

14.19 Cheatsheet 2

Rowe Chen Zhong

Arrow's Theorem

Universal Domain: Admits all of Π^n . **Pareto Efficiency:** $\forall a \neq b$ if all $a \succ_i b$ then $a \succ b$. **Independence of Irrelevant Alternatives:** $a \succ b$ is a function of $(a \succ_i b)_i$ only.

Arrow's Theorem: If $n \geq 3$, UD, PE, and IIA, then the only social welfare function is a dictatorship.

Social Choice Functions and Mechanisms

Revelation Principle: Any SCF which can be implemented by a Bayesian-Nash equilibrium can also be implemented by a BNIC mechanism. Any SCF which can be implemented in dominant-strategies can also be implemented by a SP mechanism.

Pareto Efficient if for any state of the world, it chooses a social outcome such that there exists no other social outcome that would make everybody weakly better off and at least one agent strictly better off. **Dictatorial** if there exists an agent i such that for all states $t \in T$, $\phi(t)$ is the top social alternative of i when his type is t_i .

Gibbard-Satterthwaite Theorem: If $|A| \geq 3$ is finite and types of agents are such that the utility functions represent *all strict preference rankings* on A , then any Pareto-efficient and strategy-proof mechanism is necessarily dictatorial.

Notably, utilities which allow for indifference between alternatives are not covered by this theorem.

Vickrey-Clarke-Groves Mechanism

Each agent receives a monetary transfer designed to incentivize truthful reporting of their preferences. The transfer for agent i is determined as follows:

$$y_i(t) = \sum_{j \neq i} v_j(t_j, d^*(t)) - \sum_{j \neq i} v_j(t_j, d^*(t_{-i}))$$

where: - $d^*(t)$ is the efficient outcome when all agents' types are reported. - $d^*(t_{-i})$ is the efficient outcome when agent i is excluded from the decision-making process (i.e., $t_{-i} = (t_1, t_2, \dots, t_{i-1}, t_{i+1}, \dots, t_n)$). - The transfer $y_i(t)$ represents the difference in the welfare of the other agents when agent i is included versus when agent i is excluded. This is called the **marginal contribution** of agent i to society.

The utility that agent i derives from participating in the mechanism is:

$$u_i(t_i, t) = v_i(t_i, d^*(t)) + y_i(t)$$

This utility consists of two parts: the direct utility from the chosen outcome $d^*(t)$, and the transfer $y_i(t)$. **Truthful Reporting as a Dominant Strategy** In the VCG mechanism, truthful reporting of an agent's type t_i maximizes their utility, as their payment y_i is designed to account for their contribution to the overall welfare. Thus, truth-telling is a **dominant strategy** for each agent. The mechanism is **efficient** because it is strategy-proof (it is not GTO to deviate from telling the truth if nobody else deviates) and conditioned on truth-telling it selects the utilitarian alternative (trivially).

Groves: Adding some $x_i(t_{-i})$ to the transfer rule y_i does not affect the efficiency of VCG; hence all of these are strategy-proof. **Holmstrom-Green-Laffont:** If $d: T \rightarrow A$ is efficient, (d, y) is strategy-proof, and all type spaces are complete, then the transfer rule must be VCG (plus constants).

Housing Market

1. **Housing Market** (Shapley - Scarf, 1974):

$$\langle I, H, \succ, \mu \rangle$$

where: • I : set of agents. • H : set of houses with $|H| = |I|$. • \succ : list of strict preferences over houses. • μ : initial endowment matching. • **Pareto Efficiency:** Matching μ is Pareto efficient if there is no matching ν such that:

$$\mu(i) \succeq_i \nu(i) \quad \forall i \in I, \quad \text{and} \quad \mu(i) \succ_i \nu(i) \quad \text{for some } i \in I$$

2. **Efficiency Equilibrium:** A matching μ is efficient if there exists a price vectors π such that $\forall i, j \quad \mu(j) \prec_i \mu(i)$ implies $p(\mu(j)) > p(\mu(i))$ i.e. nobody can afford a better house.

A **competitive equilibrium** is just an efficient matching ν which respects the initial endowment; $p(\nu(i)) = p(\mu(i))$ for all i .

3. **Core of the Housing Market:** A matching η is in the core if there is no coalition $T \subseteq I$ and matching $\nu \neq \mu$ such that:

$$\nu(i) \in \{h_i\}_{i \in T}, \quad \nu(i) \succeq_i \eta(i) \quad \forall i \in T$$

i.e. a subset which can weakly improve through internal collusion. (If we only consider subsets of size 1, then this is called **individually rational**).

4. **Gale's Top Trading Cycles (TTC) Algorithm:** • Agents point to the owner of their favorite house. • A cycle is formed, and each agent in the cycle is assigned their most preferred house. • This process is repeated until all agents are assigned a house.

5. **Core Properties:** • The outcome of Gale's TTC algorithm is the unique matching in the core. • This matching is the unique competitive allocation; there exists a price vector π that supports this outcome and respects the initial endowment. • A direct mechanism is **strategy-proof** if truth-telling is a dominant strategy. Core is SP.
6. **Ma 1994:** Core is the only mechanism which is PE, IR, and SP.
7. **Serial Dictatorship:** A serial dictatorship mechanism ϕ_f assigns houses based on a priority ordering f . The agent with highest priority gets their top choice, followed by others in sequence:

$$\phi_f[\succ] = \mu.$$

SD is obviously PE and SP.

Housing Lotteries

Deep Result (Abdulkadiroğlu & Sönmez, 1998)

Lotteries

- **Definition: Lottery**

– A *lottery* is a probability distribution over feasible assignments, formally denoted by:

$$L = \sum \alpha_\mu \cdot \mu$$

where $\alpha_\mu \in [0, 1]$ is the probability weight of assignment $\mu: I \rightarrow H$.

– Lotteries can be represented as matrices $P = [p_{ih}]$, where p_{ih} is the probability of agent i receiving house h .
– The random assignment induced by lottery $\sum \alpha_\mu \cdot \mu$ is:

$$P = \sum \alpha_\mu \cdot \pi(\mu)$$

where $\pi(\mu)$ is the corresponding permutation matrix.

- There are various ways to define efficiency over random allocations:
 - A lottery is *ex-post efficient* if it assigns positive weight only to Pareto-efficient assignments.
 - Two lotteries are *equivalent* if they induce the same random assignments.
- **Theorem: Birkhoff-von Neumann Theorem**
 - Any bistochastic matrix can be expressed as a convex combination of permutation matrices.

Cardinal Treatment

- Assign Von Neumann-Morgenstern utility functions.
- *Ex-ante efficiency* is defined by the expectation of this utility function.
- In the sequel, we discard these cardinal utility functions.

Ordinal Treatment

- Probability distributions over allocations can be strictly preferred by agents.
- A distribution p is preferable to distribution q over H if *every* compatible utility function would prefer p to q .
- We use the poset induced by lexicographic ordering via the ordinal preference function.
- **Definition: Stochastic Dominance and Ordinal Efficiency**
 - Given two random consumptions P_i and Q_i with strict preference \succ_i :

$$\forall h \in H, \quad \sum_{h' \succeq_i h} p_{ih'} \geq \sum_{h' \succeq_i h} q_{ih'}$$
 - A random assignment P is *ordinally efficient* if it is not stochastically dominated by any other random assignment under preference ordering \succ .

Probabilistic Serial (PS) Mechanism

- To realize an OE mechanism:
 - Treat each house as a real-valued reservoir.
 - Agents "eat" from their top choices at the same rate as houses are consumed, moving to secondary choices as needed.
- This mechanism is greedily OE.
- The proof is tedious but conceptually straightforward.
- **Note on SP**
 - SP is not well-defined here since ordinal preferences do not induce a total ordering over probability distributions.
 - Assigning a VNM utility function, PS may not be SP.
 - If a single agent i 's utility function u_i is fixed, there exists some N such that if $n \geq N$, it is dominant for i to truthfully reveal preferences under SP.

Marriage Problem

A *marriage problem* is a triple (M, W, R) where:

- M is a set of men,
- W is a set of women,

- $R_m \in \Pi^{W \cup \{m\}}$ and $R_w \in \Pi^{M \cup \{w\}}$ are (non-strict) preference functions. Here, mR_w implies that man m finds woman w unacceptable. A matching is a permutation $\mu: M \cup W \rightarrow M \cup W$ that satisfies the obvious conditions.

Types of Matchings:

- A matching μ is *Pareto efficient* in the usual sense.
- Additional terms that generalize IR:
 - A matching μ is *blocked* by $i \in M \cup W$ if $iP_i\mu(i)$ (i.e., i does not find the matching acceptable). If μ is not blocked by anyone, it is *individually rational* (IR).
 - A matching μ is *blocked by a pair* $(m, w) \in M \times W$ if $wP_m\mu(m)$ and $mP_w\mu(w)$.
 - A matching is *stable* if it is not blocked by any individual or pair (essentially, IR against individuals and male-female pairs).

Man-Proposing Deferred Acceptance Algorithm (MPDAA)

- Each man records a set of women who have rejected him.
- Each woman records a single man who is her current most preferred proposal.
- The process:
 - While true:
 - Each man m who was rejected in the previous step proposes to his top acceptable choice who has not rejected him yet.
 - Each woman holds her most preferred acceptable proposal to date and rejects the rest. In particular, a woman may reject a man she previously accepted.
 - Terminate if no man is rejected in this step.

Note: Women weakly improve throughout the process, while men weakly get worse off. Thus, the algorithm terminates due to monovariance. MPDAA produces a valid matching.

Proof: μ is stable because it is clearly IR, and a woman would not leave a better match for a worse one, meaning the matching is never blocked by a pair.

Stable Matchings

1. MPDAA produces a *man-optimal stable matching* μ^M , which is maximal in the poset induced by men.
2. It is also minimal in the poset induced by women.
3. The *rural hospitals theorem* states that the set of agents matched is the same across all stable matchings.

Maximality and minimality are straightforward arguments, and the rural hospitals theorem follows by sandwiching these two results.

Lattice Theorem

The *join* of two stable matchings μ and μ' is $\mu \vee_M \mu'$, which assigns each man his more preferred assignment under μ and μ' , and each woman her less preferred assignment under μ and μ' . The *meet* is defined analogously.

Result: The join and meet of two stable matchings are stable matchings.

You cannot improve strictly for all men

Even if blocking pairs are allowed, there is no IR matching ν such that $\nu(m)P_m\mu^M(m)$ for all $m \in M$. In other words, it is impossible to make all men strictly better off.

Mechanisms

A mechanism φ is *stable* if $\varphi(R)$ is stable for all R (where R is the set of responses). Similarly for Pareto efficiency (PE) and individual rationality (IR). Note that stable implies PE and IR. Define ϕ^M and ϕ^W as selecting the man-optimal and woman-optimal stable matchings for each problem.

There exists no stable and strategy-proof (SP) mechanism.

Proof (by counterexample): This can be shown with a simple example involving only two men and two women in opposing preferences; a man can defect by reporting that his less favored woman is unacceptable.

ϕ^M is strategy-proof for men.

College Admissions Problem

A *college admissions problem* is a four-tuple (C, S, q, R) where:

- $C = \{c_1, \dots, c_m\}$ are colleges
- $S = \{s_1, \dots, s_n\}$ are students
- $q = (q_{c_1}, \dots, q_{c_m})$ are college capacities
- $R = (R_{c_1}, \dots, R_{c_m}, R_{s_1}, \dots, R_{s_n})$ are preferences.
 - Each R_s is a preference relation over $C \cup \{\emptyset\}$
 - Each R_c is a preference relation over 2^S
 - P_c, P_s are strict preferences derived from R_c and R_s .

A college preference R_c is *responsive* if:

- For each $T \subset S$ with $|T| < q_c$ and every $s \in S \setminus T$, $(T \cup \{s\})P_c T \iff \{s\}P_c \emptyset$.
- Along the $(1, \dots, 1)$ slices of the cube, it is locally a prism: for every $T \subset S$ with $|T| < q_c$ and all $s, s' \in S \setminus T$, $(T \cup \{s\})P_c (T \cup \{s'\}) \iff \{s\}P_c \{s'\}$. We assume college preferences are always responsive.

Rigidity

- This does not simplify the description substantially. Knowing that $\emptyset \prec \{1\} \prec \{2\} \prec \{3\}$ still induces a non-trivial Hasse diagram where $\{1, 2\}$ and $\{3\}$ are incomparable. Along $(1, \dots, 1)$ slices, it is fixed by the level-1 permutation.

- However, for simplicity, the theory proceeds as if the "important parts" of the preference function are determined by preferences on $\emptyset, \{s_1\}, \dots, \{s_n\}$ only.

A *matching* is a correspondence $\mu: C \cup S \rightarrow C \cup S$. This can be viewed as the union μ of two functions $\mu_C: C \rightarrow 2^S$ and $\mu_S: S \rightarrow 2^C$ such that $|\mu(c)| \leq q_c$, $|\mu(s)| \leq 1$, and $s \in \mu(c) \iff \mu(s) \in c$ (forming a bipartite graph with each c having degree $\leq q_c$ and each s having degree ≤ 1).

One defines notions of *blocked by a (college, student, pair)*, individually rational, and stable. These are based on small finite-case coalitions:

- μ is blocked by c if $\exists s \in \mu(c)$ such that $\emptyset P_c s$. Colleges can reject students individually.
- μ is blocked by s if $\emptyset P_s \mu(s)$.
- μ is blocked by (c, s) if:
 - $cP_s \mu(s)$ and $s' \in \mu(c)$ such that $\{s\}P_c \{s'\}$, i.e., a college would replace a student with a preferred one.
 - or $|\mu(c)| < q_c$ and $\{s\}P_c \emptyset$.

CPDAA and SPDAA

- CPDAA: A college with r standing offers proposes to its top $q_c - r$ remaining acceptable students who have not rejected them yet. Students hold most preferred acceptable offers.
- SPDAA: Analogous.

College to Marriage

- Divide each college into q_{c_i} slots and match. Given a marriage, a college admission is obtained, and vice versa.

Stability Lemma

- A college admission is stable if and only if its corresponding marriage is stable.

The admissions problem as presented may lack intuitive appeal:

- There is a lattice of stable assignments.
- SPDAA produces μ^S and CPDAA produces μ^C , which are student-optimal and college-optimal.
- These are maxima and minima on the stable lattice.
- No IR matching makes each student strictly better off than μ^S .
- No stable and SP mechanism exists.
- Truth-telling is dominant for students under SPDAA.

* New Results

- A college not filling all positions at some stable matching receives the same set of students in every matching.
- If μ and ν are two stable matchings, then either:
 - For all $s \in \mu(c) \setminus \nu(c)$ and $s' \in \nu(c) \setminus \mu(c)$, $\{s\}P_c \{s'\}$.
 - For all $s \in \mu(c) \setminus \nu(c)$ and $s' \in \nu(c) \setminus \mu(c)$, $\{s'\}P_c \{s\}$.
- No stable mechanism exists where truth-telling is dominant for all colleges.

School Choice Problem

The *school choice problem* differs from college admissions in that colleges (public schools) have publicly known preferences, often due to regulations. Consequently, the mechanism does not need to be strategy-proof against the schools.

- **SPDAA (Gale-Shapley Student Optimal Stable Mechanism)**
 - *Stable:* No college would reject a student in favor of another who also wants to attend.
 - *Pareto dominates* any other stable mechanism.
 - *Strategy-proof* for students.
- However, it is not Pareto efficient for students in general.

Example (Hastily Written)

- Schools 1, 2, 3 and students a, b, c with preferences: $1acb, 2bac, 3bac, a213, b123, c123$.

Alternatively, one can focus on student preferences, using school preferences as priorities for the TTC algorithm:

- The TTC algorithm (student-preferring version) allows each student to point to a preferred school in cycles, achieving IR, PE, and SP outcomes.

Kelso-Crawford Auction

This is similar to the Firm-Proposing Deferred Acceptance Algorithm. The process is as follows:

1. Initialize all salaries to 0.
2. Repeat the following steps:
 - a) Each firm makes an offer to its most preferred set of workers under $D^f(s)$. *Offers cannot be withdrawn; the argmax is solely over the remaining workers.*
 - b) Each worker evaluates all offers received and holds the best acceptable offer.
 - c) Terminate if no rejections are made, then increase the salaries by an infinitesimal amount.

Notes:

- In step 1, even if firms were allowed to withdraw offers, they would not be incentivized to do so (locally), due to the property of gross substitution.
- The worker's best offer (the best salary-firm utility) is non-decreasing throughout the process.

The Kelso-Crawford auction terminates, producing an allocation in the core that is weakly preferred by every firm to every other core allocation.