$, I would have to claim a higher type. But clearly then, the minimum type I have to claim to win is above my real type. In this case, $q\theta_i - q\tilde{\theta} < 0$. What about the other way around?

Applications. Suppose the seller has a belief about Θ, but knows nothing about how the buyers believe. Then, he can still implement the revenue maximizing auction. Let $\psi_i(\theta_i)$ be the buyer's "virtual" type as we found earlier. This can be calculated because the seller has a belief about Θ. However, note that the raffle auction is not canonical.

Suppose the seller knows nothing about Θ. He can still implement the welfare maximizing auction. Let $\psi_i = \theta_i$.

Nothing is lost in allocation by using DSIC!

Dominant Strategy Public Goods

We already know the following result. An incentive compatible, individually rational first-best mechanism does not exist except in trivial situations. Thus, one cannot exist using dominant strategy mechanisms. What can be achieved is further limited.

Lemma. *Characterization of Deterministic DSIC Mechanisms*

A deterministic mechanism is DSIC if and only if for every θ_{-i} there is some $\tilde{\theta}_i$ such that for any $\theta_i \geq \tilde{\theta}_i$, $q = 1$ and $t_i\left(\theta_i, \theta_{-i}\right) - t_i\left(\underline{\theta}, \theta_{-i}\right) = \tilde{\theta}_i$, and for $\theta_i < \tilde{\theta}_i$, $q = 0$.

Lemma. *A class of DSIC and EPIR mechanisms.*

$q_i\left(\theta\right) = 1$ *if* $\sum_I \psi_i\left(\theta_i\right) \geq c$

$q_i\left(\theta\right) = 0$ *otherwise.*

$t_i = q_i\left[min\left\{\tilde{\theta}_i | q_i\left(\tilde{\theta}_i, \theta_{-i}\right) = 1\right\}\right]$

Exercise. Is it individually rational? Check $\underline{\theta}$. Is it incentive compatible?

Example. Let $\psi_i = \theta_i$. This is the pivot mechanism.

Recall that for two people, the utility-maximizing mechanism had, a rule $\theta_1 + \theta_2 \geq s$. Can this be achieved by a DSIC, EPIR mechanism? Yes, but we cannot get budget balance. Why? These mechanisms require a "discount." In general, people are only required to pay the minimum they have to get their desired result. This isn't a problem in auctions. All the money goes to the seller!

Proposition. *Characterization of deterministic DSIC, Exactly-Budget Balances, EPIR Mechanisms*

A deterministic dominant strategy, exactly-budget balanced and ex-post individually rational mechanism has the form:

$q = 1$ *if* $\theta_i \geq \tau_1$ *and* $\theta_2 \geq \tau_2$ *and* $\tau_1 + \tau_2 = c$.

This is an extremely limited form and vastly reduces the available welfare!

Dominant Strategy Bilateral Trade

Proposition. *Characterization for Deterministic DSIC Mechanisms*

A deterministic mechanism for bilateral trade is DSIC if and only if.

For the buyer, and any type of the seller, θ_s, there is some $\hat{\theta}_b$ such that for any $\tilde{\theta}_b \geq \hat{\theta}_b$, $q\left(\tilde{\theta}_b, \theta_s\right) = 1$ and $t_b\left(\tilde{\theta}_b, \theta_s\right) - t_b\left(\underline{\theta}_b, \theta_s\right) = \hat{\theta}_b$

That is, there is some type the buyer can claim where trade will take place and he will pay $\tilde{\theta}_b$ more in transfer.

For the seller, and any type of the buyer, θ_b, there is some $\hat{\theta}_s$ such that for any $\tilde{\theta}_s \leq \hat{\theta}_s$, $q\left(\tilde{\theta}_s, \theta_b\right) = 1$ and $t_s\left(\tilde{\theta}_s, \theta_b\right) - t_s\left(\bar{\theta}_s, \theta_s\right) = \hat{\theta}_s$

That is, there is some type the seller can claim where trade will take place and he will receive $\tilde{\theta}_s$ more in transfer.

Proposition. *A class of DSIC and EPIR mechanisms.*

$q = 1$ *if* $\psi_b\left(\theta_b\right) \geq \psi_s\left(\theta_s\right)$
$t_s = q\left[max\left\{\tilde{\theta}_s | q\left(\tilde{\theta}_s, \theta_b\right) \geq 1\right\}\right]$
$t_b = q\left[min\left\{\tilde{\theta}_b | q\left(\tilde{\theta}_b, \theta_s\right) \geq 1\right\}\right]$

Example. Let $\psi_b = \theta_b$ and $\psi_s = \theta_s$. This is the pivot mechanism.

But it is not budget balanced! Recall that for 2 people, the second-best had sale when $\theta_s - \theta_b \geq \frac{1}{4}$. Let $\psi_b = \theta_b - \frac{1}{4}$ and $\psi_s = \theta_s$.

Definition. Posted Price Mechanism

There is a fixed p. If $\theta_s < p < \theta_b$ then trade takes place and $t = p$ for both. Otherwise, it doesn't and there is no payment.

Proposition. *A deterministic DSIR, EPIR, Exactly-Budget Balanced Mechanism is a Posted Price.*

In fact, posted-price turns out to be the welfare maximizing mechanism among the DSIC, EPIR class even if budget balance is weakened to Ex-post budget balance.

This is the result of Hagerty, Kathleen M., and William P. Rogerson. "Robust trading mechanisms." Journal of Economic Theory 42.1 (1987): 94-107.

Non-Transferable Utility

Until now we have considered utility functions of the type $U_i\left(a, \theta_i\right) + m_i$ where a is an outcome and m_i is some amount of money. We have assumed that the individuals have additive and linear (risk neutral) preference for money (or some other kind of transfer). When that is violated, or when there is no money to transfer, we can take the extreme view of simply making money part of the outcomes, if it exists and letting the utility functions simply be $U_i\left(a, \theta_i\right)$.

However, since there is no money to use for comparison here, the utility values are meaningless. We will instead assume everyone has a preference relation R_i over the set A of outcomes.

Definition. *Unrestricted Domain*

$\mathscr{R}$ is the set of possible preference profiles over A (all linear orderings).

Definition. Direct Mechanism

A mapping $f : \mathscr{R}^N \to A$. Everyone reports their preferences and then an outcome is chosen.

Example. *Voting*

A might be some set of candidates to choose from. f is a decision rule about how to choose candidates. "Voting Problem."

Example. *Matching*

Examples. Suppose A is the set of possible pairings of Men and Women in a group. This is close to "Matching," except that people can have arbitrary preferences over all possible sets of pairs.

Example. *Generalized Versions of Old Problems*

A might be a combination of distributions of money and distribution of some good. For instance, who gets a car and how much money each person ends up with.

Definition. *Social Choice Function*

The decision rules f are called "Social Choice Functions."

Definition. *Incentive Compatibility*

A social choice function is said to be dominant strategy incentive compatible or "Strategy-Proof" if for all R_i and R_{-i}:

$$f\left(R_i, R_{-i}\right) R_i f\left(R_i', R_{-i}\right)$$

That is, the player has incentive to tell the truth about his preferences regardless of other preferences.

We would like to operate in this very abstract realm of the unrestricted domain. **But we can't.**

Definition. *Dictatorial Mechanisms*

f is dictatorial if there is some i such that for all $R \in \mathscr{R}^N$ and $a \in A$, $f(R)\ R_i\, a$.

i is said to be a dictator.

Theorem. *Gibbard-Satterthwaite Theorem*

In the unrestricted domain $\mathscr{R}$, if A has at least 3 elements then f is Strategy-Proof if and only if it is dictatorial.

The *if* is clear. The *only if* deserves some attention. Suppose f is strategy-proof.

Below, I give some intuition for the proof. For more, look at the paper here: https://people.cs.pitt.edu/~kirk/CS1699Fall2014/gibbard-sat.pdf.

Lemma. *Unanimity*

Any strategy proof social choice function respects unanimity.

Proof. Without loss of generality, there must be an outcome where x is chosen. Start there. Raise x in player $1's$ profile, another outcome cannot be chosen or 1 would never report this. Thus x can be raised to the top of $1's$ profile and x is still chosen. This is true for everyone so that if x is on top of everyone's list, then x must be chosen. Now reorder 1's profile below x. The choice must remain x or 1 would never report it. This is true for everyone so the social choice function respects unanimity. □

Proof for $N = 2$ and $M = 3$.

Sudoku Style.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x				
xzy	x	x				
yxz			y	y		
yzx			y	y		
zxy					z	z
zyx					z	z

By Unanimity

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x		Not z		
xzy	x	x				
yxz			y	y		
yzx			y	y		
zxy					z	z
zyx					z	z

Or else player 1 would lie to get x.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x		x		
xzy	x	x				
yxz			y	y		
yzx			y	y		
zxy					z	z
zyx					z	z

Choose this to be x.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x		x		
xzy	x	x				
yxz			y	y		
yzx			y	y		
zxy				**Not x, y**	z	z
zyx					z	z

You cannot give someone their worst match when a better match is unilaterally achievable.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x		x		
xzy	x	x		**x**		
yxz			y	y		
yzx			y	y		
zxy				z	z	z
zyx				**z**	z	z

When everything is achievable in your column/row, you must get your favorite in that column/row.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x	**x**	x	**x**	**x**
xzy	x	x	**x**	x	**x**	**x**
yxz			y	y		
yzx			y	y		
zxy				z	z	z
zyx				z	z	z

If you ever get your worst, you must always get your worst in that column/row.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x	x	x	x	x
xzy	x	x	x	x	x	x
yxz			y	y		**Not z, Not x**
yzx			y	y		
zxy				z	z	z
zyx				z	z	z

You can't give someone their worst when better is unilaterally achievable.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x	x	x	x	x
xzy	x	x	x	x	x	x
yxz			y	y		y
yzx			y	y		**y**
zxy				z	z	z
zyx				z	z	z

When everything is achievable in your column/row, you must get your favorite in that column/row.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x	x	x	x	x
xzy	x	x	x	x	x	x
yxz			y	y		y
yzx			y	y		y
zxy				z	z	z
zyx		**Not x, Not y**		z	z	z

You can't give someone their worst when better is unilaterally achievable.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x	x	x	x	x
xzy	x	x	x	x	x	x
yxz			y	y		y
yzx			y	y		y
zxy				z	z	z
zyx	**Not x, Not y**	z	**Not x, Not y**	z	z	z

If someone gets their second favorite, you can't offer their favorite.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x	x	x	x	x
xzy	x	x	x	x	x	x
yxz			y	y		y
yzx			y	y		y
zxy			**z**	z	z	z
zyx	z	z	z	z	z	z

When everything is achievable in your column/row, you must get your favorite in that column/row.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x	x	x	x	x
xzy	x	x	x	x	x	x
yxz			y	y		y
yzx			y	y		y
zxy	**z**	**z**	z	z	z	z
zyx	z	z	z	z	z	z

If you ever get your worst, you must always get your worst in that column/row.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x	x	x	x	x
xzy	x	x	x	x	x	x
yxz			y	y		y
yzx	**Not x, Not z**		y	y		y
zxy	z	z	z	z	z	z
zyx	z	z	z	z	z	z

You can't give someone their worst when better is unilaterally achievable.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x	x	x	x	x
xzy	x	x	x	x	x	x
yxz	**y**		y	y		y
yzx	y		y	y		y
zxy	z	z	z	z	z	z
zyx	z	z	z	z	z	z

When everything is achievable in your column/row, you must get your favorite in that column/row.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x	x	x	x	x
xzy	x	x	x	x	x	x
yxz	y	**Not x, Not z**	y	y	**Not x, Not z**	y
yzx	y	**Not x, Not z**	y	y	**Not x, Not z**	y
zxy	z	z	z	z	z	z
zyx	z	z	z	z	z	z

If someone gets their second favorite, you can't offer their favorite.

	xyz	xzy	yxz	yzx	zxy	zyx
xyz	x	x	x	x	x	x
xzy	x	x	x	x	x	x
yxz	y	y	y	y	y	y
yzx	y	y	y	y	y	y
zxy	z	z	z	z	z	z
zyx	z	z	z	z	z	z

Single-Peaked Preferences

Since we can't hope to make much progress working in the abstract universe, we are left with two choices:

1. Relax the solution concept.
2. Work in a "restricted domain."

The first is kind of difficult. There is a short section in the book on it, but it is for a very limited case. Instead, we focus on relaxing the domain. Most interesting questions can operate in a relaxed domain. For instance, two-sided matching.

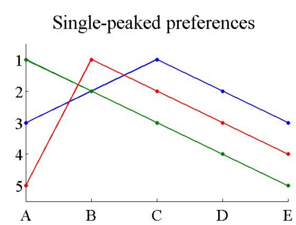
Definition. *Single-Peaked Preferences*

$A = \{1, 2, ..., K\}$ outcomes are labeled.

Each person likes some $k(i)$ best. $k(i)\, R_i a$ for all a.

$k(i) + j\, R_i\, k(i) + j + 1$ and $k(i) - j\, R_i\, k(i) - j - 1$

Example. Candidates on a Left-Right Spectrum.



If there are an even number of agents, create a fake one with fake preferences and repeat (this breaks the tie).

Proposition. *DSIC, Non-Dictatorial Rules Exist for Single-Peaked Preferences*

Definition. *Median Voter Rule*

f chooses the median of $\{k(1), k(2), ..., k(N)\}$.

Proof. Suppose N is odd (even proof is similar with small adjustment). Order the players in terms of their peak. Then the median is $k\left(\frac{N+1}{2}\right)$. Anyone below the median prefers a value at least as low as $k\left(\frac{N+1}{2}\right)$ but, anyone below the median can only raise the median. The opposite is true for anyone above the median. The median person already gets his/her favorite. $\square$

Matching/Market Design

Two-Sided Matching

Marriage Market

Notation. Men: $m_1, m_2, ..., m_n$. Women: $w_1, w_2, ..., w_n$

Each has preferences over the other side.

Notation. Let $P()$ represent the preferences of a player. $P(m_1) = w_1, w_2, m_1, w_3, ..., w_p$

Assumption. ***Preferences are strict.***

Definition. *Matching*

A matching μ is a mapping such that $\mu(m_i) \in W \cup \{m_i\}$ and vice versa. The matching must be of "order two" such that $\mu(\mu(m_i)) = m_i$ and $\mu(\mu(w_i)) = w_i$.

This guarantees no two people are matched to the same person.

Mechanism Design Perspective

The set of all possible μ is the set of outcomes A. Here, we don't operate in the universal domain. Each individual only has preferences over their own mate, not over the entire matching. That is, m_1 doesn't care about who m_2's spouse is. His utility is the same under any matching where he has the same spouse.

We want to implement some f on A over this domain of preferences. But what f?

Definition. *Individually Rational*

μ is individually rational if $\mu(m_i) \succ_{m_i} m_i$ and $\mu(w_i) \succ_{w_i} w_i$. That is, everyone likes their partner better than being alone.

Definition. *Stable*

μ is stable if there does not exist a pair $\tilde{m}$ and $\tilde{w}$ such that. $\tilde{w} \succ_{\tilde{m}} \mu(\tilde{m})$ and $\tilde{m} \succ_{\tilde{w}} \mu(\tilde{w})$. That is, there is not a pair that prefer to be paired with each other than their match under μ. These pairs are called "blocking pairs" since the block the stability of μ.

Proposition. *Stability Implies Pareto Efficiency*

Pareto improvement implies at least one person i is better off under $\tilde{\mu}$ than in (stable) μ. But in $\tilde{\mu}$, i's partner (j) must be worse off, since otherwise, i and j would be a blocking pair in μ.

In two-sided matching we would like f to select a stable and individually rational matching from the given preference profiles.

Exploring Stability

(Assume everyone likes being matched better than being alone).

M preferences	w_1	w_2	w_3	W preferences	m_1	m_2	m_3
m_1	1	3	2	w_1	1	2	3
m_2	2	1	3	w_2	1	2	3
m_3	3	2	1	w_3	1	2	3

Suppose μ is $\{(m_1, w_3), (m_2, w_2), (m_3, w_1)\}$. This is not stable. m_1, w_1 is a blocking pair.

Suppose μ is $\{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$. This is stable. There is no blocking pair.

In fact, this is the unique stable outcome. However, sometimes there are multiple stable outcomes.

M preferences	w_1	w_2	w_3	W preferences	m_1	m_2	m_3
m_1	1	2	3	w_1	2	1	3
m_2	2	1	3	w_2	1	2	3
m_3	3	2	1	w_3	1	2	3

This has two stable matchings

$\{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$

$\{(m_1, w_2), (m_2, w_1), (m_3, w_3)\}$

Notice that in these two stable matches, the men can all agree that $\{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$ is at least as good as $\{(m_1, w_2), (m_2, w_1), (m_3, w_3)\}$. Notice that the women agree on the opposite. This is called the lattice structure of stable matchings.

Proposition. *Lattice Structure*

Among all stable matches we can order them. $\mu_1 \succ_M \mu_2$ if all men agree that μ_1 is at least as good. $\succ_W$ will be defined analogously. It turns out that $\succ_M$ is always the opposite of $\succ_W$!!! If the men agree μ_1 is better than the women will agree that it is worse.

Notation. Let's refer to μ_M as the male-optimal and μ_W as the women-optimal stable match.

There are two questions to ask, *does a stable match always exist* and, *is there a mechanism that chooses a stable match from a preference profile*. Both can be answered at once.

Gale-Shapley Mechanism (Algorithm)

Let's start with the "Male's Propose" version. The women line up. The men go and stand in front of their favorite women. The woman keeps her favorite man who is in front of her. Any rejected man then moves to the next favorite woman whom he hasn't already stood in front of. At any point, a man can refuse to move on, and remain alone. After all the proposals have been made, any pairs are matched and anyone left over remains alone.

Lemma. *The algorithm ends.*

At some point, the males will have made proposals to everyone.

Proposition. *The match is stable.*

Proof. Suppose otherwise. There is a blocking pair. But $\tilde{m}$ must have visited $\tilde{w}$ before his current match. She rejected him then and so must now be matched with someone whom she likes at least as well as $\tilde{m}$. Thus, this can't be a blocking pair.

□

Corollary. *Every MM has a stable outcome.*

Proof. Because GS always completes. □

Proposition. *The match is male-optimal.*

Proof. Suppose otherwise. There is a man $\tilde{m}$ who was the first to be rejected by his "optimal" woman $\tilde{w}$. That means there is a man $\tilde{n}$ who $\tilde{w}$ rejected $\tilde{m}$ for at the first instance. $\tilde{w}$ likes $\tilde{n}$ more than $\tilde{m}$. But $\tilde{n}$ must like $\tilde{w}$ at least as well as his optimal as well, since when he proposed to $\tilde{w}$, he had not yet been rejected by *his* optimal woman. Thus $\tilde{n}$ likes $\tilde{w}$ better than his optimal woman. This implies that the proposed optimal stable match is not stable at all. $\tilde{n}$ and $\tilde{w}$ are a blocking pair. □

If we switch the proposing sides, we get the woman-optimal match.

Thus, Gale-Shapley "implements" the social choice function f which chooses the proposer-optimal stable match given any preference profile. The question is. Is it DSIC?

Proposition. *M-GS is not DSIC for W*

Proof. Suppose everyone tells the truth, and there are at least two stable matches. Then any women can truncate her preferences by saying that being alone is better than being matched with anyone worse than her W-optimal partner. This truncation cannot create any new stable matchings. *Intuition:* If this woman was involved in a blocking pair with someone the truncated for some match, she will still block that match by threatening to be alone. The W-optimal stable with original preferences is still stable. It is not the only stable match and GS must produce it! □

Proposition. *Matchings and Stable Outcome*

Anyone matched must be matched in any stable outcome.

Proposition. *There is no DSIC Mechanism to Implement stable f.*

Proof. Any f that always produces a stable outcome will have to produce the unique stable outcome if there is one. If there is more than one stable outcome, there is someone who prefers another stable outcome to the one that is chosen, but he/she can achieve that outcome by truncating his/her preferences. □

Proposition. *There **are** DSIC Mechanisms that Implement Pareto Efficient f*

Example. Sequential Dictator-takes advantage of the indifferences in preferences.

M preferences	w_1	w_2	w_3	W preferences	m_1	m_2	m_3
m_1	1	2	3	w_1	2	1	3
m_2	2	1	3	w_2	1	2	3
m_3	3	2	1	w_3	1	2	3

This has two stable matchings

$\{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$

$\{(m_1, w_2), (m_2, w_1), (m_3, w_3)\}$

What is S.D. (m_1, m_2, m_3) Outcome? $\{(m_1, w_1), (m_2, w_2), (m_3, w_3)\}$ [Stable]

What is S.D. (w_1, w_2, w_3) Outcome? $\{(m_2, w_1), (m_1, w_2), (m_3, w_3)\}$ [Stable]

What is S.D. (w_3, w_2, w_1) Outcome? $\{(m_3, w_1), (m_2, w_2), (m_1, w_3)\}$ [Unstable] (who blocks?)

Roommates

The roommate's problem is similar to marriage problem, except that everyone can be matched to anyone. Each r_i has preferences over entire set (not just one side).

Notation. Let $P()$ represent the preferences of a player. $P(r_1) = r_2, r_3, r_4, r_1$

Assumption. Preferences are strict.

Definition. *Matching*

A matching μ is a mapping such that $\mu(r_i) \in R$ and vice versa. The matching must be of "order two" such that $\mu(\mu(r_i)) = r_i$.

This guarantees no two people are matched to the same person.

This problem is much harder and has received much less attention!

Definition. *Individually Rational*

μ is individually rational if $\mu(r_i) \succ_{m_i} r_i$. That is, everyone likes their partner better than being alone.

Definition. *Stable*

μ is stable if there does not exist a pair r_i and r_j such that. $r_j \succ_{r_i} \mu(r_i)$ and $r_i \succ_{r_j} \mu(r_i)$. That is, there is not a pair that prefer to be paired with each other than their match under μ. These pairs are called "blocking pairs" since the block the stability of μ.

Proposition. *There is not always a stable outcome.*

$$P(r_1) = r_2, r_3, r_4, r_1$$

$$P(r_2) = r_3, r_1, r_4, r_2$$

$$P(r_3) = r_1, r_2, r_4, r_3$$

$$P(r_4) = r_1, r_2, r_3, r_4$$

In any outcome, someone is paired with r_4. Whoever that person is, is someone's favorite roommate. That person is willing to save him from r_4.

Proposition. *There may be a stable outcome.*

$$P(r_1) = r_2, r_3, r_4, r_1$$

$$P(r_2) = r_4, r_1, r_3, r_2$$

$$P(r_3) = r_2, r_1, r_4, r_3$$

$$P(r_4) = r_1, r_2, r_3, r_1$$

$\{1,3\}, \{2,4\}$ is the unique stable outcome.

Algorithm. *Irving's Algorithm*

Stage 1.

Step 1. Everyone proposes to their favorite. Anyone with no proposals accepts. Anyone with two proposals accepts their favorite and rejects the other.

Step 2. Upon being rejected, a proposer immediately moves to their next favorite.

Step 3. When everyone has a proposal that has been accepted, stage 1 is done. If there is someone whom everyone has rejected, there is no stable match.

Stage 2.

In a stable match, no one can be matched with anyone less desirable than their currently accepted roommate. Cross off everyone below that, and do it symmetrically. If B is unacceptable for A then also cross out A for B.

Stage 3.

Remove preference cycles.

Start with someone who has two or more choices. Who is their second favorite? Who is that person's least favorite? Who is that person's second favorite? We look for a cycle in the "last favorites."

Example. *Irving's Algorithm-No Stable Match Exists*

4	6	2	5	3
6	3	5	1	4
4	5	1	6	2
2	6	5	1	3
4	2	3	6	1
5	1	4	2	3

1 proposes to 4 who accepts.

2 proposes to 6 who accepts.

3 proposes to 4 who rejects.

3 proposes to 5 who accepts.

4 proposes to 2 who accepts.

5 proposes to 4 who accepts (rejecting 1)

1 proposes to 6 who accepts (rejecting 2)

2 proposes to 3 who accepts

6 proposes to 5 who rejects

6 proposes to 1 who accepts

4	**6**	2	5	3
6	**3**	5	1	**4**
4	5	**1**	6	**2**
2	6	**5**	1	3
4	2	**3**	6	1
5	**1**	4	2	3

Now cross off:

4	**6**	2	5	3
6	**3**	5	1	**4**
	5	**1**	6	**2**
2	6	**5**	1	
4	2	**3**	6	1
5	**1**	4	2	3

	6			
	3	5		**4**
	5			**2**
2		**5**		
4	2	**3**		
	1			

6		
3	5	4
5	2	
2	5	
4	2	3
1		

Now preference cycles.

6		
3	5	4
5	2	
2	5	
4	2	3
1		

2 to 5

3 to 2

4 to 5

3

We have a cycle

Remove 2,4 and 5,3

6	
3	5
2	
5	
4	2
1	

2 to 5

2

Remove 2,5

6
3
2
5
4
1

(1,6); (2,3); (4,5)

4	6	2	5	3
6	3	5	1	4
4	5	1	6	2
2	6	5	1	3
4	2	3	6	1
5	1	4	2	3

Phase 1 leads to:

5	6
6	4
4	5

1 to 3

2 to 1

3 to 2

1 to 3

Remove 3,2 and 1,3 and 2,1

5	6
6	4
4	5

There is no stable match.

Definition. *Individually Rational.*

μ is Pareto efficient if $\mu(r_i) \succ_{r_i} r_i$. That is, everyone likes their partner better than being alone.

Definition. *Pareto Efficient.*

μ is Pareto efficient if there does not exist $\tilde{\mu}$ such that for all i: $\mu(r_i) \succ_{r_i} \tilde{\mu}(r_i)$.

Pareto Efficient Mechanisms.

Example. *Rotating Dictator [RD]*

They go in order, picking their favorites.

Proposition. *RD is Pareto Efficient*

Proof. Suppose there is a Pareto improvement $\tilde{\mu}$. There is some player i who was the earliest to propose who has a different roommate. But he already has his favorite from the people who were not removed before he picked and since he is the first, everyone before him must be on the same teams. □

Proposition. *RD is not Individually Rational.*

Example. *Rotating Representative [RR]*

They go in order. Proposing to their favorites who accept or reject.

Proposition. *RR is Pareto Efficient*

Proof. Suppose the outcome μ resulting from the Rotating Representative mechanism is Pareto-suboptimal. There is a $\tilde{\mu}$ that is a Pareto-improvement over μ. Of those on a different team in $\tilde{\mu}$, there is an i who was the first to propose in *RR*. Suppose $\tilde{\mu}$ includes i being matched with at least one partner not available when i proposed under *RR*, then there is another player who proposed before i who is on a different team under $\tilde{\mu}$. This contradicts that i is the first such proposer. Thus, the team that i receives in $\tilde{\mu}$ were available when i proposed under *RR*. i must have proposed to this team and it must have been rejected. Thus, at least one person on this team prefers $\mu$to $\tilde{\mu}$ contradicting that it is a Pareto-improvement. Thus, μ is Pareto-optimal. □

Many-to-One Matching

Definition. Colleges

$C = \{C_1, C_2, ..., C_n\}$ with space for $q_1, q_2, ..., q_n$ students each.

Students

$S = \{s_1, s_2, ..., s_m\}$

Preferences

Assumption. *Preferences*

Students have a strict ordering over $C \bigcup \{s\}$*.*

Colleges have a strict ordering over $S \cup \{c\}$ *where* $\{c\}$ *represents that any students lower in the ordering are worse than having unfilled quota.*

Definition. *Matching*

A matching μ is function that pairs each student with no more than one college and each college with no more than q_c students.

Assumption. *$\mu(s)$ is the students' college (or remaining alone)*

$\mu(C)$ is the set of students attending the college and possibly some unfilled slots: indicated by redundant inclusions of C.

Definition. *Pairwise Stability*

A matching is ***pairwise stable*** if it is not blocked by an individual (a student wanting to remain alone or a college wanting to remove a student without replacement) or a student college blocking pair (s, C) such that $C \succeq_s \mu(s)$ and $s \succeq_C \left[min_{\succeq_C}(\mu(C))\right]$ where $min_{\succeq}$ indicates the element of $\mu(C)$ which is minimal according to the ordering $\succeq_C$.

Definition. Group Stable

A matching μ is ***group stable*** if it is not blocked by a coalition A where $A \subseteq S \bigcup C$ such that there is another μ' in which for every $s \in A$, $\mu'(s) \in A$ and $\mu'(s) \succeq_s \mu(s)$ and $s \in \mu'(C)$ then either $s \in A$ or $s \in \mu(C)$ and $\mu'(C) \succeq_C \mu(C)$.

Proposition. *A Matching is* **Group Stable** *if and only if it is* **Pairwise Stable.**

Proof. *Only if is trivial. We prove if. Suppose otherwise. There is a pairwise stable match that is not group stable. Then there is a blocking coalition A under μ'. Pick $C \in A$. C likes its students under μ' better than under μ. That means there is at least some student in A that C likes better than at least one of the current students (or open slots) and that student likes C better than his current match. This contradicts pairwise-stability.* □

Algorithm. *A Stable Matching Algorithm (1951)*

Make Student-College pairs in the following way.

1. Find all pairs such that the student lists a college first and the college lists the student in its top q. If there is none, look for instances where the student listed a college second and the college listed the student in the top q. Keep doing this until you find at least 1 match.

2. The students who are matched are guaranteed to do this well. Cross of any worse college from their list and cross them from those college's list. Then repeat step 1 until no new matches are found.

Example. *Three Colleges, Four Students.*

C_1 has two slots and the others have 1 slot.

$s_1 : 3, 1, 2$

$s_2 : 2, 1, 3$

$s_3 : 1, 3, 2$

$s_4 : 1, 2, 3$

$C_1 : 1, 2, 3, 4$

$C_2 : 1, 2, 3, 4$

$C_3 : 3, 1, 2, 4$

Round 1. Match s_1, s_2 and C_1.

Cross off 3 from s_2's list and 2 from s_1's list.

$s_1 : 3, 1$

$s_2 : 2, 1$

$s_3 : 1, 3, 2$

$s_4 : 1, 2, 3$

$C_1 : 1, 2, 3, 4$

$C_2 : 2, 3, 4$

$C_3 : 3, 1, 4$

Round 2. Match s_2 with C_2

$s_1 : 3, 1$

$s_2 : 2$

$s_3 : 1, 3, 2$

$s_4 : 1, 2, 3$

$C_1 : 1, 3, 4$

$C_2 : 2, 3, 4$

$C_3 : 3, 1, 4$

Round 3. Match s_3 with C_1

$s_1 : 3, 1$

$s_2 : 2$

$s_3 : 1$

$s_4 : 1, 2, 3$

$C_1 : 1, 3, 4$

$C_2 : 2, 4$

$C_3 : 1, 4$

Round 4. Match s_1 with C_3

$s_1 : 3$

$s_2 : 2$

$s_3 : 1$

s_4 : 1, 2, 3
C_1 : 3, 4

C_2 : 2, 4

C_3 : 1, 4
Round 5. Match s_4 with C_1

s_1 : 3

s_2 : 2

s_3 : 1

s_4 : 1
C_1 : 3, 4

C_2 : 2

C_3 : 1
Match is s_1, C_3, s_2, C_2, s_3, s_4, C_1

Proposition. *The Algorithm Is Stable*

Proof. Suppose otherwise, there is at least one s and C who like each other better. But, for the match to finish C must be matched with its top q in the modified preferences. Which means that C was removed from s preferences at some point, implying that s was matched to a college better than C. But since s's matching can only get better, this creates a contradiction. □

Example. *A Stable Matching Algorithm (Redux)*

Try Gale-Shapley (hospitals propose) in the matching market:

C_1, s_1, s_2. C_2, s_1. C_3, s_3

C_1, s_1, s_2. C_2, s_2. C_3, s_3

C_1, s_1, s_3. C_2, s_2.

C_1, s_3. C_2, s_2. C_3, s_1

C_1, s_3, s_4. C_2, s_2. C_3, s_1

This is the same outcome. In fact, the mechanisms always produce the same outcomes!

If Students-Propose:

s_1, C_3. s_2, C_2. s_3, C_1 s_4, C_1

The first is the hospital-optimal stable match, the second is the student-optimal.

Proposition. *Stable Outcomes and Pareto Efficiency (By Side)*

While the student-optimal stable match is Pareto-optimal for students, the hospital-optimal match might not be Pareto-optimal for hospitals.

Example. Consider s_2, s_4, H_1, s_1, H_2 and s_3, H_3. It is a Pareto improvement for the hospitals.

Theorem. *Rural Hospital*

When preferences are strict, the set of students hired, and the set of positions filled, is the same in any stable matching. Any college that does not fill its quota in some stable matching, gets the same set of students in any stable matching.

Proposition. *There is no Stable Mechanism that is Strategy-proof*

Proof. The accepting side can manipulate the mechanism as before. If there are two stable matches, the accepting side can pretend that all but the students it gets in the best stable match are acceptable. □