The Principal Agent model

The revelation principle

Definition. A mechanism is a message space M and a mapping $h(\cdot)$ from M to the space of outcomes which writes as $h(m) = (Q(m), t(m))$ for all m belonging to M.

Any mechanism induces an allocation rule.

Assume here quasilinear preferences: $\theta v(Q) - t$

Let:

$$m^*(\theta) \in \arg\max_{m \in M} \theta v(Q(m)) - t(m)$$

Then a mechanisms M induces the allocation rule:

$$a(\theta) = Q(m^*(\theta)), t(m^*(\theta)).$$

A *direct mechanism* is a mechanism in which $M_i = \Theta_i$: the message space for i is i's type space.

Is there loss of generality in restricting attention to direct mechanisms?

Definition. A *direct revelation mechanism* is a mapping $g(\bullet)$ from the space of types to the space of outcomes which writes as $g(\theta_i) = (q(\theta_i), T(\theta_i))$ for all θ_i.

The principal commits to offer $q(\theta_i)$ at a price $T(\theta_i)$ if the agent reports to be of type θ_i.

Definition. An agent finds it *incentive compatible* to announce his/her type in correspondence to g if and only if:

$$\theta v(q(\theta)) - T(\theta) \geq \theta' v(q(\theta')) - T(\theta')$$

Definition. A direct revelation mechanism $g(\bullet)$ is *truthful* if it is incentive compatible for the agent to announce his true type for any type.

Definition. A direct revelation mechanism $g(\bullet)$ is *individually rational* if $\theta v(q(\theta)) - T(\theta) \geq \underline{u}$ for any type.

Theorem (**Revelation Principle**). Any possible allocation rule $a(\theta)$ obtained with a mechanism $\{M, h(\bullet)\}$ can also be implemented with a truthful direct revelation mechanism.

Proof. We will show that if an outcome function is implemented by a mechanism, then it can be implemented by a direct mechanism as well.

This implies that there is no loss of generality in studying direct mechanisms.

Mechanism $\{M, h(\bullet)\}$ induces an outcome function

$$g(\theta) = Q(m^*(\theta)), T(m^*(\theta)).$$

Construct the functions $\widehat{Q} = Q \circ m^*$, $\widehat{T} = T \circ m^*$, so that:

$$\widehat{Q}(\theta), \widehat{T}(\theta) = Q(m^*(\theta)), T(m^*(\theta))$$

This is a direct mechanism implementing outcome $g(\theta)$.

Is it truthful?

To verify that $g(\theta)$ is a direct, truthful mechanisms we need to verify truthfulness. Since:

$$m^*(\theta_i) \in \arg\max_{m \in M} \theta_i v(Q(m)) - t(m)$$

We must have:

$$\theta_i v(Q(m^*(\theta_i))) - TQ(m^*(\theta_i)) \geq \theta_i v(Q(m^*(\theta_j))) - TQ(m^*(\theta_j))$$

$$\Rightarrow \theta_i v(\widehat{Q}(\theta_i)) - \widehat{T}(\theta_i) \geq \theta_i v(\widehat{Q}(\theta_j)) - \widehat{T}(\theta_j)$$

for any θ_i, θ_j.

The optimal direct mechanism with 2 types

The seller's problem can be written as:

$$\max_{T_i, q_i} \beta(T_L - cq_L) + (1-\beta)(T_H - cq_H)$$

$$s.t. \quad \begin{array}{c} \theta_L v(q(\theta_L)) - T(\theta_L) \geq \theta_L v(q(\theta_H)) - T(\theta_H) \; IC_L \\ \theta_H v(q(\theta_H)) - T(\theta_H) \geq \theta_H v(q(\theta_L)) - T(\theta_L) \; IC_H \\ \theta_H v(q(\theta_H)) - T(\theta_H) \geq 0 \; IR_H \\ \theta_L v(q(\theta_L)) - T(\theta_L) \geq 0 \; IR_L \end{array}$$

To solve this problem we proceed in steps.

Step 1

Note that IR_L and IC_H implies IR_H:

$$\begin{aligned}\theta_H v(q(\theta_H)) - T(\theta_H) &\geq \theta_H v(q(\theta_L)) - T(\theta_L) \\ &\geq \theta_L v(q(\theta_L)) - T(\theta_L) \geq 0\end{aligned}$$

Step 2

Consider a relaxed version of the problem in which we ignore IC_L:

$$\max_{T_i, q_i} \beta(T_L - cq_L) + (1-\beta)\beta(T_H - cq_H)$$

$$s.t. \quad \begin{aligned} \theta_H v(q(\theta_H)) - T(\theta_H) &\geq \theta_H v(q(\theta_L)) - T(\theta_L) \; IC_H \\ \theta_L v(q(\theta_L)) - T(\theta_L) &\geq 0 \; IR_L \end{aligned}$$

Note that the value of this program is not lower than the value of the original program.

If, once we have solved it, we can prove that indeed IC_L is satisfied at the solution, then the two values coincide.

Step 3

Note that if $\theta_L v(q(\theta_L)) - T(\theta_L) > 0$, then we can increase $T(\theta_L)$ without violating any other constraint (indeed relaxing IC_H).

This change increases the payoff, a contradiction

The case for IC_H is similar.

Step 4

Note that from IC_H we can write:

$$
\begin{aligned}
\theta_H v(q(\theta_H)) - T(\theta_H) &= \theta_H v(q(\theta_L)) - T(\theta_L) \\
&= \theta_L v(q(\theta_L)) - T(\theta_L) + (\theta_H - \theta_L) v(q(\theta_L)) \\
&= (\theta_H - \theta_L) v(q(\theta_L))
\end{aligned}
$$

Substituting this and IR_L, seller's program becomes:

$$\max_{T_i, q_i} \beta(\theta_L v(q(\theta_L)) - cq(\theta_L)) + (1-\beta)\begin{pmatrix} \theta_H v(q(\theta_H)) - cq_H \\ -(\theta_H - \theta_L)v(q_L) \end{pmatrix}$$

For the contract we solve this problem.

Note that the objective function above, W, is not necessarily concave.

Observe that:

$$W_{q_H,q_H} = (1-\beta)(\theta_H v''(q(\theta_H))) < 0$$
$$W_{q_H,q_L} = W_{q_L,q_H} = 0$$

So this problem is concave if the hessian is negative (semi-)definite:

$$W_{q_L,q_L} = \beta(\theta_L v''(q(\theta_L))) - (1-\beta)(\theta_H - \theta_L)v''(q_L) < 0$$

This is however no always the case.

It is the case if β is high enough, or $\theta_H - \theta_L$ is small enough.

We will see more examples below.

We assume here that concavity is satisfied.

Our focs are:

$$\begin{aligned}\theta_H v'(q(\theta_H)) &= c\\ \theta_L v'(q(\theta_L)) &= \frac{c}{1-\left(\frac{1-\beta}{\beta}\frac{\theta_H-\theta_L}{\theta_L}\right)}\end{aligned}$$

Note that: $q(\theta_H) > q(\theta_L)$.

For this to be a solution we need to verity that IC_L is satisfied.

From the binding IC_H we have:

$$\begin{aligned}\theta_H v(q(\theta_H)) - T(\theta_H) &= \theta_H v(q(\theta_L)) - T(\theta_L) \\ &\rightarrow \theta_H[v(q(\theta_H)) - v(q(\theta_L))] = T(\theta_H) - T(\theta_L) \\ &\rightarrow \theta_L[v(q(\theta_H)) - v(q(\theta_L))] \leq T(\theta_H) - T(\theta_L)\end{aligned}$$

Implying:

$$\theta_L v(q(\theta_L)) - T(\theta_L) \geq \theta_L v(q(\theta_H)) - T(\theta_H) \quad IC_L$$

We conclude that the solution of the relaxed problem is a solution of the original problem.

Why did I need to wait until now to establish IC_L?

Because I needed to show $q(\theta_H) \geq q(\theta_H)$ for the argument.

Solution then is:

$$
\begin{aligned}
\theta_H v'(q(\theta_H)) &= c \\
\theta_L v'(q(\theta_L)) &= \frac{c}{1 - \left(\frac{1-\beta}{\beta}\frac{\theta_H - \theta_L}{\theta_L}\right)}
\end{aligned}
$$

Note:

- High types buy more than low types.
- High types buy the efficient quantity; low types less than efficient.
- The low type receives a surplus of zero; the high type receives a positive surplus.

Continuous types

Let us now assume that we have a continuum of types $\theta \in [0,1]$ (without loss of generality)

The distribution of types is F.

The utility function is $u(q,\theta)$ with $u_\theta(q,\theta) > 0$, $u_{\theta q}(q,\theta) > 0$ or alternatively $u_\theta(q,\theta) < 0$, $u_{\theta q}(q,\theta) < 0$

A direct mechanism is now a function $h(\theta) = (q(\theta), T(\theta))$

A direct mechanism is incentive compatible if:

$$u(q(\theta),\theta) - T(\theta) \geq u(q(\theta'),\theta) - T(\theta') \text{ for all } \theta,\theta'$$

The optimal contract can now be written as:

$$\max_{T,q} \int T(\theta) - C(q(\theta))dF(\theta)$$

$$s.t. \quad \begin{array}{c} u(q(\theta),\theta) - T(\theta) \geq u(q(\theta'),\theta) - T(\theta') \text{ for all } \theta,\theta' \\ u(q(\theta),\theta) - T(\theta) \geq 0 \text{ for all } \theta \end{array}$$

We first study the constraint set, then the solution of this problem.

Implementability

A direct mechanism $h = (q, T)$ is compact valued if

$$\left\{(q, T) \text{ s.t. } \exists\theta' \text{ such that } q, T = (q(\theta'), T(\theta'))\right\}$$

is compact.

We now show that if $u_{\theta q}(q, \theta) > 0$ and a direct mechanism $h(\theta)$ is compact valued then $h(\theta)$ is incentive compatible if and only if:

$$U(\theta'') - U(\theta') = \int_{\theta'}^{\theta''} u_\theta(q(x), x)dx \text{ for any } \theta'', \theta' \text{ s.t. } \theta' < \theta''$$

and $q(\theta)$ is non decreasing

Necessity

$IC(\theta';\theta)$ constraint implies:

$$\begin{aligned} U(\theta) &= u(q(\theta),\theta) - T(\theta) \geq u(q(\theta'),\theta) - T(\theta') \\ &= U(\theta') + [u(q(\theta'),\theta) - u(q(\theta'),\theta')] \end{aligned}$$

Or:

$$U(\theta) - U(\theta') \geq [u(q(\theta'),\theta) - u(q(\theta'),\theta')]$$

Similarly $IC(\theta;\theta')$ implies:

$$U(\theta') - U(\theta) \geq [u(q(\theta),\theta') - u(q(\theta),\theta)]$$
$$\Rightarrow U(\theta) - U(\theta') \leq [u(q(\theta),\theta) - u(q(\theta),\theta')]$$

We have:

$$u(q(\theta'),\theta) - u(q(\theta'),\theta') \leq U(\theta) - U(\theta') \leq u(q(\theta),\theta) - u(q(\theta),\theta')$$

The single crossing condition implies that $q(\theta) \geq q(\theta')$ for $\theta \geq \theta'$.

Moreover we have:

$$\frac{u(q(\theta'),\theta) - u(q(\theta'),\theta')}{\theta - \theta'} \leq \frac{U(\theta) - U(\theta')}{\theta - \theta'} \leq \frac{u(q(\theta),\theta) - u(q(\theta),\theta')}{\theta - \theta'}$$

Taking the limit as $\theta - \theta' \to 0$, we have:

$$U'(\theta) = u_\theta(q(\theta),\theta)$$

at all points of continuity of $q(\theta)$.

Now observe that:

- given that h is compact valued;
- u is continuous.

Then $U(\theta)$ must be continuous by the theorem of the maximum since:

$$U(\theta) = \max_{\theta' \in [0,1]} \{u(q(\theta'), \theta) - T(\theta')\}$$

Since $U(\theta)$:

- is continuous over a compact set;
- with bounded derivative (at all point of existence).

Then the fundamental theorem of calculus implies that it is integrable.

Sufficiency

Assume:

$U(\theta'') - U(\theta') = \int_{\theta'}^{\theta''} u_\theta(q(x),x)dx$ for any θ'',θ' s.t. $\theta' < \theta''$

and $q(\theta)$ is non decreasing

If the mechanism is not IC then there must be a θ and a θ' such that

$$\begin{aligned} U(\theta') + u(q(\theta'),\theta) - u(q(\theta'),\theta') &= u(q(\theta'),\theta) - T(\theta') \\ &\geq u(q(\theta),\theta) - T(\theta) = U(\theta) \end{aligned}$$

and the reverse.

So we can write:

$$
\begin{aligned}
u(q(\theta'),\theta) - u(q(\theta'),\theta') &> U(\theta) - U(\theta') \\
&= u(q(\theta),\theta) - u(q(\theta'),\theta') \\
&= \int_{\theta'}^{\theta} u_\theta(q(x),x)dx
\end{aligned}
$$

Or:

$$\int_{\theta'}^{\theta} u_\theta(q(\theta'),x)dx > \int_{\theta'}^{\theta} u_\theta(q(x),x)dx$$

That is:

$$\int_{\theta'}^{\theta} [u_\theta(q(\theta'), x) - u_\theta(q(x), x)]dx > 0$$

But using the monotonicity of $q(x)$, we have:

$$u_\theta(q(\theta'), x) - u_\theta(q(x), x) \leq u_\theta(q(\theta'), x) - u_\theta(q(\theta'), x) = 0$$

a contradiction.

Solving the seller's problem

It follows that the optimal contract is:

$$\max_{T,q} \int [T(\theta) - C(q(\theta))] dF(\theta)$$

$$s.t. \begin{cases} U(\theta) = \int_0^\theta u_\theta(q(x), x)dx \\ q(\theta) \text{ non decreasing} \\ u(q(0), 0) - T(0) = 0 \end{cases}$$

Note that $T(\theta) - C(q(\theta)) = S(q(\theta), \theta) - U(\theta)$.

So we can write it as:

$$\max_{U,q} \int [S(q(\theta), \theta) - U(\theta)] dF(\theta)$$

$$s.t. \quad U(\theta) = \int_0^\theta u_\theta(q(x), x) dx$$

$q(\theta)$ non decreasing and $U(0) = 0$

We can substitute the first constraint in the profit function.

We obtain:

$$\pi(q) = \max_{U,q} \int [S(q(\theta),\theta) - U(\theta)] dF(\theta)$$

$$= \max_{U,q} \int \left[S(q(\theta),\theta) - \int_0^{\theta} u_{\theta}(q(x),x) dx \right] f(\theta) d\theta$$

Remember that by integration by parts we have:

$$\int_a^b kz'dx = [kz]_a^b - \int_a^b k'zdx$$

Let us apply this to:

$$-\int_0^1 \left[\int_0^\theta u_\theta(q(x),x)dx \right] f(\theta)d\theta$$

Letting

$$z = -[1 - F(\theta)] \text{ so } z' = F'(\theta) = f(\theta)$$

$$\text{and } k = \int_0^\theta u_\theta(q(x),x)dx \text{ so } k' = u_\theta(q(x),x).$$

We have:

$$\begin{aligned}
EU(\theta) &= \int_0^1 U(\theta)dF(\theta) \\
&= \int_0^1 \int_0^\theta u_\theta(q(x),x)dx \bullet F'(\theta)d\theta \\
&= -[U(\theta)[1 - F(\theta)]_0^1 + \int_0^1 u_\theta(q(x),x) \bullet [1 - F(\theta)]d\theta \\
&= U(0) + E\left[u_\theta(q(\theta),\theta)\frac{1 - F(\theta)}{f(\theta)} \right]
\end{aligned}$$

So the problem becomes:

$$\max_{q} \int \left[S(q(\theta),\theta) - u_{\theta}(q(\theta),\theta)\frac{1-F(\theta)}{f(\theta)} - U(0) \right] dF(\theta)$$

$s.t.\ q(\theta)$ non decreasing and $U(0) = 0$

This problem is not necessarily concave and does not necessarily an interior solution.

In the following we assume that:

$$\Phi(q,\theta) = S(q,\theta) - u_\theta(q,\theta)\frac{1 - F(\theta)}{f(\theta)}$$

is quasiconcave in q and has a unique interior maximum.

Sufficient conditions for quasi concavity are:

$S(q,\theta)$, typically uncontroversial

$u_\theta(q,\theta)$ not too concave

The focs are:

$$S'(q(\theta),\theta) - u_\theta(q(\theta),\theta)\frac{1-F(\theta)}{f(\theta)} = 0$$

Assume that $u(q,\theta) = q\theta$ and $C(q) = \frac{q^2}{2}$. Then we have:

$$S'(q(\theta),\theta) - u_\theta(q(\theta),\theta)\frac{1-F(\theta)}{f(\theta)} = \theta - q(\theta) - \frac{1-F(\theta)}{f(\theta)}q(\theta)$$

Note that under these assumptions $\Phi(q,\theta)$ is concave and has a unique interior maximum:

$$q(\theta) = \theta - \frac{1-F(\theta)}{f(\theta)}$$

To prove that this is a solution we need to verify monotonicity.

A necessary and sufficient condition for monotonicity of the solution of the relaxed problem is that $\Phi_{q\theta}(q,\theta) \geq 0$ for all q,θ.

To see this differentiate the foc and obtain:

$$\begin{aligned} \Phi_{qq}(q,\theta)dq + \Phi_{q\theta}(q,\theta)d\theta &= 0 \\ \rightarrow \frac{dq}{d\theta} &= -\frac{\Phi_{q\theta}(q,\theta)}{\Phi_{qq}(q,\theta)} \end{aligned}$$

A sufficient condition for this is that $u_{q\theta} \geq 0$ and $u_{q\theta\theta}(q,\theta) \leq 0$ and that types satisfy the monotone hazard rate condition, that is: $\frac{f}{1-F}$ non decreasing.

In the example seen above we have $u = \theta q$, $u_{q\theta} = 1$, $u_{q\theta\theta} = 0$ so the MHRC alone is sufficient.

What have we learned?

There is a trade off between efficiency and incentives:

$$S(q(\theta),\theta) - \frac{1-F(\theta)}{f(\theta)} u_\theta(q(\theta),\theta)$$

This leads to quantities that are distorted, lower than efficient.

The previous formulation of surplus is very similar to the formulation with discrete types:

$$S(q_i,\theta_i) - \frac{1-P_i}{p_i}[u(q_{i-1},\theta_i) - u(q_i,\theta_i)]$$

We still have no distortion at the top, but now this concerns a measure zero of types (only the highest type).