PERSPECTIVES ON MECHANISM DESIGN IN ECONOMIC THEORY

Roger Myerson, 8 Dec 2007

http://home.uchicago.edu/~rmyerson/research/nobelnts.pdf

The scope of economics has changed

In Xenophon’s original **Oeconomicus** (c 360 BCE) a model citizen

- motivates workers on his farms in the country
- maintains his political status in the city to keep his farm.

Agents’ incentives & political institutions:
concerns of economics then and now, but not always....

Focus of economics c. 1800: **production & allocation of material goods**.
Classical economic problem: Limited resources, unlimited desires.
Result: Free trade can achieve **allocative efficiency**
(reallocating resources cannot improve everyone’s welfare).

Analysis of individuals’ incentives began as a tool for explaining supply and demand in price theory (from Cournot 1838).

Game theory began with mathematicians studying optimal decisions in more general competitive frameworks. (Borel 1921,
von Neumann 1928, vN & Morgenstern 1944, Nash, 1951)

Beyond price theory: socialism v. capitalism

In early 20th century **inconclusiveness of debates** about socialism versus capitalism (Barone, Lange; Mises, Hayek) showed limits of price theory for evaluating non-price institutions.

“The economic problem of society is not merely a problem of how to allocate 'given' resources... It is rather a problem of how to secure the best use of resources known to any of the members of society, for ends whose relative importance only these individuals know. ...it is a problem of the utilization of knowledge not given to anyone in its totality.” Hayek, 1945

Leo Hurwicz took up Hayek's challenge.

The pivotal moment: **Hurwicz (1972)** examined incentives to communicate information, introduced **incentive compatibility**.

A breakthrough after Hurwicz (1972)

Hurwicz's **mechanisms** are plans for how social decisions should depend on people's information (mapping information to decisions).

Harsanyi (1967) provided the **general Bayesian model of games** and equilibria when people have different information.

Changing the coordination mechanism changes the game people play.

So Hurwicz's mechanism design became the **theory of game design**.

The revelation principle (*found independently 1977-81 by Dasgupta Hammond Maskin, Harris Townsend, Holmstrom, Myerson, Rosenthal; building on ideas of Gibbard 1973 and Aumann 1974*) characterizes the outcomes of all possible equilibria of all games that can be designed with different coordination mechanisms.

This feasible set satisfies **incentive constraints**, which say that people will not share private information or exert hidden efforts without appropriate incentives.

Adverse selection and moral hazard

When individuals have private information and choose hidden actions, social planners face two kinds of incentive constraints:

Informational incentive constraints (adverse selection): individuals need incentives to report their private information honestly.

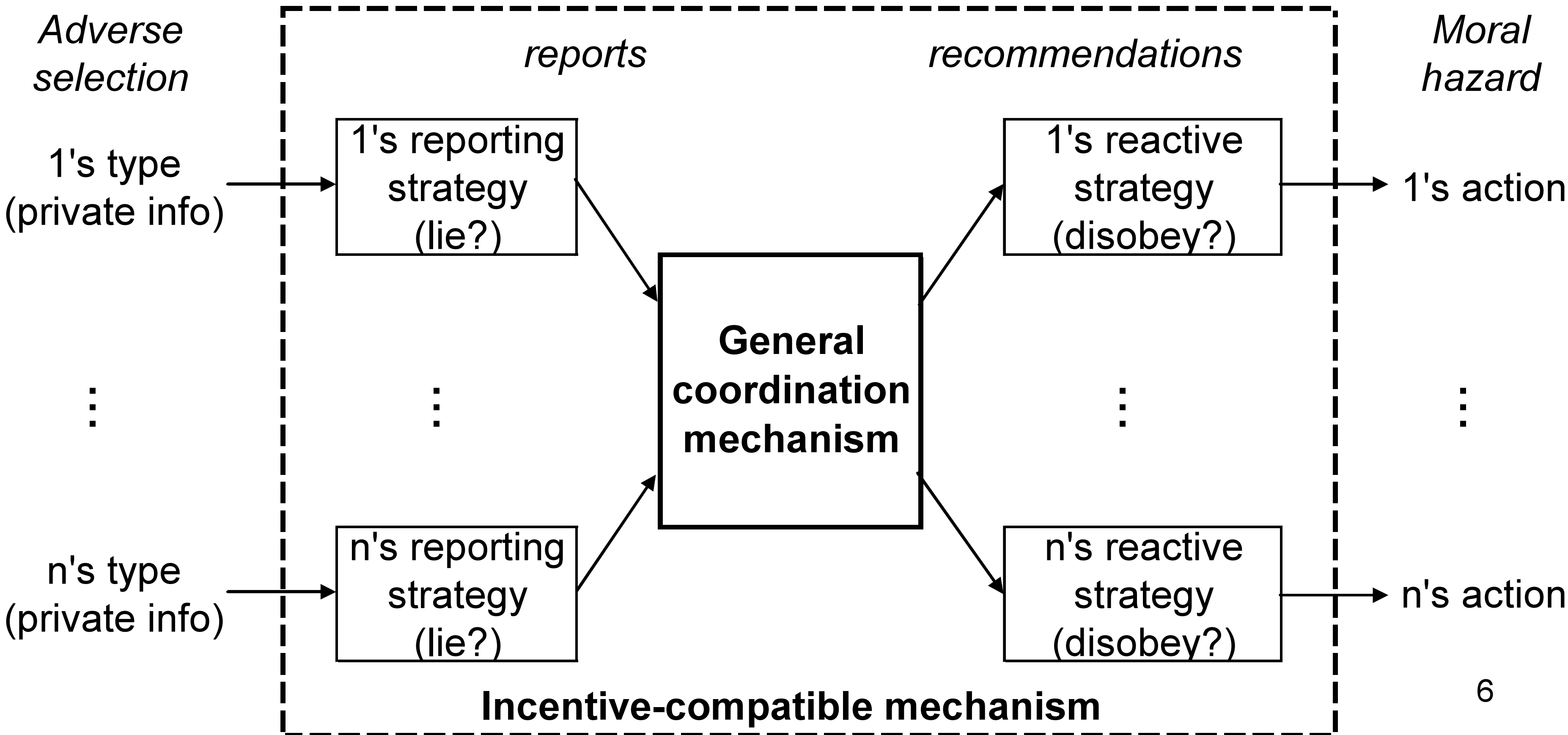
Strategic incentive constraints (moral hazard): individuals need incentives to act obediently according to the plan.

In an **incentive-compatible coordination plan**, individuals send confidential reports to a central mediator, who then confidentially recommends their actions under the plan, such that it is an equilibrium for everyone to report honestly and act obediently.

Revelation principle

Without loss of generality, a trustworthy mediator can plan to make honesty and obedience the best policy for everyone.

For **any coordination plan, any equilibrium** of people's (dishonest) reporting and (disobedient) reactions is **equivalent to an incentive-compatible** plan that makes honesty-and-obedience an equilibrium.



Examples of Mechanism Design

We will consider three simple examples, to illustrate the fundamental importance of incentive constraints in the economy:

- Trading example: adverse selection problems in sale of one object by one seller to one potential buyer.
  Incentives to bargain for a better price can deter allocatively efficient trades.

- Production example: moral hazard in management
  Incentives for good management may require that that manager has a valuable stake in the business. (The model Hayek sought?)

- Politics and the economy: moral hazard in the government.
  Capital investors may require credible political guarantees against the ruler's temptation to expropriate them.

Trading example: one seller, one buyer, one object

Each knows own private value of the object which may be
$0 (weak) or $80 (strong) for seller, each with probability 1/2,
$100 (weak) or $20 (strong) for buyer, each with probability 1/2.

Trade would be mutually beneficial unless both are strong,
but what should be the price?

A mediator who assists them must plan how their transaction may depend on the information that they reveal. This is a mechanism.

	Buyer's value [strong] $20	Buyer's value [weak] $100
Seller's value [strong] $80	0, *	1, $90
Seller's value [weak] $0	1, $10	1, $50

P(trade), E(price if trade)

Split-the-difference mediation plan

Failure of incentive compatibility for split-the-difference

This might seem a fair way to get **beneficial trade with probability 3/4**, but unfortunately it is **not incentive compatible**: honesty by both traders is not an equilibrium of this game.

If the buyer were expected to be honest, then a weak seller could get higher expected profits by claiming to be strong.

(Similarly, weak buyer would lie if she expected the seller to be honest.)

P(trade) = 3/4?

.5(10)+.5(50) < .5(90)

Buyer's value: [strong] \$20 → "\$20"; [weak] \$100 → "\$100" (or → "\$20")

Seller's value: [strong] \$80 → "\$80"; [weak] \$0 → "\$0" (or → "\$80")

	"\$20"	"\$100"
"\$80"	0, *	1, \$90
"\$0"	1, \$10	1, \$50

P(trade), E(price if trade)

Split-the-difference mediation plan or mechanism

Symmetric mediation plans

Given similarity of seller and buyer here, let's treat them symmetrically.

Let **q** denote the conditional probability of trade when one trader is weak and the other is strong. So $0 \le q \le 1$.

Let **y** be the weak type's profit margin in trading with a strong opponent.

Incentive constraints: $y \le 20$ (participation), $q \le 25/(50-y)$ (honesty).

Incentive compatibility:
$.5qy+.5(50) \ge .5q(100-y)$
Participation: $y \ge 20$

Buyer's value: [strong] \$20 → "\$20"; [weak] \$100 → "\$100" (or → "\$20")

Seller's value: [strong] \$80 → "\$80"; [weak] \$0 → "\$0" (or → "\$80")

Seller's value		"\$20"	"\$100"
[strong] \$80	"\$80"	0, *	q, \$100-y
[weak] \$0	"\$0"	q, \$y	1, \$50

P(trade), E(price if trade)

General symmetric mediation plan or mechanism

An incentive-efficient plan (ex ante)

Incentive constraints $q \leq 25/(50-y)$ and $y \leq 20$ imply that the **largest feasible probability of trade** is achieved by letting $y = 20$, $q = 5/6$.

This mechanism is **not allocatively efficient**, as it yields a positive (1/12) probability of failing to achieve a mutually beneficial trade,

It maximizes total expected gains of trade subject to incentive constraints.

So it is an **incentive-efficient mechanism**: no other incentive-compatible mechanism gives greater expected higher expected gains to both traders (evaluating expectations **ex ante**: before they learn their types).

y=20, q=5/6:			Buyer's value	
	Seller's value		\$20 [s]	\$100 [w]
EU(str)=\$0	[s]	\$80	0, *	5/6, \$80
EU(wk)=\$33.33	[w]	\$0	5/6, \$20	1, \$50
			P(trade), E(price if trade)	

P(trade)=2/3 (max!)

(.5)(5/6)20+(.5)50 = (.5)(5/6)80

But strong types never get any profit from trade in this plan.

Other interim incentive-efficient trading plans

The ex-ante incentive-efficient plan gives no profit to strong types.

Other symmetric plans with q=25/(50-y), y<20, are better for strong types and are **interim incentive efficient**: in the sense that no other incentive-compatible plan would be preferred by every type of every individual (Holmstrom Myerson, 1983)

(**Interim**: evaluate expected payoffs for each person given own type only.)

y=10, q=5/8:

	Seller's value		Buyer's value $20 [s]	Buyer's value $100 [w]
EU(str)=$3.125	[s]	$80	0, *	5/8, $90
EU(wk)=$28.125	[w]	$0	5/8, $10	1, $50

P(trade), E(price if trade)

y=0, q=1/2:

	Seller's value		Buyer's value $20 [s]	Buyer's value $100 [w]
EU(str)=$5	[s]	$80	0, *	1/2, $100
EU(wk)=$25	[w]	$0	1/2, $0	1, $50

Neutral barg soln

P(trade), E(price if trade)

(Latter is a generalized Nash bargaining solution of Myerson 1984.)

What if we don't worry about incentives for honesty?

Some may lie. Split-the-difference has three reporting equilibria:

(1) Seller always claims strong, buyer is honest.

Seller's value		Buyer's value $20 [s]	Buyer's value $100 [w]
[s]	$80	0, *	1, $90
[w]	$0	0, *	1, $90

P(trade), E(price if trade)

(2) Buyer always claims strong, seller is honest.

Seller's value		Buyer's value $20 [s]	Buyer's value $100 [w]
[s]	$80	0, *	0, *
[w]	$0	1, $10	1, $10

P(trade), E(price if trade)

(3) Both lie randomly with probability 0.6 if weak.

Seller's value		Buyer's value $20 [s]	Buyer's value $100 [w]
[s]	$80	0, *	0.4, $90
[w]	$0	0.4, $10	0.64, $50

P(trade), E(price if trade)

Each equilibrium is equivalent to an incentive-compatible mediation plan.
But these do not seem so fair or efficient!

General incentive constraints without symmetry for this example.

$y_{j,k}$ = (Expected price if trade with seller-type j & buyer-type k),

$0 \leq q_{j,k}$ = (Probability of trade if seller-type is j & buyer-type is k) ≤ 1.

Participation constraints:

$0.5q_{80,20}(y_{80,20} - 80) + 0.5q_{80,100}(y_{80,100} - 80) \geq 0,$

$0.5q_{0,20}(y_{0,20} - 0) + 0.5q_{0,100}(y_{0,100} - 0) \geq 0,$

$0.5q_{80,20}(20 - y_{80,20}) + 0.5q_{0,20}(20 - y_{0,20}) \geq 0,$

$0.5q_{80,100}(100 - y_{80,100}) + 0.5q_{0,100}(100 - y_{0,100}) \geq 0.$

Informational incentive constraints:

$0.5q_{80,20}(y_{80,20} - 80) + 0.5q_{80,100}(y_{80,100} - 80)$
$\quad \geq 0.5q_{0,20}(y_{0,20} - 80) + 0.5q_{0,100}(y_{0,100} - 80),$

$0.5q_{0,20}(y_{0,20} - 0) + 0.5q_{0,100}(y_{0,100} - 0)$
$\quad \geq 0.5q_{0,20}(y_{80,20} - 0) + 0.5q_{80,100}(y_{80,100} - 0),$

$0.5q_{80,20}(20 - y_{80,20}) + 0.5q_{0,20}(20 - y_{0,20})$
$\quad \geq 0.5q_{80,100}(20 - y_{80,100}) + 0.5q_{0,100}(20 - y_{0,100}),$

$0.5q_{80,100}(100 - y_{80,100}) + 0.5q_{0,100}(100 - y_{0,100})$
$\quad \geq 0.5q_{80,20}(100 - y_{80,20}) + 0.5q_{0,20}(100 - y_{0,20}).$

The Coase Theorem and incentive constraints

Coase (1960): If there were no transactions costs, then free trade could achieve allocative efficiency from any initial allocation of property rights.

Incentive constraints are a fundamental source of transactions costs.
Participation constraints express initial property rights.
(Samuelson, 1985)

Reassigning the initial ownership of the object makes it easy for a mediator to guarantee that the object ends up with the individual who values it most! (Trade if and only if both are willing to trade at 50.)

	Seller's value	
Buyer's value	$20	$100
$80	1, $50	0, *
$0	0, *	0, *

P(trade), E(price if trade)

Moral hazard in production (Tirole 2006)

A project's probability of success depends on manager's hidden effort.

To deter abuse of power, **manager must have stakes** to lose in failure.

Under socialist egalitarianism, who has stakes commensurate with the temptations in managing industrial concentrations of capital?

p_G = P(success if act good) = 1/2, p_B = P(success if act bad) = 1/4,
K = (capital input) = 100, R = (returns if success) = 240,
B = (agent's private benefit of bad action) = 30. *So* $p_G R > K > p_B R+B$.

Given agent's collateral A < 60, choose w = (wage if success) $\geq$ –A
to maximize expected social profit $V = p_G(R - w) + (1-p_G)A - K$
subject to: $p_G w - (1-p_G)A \geq 0$, [*participation*]
$p_G w - (1-p_G)A \geq B + p_B w - (1-p_B)A$. [*G-obedience*]

Solution: w = 120 – A, and so V = A –40.
$V \geq 0$ is not feasible unless the agent has collateral $A \geq 40$.
The agent gets moral-hazard rents worth $p_G w-(1-p_G)A = 60-A$.

For egalitarianism, punish managers who fail?

In this example, even if we allow punishment of managers who fail, the investing society cannot expect to profit from the investment unless the manager has substantial assets to lose (A>20).

The state could profitably motivate managers by punishing failure with A=0 if there were no participation constraint (managers coercively recruited).

This example may formalize Hayek's intuition against socialism.

$p_G = 1/2$, $p_B = 1/4$, $K = 100$, $R = 240$, $B = 30$, $A < 60$.

Choose w = (wage if success) $\geq -A$ and z = (punishment if fail) ≥ 0
to maximize expected social profit $V = p_G(R - w) + (1-p_G)A - K$
subject to: $p_G w - (1-p_G)(A+z) \geq 0$, [*participation*]
$p_G w - (1-p_G)(A+z) \geq B + p_B w - (1-p_B)(A+z)$. [*G-obedience*]

Solution: $z = 60-A$, $w = 60$, and so $V = 0.5A - 10$.
$V \geq 0$ is not feasible unless the agent has collateral $A \geq 20$.

Without participation constraint: $w=0$, $z \geq 120$ motivates G, yields $V=20$.

The analogous adverse-selection example

A project's probability of success depends on the manager's hidden type, good or bad.

Socialist monopoly of capital can facilitate honest communication, as bad agents cannot gain from imitating good if nobody gets profits. (Dewatripont Maskin 1993)

Collectivizing property can ameliorate adverse-selection problems, but it can exacerbate moral-hazard problems.

Given $p_G R > K > p_B R$ [*E(Return*| *Good type) > Kapital > E(Return*| *Bad)*]

Choose (q_G, q_B, w_G, w_B) to maximize expected social profit:

$$V = \alpha\, q_G[p_G(R - w_G) + (1-p_G)A - K] + (1-\alpha)q_B[p_B(R - w_B) + (1-p_B)A - K]$$

subject to: $w_G \geq -A$, $w_B \geq -A$, $0 \leq q_G \leq 1$, $0 \leq q_B \leq 1$, [*resources*]

$q_G[p_G w_G - (1- p_G)A] \geq 0$, $q_B[p_B w_B - (1- p_B)A] \geq 0$, [*participation*]

$q_G[p_G w_G - (1- p_G)A] \geq q_B[p_G w_B - (1- p_G)A]$, [*honesty-G*]

$q_B[p_B w_B - (1- p_B)A] \geq q_G[p_B w_G - (1- p_B)A]$. [*honesty-B*]

In socialism, the ideal $q_G=1$, $q_B=0$ is feasible even if $A=0$, with $w_G=0=w_B$.

In capitalism, competitive lending implies $V=0$ in equilibrium, but then the ideal $q_G=1$, $q_B=0$ is not feasible if agent's collateral A is small.

Moral hazard at the center of government

Costs of unrestrained central power can be understood from models of moral hazard at the center of government.

To encourage investments that increase his tax base, even a ruler may prefer to create political guarantees of private property rights, even when such liberalization entails a risk of his losing power.

Incentives for such liberalization may depend on natural resources.

$Y(K)$ = (output flow if K invested) = $(K+n)^{0.5}$, n = (natural resources) = 12,
r = (interest rate) = 0.05, b = (basic political-risk rate) = 0.05,
a = (additional risk per liberalization) = 0.05. [K is durable, mobile.]

Choose K = (capitalist investment) and λ = (liberalization) to
maximize the ruler's expected value $V = (Y(K) - rK)/(r+b+a\lambda)$
subject to $V \geq (1-\lambda)(K + Y(0)/(r+b))$. [*no incentive to expropriate*]

With n=12, ruler's optimal regime is: $\lambda = 0.504$, $K = 52.4$. ($\lambda=0 \Rightarrow K=0$.)
With $n = 0$, optimal regime becomes $\lambda = 0$, $K = 44.44$.
With $n = 25$, optimal regime becomes $\lambda = 0$, $K = 0$.

Conclusions

- Mechanism design added incentive constraints to resource constraints in our definition of the economic problem.
- Incentive constraints can explain failures of classical allocative efficiency. But we can identify good social rules that are incentive efficient.
- The cases for collectivism or private ownership may depend on trade-offs between moral-hazard and adverse-selection incentive problems.
- Moral-hazard problems are fundamental in any institution. Institutional rules are enforced by actions of leaders and officials who must be motivated by an expectation of rewards and privileges as long as they fulfill their institutional responsibilities.
- Economics was once defined by the allocation of material goods, but now it covers all questions about incentives in institutions. With mechanism design and game theory, we can analyze competitive incentive problems in both markets and politics.

We have returned to the breadth of vision of first ancient economists.