Lecture 16

Monday, 31 October 2022 11:01 AM

Mechanism Design - contd.

Recall: m alternatives A, n voters N
each voter i has total order π_i over A
$\pi_i(a) > \pi_i(b) \Rightarrow i$ prefers a to b
(also called "preference")
$\pi = (\pi_1, \pi_2, \ldots, \pi_n)$, $\pi_{-i} = (\pi_1, \ldots, \pi_{i-1}, \pi_{i+1}, \ldots, \pi_n)$

A Social Welfare Fn (SWF) $F: (\pi_1, \ldots, \pi_n) \to \sigma$
where σ is a total order over A
A Social Choice Fn (SCF) $f: (\pi_1, \ldots, \pi_n) \to A$

SWFs:

① F is unanimous if:
$\exists a, b\ \forall i\ \pi_i(a) > \pi_i(b) \Rightarrow \sigma(a) > \sigma(b)$

② F is independent of irrelevant alternatives (IIA) if:
$\forall \pi, \pi', a, b\ \left(\pi_i(a) > \pi_i(b) \Leftrightarrow \pi_i'(a) > \pi_i'(b) \right.$
$\left. \Rightarrow \sigma(a) > \sigma(b) \Leftrightarrow \sigma'(a) > \sigma'(b) \right)$
where $\sigma = F(\pi)$, $\sigma' = F(\pi')$

③ F is a dictatorship if $\exists k: \forall \pi\ \pi_k(a) > \pi_k(b)$
$\Rightarrow \sigma(a) > \sigma(b)$
where $\sigma = F(\pi)$

Arrow's Theorem: If F is unanimous, IIA, and $|A| \geq 3$ then F must be a dictatorship

Today: What about SCFs?

The property we want SCFs to satisfy is called incentive compatibility: agents should want to truthfully reveal their preferences.

Defn: f is IC if $\forall i, \forall \pi_i, \forall \pi_{-i}, \pi_i'$,
$\pi_i(f(\pi_i, \pi_{-i})) \geq \pi_i(f(\pi_i', \pi_{-i}))$

i.e., truthtelling should be a weakly dominant strategy (informally) for all voters.

(Compare w/ equilibrium defn: $s = (s_1, \ldots, s_n)$ is an eq. if
$\forall i, \forall s_i'$, $u_i(s_i, s_{-i}) \geq u_i(s_i', s_{-i})$)

Claim: If f is IC and f is onto A, then f is unanimous, i.e.,
$\forall \pi$ s.t. $\exists a\ \forall i, \forall b \neq a$, $\pi_i(a) > \pi_i(b)$
$\Rightarrow f(\pi) = a$

Proof: Suppose not. $\exists \pi: \exists a\ \forall i\ \forall b \neq a\ \pi_i(a) > \pi_i(b)$ & $f(\pi) \neq a$.
Since f is onto A, $\exists \tau: f(\tau) = a$.
Let: $\pi^{(0)} = \tau$
$\pi^{(1)} = (\pi_1, \tau_2, \ldots, \tau_n)$
$\pi^{(i)} = (\pi_1, \ldots, \pi_i, \tau_{i+1}, \ldots, \tau_n)$
$\pi^{(n)} = (\pi_1, \ldots, \pi_n)$
$f(\pi^{(0)}) = a$, $f(\pi^{(n)}) \neq a$
Then $\exists k \in [n]: f(\pi^{(k-1)}) = a$, $f(\pi^{(k)}) \neq a$
$\pi^{(k-1)} = (\pi_1, \ldots, \pi_{k-1}, \tau_k, \ldots, \tau_n) \longrightarrow a$
$\pi^{(k)} = (\pi_1, \ldots, \pi_{k-1}, \pi_k, \tau_{k+1}, \ldots, \tau_n) \not\to a$
In π_k, $\pi_k(a) > \pi_k(b)\ \forall b \neq a$
But $f(\pi^{(k)}) \neq a$
$f(\pi^{(k-1)}) = a$
Thus, $\pi_k\left(f(\tau_k, \pi^{(k-1)}_{-k})\right) > \pi_k\left(f(\pi_k, \pi^{(k-1)}_{-k})\right)$
Hence, f is not IC. ∎

Defn. f is a dictatorship if $\exists k$ s.t. $\forall \pi$, $\pi_k(a) > \pi_k(b)\ \forall b \neq a$
$\Rightarrow f(\pi) = a$

Theorem (Gibbard-Satterthwaite Impossibility Theorem):
If SCF f is unanimous, $|A| \geq 3$, and f is onto, then f is a dictatorship.

Defn: Given π_i, $S \subseteq A$, π_i^S is defined as:
$\pi_i^S(a) > \pi_i^S(b)$ if ① either $a \in S$, $b \notin S$
② (either $a, b \in S$ or $a, b \notin S$) and $\pi_i(a) > \pi_i(b)$

E.g.: $\pi_i = (b, f, e, a, c, d)$, $S = \{a, b, c\}$
$\pi_i^S = (b, a, c, f, e, d)$

Assume $|A| \geq 3$ henceforth.

Claim: If f is IC & onto A, then $\forall \pi$, $S \subseteq A$, $f(\pi^S) \in S$
(note that this generalizes previous claim on unanimity)
(proof v. similar, skipped)

Example:

$\pi_1 = (b, f, a, e, c, d)$ $\pi_1^S = (b, a, c, f, e, d)$
$\pi_2 = (b, a, d, e, c, f)$ $\pi_2^S = (b, a, c, d, e, f)$
$\pi_3 = (a, b, c, d, e, f)$ $\pi_3^S = (a, b, c, d, e, f)$
$S = \{a, b, c\}$ $f(\pi^S) \in S$

Claim: If f is IC & onto A, then $\forall \pi$, $T \subseteq S \subseteq A$
if $f(\pi^S) \in T\ (\subseteq S)$, then $f(\pi^T) = \pi^S$
(prove yourself)

For example above, say $f(\pi^S) = b$.
Then $f(\pi^{\{a,b\}}) = f(\pi^{\{b,c\}}) = f(\pi^{\{b\}}) = b$

Corollary: If f is IC & onto A, then $\forall \pi$, $S \subseteq A$
$f(\pi^S) = a \Rightarrow \forall T \subseteq S: a \in T$, $f(\pi^T) = a$

To prove the G-S theorem on SCFs, we will use Arrow's theorem on SWFs.
Suppose $\exists$ an SCF f that is IC, onto A, & not a dictatorship.
We will construct a SWF F that is IIA, unanimous, & not a dictatorship.

Given any IC, onto SCF f, construct F as follows:
given π, a, b
$F(\pi)(a) > F(\pi)(b)$ if $f(\pi^{\{a,b\}}) = a$

Claim: F thus constructed is a total order
Proof: We need to show that F is transitive, i.e., $\forall \pi, a, b, c$,
$F(\pi)(a) > F(\pi)(b)$ & $F(\pi)(b) > F(\pi)(c)$
$\Rightarrow F(\pi)(a) > F(\pi)(c)$
or, $f(\pi^{\{a,b\}}) = a$, $f(\pi^{\{b,c\}}) = b$
$\Rightarrow f(\pi^{\{a,c\}}) = a$
Consider $f(\pi^{\{a,b,c\}})$
$\neq c$, since $f(\pi^{\{b,c\}}) = b$
$\neq b$, since $f(\pi^{\{a,b\}}) = a$
hence $f(\pi^{\{a,b,c\}}) = a$
But then $f(\pi^{\{a,c\}}) = a$. ∎

Hence, F thus constructed is a SWF.

Claim: If f is IC, onto A, & not a dictatorship, then F is unanimous, IIA, & not a dictatorship.

Proof:

① F is unanimous: $\forall i\ \pi_i(a) > \pi_i(b) \Rightarrow F(\pi)(a) > F(\pi)(b)$
Consider $f(\pi^{\{a,b\}})$. In $\pi^{\{a,b\}}$, each agent i has $\pi_i: (a, b, \ldots)$
Hence by the first claim (on unanimity for SCFs),
$f(\pi^{\{a,b\}}) = a$. Hence $F(\pi)(a) > F(\pi)(b)$.

② F is IIA: $\forall \pi, \forall \pi', a, b: \left(\forall i,\ \pi_i(a) > \pi_i(b) \Leftrightarrow \pi_i'(a) > \pi_i'(b)\right.$
$\left.\Rightarrow \sigma(a) > \sigma(b) \Leftrightarrow \sigma'(a) > \sigma'(b)\right)$
(prove yourself, use incremental changes as in previous proofs)

③ F is not a dictatorship: $\forall k\ \exists \pi, a, b: \pi_k(a) > \pi_k(b)$
& $F(\pi)(b) > F(\pi)(a)$
Fix k.
Since f is not a dictatorship, $\exists \pi, a, b$ s.t. $\pi_k(a) > \pi_k(b)$
but $f(\pi) = b$,
Then consider $\pi^{\{a,b\}}$. By claim, since $\{a,b\} \subseteq A$ & $f(\pi^A) = b$,
$f(\pi^{\{a,b\}}) = b$. Note that $\pi_k^{\{a,b\}}(a) > \pi_k^{\{a,b\}}(b)$
But $F(\pi^{\{a,b\}})(b) > F(\pi^{\{a,b\}})(a)$. Hence k cannot be a dictator for F.

This completes proof of the G-S theorem.

Thus, if agents can only express ordinal preferences, we cannot get IC mechanisms.

Cardinal Mechanisms with Money

We assume now that agents have "cardinal" values for the alternatives:
$\forall i,\ v_i: A \to \mathbb{R}$

Further, agents can be compensated w/ money, i.e., they have quasi linear utilities:
$u_i: A \times \mathbb{R} \to \mathbb{R}$
$u_i(a, p) = v_i(a) - p$

Problem: Single good auction.
A single item is to be given to one of n bidders
Further, $v_i(a) = w_i$ if $a = i$
$= 0$ o.w.
(i.e., agent i gets value w_i if it gets the item, and 0 otherwise).

Design a mechanism that takes as input bids from the agents, and
① is IC,
② maximizes $SW = \sum_{i \in N} v_i(a)$, where a is chosen alternative.

$M: \mathbb{R}^n \longrightarrow A \times \mathbb{R}^n$
($\mathbb{R}^n$: n bids; A: outcome; $\mathbb{R}^n$: payments)

IC: It should be a weakly dominant strategy for each agent to truthfully bid their value w_i.

Possible mechanisms:

a. Take bids from agents, give good to highest bidder
b. Give good to highest bidder, winning bidder pays its bid, others pay zero.
c. Give good to highest bidder. Winning bidder pays second highest bid, others pay zero.

Let b_i be i^{th} agent's bid, w_i be its true value.
$b = (b_1, \ldots, b_n)$
$M(b) = k$, where $k \in \arg\max \{b_j\}$
agent k pays $\max_{j \neq k} \{b_j\}$, others pay zero.

Claim: The described mechanism is IC and maximizes social welfare
Proof: Sufficient to show mechanism is IC.
Fix any agent i. Let b_{-i} be other bids.
Let $p = \max_{j \neq i} b_j$.

① Suppose $w_i \geq p$. Then by bidding w_i, utility is $w_i - p$.
if $b_i < p$, utility is zero.
if $b_i \geq p$, utility is $w_i - p$

② Suppose $w_i < p$. Then by bidding w_i, utility is 0.
if $b_i < p$, utility is zero.
if $b_i \geq p$, utility is $w_i - p < 0$.

This is known as Vickrey's Second Price auction.