Notes on Market Design

Justin Hadad

University of Oxford, Department of Economics. These notes were prepared for *MATH 272: Market Design*, a course offered by Alex Teytelboym at the University of California, Berkeley.
All errors are my own. This document is incomplete.

Contents

Part 1

Microeconomic foundations

CHAPTER 1

Individual and social choice

1. Individual preferences and choice

Markets involve agents (which can be individuals or entities) making choices among a set of alternatives. These agents have preferences over some small (finite) or large (sometimes infinite) set of alternatives. Call this set of alternatives X. To capture how agents compare different alternatives, we use a binary relation $\succsim$ which represents the agent's preferences. Specifically, for any two alternatives x and y in X, $x \succsim y$ means that the agent either prefers x over y or is indifferent between them. We can further define two additional relations based on $\succsim$: $x \succ y$ means the agent strictly prefers alternative x over y (i.e., $x \succsim y$ and $y \not\succsim x$), and $x \sim y$ means the agent is indifferent between alternatives x and y (i.e., $x \succsim y$ and $y \succsim x$).

Throughout essentially all of market design, we impose a notion of *rationality*. That is, we assume agents can compare all alternatives, and that their preferences are consistent, in a no-cycles sense (transitivity).

Definition 1.1. *The preference relation $\succsim$ is rational if it possesses the following two properties: (i) completeness: for all $x, y \in X$, we have that $x \succsim y$ or $y \succsim x$ (or both); (ii) transitivity: for all $x, y, z \in X$, if $x \succsim y$ and $y \succsim z$, then $x \succsim z$.*

The preference relation is an example of a weak order (or total preorder) that describes preferences for certain choices over others. The preference relation is ordinal, but sometimes cardinal utility functions are useful.

Definition 1.2. *A function $u : X \to \mathbb{R}$ is a utility function representing preference relation $\succsim$ if, for all $x, y \in X$,*

$$x \succsim y \Leftrightarrow u(x) \geq u(y).$$

Utility functions give an understanding of how much agents prefer some alternatives over others. To connect the previous two definitions: a preference relation can only be represented by a utility function if it is rational.

Proposition 1.1. *A preference relation $\succsim$ can be represented by a utility function only if it is rational.*

A sketch of the proof follows. We must show that if a utility function represents $\succsim$ then $\succsim$ must be complete and transitive. To show completeness: because $u(\cdot)$ is a real-valued function defined on X, it must be that for any $x, y \in X$, either $u(x) \geq u(y)$ or $u(y) \geq u(x)$. But because $u(\cdot)$ is a utility function representing $\succsim$, this implies either that $x \succsim y$ or that $y \succsim x$. Then to show transitivity: suppose $x \succsim y$ and $y \succsim z$. Because $u(\cdot)$ represents $\succsim$, we must have $u(x) \geq u(y)$ and $u(y) \geq u(z)$. Therefore, $u(x) \geq u(z)$. Because $u(\cdot)$ represents $\succsim$, this implies $x \succsim z$.

If the set of alternatives is finite, any preference relation can be represented by a utility function. This is a consequence of the fact that we can always assign numerical values to each alternative in a finite set to reflect their ordering. For infinite sets of alternatives, if the preference relation is continuous (i.e., preserved under limits), then it can be represented by a continuous utility function. Specifically, the preference relation needs to ensure that the upper and lower contour sets are closed. This means that if a sequence of alternatives converges to a limit, and if all the alternatives in the sequence are in a contour set, then the limit must also be in the same contour set.

Definition 1.3. *The preference relation $\succsim$ on X is continuous if it is preserved under limits. That is, for any sequence of pairs $\{(x^n, y^n)\}_{n=1}^{\infty}$ with $x^n \succsim y^n$ for all n, $x = \lim_{n\to\infty} x^n$, and $y = \lim_{n\to\infty} y^n$, we have $x \succsim y$.*

Preservation under limits implies that if a sequence of alternatives gets closer and closer to being equally preferable to some alternative x, then the limit-alternative should be equally preferable to x.

1.1. Choice. In another approach to the theory of decision making, choice behavior is the primitive object. Let $\mathscr{B}$ be the set of nonempty subsets of X; that is, every element of $\mathscr{B}$ is a set $B \subset X$. We call the elements $B \in \mathscr{B}$ budget sets. The budget sets in $\mathscr{B}$ should be thought of as an exhaustive listing of all the choice experiments. Choice rules assign a choice from a budget set.

Definition 1.4. *$C(\cdot)$ is a choice rule (a correspondence) that assigns a nonempty set of chosen elements $C(B) \subset B$ for every budget set $B \in \mathscr{B}$.*

The elements of $C(B)$ are the acceptable alternatives in B. From choice rules we say that agents *reveal their preferences*, so instead of starting from a theoretical weak ordering, we can start from a series of choices that agents make over real budget sets, and then generate preferences.

Definition 1.5. *Given a choice structure $(\mathscr{B}, C(\cdot))$ the revealed preference relation $\succsim^*$ is defined by $x \succsim^* y \Leftrightarrow$ there is some $B \in \mathscr{B}$ such that $x, y \in B$ and $x \in C(B)$.*

The *weak axiom of revealed preference* reflects the expectation for consistency in consumer choice.

Definition 1.6. *The choice structure $(\mathscr{B}, C(\cdot))$ satisfies the weak axiom of revealed preference if the following property holds: If for some $B \in \mathscr{B}$ with $x, y \in B$ we have $x \in C(B)$, then for any $B' \in \mathscr{B}$ with $x, y \in B'$ and $y \in C(B')$, we must also have $x \in C(B')$.*

We obtain a simpler statement of the weak axiom by defining a revealed preference relation $\succsim^*$ from the observed choice behavior in $C(\cdot)$.

Proposition 1.2. *Suppose that $\succsim$ is a rational preference relation. Then the choice structure generated by $\succsim$, $(\mathscr{B}, C^*(\cdot, \succsim))$, satisfies the weak axiom.*

Proof. Suppose that for some $B \in \mathscr{B}$, we have $x, y \in B$ and $x \in C^*(B, \succsim)$. By the definition of $C^*(B, \succsim)$, this implies $x \succsim y$. To check whether the weak axiom holds, suppose that for some $B' \in \mathscr{B}$ with $x, y \in B'$, we have $y \in C^*(B', \succsim)$. This implies that $y \succsim z$ for all $z \in B'$. But we already know that $x \succsim y$. Hence, by transitivity, $x \succsim z$ for all $z \in B'$, and so $x \in C^*(B', \succsim)$. This is the conclusion of the weak axiom. □

2. Social choice

2.1. Using a Social Choice Function. So far we discussed an agent with a preference relation $\succsim$. Now, assume there are a set of alternatives X and a finite set of agents I. The set of preference relations $\succsim$ on X is denoted $\mathscr{R}$. We also designate by $\mathscr{P}$ the subset of $\mathscr{R}$ consisting of the preference relations $\succsim \in \mathscr{R}$ with the property that no two distinct alternatives are indifferent for $\succsim$.

Definition 1.7. *Given any subset $\mathscr{A} \subset \mathscr{R}^I$, a social choice function $f : \mathscr{A} \to X$ defined on $\mathscr{A}$ assigns a chosen element $f(\succsim_1, \ldots, \succsim_I) \in X$ to every profile of individual preferences in $\mathscr{A}$.*

We intend to show an impossibility of satisfying a set of normatively-appealing properties. First restrict that the social choice function must be *(weakly) Paretian* and *monotonic*. Weak Pareto implies that if everyone agrees that one option is better, the system should respect that collective preference; monotonicity implies that if more people start preferring one option over another, it doesn't make sense for that option to be ranked lower than before.

Definition 1.8. *The social choice function $f : \mathscr{A} \to X$ defined on $\mathscr{A} \subset \mathscr{R}^I$ is weakly Paretian if for any profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$ the choice $f(\succsim_1 \ldots, \succsim_I) \in X$ is a weak Pareto optimum. That is, if for some pair $\{x, y\} \subset X$ we have that $x \succ_i y$ for every i, then $y \neq f(\succsim_1, \ldots, \succsim_I)$.*

To define monotonicity we need a preliminary concept.

DEFINITION 1.9. *The alternative $x \in X$ maintains its position from the profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{R}^I$ to the profile $(\succsim_1', \ldots, \succsim_I') \in \mathscr{R}^I$ if*

$$x \succsim_i y \text{ implies } x \succsim_i' y$$

for every i and every $y \in X$.

DEFINITION 1.10. *The social choice function $f : \mathscr{A} \to X$ defined on $\mathscr{A} \subset \mathscr{R}^I$ is monotonic if for any two profiles $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}, (\succsim_1', \ldots, \succsim_I') \in \mathscr{A}$ with the property that the chosen alternative $x = f(\succsim_1 \ldots, succsim_I)$ maintains its position from $(\succsim_1, \ldots, \succsim_I)$ to $(\succsim_1', \ldots, \succsim_1')$, we have that $f(\succsim_1', \ldots, \succsim_I') = x$.*

There are obviously social choice functions that are weakly Paretian and monotonic.[1] We now define *dictatorship*:

DEFINITION 1.11. *An agent $h \in I$ is a dictator for the social choice function $f : \mathscr{A} \to X$ if, for every profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}, f(\succsim_1, \ldots, \succsim_I)$ is a most preferred alternative for $\succsim_h$ in X; that is, $f(\succsim_1, \ldots, \succsim_I) \in \{x \in X : x \succsim_h y \text{ for every } y \in X\}$.*

But there is an impossibility:

THEOREM 1.1 (Arrow's Impossibility Theorem for Social Choice Functions). *If there are at least three alternatives in X, then any social welfare function which takes an unrestricted domain and which is weak Pareto and monotonic must have a dictator.*

The "dictator" part might sound harsh, but what it's essentially saying is that the only way to satisfy all these conditions is to have a system where one person's preference rules. This doesn't mean that in real-world voting systems there is always a dictator; rather, it shows the limitations and trade-offs inherent in trying to create a "fair" voting system.

2.2. Using a Social Welfare Functional. Rather than generating a particular element of a preference relation, we can generate a social ordering: a preference ordering for society that is a complete and transitive ordering over X. We do this via a social welfare functional.

DEFINITION 1.12. *A social welfare functional defined on a given subset $\mathscr{A} \subset \mathscr{R}^I$ is a rule $F : \mathscr{A} \to \mathscr{R}$ that assigns a rational preference relation $F(\succsim_1, \ldots, \succsim_I) \in \mathscr{R}$, interpreted as the social preference relation, to any profile of individual rational preference relations $(\succsim_1 \ldots, \succsim_I)$ in the admissible domain $\mathscr{A} \subset \mathscr{R}^I$.*

As before, there are normatively-appealing properties that we hope to satisfy.

DEFINITION 1.13 (Pareto "WP"). *The social welfare functional $F : \mathscr{A} \to \mathscr{R}$ is Paretian if, for any pair of alternatives $\{x, y\} \subset X$ and any preference profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$, we have that x is socially preferred to y, that is, $xF_p(\succsim_1, \ldots, \succsim_I)\, y$, whenever $x \succ_i y$ for every i.*

DEFINITION 1.14 (Independence of Irrelevant Alternatives "IIA"). *The social welfare functional $F : \mathscr{A} \to \mathscr{R}$ defined on the domain $\mathscr{A}$ satisfies the pairwise independence condition (or the independence of irrelevant alternatives condition) if the social preference between any two alternatives $\{x, y\} \subset X$ depends only on the profile of individual preferences over the same alternatives. Formally, for any pair of alternatives $\{x, y\} \subset X$, and for any pair of preference profiles $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$ and $(\succsim_1', \ldots, \succsim_I') \in \mathscr{A}$ with the property that, for every i,*

$$x \succsim_i y \Leftrightarrow x \succsim_i' y \text{ and } y \succsim_i x \Leftrightarrow y \succsim_i' x,$$

And now we have another notion of impossibility:

[1]Consider, e.g., a pairwise majority voting social decision function defined on a domain of "single-peaked" preferences, where each agent has a most preferred outcome and any outcome closer to this outcome is preferred to one that is farther away.

Theorem 1.2 (Arrow's Impossibility Theorem). *Suppose that the number of alternatives is at least three and that the domain of admissible individual profiles, denoted $\mathscr{A}$, is either $\mathscr{A} = \mathscr{R}^I$ or $\mathscr{A} = \mathscr{P}^I$. Then every social welfare functional $F : \mathscr{A} \to \mathscr{R}$ that is Paretian and satisfies the pairwise independence condition is dictatorial in the following sense: There is an agent h such that, for any $\{x, y\} \subset X$ and any profile $(\succsim_1, \ldots, \succsim_I) \in \mathscr{A}$, we have that x is socially preferred to y, that is, $xF_p\,(\succsim_1, \ldots, \succsim_I)\,y$, whenever $x \succ_h y$.*

Proof. We present a proof from Yu (2012). The set of preference profiles $\mathscr{A}$ has a typical element denoted as an ordered list $\succsim = (\succsim_1, \ldots, \succsim_I)$. Let individuals numbered $1, 2, \ldots, I$ each have rational preferences over alternatives $A = a_1, \ldots, a_M$ where $M \geq 3$. For ease (it will become clear why we do this below), say $F\,(\succsim_1, \ldots, \succsim_I) \in \mathscr{R}$ assigns rational social preferences $\trianglerighteq$ over A.

Let the set of preference profiles be unrestricted (either $\mathscr{A} = \mathscr{R}^I$ or $\mathscr{A} = \mathscr{P}^I$) and suppose $F : \mathscr{A} \to \mathscr{R}$ satisfies WP and IIA. Now, consider an arbitrary $\succsim$ in which $a_i \succ_h a_j$ for all h, and then swap the position of the two alternatives sequentially from 1 to I. By WP we start with $a_i \triangleright a_j$ and end with $a_j \triangleright a_i$. We call the first voter whose swap invalidates $a_i \triangleright a_j$ the (i, j)-pivotal voter and denote her h_{ij}, which by IIA is independent of $\succsim$.

Now consider $\succsim'$ and $\succsim''$ with the depicted rankings:

	1	$\cdots$	$h_{ij}-1$	h_{ij}	$h_{ij}+1$	$\cdots$	I
$\vec{\succsim}'$	a_j	$\cdots$	a_j	a_i	a_i	$\cdots$	a_i
	a_k	$\cdots$	a_k				
	a_i	$\cdots$	a_i	a_j	a_j	$\cdots$	a_j
				a_k	a_k	$\cdots$	a_k
$\vec{\succsim}''$	$\square$	$\cdots$	$\square$				
	$\square a_j$	$\ldots$	$\square a_j$	a_j	a_i	$\cdots$	a_i
	$\square$	$\ldots$	$\square$		$\square$	$\ldots$	$\square$
	a_i	$\ldots$	a_i	a_i	$\square a_j$	$\ldots$	$\square a$
				a_k	$\square$	$\ldots$	$\square$

We must have $a_i \triangleright a_j \triangleright a_k$, where the first relation is by the definition of h_{ij} and the second by WP. For $\succsim''$, squares denote possible positions of a_k, with indifference drawn by putting alternatives at the same level. We have $a_j \trianglerighteq a_i \triangleright a_k$, where the first is by the definition of h_{ij} and the second by IIA, as individual preferences over a_i and a_k are the same in $\succsim'$ and $\succsim''$. Focusing on a_j and a_k, we conclude by IIA that h_{ij} "dictates" $a_j \triangleright a_k$; that is, $a_j \succ_{h_{ij}} a_k$ implies $a_j \triangleright a_k \quad \forall i \neq j \neq k$.

The above says that $a_j \triangleright a_k$ should not change while swapping as long as h_{ij} ranks j above k, so $h_{jk} \geq h_{ij}$. For h_{kj}, the ranking of the two alternatives should become $a_j \triangleright a_k$ no later than h_{ij} makes the change, so $h_{kj} \leq h_{ij}$. We have $h_{jk} \geq h_{ij} \geq h_{kj}$. As j and k are distinct and arbitrary, $h_{kj} \geq h_{jk}$ also holds, implying $h_{jk} = h_{kj} = h_{ij}$, which can be easily extended to any such h_{ts} where $t \neq s$. But by the fact that h_{ij} "dictates" $a_j \triangleright a_k$, we have that this unique pivotal voter is a dictator. $\square$

Some examples of violations follow:

Proposition 1.3. *The Pareto principle conflicts with some notion of rights.*

Proof. Two agents, prude (P) and lewd (L). Three alternatives: p (P reads), ℓ (L reads), n (neither reads). Preferences are as follows:

P	L
n	p
p	ℓ
ℓ	n

Under minimal rights, each can reject one of 'no one reading the book', 'themselves reading the book'. L rejects n and P rejects p. So ℓ selected but p dominates ℓ. $\square$

Before the example to follow, we must define the *Borda Count*. Each agent ranks $1, ..., N$ alternatives, and her nth favorite ranking gets a score of n (for $1 \leq n \leq N$). Lowest scores are preferred.

PROPOSITION 1.4. *The Borda Count mechanism violates IIA.*

PROOF. Consider preferences:

State	Agent 1	Agent 2	**Borda Score**
x	1	2	3
y	3	1	4
z	2	3	5

So xPy and yPz. Suppose instead that preferences are:

State	Agent 1	Agent 2	**Borda Score**
x	1	3	4
y	2	1	3
z	3	2	5

So yPx and xPz. This violates IIA, as agents' preferences over $\{x, y\}$ have not changed. □

2.3. Using a Social Welfare Function. We also, at times, will want to map utilities into the real numbers. The type of function that maps utilities into the utility possibility set is called a *(Bergon-Samuelson) social welfare function.*

DEFINITION 1.15. *A (Bergson-Samuelson) social welfare function is a mapping $W : U \to \mathbb{R}$, i.e.,*

$$W(u_1, \ldots, u_H).$$

Denote $u^* = (u_1^*, \ldots, u_H^*)$ as a social optimum if it maximizes W over $u \in U$.

Below are some examples of social welfare functions.

Social Welfare Function	**Equation**
Utilitarian	$W = u_1 + \ldots + u_H$
Generalized Utilitarian	$\sum_h g(u_h)$, g increasing, concave
Maximin ("Rawlsian")	$W = \min\{u_1, \ldots, u_H\}$
Constant Elasticity of Substitution (CES)	$W_\rho = \begin{cases} (\sum_h u_h^{1-\rho})^{1/(1-\rho)} & \rho \neq 1 \\ \sum_h \ln u_i & \rho = 1 \end{cases}$

Utilitarian ($\rho = 0$) and Rawlsian ($\rho \to \infty$) are special cases of the CES social welfare function.

CHAPTER 2

General equilibrium

1. Consumer theory

1.1. Establishing the Walrasian budget set. Agents choose consumption levels of L commodities indexed by $\ell = 1, \ldots, L$, where L is finite. Generally we use a commodity vector or bundle to list the amounts consumed of each commodity $x = [x_1, ..., x_L] \in R^L$ where R^L is the commodity space. The simplest consumption set is $X = \mathbb{R}^L_+ = \{x \in \mathbb{R}^L : x_\ell \geq 0 \text{ for } \ell = 1, \ldots, L\}$.

Note $\mathbb{R}^L_+$ is convex: if $x, x' \in \mathbb{R}^L_+$, then the bundle $x'' = \alpha x + (1-\alpha)x' \in \mathbb{R}^L_+$ for any $x \in [0, 1]$. We then make two key assumptions:

- The L commodities are all traded in the market at dollar prices that are publicly quoted; this is the principle of completeness.
- Consumers are price takers.

Consider a price vector exists $p = [p_1, ...,_L] \in R^L$ which gives the cost of each commodity. Prices can be negative but we usually restrict to $p \gg 0$.

Definition 2.1. *The Walrasian (or competitive) budget set $B_{p,w} = \{x \in \mathbb{R}^L_+ : p \cdot x \leq W\}$ is the (convex) set of all feasible consumption bundles for the consumer who faces market prices p and has wealth w.*

The Walrasian budget set $B_{p,w}$ is a convex set: That is, if bundles x and x' are both elements of $B_{p,w}$, then the bundle $x'' = \alpha x + (1-\alpha)x'$ is also.

The consumer's Walrasian demand correspondence $x(p, w)$ assigns a set of chosen consumption bundles for each price-wealth pair (p, w). We make two assumptions:

- Walrasian demand correspondence is homogeneous of degree zero; that is, $x(\alpha p, \alpha w) = x(p, w)$ for any p, w and $\alpha > 0$.
- Walrasian demand satisfies Walras' Law; that is, for every $p \gg 0$ and $w > 0$, we have $p \cdot x = w$ for all $x \in x(p, w)$.

1.2. Making assumptions on preferences. We seek results about the bundles consumers purchase. We need to establish assumptions on consumer behavior. Recall the preference relation is rational if it is complete and transitive.

Definition 2.2. *The preference relation $\succsim$ on X is rational if it possesses the following two properties:*

- *Completeness. For all $x, y \in X$, we have $x \succsim y$ or $y \succsim x$ (or both).*
- *Transitivity. For all $x, y, z \in X$, if $x \succsim y$ and $y \succsim z$, then $x \succsim z$.*

The first class of assumptions we make are *desirability* assumptions.

Definition 2.3 (Monotonicity). *The preference relation $\succsim$ on X is monotone if $x \in X$ and $y \gg x$ implies $y > x$. It is strongly monotone if $y \geq x$ and $y \neq x$ imply that $y > x$.*

Strong monotonicity says that if y is larger than x for some commodity and is no less for any other, then y is strictly preferred to x. For much of the theory, however, a weaker desirability assumption of local nonsatiation suffices:

Definition 2.4 (Local Nonsatiation). *The preference relation $\succsim$ on X is locally nonsatiated if for every $x \in X$ and every $\varepsilon > 0$, there is $y \in X$ such that $\|y - x\| \leq \varepsilon$ and $y > x$.*

It is easy to show that if $\succsim$ is strongly monotone, then it is monotone, and if $\succsim$ is monotone, then it is locally nonsatiated.

The second class of assumptions pertain to *convexity*.

Definition 2.5 (Convexity). *The preference relation $\succsim$ on X is convex if for every $x \in X$, the upper contour set $\{y \in X : y \succsim x\}$ is convex; that is, if $y \succsim x$ and $z \succsim x$, then $\alpha y + (1-\alpha)z \succsim x$ for any $\alpha \in [0,1]$.*

With convex preferences, from any initial consumption situation x, and for any two commodities, it takes increasingly larger amounts of one commodity to compensate for successive unit losses of the other. We call this the "diminishing marginal rates of substitution." We make the inequality strict for strict convexity.

Definition 2.6 (Quasilinear Preferences). *The preference relation $\succsim$ on $X = (-\infty, \infty) \times \mathbb{R}_+^{L-1}$ is quasilinear with respect to commodity 1 (called, in this case, the numeraire commodity) if*

- *All the indifference sets are parallel displacements of each other along the axis of commodity 1. That is, if $x \sim y$, then $(x + \alpha e_1) \sim (y + \alpha e_1)$ for $e_1 = (1, 0, \dots, 0)$ and any $\alpha \in \mathbb{R}$.*
- *Good 1 is desirable; that is, $x + \alpha e_1 > x$ for all x and $\alpha > 0$.*

We look to see if we can ensure the existence of a utility function from a preference relation. The assumption we need is continuity.

Definition 2.7. *The preference relation $\succsim$ on X is continuous if it is preserved under limits. That is, for any sequence of pairs $\{(x^n, y^n)\}_{n=1}^{\infty}$ with $x^n \succsim y^n$ for all n. $x = \lim_{n\to\infty} x^n$, and $y = \lim_{n\to\infty} y^n$, we have $x \succsim y$.*

We use the upper contour set of bundle x to refer to the set of all bundles that are at least as good as $x : \{y \in X : y \succsim x\}$. The lower contour set of x is the set of all bundles that x is at least as good as: $\{y \in X : x \gtrsim y\}$. We also say a preference relation is continuous if the upper and lower countour sets are closed; that is, they include their boundaries.

Proposition 2.1. *Suppose that the rational preference relation $\succsim$ on X is continuous. Then there is a continuous utility function $u(x)$ that represents $\succsim$.*

From now on, we assume that the consumer's preference relation is continuous and hence can be represented by a continuous utility function.

Note that restrictions on preferences translate into restrictions on the form of utility functions. The property of monotonicity, for example, implies that the utility function is increasing: $u(x) > u(y)$ if $x \gg y$. The property of convexity of preferences, on the other hand, implies that $u(\cdot)$ is quasiconcave, and similarly, strict convexity of preferences implies strict quasiconcavity of $u(\cdot)$. The utility function $u(\cdot)$ is quasiconcave if the set $\{y \in \mathbb{R}_+^L : u(y) \geq u(x)\}$ is convex for all x or, equivalently, if $u(\alpha x + (1-\alpha)y) \geq \text{Min}\{u(x), u(y)\}$ for any x, y and all $\alpha \in [0,1]$.

Moving forward, we assume the consumer has a rational, continuous, and locally nonsatiated preference relation, and $u(x)$ is a continuous utility function representing these preferences. The consumption set is $X = \mathbb{R}_+^L$.

1.3. Utility maximization. The consumer's problem of choosing her most preferred consumption bundle given prices $p \gg 0$ and wealth level $w > 0$ can now be stated as the following utility maximization problem:

$$\max_x \{u(x) \mid p \cdot w = 1\}.$$

Proposition 2.2. *Suppose that $u(\cdot)$ is a continuous utility function representing a locally nonsatiated preference relation $\succsim$ defined on the consumption set $X = \mathbb{R}_+^L$. Then the Marshallian / Walrasian demand correspondence $x(p, w)$ possesses the following properties:*

1. *Homogeneity of degree zero in $(p, w) : x(\alpha p, \alpha w) = x(p, w)$ for any p, w and scalar $\alpha > 0$.*
2. *Walras law: $p \cdot x = w$ for all $x \in x(p, w)$.*

(3) *Convexity/uniqueness: If $\succsim$ is convex, so that $u(\cdot)$ is quasiconcave, then $x(p,w)$ is a convex set. Moreover, if $\succsim$ is strictly convex, so that $u(\cdot)$ is strictly quasiconcave, then $x(p,w)$ consists of a single element.*

PROOF. **Homogeneity**: For any scalar $\alpha > 0$,

$$\left\{x \in \mathbb{R}_+^L : \alpha p \cdot x \leq \alpha w\right\} = \left\{x \in \mathbb{R}_+^L : p \cdot x \leq w\right\};$$

that is, the set of feasible consumption bundles in the utility maximization problem does not change when all prices and wealth are multiplied by a constant $\alpha > 0$. The set of utility-maximizing consumption bundles must therefore be the same in these two circumstances, and so $x(p,w) = x(\alpha p, \alpha w)$.

Walras' law: follows from local nonsatiation.

Convexity/uniqueness: Suppose that $u(\cdot)$ is quasiconcave and that there are two bundles x and x', with $x \neq x'$, both of which are elements of $x(p,w)$. To establish the result, we show that $x'' = \alpha x + (1-\alpha)x'$ is an element of $x(p,w)$ for any $\alpha \in [0,1]$. To start, we know that $u(x) = u(x')$. Denote this utility level by u^*. By quasiconcavity, $u(x'') \geq u^*$. In addition, since $p \cdot x \leq w$ and $p \cdot x' \leq w$, we also have

$$p \cdot x'' = p \cdot [\alpha x + (1-\alpha)x'] \leq w.$$

Therefore, x'' is a feasible choice in the utility maximization problem because $B_{p,w}$ is a convex set. Thus, since $u(x'') \geq u^*$ and x'' is feasible, we have $x'' \in x(p,w)$. This establishes that $x(p,w)$ is a convex set if $u(\cdot)$ is quasiconcave. The strict case follows directly. □

1.4. Expenditure minimization. Still assume that $u(\cdot)$ is a continuous utility function representing a locally nonsatiated preference relation $\succsim$ defined on the consumption set $\mathbb{R}_+^L$. Now, instead of maximizing utility given prices and wealth, a consumer can minimize expenditure given prices and a desired utility level:

$$\min_x \{r \cdot x \mid u(x) = u\}.$$

We call the resulting correspondence Hicksian demand, $h(p,u)$.

PROPOSITION 2.3. *Suppose that $u(\cdot)$ is a continuous utility function representing a locally nonsatiated preference relation $\succsim$ defined on the consumption set $X = \mathbb{R}_+^L$. Then for any $p \gg 0$, the Hicksian demand correspondence $h(p,u)$ possesses the following properties:*

(1) *Homogeneity of degree zero in $p : h(\alpha p, u) = h(p,u)$ for any p, u and $\alpha > 0$.*
(2) *No excess utility: For any $x \in h(p,u), u(x) = u$.*
(3) *Convexity/uniqueness: If $\succsim$ is convex, then $h(p,u)$ is a convex set; and if $\succsim$ is strictly convex, so that $u(\cdot)$ is strictly quasiconcave, then there is a unique element in $h(p,u)$*

.

The proof for (1) and (3) follows that for the utility maximization problem; (2) follows from continuity of $u(\cdot)$. Suppose there exists an $x \in h(p,u)$ such that $u(x) > u$. Consider a bundle $x' = \alpha x$, where $\alpha \in (0,1)$. By continuity, for α close enough to $1, u(x') \geq u$ and $p \cdot x' < p \cdot x$, contradicting x being optimal in the EMP with required utility level u.

The value that expenditure minimization takes (that is, the money the consumer spends to reach a certain utility) defines the expenditure function, $e(p,u)$.

PROPOSITION 2.4. *Suppose that $u(\cdot)$ is a continuous utility function representing a locally nonsatiated preference relation $\succsim$ defined on the consumption set $X = \mathbb{R}_+^L$. The expenditure function $e(\rho, u)$ is*

(1) *Homogeneous of degree one in p.*
(2) *Strictly increasing in u and nondecreasing in p, for any ℓ.*
(3) *Concave in p.*
(4) *Continuous in p and u.*

THEOREM 2.1 (The Duality Theorem). *Let K be a nonempty closed set, and let $\mu_K(\cdot)$ be its support function. Then there is a unique $\bar{x} \in K$ such that $\bar{p} \cdot \bar{x} = \mu_K(\bar{p})$ if and only if $\mu_K(\cdot)$ is differentiable at $\bar{p}$. Moreover, in this case,*

$$\nabla \mu_K(\bar{p}) = \bar{x}.$$

The logic is straightforward. Suppose $\mu_K(\cdot)$ is differentiable at $\bar{p}$, and consider the function $\xi(p) = p \cdot \bar{x} - \mu_X(p)$, where $\bar{x} \in K$ and $\mu_K(\bar{p}) = \bar{p} \cdot \bar{x}$. By the definition of $\mu_K(\cdot), \xi(p) = p \cdot \bar{x} - \mu_K(p) \geq 0$ for all p. We also know that $\xi(\bar{p}) = \bar{p} \cdot \bar{x} - \mu_K(\bar{p}) = 0$. So the function $\xi(\cdot)$ reaches a minimum at $p = \bar{p}$, so its partial derivatives at $\bar{p}$ must all be zero. This implies the result: $\nabla \xi(\bar{p}) = \bar{x} - \nabla \mu_K(\bar{p}) = 0$.

1.5. Duality.

PROPOSITION 2.5. *Suppose that $u(\cdot)$ is an unbounded continuous utility function representing a locally nonsatiated preference relation $\succsim$ defined on the consumption set $X = \mathbb{R}_+^L$ and that the price vector is $p \gg 0$. We have*

1. *If x^* is optimal in the utility maximization problem when wealth is $w > 0$, then x^* is optimal in the expenditure minimization problem when the required utility level is $u(x^*)$. Moreover, the minimized expenditure level in the expenditure minimization problem is exactly w.*
2. *If x^* is optimal in the expenditure minimization problem when the required utility level is $u > u(0)$, then x^* is optimal in the utility maximization problem when wealth is $p \cdot x^*$. Moreover, the maximized utility level in the utility maximization problem is exactly u.*

PROOF. We prove both claims. For the first, suppose that x^* is not optimal in the EMP with required utility level $u(x^*)$. Then there exists an x' such that $u(x') \geq u(x^*)$ and $p \cdot x' < p \cdot x^* \leq w$. By local nonsatiation, we can find an x'' very close to x' such that $u(x'') > u(x')$ and $p \cdot x'' < w$. But this implies that $x'' \in B_{p,w}$ and $u(x'') > u(x^*)$, contradicting the optimality of x^* in the UMP. Thus, x^* must be optimal in the EMP when the required utility level is $u(x^*)$, and the minimized expenditure level is therefore $p \cdot x^*$. Finally, since x^* solves the UMP when wealth is w, by Walras' law we have $p \cdot x^* = w$.

For the second, since $u > u(0)$, we must have $x^* \neq 0$. Hence, $p \cdot x^* > 0$. Suppose that x^* is not optimal in the UMP when wealth is $p \cdot x^*$. Then there exists an x' such that $u(x') > u(x^*)$ and $p \cdot x' \leq p \cdot x^*$. Consider a bundle $x'' = \alpha x'$ where $\alpha \in (0,1)$ (x'' is a "scaled-down"■ version of x'). By continuity of $u(\cdot)$, if α is close enough to 1 , then we will have $u(x'') > u(x^*)$ and $p \cdot x'' < p \cdot x^*$. But this contradicts the optimality of x^* in the EMP. Thus, x^* must be optimal in the UMP when wealth is $p \cdot x^*$, and the maximized utility level is therefore $u(x^*)$. □

PROPOSITION 2.6 (Shephard's Lemma). *Suppose that $u(\cdot)$ is a continuous utility function representing a locally nonsatiated and strictly convex preference relation $\succsim$ defined on the consumption set $X = \mathbb{R}_+^L$. For all p and u, the Hicksian demand $h(p,u)$ is the derivative vector of the expenditure function with respect to prices:*

$$h(p,u) = \nabla_p e(p,u)$$

Could prove this using first order conditions or the envelope theorem. We can use Shephard's Lemma to derive an important economic equation that separates substitution effects and income effects.

THEOREM 2.2 (Slutksy). *Suppose that $u(\cdot)$ is a continuous utility function representing a locally nonsatiated and strictly convex preference relation $\succsim$ defined on the consumption set $X = \mathbb{R}_+^L$. Then for all (p,w), and $u = v(p,w)$. we have*

$$\frac{\partial x_\ell(p,w)}{\partial p_k} = \frac{\partial h_\ell(p,u)}{\partial p_k} - x_k \frac{\partial x_\ell(p,w)}{\partial w} \tag{1}$$

PROOF. At the optimum bundle

$$x_\ell(p,w) = h_\ell(p,u)$$

and, given optimizing behavior, we can substitute the expenditure function for $w = e(p,u)$:

$$x_\ell(p, e(p,u)) = h_\ell(p,u)$$

Differentiate and we have

$$\frac{\partial h_\ell(p,u)}{\partial p_k} = \frac{\partial x_\ell(p,w)}{\partial p_k} + \frac{\partial x_\ell(p,e(p,u))}{\partial w}\frac{\partial e(p,u)}{\partial p_k}$$

Then by Shephard's Lemma

$$\frac{\partial h_\ell(p,u)}{\partial p_k} = \frac{\partial x_\ell(p,w)}{\partial p_k} + \frac{\partial x_\ell(p,e(p,u))}{\partial w}x_k$$

And we get Slutsky's equation

$$\frac{\partial x_\ell(p,w)}{\partial p_k} = \frac{\partial h_\ell(p,u)}{\partial p_k} - \frac{\partial x_\ell(p,e(p,u))}{\partial w}x_k$$

□

2. General equilibrium

Buying and selling generally—and profoundly—balances. Quote from Kenneth Arrow: "This experience of balance is indeed so widespread that it raises no intellectual disquiet among laymen; they take it so much for granted that they are not disposed to understand the mechanism by which it occurs."

We often prove existence of equilibria using fixed-point theorems. First we need to define upper hemicontinuity.

Definition 2.8. *Given $A \subset \mathbb{R}^N$ and the closed set $Y \subset \mathbb{R}^K$, the correspondence $f : A \to Y$ is upper hemicontinuous if it has a closed graph and the images of compact sets are bounded, that is, for every compact set $B \subset A$ the set $f(B) = \{y \in Y : y \in f(x) \text{ for some } x \in B\}$ is bounded.*

Theorem 2.3 (Kakutani's Fixed Point Theorem). *Suppose X is a nonempty, compact, convex subset of R^n for some integer n. Suppose that $f : X \rightrightarrows X$ is a correspondence from X to (subsets of) X that is upper hemi-continuous, and convex and nonempty valued. Then $\exists x \in X$ such that $x \in f(x)$.*

Theorem 2.4. *Let utility functions satisfy continuity, monotonicity and concavity; agents have strictly positive endowments; and Marshallian demands be upper semi-continuous in prices. Then Kakutani's fixed point corresponds exactly to a Walrasian equilibrium.*

To prove existence using Kakutani's, we need another result:

Proposition 2.7 (Berge's Theorem). *Suppose we have a continuous correspondence $C : Q \rightrightarrows R^N$, with $c(q)$ compact and non-empty for all $q \in Q$ and a continuous function $f : R^N \to R$. Consider a maximization problem* $\max f(x)$ *s.t. $x \in c(q)$. The maximizer correspondence will be upper-hemi-continuous. The value function will be continuous.*

Existence using Kakutani's fixed point. (***Note:*** Monotonicity could be weakened to local nonsatiation, and concavity to quasi-concavity.) We reproduce a proof of existence from Levin's notes on general equilibrium. x^i is defined to be agent i's Marshallian demand given the price vector p, and e^i her endowment.

Normalize prices wlog; define

$$\Delta = \left\{p \in \mathbb{R}^L_+ : p_1 + \ldots + p_L = 1\right\}$$

to be the set of price vectors that sum to one (the price simplex).

Then define Marshallian demands in such a way that they are upper semi-continuous in prices. In order to say agents' demand correspondences are upper semi-continuous using Berge's Theorem, we need $\mathcal{B}^i(p)$ to be compact-valued at prices on the boundary of Δ. Therefore we define the compact set $T = \left\{x \in \mathbb{R}^L_+ : x \le 2\sum_{i\in\mathcal{I}} e^i\right\}$ and consider for each agent i the correspondence

$$\psi^i(p) = \arg\max_{c \in \mathcal{B}^i(p)\cap T} u^i(c)$$

which is non-empty and upper semi-continuous for all i, and because $u(\cdot)$ is concave, we have that ψ^i is convex-valued. Note that we assumed $e^i \gg 0$ so that the budget correspondence $\mathcal{B}^i(p)$ is continuous everywhere.

Then define the aggregate demand correspondence

$$\Psi^D(p) = \sum_{i\in\mathcal{I}} \psi^i(p) = \left\{ x : \exists x^1 \in \psi^1(p), \ldots, x^I \in \psi^I(p) \text{ s.t. } x = \sum_{i\in\mathcal{I}} x^i \right\}$$

which is nonempty, convex-valued, and upper semi-continuous in prices, given what we know about $\psi^1, \ldots, \psi^I$.

Now introduce an agent who chooses new prices to maximize the value of excess aggregate demand at the *old* prices. Define the correspondence of this agent (sometimes called the "price player") as $\Psi^P : T \rightrightarrows \Delta$, where

$$\Psi^P = \arg\max_{p\in\Delta^{L-1}} p \cdot (x - e).$$

where $e = \sum_{i\in\mathcal{I}} e^i$ is the aggregate endowment. This means the price player chooses new prices to maximize the value of the aggregate excess demand, at the old prices. Note Ψ^P is nonempty, convex-valued, and upper semi-continuous.

Now define $\Psi : \Delta \times T \rightrightarrows \Delta \times T$ by

$$\Psi(p, x) = \left(\Psi^P(x), \Psi^D(p)\right).$$

Because the product of non-empty and convex-valued upper semi-continuous correspondences is itself non-empty, convex-valued and upper semi-continuous, we can apply Kakutani's theorem. Thus we have a fixed point $(p^*, x^*) \in \Psi(p^*, x^*)$. We now show this is a Walrasian equilibrium.

Because $x^* \in \Psi^D(p)$, there exist $x^{1*}, \ldots, x^{I*}$ summing to x^* with the property that $x^{i*} \in \arg\max_{c\in\mathcal{B}^i(p^*)\cap T} u^i(c)$. For the individual optimization part of equilibrium, we need to verify that $x^{i*} \in \arg\max_{c\in\mathcal{B}^i(p^*)} u^i(c)$. For this, note that $p^* \in \Psi^P(x^*)$:

$$0 \geq p^* \cdot (x^* - e) \geq p \cdot (x^* - e) \qquad \text{for all } p \in \Delta.$$

The latter inequality implies that $x^* - e \leq 0$ and in particular $x^{i*} < 2e$. Therefore $x^{i*} \in \arg\max_{c\in\mathcal{B}^i(\bar{p})} u^i(c)$ because if there existed $c \in \mathcal{B}^i(p)$ with $u^i(c) > u^i\left(x^{i*}\right)$ then for small $\lambda > 0, \lambda c + (1-\lambda)x^{i*} \in \mathcal{B}^i(p) \cap T$ and by concavity of $u^i, u^i(\lambda c + (1-\lambda)x^{i*}) > u^i\left(x^{i*}\right)$, a contradiction.

Now we show the market clears; that is, that $x^* = e$. We use Walras's Law—that budget constraints imply excess demand sums to zero—to claim that $p^* \cdot x^* = p^* \cdot e$. Therefore if $x^*_l - e_l < 0$ for some good l, we must have $p^*_l = 0$ by the price player's optimization. But then we can simply replace x^{*1}_l by $x^{*1}_l - (x^*_l - e_l)$ and get market clearing.

□

CHAPTER 3

Game theory

Readings due: MWG 6B, 7B, 7D-7E, 8A-8E

1. Preferences under uncertainty

Definition 3.1. *A simple lottery L is a list $L = (p_1, \ldots, p_N)$ with $p_n \geq 0$ for all n and $\sum_n p_n = 1$, where p_n is interpreted as the probability of outcome n occurring.*

Definition 3.2. *The preference relation $\succsim$ on the space of simple lotteries $\mathscr{L}$ is continuous if for any $L, L', L'' \in \mathscr{L}$, the sets*

$$\{\alpha \in [0,1] : \alpha L + (1-\alpha)L' \succsim L''\} \subset [0,1]$$

and

$$\{\alpha \in [0,1] : L'' \succsim \alpha L + (1-\alpha)L'\} \subset [0,1]$$

are closed.

We put more structure on $U(\cdot)$ via the independence axiom.

Definition 3.3. *The preference relation $\succsim$ on the space of simple lotteries $\mathscr{L}$ satisfies the independence axiom if for all $L, L', L'' \in \mathscr{L}$ and $\alpha \in (0,1)$ we have $L \succsim L'$ if and only if $\alpha L + (1-\alpha)L'' \succsim xL' + (1-\alpha)L''$.*

Now the expected utility representation of preferences.

Definition 3.4. *The utility function $U : \mathscr{L} \to \mathbb{R}$ has an expected utility form if there is an assignment of numbers $(u_1, \ldots, u_N)$ to the N outcomes such that for every simple lottery $L = (p_1, \ldots, p_N) \in \mathscr{L}$ we have*

$$U(L) = u_1 p_1 + \cdots + u_N p_N.$$

Proposition 3.1. *Suppose that the rational preference relation $\succsim$ on the space of lotteries $\mathscr{L}$ satisfies the continuity and independence axioms. Then $\succsim$ admits a utility representation of the expected utility form. That is, we can assign a number u_n to each outcome $n = 1, \ldots, N$ in such a manner that for any two lotteries $L = (p_1, \ldots, p_N)$ and $L' = (p'_1, \ldots, p'_N)$, we have $L \succsim L$ if and only if $\sum_{n=1}^N u_n p_n \geq \sum_{n=1}^N u_n p'_n$.*

Now we can begin discussing complete and incomplete information games.

2. Games with complete information

Games are settings of strategic interdependence. Define a game as $\Gamma_N = [I, \{S_i\}, \{u_i(\cdot)\}]$: I players, set of strategies S_i for each $i \in I$, $u(\cdot)$ utility functions.

Definition 3.5. *A strategy $s_i \in S_i$ is a strictly dominant strategy for player i in game $\Gamma_N = [I, \{S_i\}, \{u_i(\cdot)\}]$ if for all $s'_i \neq s_i$, we have*

$$u_i(s_i, s_{-i}) > u_i(s'_i, s_{-i})$$

for all $s_{-i} \in S_{-i}$.

Definition 3.6. *A strategy $s_i \in S_i$ is strictly dominated for player i in game $\Gamma_N = [I, \{S_i\}, \{u_i(\cdot)\}]$ if there exists another strategy $s_i' \in S_i$ such that for all $s_{-i} \in S_{-i}$,*

$$u_i(s_i', s_{-i}) > u_i(s_i, s_{-i}).$$

In this case, we say that strategy s_i' strictly dominates strategy s_i.

Definition 3.7. *A strategy $s_i \in S_i$ is weakly dominated in game $\Gamma_N = [I, \{S_i\}, \{u_i(\cdot)\}]$ if there exists another strategy $s_i' \in S_i$ such that for all $s_{-i} \in S_{-i}$,*

$$u_i(s_i', s_{-i}) \geq u_i(s_i, s_{-i}),$$

with strict inequality for some s_{-i}. In this case, we say that strategy s_i' weakly dominates strategy s_i. A strategy is a weakly dominant strategy for player i in game $\Gamma_N = [I, \{S_i\}, \{u_i(\cdot)\}]$ if it weakly dominates every other strategy in S_i.

2.1. Allowing for Mixed Strategies.

Definition 3.8. *A strategy $\sigma_i \in \Delta(S_i)$ is strictly dominated for player i in game $\Gamma_N = [I, \{\Delta(S_i)\}, \{u_i(\cdot)\}]$ if there exists another strategy $\sigma_i' \in \Delta(S_i)$ such that for all $\sigma_{-i} \in \prod_{j \neq i} \Delta(S_j)$,*

$$u_i(\sigma_i', \sigma_{-i}) > u_i(\sigma_i, \sigma_{-i}).$$

In this case, we say that strategy σ_i' strictly dominates strategy σ_i. A strategy σ_i is a strictly dominant strategy for player i in game $\Gamma_N = [I, \{\Delta(S_i)\}, \{u_i(\cdot)\}]$ if it strictly dominates every other strategy in $\Delta(S_j)$.

Definition 3.9. *Player i 's pure strategy $s_i \in S_i$ is strictly dominated in game $\Gamma_N = [I, \{\Delta(S_i)\}, \{u_i(\cdot)\}]$ if and only if there exists another strategy $\sigma_i' \in \Delta(S_i)$ such that*

$$u_i(\sigma_i', s_{-i}) > u_i(s_i, s_{-i})$$

for all $S_{-j} \in S_{-i}$.

Definition 3.10 (Best Response). *In game $\Gamma_N = [I, \{\Delta(S_i)\}, \{u_i(\cdot)\}]$, strategy σ_i is a best response for player i to his rivals' strategies σ_{-i} if*

$$u_i(\sigma_i, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i})$$

for all $\sigma_i' \in \Delta(S_i)$. Strategy σ_i is never a best response if there is no σ_{-j} for which σ_i is a best response.

Now we can discuss Nash Equilibria.

Definition 3.11 (Nash Equilibrium). *A strategy profile $s = (s_1, \ldots, s_I)$ constitutes a Nash equilibrium of game $\Gamma_N = [I, \{S_i\}, \{u_i(\cdot)\}]$ if for every $i = 1, \ldots, I$,*

$$u_i(s_i, s_{-i}) \geq u_i(s_i', s_{-i})$$

for all $s_i' \in S_i$.

Definition 3.12. *A mixed strategy profile $\sigma = (\sigma_1, \ldots, \sigma_l)$ constitutes a Nash equilibrium of game $\Gamma_N = [I, \{\Delta(S_i)\}, \{u_i(\cdot)\}]$ if for every $i = 1, \ldots, l$,*

$$u_i(\sigma_i, \sigma_{-i}) \geq u_i(\sigma_i', \sigma_{-i})$$

for all $\sigma_i' \in \Delta(S_i)$.

Under fairly broad assumptions we can prove that Nash equilbria always exist. First we need a key technical result.

Proposition 3.2. *If the sets $S_1, \ldots, S_I$ are nonempty, S_i is compact and convex, and $u_i(\cdot)$ is continuous in $(s_1, \ldots, s_I)$ and quasiconcave in s_i, then player i 's best-response correspondence $b_i(\cdot)$ is nonempty, convex-valued, and upper hemicontinuous.*

Proof. Note first that $b_i(s_{-i})$ is the set of maximizers of the continuous function $u_i(\cdot, s_{-i})$ on the compact set S_i and so it is nonempty. The convexity of $b_i(s_{-i})$ follows because the set of maximizers of a quasiconcave function [here, the function $u_i(\cdot, s_{-i})$] on a convex set (here, S_i) is convex. Finally, for upper hemicontinuity, we need to show that for any sequence $\left(s_i^n, s_{-i}^n\right) \rightarrow (s_i, s_{-i})$ such that $s_i^n \in b_i\left(s_{-i}^n\right)$ for all n, we have $s_i \in b_i(s_{-i})$. To see this, note that for all $n, u_i\left(s_i^n, s_{-i}^n\right) \geq u_i\left(s_i', s_{-i}^n\right)$ for all $s_i' \in S_i$. Therefore, by the continuity of $u_i(\cdot)$, we have $u_i(s_i, s_{-i}) \geq u_i(s_i', s_{-i})$. □

Proposition 3.3. *A Nash equilibrium exists in game $\Gamma_N = [I, \{S_i\}, \{u_i(\cdot)\}]$ if for all $i = 1, \ldots, I$. (i) S_i is a nonempty, convex, and compact subset of some Euclidean space $\mathbb{R}^M$. (ii) $u_i(s_1, \ldots, s_l)$ is continuous in $(s_1, \ldots, s_l)$ and quasiconcave in s_i.*

Proof. Define the correspondence $b : S \rightarrow S$ by

$$b(s_1, \ldots, s_l) = b_1(s_{-1}) \times \cdots \times b_l(s_{-I}).$$

Note that $b(\cdot)$ is a correspondence from the nonempty, convex, and compact set $S = S_1 \times \cdots \times S_I$ to itself; $b(\cdot)$ is too a nonempty, convex-valued, and upper hemicontinuous correspondence. Thus, all the conditions of the Kakutani fixed point theorem are satisfied. Hence there exists a fixed point for this correspondence, a strategy profile $s \in S$ such that $s \in b(s)$. The strategies at this fixed point constitute a Nash equilibrium because by construction $s_i \in b_i(s_{-i})$ for all $i = 1, \ldots, I$. □

And we can additionally say that mixed strategy equilibria always exist.

Proposition 3.4. *Every game $\Gamma_N = [I, \{\Delta(S_i)\}, \{u_i(\cdot)\}]$ in which the sets $S_1, \ldots, S_1$ have a finite number of elements has a mixed strategy Nash equilibrium.*

Proof. The game $\Gamma_N = [I, \{\Delta(S_i)\}, \{u_i(\cdot)\}]$, viewed as a game with strategy sets $\{\Delta(S_i)\}$ and payoff functions $u_i(\sigma_1, \ldots, \sigma_I) = \sum_{s \in S}\left[\prod_{k=1}^I \sigma_k(s_k)\right] u_i(s)$ for all $i = 1, \ldots, I$, satisfies all the assumptions of Proposition 8.D.3. Hence, Proposition 8.D.2 is a direct corollary of that result. □

3. Games with incomplete information

Each player i has a payoff function $u_i(s_i, s_{-i}, \theta_i)$, where $\theta_i \in \Theta_i$ is a random variable chosen by nature that is observed only by player i. The joint probability distribution of the values of θ_i is given by $F(\theta_1, \ldots, \theta_l)$, which is common knowledge among the players. Letting $\Theta = \Theta_1 \times \cdots \times \Theta_l$,

Definition 3.13. *A Bayesian game is summarized by the data $\left[L, \left\{S_{ij}', \{u_i(\cdot)\}, \Theta, F(\cdot)\right]\right.$.*

Definition 3.14. *A (pure strategy) Bayesian Nash equilibrium for the Bayesian game denoted by $[I, \{S_i\}, \{u_i(\cdot)\}, \Theta, F(\cdot)]$ is a profile of decision rules $(s_1(\cdot), \ldots, s_I(\cdot))$ that constitutes a Nash equilibrium of game $\Gamma_N = [I, \{\mathscr{S}_i\}, \{\tilde{u}_i(\cdot)\}]$. That is, for every $i = 1, \ldots, I$,*

$$\tilde{u}_i(s_i(\cdot), s_{-i}(\cdot)) \geq \tilde{u}_i(s_i'(\cdot), s_{-i}(\cdot))$$

for all $s_i'(\cdot) \in \mathscr{S}_i$, where $\tilde{u}_i(s_i(\cdot), s_{-i}(\cdot)) = E_\theta[u_i(s_1(\theta_1), \ldots, s_I(\theta_I), \theta_i)]$.

Proposition 3.5. *A profile of decision rules $(s_1(\cdot), \ldots, s_1(\cdot))$ is a Bayesian Nash equilibrium in Bayesian game $[I, \{S_i\}, \{u_i(\cdot)\}, \Theta, F(\cdot)]$ if and only if, for all i and all $\vec{\theta}_i \in \Theta_i$ occurring with positive probability*

$$E_{\theta_{-i},}\left[u_i\left(s_i\left(\bar{\theta}_i\right), s_{-i}(\theta_{-i}), \bar{\theta}_i\right) \mid \bar{\theta}_i\right] \geq E_{\theta_{-},}\left[u_i\left(s_i', s_{-i}(\theta_{-i}), \bar{\theta}_i\right) \mid \bar{\theta}_i\right]$$

for all $s_i' \in S_i$, where the expectation is taken over realizations of the other players' random variables conditional on player i 's realization of his signal $\bar{\theta}_i$.

Proof. For necessity, note that if the above did not hold for some player i for some $\bar{\theta}_i \in \Theta_i$ that occurs with positive probability, then player i could do better by changing his strategy choice in the event he gets realization $\bar{\theta}_i$, contradicting $(s_1(\cdot), \ldots, s_t(\cdot))$ being a Bayesian Nash equilibrium. In the other direction, if condition (8.E.2) holds for all $\bar{\theta}_i \in \Theta_i$ occurring with positive probability, then player i cannot improve on the payoff he receives by playing strategy $s_i(\cdot)$. □

CHAPTER 4

Mechanism design

Readings due: MWG 23B-23D

We present material from MWG; for a more discursive form of this content, we strongly recommend Ludvig Sinander's course notes on mechanism design.

There are I agents, indexed by $i = 1, \ldots, I$ who must make a collective choice from some set X of possible alternatives.

Agents' preferences depend on the realizations of $\theta = (\theta_1, \ldots, \theta_I)$; the agents may want the collective decision to depend on θ. Recall incomplete information games, types, social choice functions, which is redefined below.

Definition 4.1 (Social choice function). *A social choice function is a function $f : \Theta_1 \times \cdots \times \Theta_I \to X$ that, for each possible profile of the agents' types $(\theta_1, \ldots, \theta_I)$, assigns a collective choice $f(\theta_1, \ldots, \theta_I) \in X$.*

Definition 4.2 (Paretian social choice function). *The social choice function $f : \Theta_1 \times \cdots \times \Theta_I \to X$ is ex post efficient (or Paretian) if for no profile $\theta = (\theta_1, \ldots, \theta_I)$ there is an $x \in X$ such that $u_i(x, \theta_i) \geq u_i(f(\theta), \theta_i)$ for every i, and $u_i(x, \theta_i) > u_i(f(\theta), \theta_i)$ for some i.*

Let the density over the possible realizations of $\theta \in \Theta_1 \times \cdots \times \Theta_1$ be $\phi(\cdot)$. The probability density $\phi(\cdot)$ as well as the sets $\Theta_1, \ldots, \Theta_t$ and the utility functions $u_i(\cdot, \theta_i)$ are common knowledge but the value of each agent i's type is observed only by i.

Definition 4.3. *A mechanism $\Gamma = (S_1, \ldots, S_I, g(\cdot))$ is a collection of I strategy sets $(S_1, \ldots, S_I)$ and an outcome function $g : S_1 \times \cdots \times S_I \to X$.*

The mechanism Γ and possible types $(\Theta_1, \ldots, \Theta_I)$, probability density $\phi(\cdot)$, and Bernoulli utility functions $(u_1(\cdot), \ldots, u_I(\cdot))$ define a Bayesian game of incomplete information.

Definition 4.4. *The mechanism $\Gamma = (S_1, \ldots, S_I, g(\cdot))$ implements social choice function $f(\cdot)$ if there is an equilibrium strategy profile $(s_1^*(\cdot), \ldots, s_I^*(\cdot))$ of the game induced by Γ such that $g(s_1^*(\theta_1), \ldots, s_I^*(\theta_I)) = f(\theta_1, \ldots, \theta_I)$ for all $(\theta_1, \ldots, \theta_I) \in \Theta_1 \times \cdots \times \Theta_I$.*

There are two prominent solution concepts: *Bayesian Nash* and *dominant strategy equilibrium.* We present both.

Definition 4.5. *A direct revelation mechanism is a mechanism in which $S_i = \Theta_i$ for all i and $g(\theta) = f(\theta)$ for all $\theta \in \Theta_1 \times \cdots \times \Theta_I$.*

Definition 4.6. *The social choice function $f(\cdot)$ is truthfully implementable (or incentive compatible) if the direct revelation mechanism $\Gamma = (\Theta_1, \ldots \Theta_I, f(\cdot))$ has an equilibrium $(s_1^*(\cdot), \ldots, s_I^*(\cdot))$ in which $s_i^*(\theta_i) = \theta_i$ for all $\theta_i \in \Theta_i$ and all $i = 1, \ldots, I$; that is, if truth telling by each agent i constitutes an equilibrium of $\Gamma = (\Theta_1, \ldots, \Theta_I, f(\cdot))$.*

1. Bayesian Nash mechanism design

Definition 4.7. *The strategy profile $s^*(\cdot) = (s_1^*(\cdot), \ldots, s_I^*(\cdot))$ is a Bayesian Nash equilibrium of mechanism $\Gamma = (S_1, \ldots, S_I, g(\cdot))$ if, for all i and all $\theta_i \in \Theta_i$,*

$$E_{\theta_{-i}}\left[u_i\left(g\left(s_i^*(\theta_i), s_{-i}^*(\theta_{-i})\right), \theta_i\right) \mid \theta_i\right] \geq E_{\theta_{-i}}\left[u_i\left(g\left(\hat{s}_i, s_{-i}^*(\theta_{-i})\right), \theta_i\right) \mid \theta_i\right]$$

for all $\hat{s}_i \in S_i$.

DEFINITION 4.8. *The mechanism* $\Gamma = (S_1, \ldots, S_l, g(\cdot))$ *implements the social choice function* $f(\cdot)$ *in Bayesian Nash equilibrium if there is a Bayesian Nash equilibrium of* $\Gamma, s^*(\cdot) = (s_1^*(\cdot), \ldots, s_l^*(\cdot))$, *such that* $g(s^*(\theta)) = f(\theta)$ *for all* $\theta \in \Theta$.

DEFINITION 4.9. *The social choice function* $f(\cdot)$ *is truthfully implementable in Bayesian Nash equilibrium (or Bayesian incentive compatible) if* $s_i^*(\theta_i) = \theta_i$ *for all* $\theta_i \in \Theta_i$ *and* $i = 1, \ldots, I$ *is a Bayesian Nash equilibrium of the direct revelation mechanism* $\Gamma = (\Theta_1, \ldots, \Theta_I, f(\cdot))$. *That is, if for all* $i = 1, \ldots, I$ *and all* $\theta_i \in \Theta_i$,

$$E_{\theta_{-i}}\left[u_i\left(f\left(\theta_i, \theta_{-i}\right), \theta_i\right) \mid \theta_i\right] \geq E_{\theta_{-i}}\left[u_i\left(f\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) \mid \theta_i\right]$$

for all $\hat{\theta}_i \in \Theta_i$.

We can only focus on if $f(\cdot)$ is truthfully implementable because of the revelation principle for Bayesian Nash equilibrium.

PROPOSITION 4.1. *Suppose that there exists a mechanism* $\Gamma = (S_1, \ldots, S_I, g(\cdot))$ *that implements the social choice function* $f(\cdot)$ *in Bayesian Nash equilibrium. Then* $f(\cdot)$ *is truthfully implementable in Bayesian Nash equilibrium.*

PROOF. If $\Gamma = (S_1, \ldots, S_l, g(\cdot))$ implements $f(\cdot)$ in Bayesian Nash equilibrium, then there exists a profile of strategies $s^*(\cdot) = (s_1^*(\cdot), \ldots, s_I^*(\cdot))$ such that $g(s^*(\theta)) = f(\theta)$ for all θ, and for all i and all $\theta_i \in \Theta_i$,

$$E_{\theta_{-i}}\left[u_i\left(g\left(s_i^*\left(\theta_i\right), s_{-i}^*\left(\theta_{-i}\right)\right), \theta_i\right) \mid \theta_i\right] \geq E_{\theta_{-i}}\left[u_i\left(g\left(s_i, s_{-i}^*\left(\theta_{-i}\right)\right), \theta_i\right) \mid \theta_i\right]$$

for all $\hat{s}_i \in S_i$. Then we have for all i and $\theta_i \in \Theta_i$,

$$E_{\theta_{-i}}\left[u_i\left(g\left(s_i^*\left(\theta_i\right), s_{-i}^*\left(\theta_{-i}\right)\right), \theta_i\right) \mid \theta_i\right] \geq E_{\theta_{-i}}\left[u_i\left(g\left(s_i^*\left(\theta_i\right), s_{-i}^*\left(\theta_{-i}\right)\right), \theta_i\right) \mid \theta_i\right]$$

for all $\hat{\theta}_i \in \Theta_i$. Since $g(s^*(\theta)) = f(\theta)$ for all θ we have for all i and all $\theta_i \in \Theta_i$

$$E_{\theta_{-i}}\left[u_i\left(f\left(\theta_i, \theta_{-i}\right), \theta_i\right) \mid \theta_i\right] \geq E_{\theta_{-i}}\left[u_i\left(f\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right) \mid \theta_i\right]$$

for all $\hat{\theta}_i \in \Theta_i$, which is the condition for $f(\cdot)$ to be truthfully implementable in Bayesian Nash equilibrium. □

2. Dominant strategy mechanism design

DEFINITION 4.10. *The strategy profile* $s^*(\cdot) = (s_1^*(\cdot), \ldots, s_I^*(\cdot))$ *is a dominant strategy equilibrium of mechanism* $\Gamma = (S_1, \ldots, S_I, g(\cdot))$ *if, for all* i *and all* $\theta_i \in \Theta_i$,

$$u_i\left(g\left(s_i^*\left(\theta_i\right), s_{-i}\right), \theta_i\right) \geq u_i\left(g\left(s_i', s_{-i}\right), \theta_i\right)$$

for all $s_i' \in S_i$ *and all* $s_{-i} \in S_{-i}$.

DEFINITION 4.11. *The mechanism* $\Gamma = (S_1, \ldots, S_I, g(\cdot))$ *implements the social choice function* $f(\cdot)$ *in dominant strategies if there exists a dominant strategy equilibrium of* $\Gamma, s^*(\cdot) = (s_1^*(\cdot), \ldots, s_I^*(\cdot))$, *such that* $g(s^*(\theta)) = f(\theta)$ *for all* $\theta \in \Theta$.

DEFINITION 4.12. *The social choice function* $f(\cdot)$ *is truthfully implementable in dominant strategies (or dominant strategy incentive compatible, or strategy-proof, or straightforward) if* $s_i^*(\theta_i) = \theta_i$ *for all* $\theta_i \in \Theta_i$ *and* $i = 1, \ldots, I$ *is a dominant strategy equilibrium of the direct revelation mechanism* $\Gamma = (\Theta_1, \ldots, \Theta_I, f(\cdot))$. *That is, if for all* i *and all* $\theta_i \in \Theta_i$

$$u_i\left(f\left(\theta_i, \theta_{-i}\right), \theta_i\right) \geq u_i\left(f\left(\hat{\theta}_i, \theta_{-i}\right), \theta_i\right)$$

for all $\hat{\theta}_i \in \Theta_i$ *and all* $\theta_{-i} \in \Theta_{-i}$.

PROPOSITION 4.2 (The Revelation Principle for Dominant Strategies). *Suppose that there exists a mechanism* $\Gamma = (S_1 \ldots, S_I, g(\cdot))$ *that implements the social choice function* $f(\cdot)$ *in dominant strategies. Then* $f(\cdot)$ *is truthfully implementable in dominant strategies.*

PROOF. If $\Gamma=(S_1,\ldots,S_1,g(\cdot))$ implements $f(\cdot)$ in dominant strategies, then there exists a profile of strategies $s^*(\cdot)=(s_1^*(\cdot),\ldots,s_I^*(\cdot))$ such that $g\left(s^*(\theta)\right)=f(\theta)$ for all θ and, for all i and all $\theta_i\in\Theta_i$,

$$u_i\left(g\left(s_i^*\left(\theta_i\right),s_{-i}\right),\theta_i\right)\geq u_i\left(g\left(\hat{s_i},s_{-i}\right),\theta_i\right)$$

for all $\hat{s_i}\in S_i$ and all $s_{-i}\in S_{-i}$. This implies, in particular, that for all i and all $\theta_i\in\Theta_i$,

$$u_i\left(g\left(s_i^*\left(\theta_i\right),s_{-i}^*\left(\theta_{-i}\right)\right),\theta_i\right)\geq u_i\left(g\left(s_i^*\left(\hat{\theta}_i\right),s_{-i}^*\left(\theta_{-i}\right)\right),\theta_i\right)$$

for all $\hat{\theta}_i\in\Theta_l$ and all $\theta_{-i}\in\Theta_{-i}$. Since $g\left(s^*(\theta)\right)=f(\theta)$ for all θ, this implies that, for all i and all $\theta_i\in\Theta_i$,

$$u_i\left(f\left(\theta_i,\theta_{-i}\right),\theta_i\right)\geq u_i\left(f\left(\hat{\theta}_i,\theta_{-i}\right),\theta_i\right)$$

for all $\hat{\theta}_i\in\Theta_i$ and all $\theta_{-i}\in\Theta_{-i}$. This is precisely the condition for $f(\cdot)$ to be truthfully implementable in dominant strategies. $\square$

Define the lower contour set of alternative x when agent i has type θ_i to be

$$L_i\left(x,\theta_i\right)=\left\{z\in X:u_i\left(x,\theta_i\right)\geq u_i\left(z,\theta_i\right)\right\}.$$

PROPOSITION 4.3. *The social choice function $f(\cdot)$ is truthfully implementable in dominant strategies if and only if for all i, all $\theta_{-i}\in\Theta_{-i}$, and all pairs of types for agent i,θ_i' and $\theta_i''\in\Theta_i$, we have*

$$f\left(\theta_i'',\theta_{-i}\right)\in L_i\left(f\left(\theta_i',\theta_{-i}\right),\theta_i'\right)\quad\text{and}\quad f\left(\theta_i',\theta_{-i}\right)\in L_i\left(f\left(\theta_i'',\theta_{-i}\right),\theta_i''\right)$$

3. Examples and Extensions

EXAMPLE 1 (Truthful Implementation of the Social Choice Function Implemented by the First-Price Sealed-Bid Auction). Each potential buyer i is allowed to submit a sealed bid, $b_i\geq 0$. The bids are then opened and the buyer with the highest bid gets the good and pays an amount equal to his bid to the seller.

Consider two potential buyers ($I=2$) and each θ_i is independently drawn from the uniform distribution on $[0,1]$. We will look for an equilibrium in which each buyer's strategy $b_i(\cdot)$ takes the form $b_i\left(\theta_i\right)=\alpha_i\theta_i$ for $\alpha_i\in[0,1]$. Suppose that buyer 2's strategy has this form, and consider buyer 1's problem. For each θ_1 he wants to solve

$$\operatorname{Max}_{b_1\geq 0}\left(\theta_1-b_1\right)\operatorname{Prob}\left(b_2\left(\theta_2\right)\leq b_1\right)$$

Since buyer 2's highest possible bid is α_2 (he submits a bid of α_2 when $\theta_2=1$), it is evident that buyer 1 should never bid more than α_2. Moreover, since θ_2 is uniformly distributed on $[0,1]$ and $b_2\left(\theta_2\right)\leq b_1$ if and only if $\theta_2\leq\left(b_1/\alpha_2\right)$, we can write buyer 1's problem as

$$\operatorname{Max}_{b_1\in[0,a_2]}\left(\theta_1-b_1\right)\left(b_1/\alpha_2\right)$$

The solution to this problem is

$$b_1\left(\theta_1\right)=\begin{cases}\frac{1}{2}\theta_1 & \text{if } \frac{1}{2}\theta_1\leq\alpha_2\\ \alpha_2 & \text{if } \frac{1}{2}\theta_1>\alpha_2\end{cases}$$

By similar reasoning,

$$b_2\left(\theta_2\right)=\begin{cases}\frac{1}{2}\theta_2 & \text{if } \frac{1}{2}\theta_2\leq\alpha_1\\ \alpha_1 & \text{if } \frac{1}{2}\theta_2>\alpha_1\end{cases}$$

Letting $\alpha_1=\alpha_2=\frac{1}{2}$, we see that the strategies $b_i\left(\theta_i\right)=\frac{1}{2}\theta_i$ for $i=1,2$ constitute a Bayesian Nash equilibrium for this auction. Thus, there is a Bayesian Nash equilibrium of this first-price sealed-bid auction that indirectly yields the outcomes specified by the social choice function $f(\theta) = (y_0(\theta), y_1(\theta), y_2(\theta), t_0(\theta), t_1(\theta), t_2(\theta))$ in which

$$\begin{aligned}
y_1(\theta) &= 1 \quad \text{if } \theta_1 \geq \theta_2; \quad = 0 \text{ if } \theta_1 < \theta_2 \\
y_2(\theta) &= 1 \quad \text{if } \theta_1 < \theta_2; \quad = 0 \text{ if } \theta_1 \geq \theta_2 \\
y_0(\theta) &= 0 \quad \text{for all } \theta \\
t_1(\theta) &= -\frac{1}{2}\theta_1 y_1(\theta) \\
t_2(\theta) &= -\frac{1}{2}\theta_2 y_2(\theta) \\
t_0(\theta) &= -\left(t_1(\theta) + t_2(\theta)\right).
\end{aligned}$$

We can go further and say this is a truthful implementation. When facing the direct revelation mechanism $(\Theta_1, \ldots, \Theta_1, f(\cdot))$ with $f(\theta) = (y_0(\theta), y_1(\theta), y_2(\theta), t_0(\theta), t_1(\theta), t_2(\theta))$, buyer 1's optimal announcement $\hat{\theta}_1$ when he has type θ_1 solves

$$\text{Max}_{\hat{\theta}_1} \left(\theta_1 - \frac{1}{2}\hat{\theta}_1\right) \text{Prob}\left(\theta_2 \leq \hat{\theta}_1\right)$$

or

$$\text{Max}_{\hat{\theta}_1} \left(\theta_1 - \frac{1}{2}\hat{\theta}_1\right) \hat{\theta}_1.$$

The first-order condition for this problem gives $\hat{\theta}_1 = \theta_1$. So truth telling is buyer 1's optimal strategy given that buyer 2 always tells the truth. A similar conclusion follows for buyer 2. Thus, the social choice function implemented by the first-price sealed-bid auction (in a Bayesian Nash equilibrium) can also be truthfully implemented (in a Bayesian Nash equilibrium) through a direct revelation mechanism.

Note that we can generate the revenue equivalence theorem through the analysis of Bayesian mechanism design.

Proposition 4.4 (The Revenue Equivalence Theorem). *Consider an auction setting with I risk-neutral buyers, in which buyer i's valuation is drawn from an interval $\left[\underline{\theta}_i, \bar{\theta}_i\right]$ with $\underline{\theta}_i \neq \bar{\theta}_i$ and a strictly positive density $\phi_i(\cdot) > 0$, and in which buyers' types are statistically independent. Suppose that a given pair of Bayesian Nash equilibria of two different auction procedures are such that for every buyer i : (i) For each possible realization of $(\theta_1, \ldots, \theta_I)$, buyer i has an identical probability of getting the good in the two auctions; and (ii) Buyer i has the same expected utility level in the two auctions when his valuation for the object is at its lowest possible level. Then these equilibria of the two auctions generate the same expected revenue for the seller.*

We prove this in the unit on auctions, though it is simple to prove (use the revelation principle to focus only on incentive compatible mechanisms).

We conclude with an impossibility: For a general class of problems there is no satisfactory social choice function in dominant strategies. For this, recall dictatorship and monotonicity.

Theorem 4.1 (Gibbard-Satterthwaite). *Suppose that X is finite and contains at least three elements, that $\mathscr{R}_i = \mathscr{P}$ for all i, and that $f(\Theta) = X$. Then the social choice function $f(\cdot)$ is truthfully implementable in dominant strategies if and only if it is dictatorial.*

Part 2

Matching without transfers

CHAPTER 5

One-to-one matching

1. Model

There are a finite set of doctors D and hospitals H. Each doctor $d \in D$ has complete, transitive, strict preferences over hospitals $H \cup \{\emptyset\}$ denoted $\succ_d$. Each hospital $h \in H$ also has preferences over $D \cup \{\emptyset\}$, denoted $\succ_h$. A doctor (hospital) is acceptable to a hospital (doctor) if $d \succ_h \emptyset$ ($h \succ_d \emptyset$).

Definition 5.1. *A matching is a function $\mu : D \cup H \to D \cup H \cup \{\emptyset\}$ such that*

- *$\mu(d) \in H \cup \{\emptyset\}$ - a doctor is matched to a hospital or is unmatched,*
- *$\mu(h) \in D \cup \{\emptyset\}$ - a hospital is matched to a doctor or is unmatched,*
- *$\mu(h) = d \Leftrightarrow \mu(d) = h \forall d \in D, h \in H$ - if a doctor is matched to particular hospital then that hospital must be matched to that doctor.*

2. Stability and deferred acceptance

Definition 5.2. *A matching μ is (pairwise) stable if it is:*

- *Individually rational: there is no agent i such that $\emptyset \succ_i \mu(i)$.*
- *Not (pairwise) blocked: there is no pair d and h, such that $d \succ_h \mu(h)$ and $h \succ_d \mu(d)$.*

Theorem 5.1 (Gale and Shapley, 1962). *There exists a stable matching in any one-to-one matching market.*

To prove this it will be helpful to define the deferred acceptance algorithm:

Definition 5.3 (Deferred Acceptance). *Consider finite sets H and D with agents $h \in H$ and $d \in D$. The (doctor-proposing) deferred acceptance algorithm is as follows:*
1. $\forall d \in D : d$ proposes to its most preferred firm to whom she hasn't already proposed.
2. $\forall h \in H : h$ tentatively accepts its best proposal and rejects the rest.
3. $\forall d \in D$: If d is rejected, she proposes to the next h on its list.
Repeat steps 2 and 3 until there are no more proposals. Return the matching and call it μ.

Now we simultaneously prove Gale and Shapley (1962) and the following:

Theorem 5.2. *The DA algorithm terminates at a stable matching.*

Proof. The resulting matching μ of DA is individually rational because at each step of the algorithm, no doctor proposes to an unacceptable hospital and no hospital holds a proposal of an unacceptable doctor.

μ is not blocked by any pair d and h because: Suppose $h \succ_d \mu(d)$ for some d and h (and d is acceptable to h). This means that d applied to h and was rejected by h at some step of DA. Since h 's tentative match only improves as the algorithm proceeds, the match $\mu(h)$ at the end of DA is still better for h than d. So h is not interested in blocking μ with d, i.e., $\mu(h) \succ_h d$. □

Definition 5.4. *A stable matching μ is doctor-optimal if it is weakly preferred by all doctors to any other stable matching μ'.*

We need an important definition before we prove the next result:

Definition 5.5. *A hospital is unachievable for a doctor if there is no stable matching in which they are matched.*

THEOREM 5.3. *The doctor-proposing DA algorithm finds the doctor-optimal stable matching.*

PROOF. We show that no doctor is ever rejected by an achievable hospital.

By induction, suppose that at some step of DA a doctor has not yet been rejected by an achievable hospital. If the hospital rejects the doctor because he is unacceptable, then this hospital is unachievable for the doctor. A contradiction. If the hospital rejects the doctor in favour of another doctor d', then this hospital is also unachievable for the doctor.

Hence, the doctor-proposing deferred acceptance algorithm produces a matching is which each doctor is matched to their most preferred achievable hospital. □

THEOREM 5.4 (Opposition of Interests). *If μ and μ' are stable matchings, then all doctors like μ at least as much as μ' if and only if all hospitals like μ' at least as well as μ.*

PROOF. Let μ and μ' be stable matchings such that $\mu \succ_D \mu'$. Let us show that $\mu' \succ_H \mu$. Suppose, towards a contradiction, that this isn't the case.

Then there must be a hospital h which prefers μ to μ'. This means that the hospital h hires a different doctor at μ' and μ. Hence, there is a doctor who has a different match $d = \mu(h)$. But, all doctors prefer μ to μ', so d and h block μ'. Therefore, μ' is not stable, a contradiction. □

THEOREM 5.5 (Lone Wolves / Rural Hospitals). *Any agent who is unmatched in one stable outcome is unmatched in all stable outcomes. That is, the set of matched agents is invariant across stable matchings.*

PROOF. Pick a matching. Using the Opposition of Interests Theorem, note that since each agent has at most one partner, we can only have both:

- that all doctors who are matched under the original matching are matched under the doctor-optimal stable matching, and
- that all hospitals who are matched under the doctor-optimal stable matching are matched under the original matching

That all doctors who are matched under the original matching are matched under the doctor-optimal stable matching and

if the same sets of doctors and hospitals are matched under each matching. □

THEOREM 5.6 (Weak Pareto Optimality). *There is no individually rational matching which is strictly preferred by all doctors to the doctor-optimal stable matching.*

Comparative statics are straightforward. Clearly adding an extra doctor makes all the doctors weakly worse off and all the hospitals weakly better off. Increasing the length of a doctor's acceptable hospital list makes all the doctors weakly worse off and all the hospital weakly better off.

3. Incentive compatibility

Consider matching as a direct mechanism: a (direct matching) mechanism $\psi \in \Psi$ is a is a mapping from the set of preference profiles to the set of matchings. (The Revelation Principle lets us focus only on direct mechanisms!) We say a mechanism is stable/efficient if it always produce a stable/efficient matching.

DEFINITION 5.6. *A mechanism ψ is strategyproof for $A \subseteq D \cup H$ if, for all $a \in A$ and any preference profile $\succ$, there is no preference profile $\succ'_a$ such that $\psi(\succ'_a, \succ_{-a})$ is strictly preferred to $\psi(\succ)$ by a.*

THEOREM 5.7. *The doctor-proposing DA algorithm is strategyproof for doctors.*

PROOF. We adapt the proofs in Theorem 10 and Theorem 11 in Hatfield and Milgrom (2005) to the one-to-one case.

Fix the preferences of other doctors and of all the hospitals, and define a hospital x to be the hospital with which d matches by submitting the set of preferences $\succsim_d$: $h_1 \succ_d h_2 \succ_d \cdots \succ_d h_n \succ_d x$. First we show that the outcome that d obtains by reporting preferences $\succsim'_d$: x is also x. Let μ

denote the matching when doctor d submits preference $\succsim_d$. If this matching is stable under the reported preferences but not stable under $\succsim_d'$, then there exists a blocking coalition. This blocking coalition must contain d, as no other doctor's preferences have changed, but that is impossible, since x is d 's favorite hospital according to the preferences $\succsim_d'$. Since μ is stable under $\succsim_d'$, the doctor-optimal stable match under $\succsim_d'$ must make every doctor (weakly) better off than at μ. In particular, doctor d must obtain x.

We know now that submitting $\succsim_d'$: x obtains x, so by the rural hospitals theorem, d is in every stable matching so long as $\succsim_d'$: x is submitted. So, every matching μ at which d is unemployed is blocked by some coalition when d submits $\succsim_d'$. Consequently, if d submits the preferences $\succsim_d^*$: $h_1 \succ h_2 \succ \cdots \succ h_n \succ x$, then every allocation at which d is unemployed is still blocked, by the same coalition and set of contracts.

Now consider some other preference ordering $\succsim_d'' : h_1' \succ h_2' \succ \cdots \succ h_n' \succ x \succ h_{n+1}' \succ \cdots \succ h_N'$ that obtain a hospital for d that is $\succsim_d''$-preferred or indifferent to x. Since the doctor-best stable allocation is one at which d gets a $\succsim_d^*$-acceptable hospital, that allocation is weakly $\succsim_d^*$-preferred to x. Finally, the doctor-optimal match when $\succsim_d^*$ is submitted is still the doctor-optimal match when $\succsim_d''$ is submitted, as d 's preferences over hospitals less preferred than x cannot be used to block a match where she receives a hospital weakly preferred to x.

□

THEOREM 5.8. *There is no stable mechanism that is strategyproof for all agents.*

Proof is simple by counterexample.

DEFINITION 5.7. *A mechanism ψ is group-strategyproof for $A \subseteq D \cup H$ if, for all $A' \subseteq A$ and any preference profile $\succ$, there is no preference profile $\succ'_{A'}$ such that $\psi(\succ'_{A'}, \succ_{-A'})$ is strictly preferred to $\psi(\succ)$ by all $a \in A'$.*

THEOREM 5.9. *The doctor-proposing DA algorithm is group-strategyproof for doctors.*

Which follows from this result: Suppose a group of agents lies about their preferences by reporting $\succ'$. Then there is no $\succ'$-stable matching that is preferred to every $\succ$-stable matching by all members of the lying group.

PROOF. We follow Hatfield and Kojima (2009). We prove the result by induction on the size of the coalition whose members misreport their preferences. We know that no coalition of size 1 can profitably misreport preferences. Suppose that no coalition of size at most $n-1$ can profitably misreport preferences, and suppose by way of contradiction that there exists a preference profile $\succsim$ and a coalition A of size n that can profitably misreport its members' preferences.

Define two sets:

- $\{h_d\}_{d \in A}$: the set of hospitals that doctors in A obtain when they report their true preferences.
- $\{h_d'\}_{d \in A}$: the set of hospitals that doctors in A obtain when they misreport their preferences such that every $d \in A$ strictly prefers h_d' to h_d under $\succsim_d$.

Let $\succsim_d'$ be preferences of d in which h_d' is the most preferred hospital to d and the ranking of every other hospital is identical to $\succsim_d$. It is easy to see each doctor in A obtains h_d' under the doctor-optimal stable mechanism when $(\succsim_A', \succsim_{-A})$ is reported. Now consider a preference profile $(\succsim_d', \succsim_{-d})$ with a fixed $d \in A$. The set of stable allocations under $(\succsim_d', \succsim_{-d})$ coincides with the set of stable allocations under $\succsim$, since:

1. **Any stable set Z' of allocations under $\succsim$ is stable under $(\succsim_d', \succsim_{-d})$.** Doctor d prefers h_d' to h_d under $\succsim_d$ and so if a set Y including a matching with d blocks Z' under $\succsim_d'$, then the set Y blocks Z' under $\succsim_d$, contradicting the stability of Z' under $\succsim$. Moreover, no blocking set of Z' that does not involve d is possible under $(\succsim_d', \succsim_{-d})$, since such a blocking set would be a blocking set under $\succsim$, again contradicting stability of Z' under $\succsim$.

2. **Any set Z' of allocations that is unstable under $\succsim$ is unstable under $(\succsim_d', \succsim_{-d})$.** By construction of $\succsim_d'$, no allocation that does not assign h_d' to d and is unstable under $\succsim$ is stable under $(\succsim_d', \succsim_{-d})$. Suppose there exists an allocation that assigns h_d' to d and is stable under

$(\succsim'_d, \succsim_{-d})$. Then, since the doctor-optimal stable mechanism assigns the most preferred contract to d that can be assigned in some stable allocation, d is assigned h'_d by the doctor-optimal stable mechanism under $(\succsim'_d, \succsim_{-d})$. This implies that d could profitably misreport her preferences by submitting $\succsim'_d$ when $\succsim$ is the true preference profile, a contradiction.

We conclude from the above that the doctor-optimal stable allocations under $\succsim$ and $(\succsim'_d, \succsim_{-d})$ coincide. Therefore every doctor $d' \in A\backslash\{d\}$ obtains $x_{d'}$ under $(\succsim'_d, \succsim_{-d})$. But this implies that the group of doctors $A\backslash\{d\}$ can profitably misreport preferences by declaring $\succsim'_{D'\backslash\{d\}}$ instead of $\succsim_{D'\backslash\{d\}}$ when the true preferences are $(\succsim'_d, \succsim_{-d})$, since by assumption every doctor $d' \in A\backslash\{d\}$ obtains $h'_{d'}$ under $(\succsim'_A, \succsim_{-A})$. This contradicts the inductive assumption that no group of size at most $n-1$ can profitably misreport its members' preferences. $\square$

CHAPTER 6

Many-to-many matching

We follow Hatfield and Kominers (2012). Although the expanding work on matching has eliminated nearly all the theoretical restrictions imposed in the early literature, two (natural) assumptions have been maintained throughout, either implicitly or explicitly:

- **acyclicity**: no agent may both buy from and sell to another agent, even through intermediaries, and
- **full substitutability**: upon being endowed with an additional item, an agent's demand for other items is lower, both in the sense of a reduced desire to buy additional items and an increased desire to sell items he currently owns.

In classical matching theory, if either condition is violated, then stable allocations cannot be guaranteed.

1. Model

There are a finite set F of of firms and a finite set X of contracts. Each contract $x \in X$ is associated with both a buyer $x_B \in F$ and a seller $x_S \in F$; there may be several contracts with the same buyer and the same seller.

For concreteness, one may suppose each contract $x \in X$ denotes the exchange of a single unit of a good from x_S to x_B. However, contracts need not use a constant unit. For example, labor markets might allow both full- and part-time job contracts.

Let $x_F \equiv \{x_B, x_S\}$ be the set of the firms associated with contract x. For a set of contracts Y, we denote

$$Y_B \equiv \bigcup_{y \in Y} \{y_B\}, \quad Y_S \equiv \bigcup_{y \in Y} \{y_S\}, \quad Y_F \equiv Y_B \cup Y_S.$$

Definition 6.1. *The contract set X is acyclic if there does not exist a cycle, i.e. a set of contracts*

$$\left\{x^1, \ldots, x^N\right\} \subseteq X$$

such that $x_B^1 = x_S^2, x_B^2 = x_S^3, \ldots, x_B^{N-1} = x_S^N, x_B^N = x_S^1$.

This condition is equivalent to the condition that there is an ordering $\triangleright$ on F such that for all $x \in X, x_S \triangleleft x_B$.

Definition 6.2. *X is exhaustive if there is a contract between any two firms, that is, if for all $f \neq f', f, f' \in F$ there exists a contract x such that $x_F = \{f, f'\}$.*

Each $f \in F$ has a strict preference relation $\succsim^f$ over sets of contracts involving f. Let $Y|_f \equiv \{y \in Y : f \in y_F\}$ be the set of contracts in Y associated with firm f.

For any $Y \subseteq X$, we first define the choice set of f as the set of contracts f chooses from Y :

$$C^f(Y) \equiv \max_{P^f} \{Z \subseteq Y : x \in Z \Rightarrow f \in x_F\}$$

It will also be convenient to define the choice function for f as a buyer when f has access to the contracts in $Y \subseteq X$ for which f is a buyer and has access to the contracts in $Z \subseteq X$ for which f is a seller. Hence we define

$$C_B^f(Y \mid Z) \equiv \left\{x \in C^f\left(\{y \in Y : y_B = f\} \cup \{z \in Z : z_S = f\}\right) : x_B = f\right\}.$$

Analogously, we define

$$C^f_S(Z \mid Y) \equiv \left\{x \in C^f\left(\{y \in Y : y_B = f\} \cup \{z \in Z : z_S = f\}\right) : x_S = f\right\}.$$

We also define the buyer- and seller-rejected sets as

$$R^f_B(Y \mid Z) \equiv Y - C^f_B(Y \mid Z),$$
$$R^f_S(Z \mid Y) \equiv Z - C^f_S(Z \mid Y).$$

2. Substitutability and aggregate demand

DEFINITION 6.3. *The preferences of $f \in F$ are same-side substitutable if for all $Y' \subseteq Y \subseteq X$ and $Z' \subseteq Z \subseteq X$,*

(1) $R^f_B\left(Y' \mid Z\right) \subseteq R^f_B(Y \mid Z)$ *and*
(2) $R^f_S\left(Z' \mid Y\right) \subseteq R^f_S(Z \mid Y)$.

This states that any contract that is rejected from a smaller offer set is also rejected from a larger one.

DEFINITION 6.4. *The preferences of $f \in F$ are cross-side complementary if for all $Y' \subseteq Y \subseteq X$ and $Z' \subseteq Z \subseteq X$,*

(1) $R^f_B(Y \mid Z) \subseteq R^f_B\left(Y \mid Z'\right)$ *and*
(2) $R^f_S(Z \mid Y) \subseteq R^f_S\left(Z \mid Y'\right)$.

We say any preference relation that is both same-side substitutable and cross-side complementary is **fully substitutable**.

The Law of Aggregate Demand states that when a firm receives more options as a buyer (while retaining the same options as a seller), its excess stock increases—the number of contracts the firm chooses as a seller does not increase more than the number of contracts the firm chooses as a buyer does.

DEFINITION 6.5. *The preferences of f satisfy the Law of Aggregate Demand if for all $Y, Z \subseteq X$ and $Y' \subseteq Y$,*

$$\left|C^f_B(Y \mid Z)\right| - \left|C^f_B\left(Y' \mid Z\right)\right| \geq \left|C^f_S(Z \mid Y)\right| - \left|C^f_S\left(Z \mid Y'\right)\right|.$$

DEFINITION 6.6. *An allocation A is stable if it is*

(1) *Individually rational: for all $f \in F, C^f(A) = A|_f$;*
(2) *Unblocked: There does not exist a nonempty blocking set $Z \subseteq X$ such that $Z \nsubseteq A$ and for all $f \in Z_F$, $Z|_f \subseteq C^f(A \cup Z)$.*

Ostrovsky (2008) introduced the following notions of chains and chain stability.

DEFINITION 6.7. *A set of contracts $\left\{x^1, \dots, x^N\right\}$ is a chain if*

(1) $x^n_B = x^{n+1}_S$ *for all $n = 1, \dots, N-1$,*
(2) $x^n_S = x^m_S$ *implies that $n = m$, and*
(3) $x^N_B \neq x^1_S$.

DEFINITION 6.8. *An allocation A is chain stable if it is individually rational and there is no chain that is a blocking set.*

When preferences are fully substitutable and the contract set is acyclic, chain stability and stability are equivalent.

Now we demonstrate that fully substitutable preferences are sufficient to guarantee the existence of a lattice of stable allocations when the contract set is acyclic, and for the standard opposition of interest results to hold.

To prove the existence of a stable allocation, we introduce the Φ operator.

Definition 6.9 (Φ Operator). *The Φ operator that generalizes the deferred acceptance algorithm is defined by*

$$\begin{aligned}\Phi_B\left(X^B,X^S\right) &\equiv X - R_S\left(X^S \mid X^B\right)\\ \Phi_S\left(X^B,X^S\right) &\equiv X - R_B\left(X^B \mid X^S\right)\\ \Phi\left(X^B,X^S\right) &\equiv \left(\Phi_B\left(X^B,X^S\right),\Phi_S\left(X^B,X^S\right)\right).\end{aligned}$$

We show that fixed points of the operator correspond to stable allocations. We will need Tarski's fixed point theorem.

3. Tarski's fixed point theorem

Definition 6.10. *A lattice $L = (X, <, \wedge, \vee)$ is complete if there are both a meet (i.e., a greatest lower bound) and a join (that is, a lowest upper bound) for any subset Y of X. These generalized meet and join operations on Y are denoted by $\bigwedge Y$ and $\bigvee Y$,*

Definition 6.11. *A function $f : X \to X$ is monotone if $x \leq y$ implies $f(x) \leq f(y)$ for any elements x, y of X.*

Theorem 6.1 (Tarski, 1955). *If $L = (X, <, \wedge, \vee)$ is a complete lattice and $f : X \to X$ is a monotone function, then $L_f := (X_f, <)$ is a nonempty, complete lattice subset of L, where $X_f := \{x \in X : f(x) = x\}$ is the set of fixed points of f.*

Proof. We reproduce the proof from Fleiner (2003). Let Y be a (possibly empty) subset of X_f. By monotonicity of f, $f(\bigwedge Y) \leq f(y) = y$ for any $y \in Y$; hence $f(\bigwedge Y) \leq \bigwedge Y$. Define

$$K := \{k \in X : k \leq f(k) \wedge \bigwedge Y\}$$

and $l := \bigvee K$. Clearly, if $x = f(x) \leq \bigwedge Y$ for a fixed point x of f, then $x \in K$ and $x \leq l$. Hence, it is enough to show that $f(l) = l$.

By definition, $k \leq l \leq y$ for any $k \in K$ and $y \in Y$. Thus, by monotonicity, $k \leq f(k) \leq f(l) \leq f(y)$. This means that $l = \bigvee K \leq \bigvee\{f(k) : k \in K\} \leq f(l) \leq \bigwedge Y$, hence, that $l \leq f(l) \leq \bigwedge Y$. Again, by monotonicity, $f(l) \leq f(f(l))$; that is $f(l) \in K$. We got that $l \leq f(l) \leq \bigvee K = l$. Thus, l is indeed the meet of Y in X_f.

Obviously, $L^{-1} = (X, \geq)$ is a complete lattice as well, and f is monotone on L^{-1}. According to the above argument, any subset Y of X_f has a $\geq$-meet in X_f, that is, a $\leq$-join in X_f. We conclude that L_f is indeed a nonempty, complete lattice subset of L. $\square$

Theorem 6.2. *Suppose that the set of contracts X is acyclic and that all firms' preferences are fully substitutable. Then if $\Phi\left(X^B,X^S\right) = \left(X^B,X^S\right)$, the allocation $X^B \cap X^S$ is stable. Conversely, if A is a stable allocation, there exist $X^B, X^S \subseteq X$ such that $\Phi\left(X^B,X^S\right) = \left(X^B,X^S\right)$ and $X^B \cap X^S = A$.*

It is clear that Φ is isotone with respect to this order if preferences are fully substitutable. Hence, by Tarski's theorem, there exists a lattice of fixed points of the operator Φ. Furthermore, if the contract set is acyclic, then these fixed points correspond to stable allocations. We prove this below.

Proof. Since the contractual set X is acyclic, there is an ordering $\lhd$ on firms such that $x_S \lhd x_B$ for all $x \in X$. Fix some labeling of firms $f_1, \ldots, f_N$ so that $f_n \lhd f_{n+1}$ for all $n = 1, \ldots, N-1$.

For the first part: Suppose that $X^B \cap X^S \equiv A$ is a fixed point, but that A is not stable. Then either A is not individually rational or A admits a blocking set Z.

If A is not individually rational, there must exist $x \in A$ such that $x \in R^f(A)$ for some $f \in F$. Then either $x \in R^f_B(A \mid A)$ and $x_B = f$ or $x \in R^f_S(A \mid A)$ and $x_S = f$. Assume the former. (The latter case is symmetric.) Then $x \in R^f_B\left(X^B \mid A\right)$ by same-side substitutability. However, every contract in the set $X^S - A$ is rejected by some firm as a seller, and so $R^f_B\left(X^B \mid A\right) = R^f_B\left(X^B \mid X^S\right)$. Hence $x \in R^f_B\left(X^B \mid X^S\right)$, and hence $x \notin X^S = \Phi_S\left(X^B,X^S\right) = X - R_B\left(X^B \mid X^S\right)$, and hence $x \notin X^B \cap X^S = A$, a contradiction.

If there exists a blocking set Z for A, consider a contract z such that $z_S \trianglelefteq y_S$ for all other $y \in Z$. By same-side substitutability, since $z \in C_S^{z_S}(Z \cup A \mid Z \cup \bar{A})$, we have that $z \in C_S^{z_S}(\{z\} \cup A \mid Z \cup A) = C_S^{z_S}(\{z\} \cup A \mid A)$ as there are no contracts in Z such that z_S is a buyer by assumption. Hence $z \in C_S^{z_S}\left(\{z\} \cup X^S \mid A\right)$ and then by cross-side complementarity, $z \in C_S^{z_S}\left(\{z\} \cup X^S \mid X^B\right)$. Hence, if $z \in X^S$, then $z \in X^B = \Phi_B\left(X^B, X^S\right) = X - R_S\left(X^S \mid X^B\right)$. But $z \notin A = X^B \cap X^S$ by assumption, and $X^B \cup X^S = X$, and so $z \in X^B$. Now consider an arbitrary contract $w \in Z$, and suppose that for all contracts $y \in Z$ such that $y_S \triangleleft w_S, y \in X^B$. By same-side substitutability, since $w \in C_S^{w_S}(Z \cup A \mid Z \cup A)$, we have that $w \in C_S^{w_S}(\{w\} \cup A \mid Z \cup A)$. Now, by induction, for any contract $y \in Z$ such that $y_B = w_S, y \in X^B$. Hence, $\{y \in Z : y_B = w_S\} \subseteq X^B$, and $A \subseteq X^B$, implying that $\{y \in Z : y_B = w_S\} \cup A \subseteq X^B$, and so by cross-side complemenarity, $w \in C_S^{w_S}\left(\{w\} \cup A \mid X^B\right)$. Therefore $w \in C_S^{w_S}\left(\{w\} \cup X^S \mid X^B\right)$. Thus, if $w \in X^S$, then $w \in X^B = \Phi_B\left(X^B, X^S\right) = X - R_S\left(X^S \mid X^B\right)$. But $w \notin A = X^B \cap X^S$ by assumption, and $X^B \cup X^S = X$, and so $w \in X^B$. Using induction then, we have that $Z \subseteq X^B$. Working symmetrically for buyers, we have that $Z \subseteq X^S$. Hence, $Z \subseteq X^S \cap X^B = A$ and hence Z is not a blocking set, a contradiction.

For the second part: Suppose that A is a stable allocation. We construct X^B and X^S iteratively over firms. Let $X^B(0) \equiv X^S(0) \equiv A$. Let

$$X^B(n) \equiv \left\{x \in \left(X - X^S(n-1)\right) : x_B = f_n\right\} \cup X^B(n-1)$$
$$X^S(n) \equiv \left\{x \in X : x \in R_S^{f_n}\left(\{x\} \cup A \mid X^B(n)\right)\right\} \cup X^S(n-1).$$

Finally, let $X^B = X^B(N)$ and $X^S = X^S(N)$. We now show that $\left(X^B, X^S\right)$ is a fixed point. We have

$$\begin{aligned}
X - R_S\left(X^S \mid X^B\right) &= X - \left(X^S - C_S\left(X^S \mid X^B\right)\right) \\
&= X - \bigcup_{n=1}^{N} R_S^{f_n}\left(\left(X^S\big|_{f_n}\right) \mid \left(X^B\big|_{f_n}\right)\right) \\
&= X - \bigcup_{n=1}^{N} R_S^{f_n}\left(\left(X^S(n)\big|_{f_n}\right) \mid \left(X^B\big|_{f_n}\right)\right) \text{ as } X^S(n)\big|_{f_n} = X^S\big|_{f_n} \\
&= X - \bigcup_{n=1}^{N} \left(X^S(n) - A\right)\Big|_{f_n} \quad \text{as } f_n \text{ has fully substitutable preferences} \\
&= X - \left(X^S - A\right) \text{ by the definition of } X^S \\
&= X^B \text{ by the definition of } X^B.
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
X - R_B\left(X^B \mid X^S\right) &= X - \left(X^B - C_B\left(X^B \mid X^S\right)\right) \\
&= X - \bigcup_{n=1}^{N} R_B^{f_n}\left(\left(X^B \mid f_n\right) \mid \left(X^S\big|_{f_n}\right)\right) \\
&= X - \bigcup_{n=1}^{N} R_B^{f_n}\left(\left(X^B(n)\big|_{f_n}\right) \mid \left(X^S\big|_{f_n}\right)\right) \text{ as } X^B(n)\big|_{f_n} = X^B\big|_{f_n} \\
&= X - \bigcup_{n=1}^{N} \left(X^B(n) - A\right)\Big|_{f_n} \quad \text{as shown below} \\
&= X - \left(X^B - A\right) \text{ by the definition of } X^B \\
&= X^S \text{ by the definition of } X^S
\end{aligned}$$

To see the fourth equality, observe that

$$\bigcup_{n=1}^{N}\left(X^{B}(n)-A|_{f_n}\right)=X^{B}-A=R_{B}\left(X^{B}(N)\right).$$

Suppose that there exists a nonempty set of contracts

$$Y\subseteq\bigcup_{n=1}^{N}\left(X^{B}(n)-A|_{f_n}\right)-\bigcup_{n=1}^{N}R_{B}^{f_n}\left(X^{B}(n)\mid X^{S}(n)\right).$$

We also know that no contract Y is rejected by a seller (assuming sellers have access to X^B as buyers) as these are contracts X^B. Hence, Y is a blocking set and hence A is not stable, a contradiction. Finally, $A|_{f_n}\cap R_B^f\left(X^B(n)\mid X^S(n)\right)=\varnothing$ as A is individually rational.

Finally, we need to show that $X^B\cap X^S=A$. First, since $X^B(0)\cap X^S(0)=A$ and $X^B(n-1)\subseteq X^B(n)$ and $X^S(n-1)\subseteq X^S(n), A\subseteq X^B(0)\cap X^S(0)$. Suppose that $z\in X^S-A$. Then $z\notin X^B$, as it could only be added in the z_B-th step and since $z\in X^S, a\in X^S\left(f_{z_B}-1\right)$. □

THEOREM 6.3. *Suppose that the set of contracts X is acyclic and that preferences are fully substitutable. Then there exists a nonempty finite lattice of fixed points $\left(X^B,X^S\right)$ of Φ which correspond to stable allocations $A=X^B\cap X^S$.*

THEOREM 6.4. *Suppose that the set of contracts X is acyclic and that preferences are fully substitutable. Then the highest fixed point $\left(\hat{X}^B,\hat{X}^S\right)$ of Φ corresponds to the buyer-optimal stable allocation $\hat{X}^B\cap\hat{X}^S$, and the lowest fixed-point $\left(\check{X}^B,\check{X}^S\right)$ of Φ corresponds to the seller-optimal stable allocation $\check{X}^B\cap\check{X}^S$.*

PROOF. For any stable allocation $\left(X^B,X^S\right)$, we have that

$$\begin{aligned}X^B\cap X^S&=X^B\cap\left(X-R_B\left(X^B\mid X^S\right)\right)\\&=X^B\cap\left(X-\left(X^B-C_B\left(X^B\mid X^S\right)\right)\right)\\&=X^B-\left(X^B-C_B\left(X^B\mid X^S\right)\right)\\&=C_B\left(X^B\mid X^S\right).\end{aligned}$$

For each firm f who is only a buyer, $C_B^f\left(X^B\mid X^S\right)=C^f\left(X^B\right)$, the firm f has a strictly larger choice set under $\left(\hat{X}^B,\hat{X}^S\right)$ than under any other stable allocation, and hence (weakly) prefers $\left(\hat{X}^B,\hat{X}^S\right)$. The proof that $\left(\check{X}^B,\check{X}^S\right)$ is the seller-optimal stable allocation is symmetric. □

CHAPTER 7

Many-to-one matching

In many-to-one matching with contracts, there exist weaker conditions on preferences that guarantee the existence of a stable allocation.

1. Model

The sets of doctors and hospitals are denoted by D and H, respectively, and the set of contracts is denoted by X. We assume only that each contract $x \in X$ is bilateral, so that it is associated with one "doctor" $x_D \in D$ and one "hospital" $x_H \in H$. When all terms of employment are fixed and exogenous, the set of contracts is just the set of doctor-hospital pairs: $X \equiv D \times H$. For the Kelso-Crawford model, a contract specifies a firm, a worker, and a wage, $X \equiv D \times H \times W$. The choice set is the singleton set of the most preferred contract:

$$C_d(X') = \begin{cases} \varnothing \text{ if } \{x \in X' \mid x_D = d, x \succ_d \varnothing\} = \varnothing \\ \{\max_{\succ_d} \{x \in X' \mid x_D = d\}\} \text{ otherwise.} \end{cases}$$

The choices of a hospital h are more complicated, because it has preferences $\succ_h$ over sets of doctors. Its chosen set is a subset of the contracts that name it, that is, $C_h(X') \subset \{x \in X' \mid x_H = h\}$. A hospital can sign only one contract with any given doctor:

$$(\forall h \in H)(\forall X' \subset X)(\forall x, x' \in C_h(X'))\, x \neq x' \Rightarrow x_D \neq x'_D$$

Let $C_D(X') = \cup_{d \in D} C_d(X')$ denote the set of contracts chosen by some doctor from set X'. The remaining offers in X' are in the rejected set: $R_D(X') = X' - C_D(X')$. Similarly, the hospitals' chosen and rejected sets are denoted by $C_H(X') = \cup_{h \in H} C_h(X')$ and $R_H(X') = X' - C_H(X')$.

Definition 7.1. *A set of contracts $X' \subset X$ is a stable allocation if*

1. $C_D(X') = C_H(X') = X'$ *and*
2. *there exists no hospital h and set of contracts $X'' \neq C_h(X')$ such that*

$$X'' = C_h(X' \cup X'') \subset C_D(X' \cup X'').$$

Definition 7.2. *Elements of X are substitutes for hospital h if for all subsets $X' \subset X'' \subset X$ we have $R_h(X') \subset R_h(X'')$.*

Theorem 7.1. *Suppose contracts are substitutes for the hospitals. Then:*

1. *The set of fixed points of F on $X \times X$ is a nonempty finite lattice, and in particular includes a smallest element $(\underline{X}_D, \underline{X}_H)$ and a largest element $(\bar{X}_D, \bar{X}_H)$;*
2. *Starting at $(X_D, X_H) = (X, \varnothing)$, the generalized Gale-Shapley algorithm converges monotonically to the highest fixed point*

$$\begin{aligned}&(\bar{X}_D, \bar{X}_H)\\ &\quad = \sup\{(X', X'') \mid F(X', X'') \geq (X', X'')\};\ \textit{and}\end{aligned}$$

3. *Starting at $(X_D, X_H) = (\varnothing, X)$, the generalized Gale-Shapley algorithm converges monotonically to the lowest fixed point*

$$\begin{aligned}&(\underline{X}_D, \underline{X}_H)\\ &\quad = \inf\{(X', X'') \mid F(X', X'') \leq (X', X'')\}\end{aligned}$$

Hatfield and Milgrom (2005) claimed — incorrectly — that substitutability of hospitals' choice functions is necessary to ensure the existence of stable outcomes in many-to-one matching with contracts. We now present more on substitutes following Hatfield and Kominers (2016).

2. Hidden substitutes

Denote by $\mathrm{d}(x)$ the doctor associated with contract x and $\mathrm{d}(Y)$ the set of doctors associated with some contract in Y, i.e., $\mathrm{d}(Y) = \cup_{y \in Y} \mathrm{d}(y)$

DEFINITION 7.3. *A completion of a many-to-one choice function C^h of hospital $h \in H$ is a choice function $\bar{C}^h$ such that for all $Y \subseteq X$, either*

- *$\bar{C}^h(Y) = C^h(Y)$, or*
- *there exist distinct $z, \hat{z} \in \bar{C}^h(Y)$ that are associated with the same doctor, i.e., $z, \hat{z} \in X_d$ for some $d \in D$.*

PROPOSITION 7.1. *If $\bar{C}$ is a completion of a profile of choice functions C, and $\bar{C}$ satisfies the irrelevance of rejected contracts condition, then any outcome stable with respect to $\bar{C}$ is stable with respect to C.*

PROOF. Assume that A is stable with respect to $\bar{C}$, and show that A is stable with respect to C. We prove the result in three steps: A is individually rational for doctors under C : As doctors have the same choice functions under $\bar{C}$ as under C, the individual rationality of A under $\bar{C}^d$ for each doctor $d \in D$ immediately implies the individually rationality of A under C^d for each doctor $d \in D$. A is individually rational for hospitals under C : The individual rationality of A for doctors implies that each doctor has at most one contract in A, i.e., $|A_d| \leq 1$ for each $d \in D$. Then, as $\bar{C}^h$ completes C^h, it follows that $C^h(A) = \bar{C}^h(A)$ for all $h \in H$ as A does not contain two (or more) contracts with any individual doctor; hence, the individual rationality of A under $\bar{C}^h$ for each hospital $h \in H$ immediately implies the individually rationality of A under C^h for each hospital $h \in H$. A is unblocked under C : Suppose that A is blocked under C by some hospital h and a blocking set $Z \subseteq X_h \backslash A$ under C. First, as Z blocks A under C, and $\bar{C}^d = C^d$ for each $d \in D$, we know that

$$Z_d \subseteq C^d(Z \cup A) = \bar{C}^d(Z \cup A) \text{ for all } d \in D$$

Now, as $\bar{C}^h$ completes C^h, we know from the definition of completability that either

- $\bar{C}^h(Z \cup A) = C^h(Z \cup A)$, or
- there exist distinct $z, \hat{z} \in W \equiv \bar{C}^h(Z \cup A)$ such that $\mathrm{d}(z) = \mathrm{d}(\hat{z})$.

In the former case, we have $Z_h \subseteq C^h(Z \cup A) = \bar{C}^h(Z \cup A)$, as Z blocks A under C; combining this fact with (4) shows that Z blocks A under $\bar{C}$, contradicting the stability of A under $\bar{C}$.

In the latter case, we note that as A is individually rational for doctors under C, we must have $\left|A_{\mathrm{d}(z)}\right| \leq 1$ for each $d \in D$. Then, as we have $z, \hat{z} \in W = \bar{C}^h(Z \cup A)$ such that $\mathrm{d}(z) = \mathrm{d}(\hat{z})$, we know that $\bar{Z} \equiv W \backslash A$ must be nonempty. Now, we have

$$\bar{C}^h(\bar{Z} \cup A) = \bar{C}^h((W \backslash A) \cup A) = \bar{C}^h(W \cup A) = \bar{C}^h((Z \cup A) \cup A) = \bar{C}^h(Z \cup A) = W \supseteq \bar{Z}$$

where the third equality follows from the fact that $\bar{C}^h$ satisfies the irrelevance of rejected contracts condition and $W = \bar{C}^h(Z \cup A)$. Combining shows that $\bar{Z}$ blocks A under $\bar{C}$, contradicting the stability of A under $\bar{C}$. □

If a choice function C^h has a completion that is substitutable, then we say that C^h is substitutably completable. If every choice function in a profile of choice functions C is substitutably completable, then we say that C is substitutably completable.

Many-to-one substitutability is not strictly necessary for stability: while substitutability is key for the existence of stable outcomes, sometimes non-substitutable many-to-one choice functions can inherit substitutable behavior from many-to-many choice functions.

THEOREM 7.2. *If the profile of choice functions C has a substitutable completion that satisfies the irrelevance of rejected contracts condition, then there exists an outcome that is stable with respect to C.*

CHAPTER 8

Matching in networks

Now, for the more general model of matching in networks, both conditions are necessary: If acyclicity fails, then there are fully substitutable preferences for each firm such that no stable allocation exists. For acyclic contract sets, if one firm has preferences that are not fully substitutable, then there exist fully substitutable preferences for the other firms such that no stable allocation exists.

We will add some terminology: for an acyclic contract set X, if $f \lhd f'$, we will say that f is **upstream** of f' and that f' is **downstream** of f.

1. Stability conditions

THEOREM 8.1. *If the set of contracts X admits a cycle $L = \{x^1, \ldots, x^N\}$ and there exists a firm $g \notin L_F$ and a contract between g and some firm in the cycle, then there exist fully substitutable preferences such that no stable allocation exists.*

PROOF. Let y be the contract between g and some member of the cycle. Without loss of generality, we suppose that $f_n \equiv x^n_S, y_S = f_1$, and $y_B = g$. Then for $n = 2, \ldots, N$, let the preferences of firm f_n be

$$P^{f_n} : \{x^{n-1}, x^n\} \succ \varnothing$$

which are fully substitutable. Let

$$P^g : \{y\} \succ \varnothing$$
$$P^{f_1} : \{x^N, y\} \succ \{x^N, x^1\} \succ \varnothing;$$

these preferences are fully substitutable. Let all other firms desire no contracts. Any set $Y \nsubseteq L \cup \{y\}$ is not stable, as it is not individually rational. Any $Z \subsetneq L \cup \{y\}$ is not stable, as it not individually rational unless $Z = \varnothing$, in which case L is a blocking set, or $Z = L$, in which case $\{y\}$ is a blocking set. Finally, $L \cup \{y\}$ is not stable, as it is not individually rational for f_1. $\square$

THEOREM 8.2. *Suppose X is exhaustive and acyclic, and there exists a firm f whose preferences are not fully substitutable, and there exist at least two firms upstream of f and two firms downstream of f. Then there exist fully substitutable preferences for the firms other than f such that no stable allocation exists.*

PROOF. If the preferences of a firm f are not same-side substitutable, then there exist contracts $x, y, z \in X|_f$ and sets of contracts $Y, Z \subseteq X$ such that $Y_B = \{f\}$ and $Z_S = \{f\}$ such that

$$y \notin C^f_B(Y \mid Z) \text{ but } y \in C^f_B(\{x\} \cup Y \mid Z) \text{ or}$$
$$z \notin C^f_S(Z \mid Y) \text{ but } z \in C^f_S(\{x\} \cup Z \mid Y).$$

Assume the former; the latter case is symmetric. There are two cases.

Case 1: $x_S \neq y_S$. By assumption, there must exist a firm g that is downstream of f and hence downstream of both x_S and y_S. Furthermore, by exhaustivity, there must exist contracts $\hat{x}$ and $\hat{y}$ with $x_S = \hat{x}_S, y_S = \hat{y}_S$ and $\hat{x}_B = \hat{y}_B = g$. Let y_S have preferences such that

$$C^{y_S}(W) = \begin{cases} (W \cap (Y \cup Z))|_{y_S} & \{y, \hat{y}\} \subseteq W \\ (W \cap (Y \cup Z \cup \{\hat{y}\}))|_{y_S} & \text{otherwise} \end{cases}.$$

That is, y_S is willing to accept any and all of the contracts the firm is associated with in $Y \cup Z - \{y\}$, and y_S wants one of y and $\hat{y}$, preferring y, and rejects all other contracts. Let x_S have preferences such that

$$C^{x_S}(W) = \begin{cases} (W \cap (Y \cup Z \cup \{\hat{x}\}))|_{x_S} & \{x, \hat{x}\} \subseteq W \\ (W \cap (\{x\} \cup Y \cup Z \cup \{\hat{x}\}))|_{x_S} & \text{otherwise} \end{cases}.$$

Let g have preferences such that

$$C^g(W) = \begin{cases} (W \cap (Y \cup Z \cup \{\hat{y}\}))|_g & \{\hat{y}, \hat{x}\} \subseteq W \\ (W \cap (Y \cup Z \cup \{\hat{y}, \hat{x}\}))|_g & \text{otherwise} \end{cases}.$$

Let every firm $\tilde{f} \in F - \{x_S, y_S, f, g\}$ have preferences such that

$$C^{\tilde{f}}(W) = (W \cap (Y \cup Z))|_{\tilde{f}}.$$

Consider any allocation A; we will show A can not be stable.

(1) Suppose $A|_f \prec_f C^f(Y \cup Z)$. If A is individually rational, then all other firms want their contracts in $C^f(Y \cup Z)$ irrespective of their other contracts, so $C^f(Y \cup Z)$ blocks A.
(2) Suppose $A|_f = C^f(Y \cup Z)$. Then $\hat{y} \in A$, as otherwise $\{\hat{y}\}$ blocks A. But then $C^f(\{x\} \cup Y \cup Z)$ blocks A.
(3) Suppose $C^f(\{x\} \cup Y \cup Z) \succ_f A\big|_f \succ_f C^f(Y \cup Z)$. In this case, if A is individually rational, then $A \subseteq \{x, \hat{x}, \hat{y}\} \cup Y \cup Z$; then $x \in A$ as otherwise we could not have $A|_f \succ_f C^{\bar{f}}(Y \cup Z)$. But then, $C^f(\{x\} \cup Y \cup Z)$ blocks A.
(4) Suppose $C^f(\{x\} \cup Y \cup Z) = A|_f$. Then if $\hat{y} \in A$, the allocation A is not individually rational for y_S, and if $\hat{x} \in A$, the allocation A is not individually rational for x_S; but this implies that $\{\hat{x}\}$ blocks A.

Case 2: $x_S = y_S \equiv d$. By assumption, there are two firms, g and h, downstream of f, and one firm, e, upstream of f, and hence upstream of g and h. Now consider the contracts v, v', w and w' such that $v_S = w_S = d, v'_S = w'_S = e, v_B = v'_B = g$ and $w_B = w'_B = h$ (which exist as X is exhaustive). Let d have preferences such that

$$C^d(W) = (W \cap ((Y - \{y\}) \cup Z))|_d \cup \tilde{C}^d(W \cap \{x, y, w, v\})$$

where $\tilde{C}^d(\tilde{W})$ is the responsive choice function over $\{x, y, w, v\}$ with quota 2 and underlying preference order $w \succ y \succ x \succ v$. Let e, g, h have preferences such that

$$C^e(W) = \begin{cases} (W \cap (Y \cup Z \cup \{v'\}))|_e & \{v', w'\} \subseteq W \\ (W \cap (Y \cup Z \cup \{v', w'\}))|_e & \text{otherwise} \end{cases}$$

$$C^g(W) = \begin{cases} (W \cap (Y \cup Z \cup \{v\}))|_g & \{v, v'\} \subseteq W \\ (W \cap (Y \cup Z \cup \{v, v'\}))|_g & \text{otherwise} \end{cases}$$

$$C^h(W) = \begin{cases} (W \cap (Y \cup Z \cup \{w'\}))|_h & \{w, w'\} \subseteq W \\ (W \cap (Y \cup Z \cup \{w, w'\}))|_h & \text{otherwise} \end{cases}.$$

Finally, let every firm $\tilde{f} \in F - \{d, e, f, g, h\}$ have preferences such that

$$C^{\tilde{f}}(W) = (W \cap (Y \cup Z))|_{\tilde{f}}.$$

Consider any allocation A; we will show A can not be stable.

(1) Suppose $A|_f \prec_f C^f(Y \cup Z)$. Then $C^f(Y \cup Z)$ blocks A, as all the firms want their contracts in $C^f(Y \cup Z)$, irrespective of other contracts.
(2) Suppose $A|_f = C^f(Y \cup Z)$. Since d does not obtain x or y, he desires both v and w. If A is stable then, $v \in A$. Furthermore, since e does not obtain v', for A to be stable, then $w' \in A$. Hence, if A is stable, $\{w', v\} \subseteq A$ and $w \notin A$. In that case, $C^f(\{x\} \cup Y \cup Z)$ blocks A.
(3) Suppose $C^f(Y \cup Z) \prec_f A\big|_f \prec C^f(\{x\} \cup Y \cup Z)$. Then $x \in A$, so $C^f(\{x\} \cup Y \cup Z) - \{x\}$ blocks A, as d will always choose y and the other firms in Y and Z will always accept offers of any and all contracts in Y.

(4) Suppose $C^f(\{x\} \cup Y \cup Z) = A|_f$. If $v' \notin A$, then $\{v'\}$ blocks A. (Note that if $v \in A$, then $\{x, y, v\} \subseteq A$, and so A is not individually rational for d.) But $v' \in A$ implies that $w' \notin A$. Hence $\{w\}$ blocks A. (Note that $w \notin A$, as then $\{w, x, y\} \subseteq A$, and so A is not individually rational for d.)

If the preferences of a firm f are not cross-side complementary, then there exists contracts $y, z \in X$ and sets of contracts $Y, Z \subseteq X - \{y, z\}$ such that $Y_B = \{f\}$ and $Z_S = \{f\}$ such that

$$y \in C_B^f(\{y\} \cup Y \mid Z) \text{ but } y \notin C_B^f(\{y\} \cup Y \mid \{z\} \cup Z) \text{ or}$$
$$z \in C_S^f(\{z\} \cup Z \mid Y) \text{ but } z \notin C_S^f(\{z\} \cup Z \mid \{y\} \cup Y).$$

Assume the latter; the former case is symmetric. By exhaustivity, there must exist a contract $w \in X$ such that $w_S = y_S$ and $w_B = z_B$. Let y_S have preferences such that

$$C^{y_S}(W) = \begin{cases} (W \cap (Y \cup Z \cup \{w\}))|_{y_S} & \{w, y\} \subseteq W \\ (W \cap (\{y\} \cup Y \cup Z \cup \{w\}))|_{y_S} & \text{otherwise} \end{cases}.$$

Let z_B have preferences such that

$$C^{z_B}(W) = \begin{cases} (W \cap (Y \cup Z \cup \{z\}))|_{z_B} & \{z, w\} \subseteq W \\ (W \cap (Y \cup Z \cup \{z, w\}))|_{z_B} & \text{otherwise} \end{cases}.$$

Finally, let every firm $\tilde{f} \in F - \{f, y_S, z_B\}$ have preferences such that

$$C^{\tilde{f}}(W) = (W \cap (Y \cup Z))|_{\tilde{f}}.$$

Consider any allocation A; we will show A can not be stable.

(1) Suppose $A|_f \prec_f C^f(Y \cup \{z\} \cup Z)$. If A is individually rational, all other firms want their contracts in $C^f(Y \cup \{z\} \cup Z)$ irrespective of their other contracts, so $C^f(Y \cup \{z\} \cup Z)$ blocks A.
(2) Suppose $A|_f = C^f(Y \cup \{z\} \cup Z)$. Then, if A is individually rational, $w \notin A$, as $z \in C^f(Y \cup \{z\} \cup Z)$. But then $C^f(\{y\} \cup Y \cup \{z\} \cup Z)$ blocks A.
(3) Suppose $C^f(\{y\} \cup Y \cup \{z\} \cup Z) \succ_f A|_f \succ_f C^f(Y \cup \{z\} \cup Z)$. In this case, $y \in A$ as otherwise $A|_f$ is available to f from $Y \cup \{z\} \cup Z$, so we could not have $A|_f \succ_f C^f(Y \cup \{z\} \cup Z)$. But then, $C^f(\{y\} \cup Y \cup \{z\} \cup Z)$ blocks A.
(4) Suppose $C^f(\{y\} \cup Y \cup \{z\} \cup Z) = A|_f$. Then if $w \in A$, A is not individually rational for z_B; if $w \notin A$, $\{w\}$ blocks A.

□

THEOREM 8.3. *Suppose that the set of contracts X is acyclic and that all firms' preferences are fully substitutable. Then an allocation A is stable if and only if it is chain stable.*

PROOF. Consider an allocation A that is not stable. If A is not individually rational, then A is not chain stable. Hence, suppose there is a blocking set Z for A. Since X is acyclic, there is an ordering of firms in Z_F such that $x_B \triangleright x_S$ for all $x \in X$. Consider a firm $f \in Z_F$ such that $f \triangleright g$ for all $g \in Z_F$, and consider one contract $y^1 \in Z$ such that $y_B^1 = f$. Now consider y_S^1. By same-side substitutability, $y^1 \in C_B^{y_B^1}\left(\{y^1\} \cup A \mid A\right)$ and $y^1 \in C_S^{y_S^1}\left(\{y^1\} \cup A \mid A \cup Z\right)$. If $y^1 \in C_S^{y_S^1}\left(\{y^1\} \cup A \mid A\right)$, then the chain $\{y^1\}$ is a blocking set and we are done. If not, then there exists a contract $y^2 \in Z$ such that $y^1 \in C_S^{y_S^1}\left(\{y^1\} \cup A \mid A \cup \{y^2\}\right)$ and $y^2 \in C_B^{y_S^1}\left(\{y^2\} \cup A \mid A \cup \{y^1\}\right)$, by same-side substitutability. (If no such y^2 exists, then preferences of y_S^1 are not fully substitutable, as there exists $y^2 \notin C_B^{y_S^1}\left(\{y^2\} \cup A \mid A \cup \{y^1\}\right)$, but $y^2 \in C_B^{y_S^1}\left(Z \cup A \mid A \cup \{y^1\}\right)$.) By an argument analogous to that we provided for y^1, we see that either the chain $\{y^1, y^2\}$ is a blocking set, or there exists $y^3 \in Z$ such that $y^2, y^3 \in C^{y_S^2}\left(\{y^1, y^2, y^3\} \cup A\right)$. Iterating this argument, we find a chain $\{y^1, y^2, \ldots, y^n\}$ blocking A, as Z is a finite set. □

2. The structure of the set of stable matchings

Definition 8.1. *The preferences of f satisfy the Law of Aggregate Supply if for all $Y, Z \subseteq X$ and $Z' \subseteq Z$,*

$$\left|C_S^f(Z \mid Y)\right| - \left|C_S^f(Z' \mid Y)\right| \geq \left|C_B^f(Y \mid Z)\right| - \left|C_B^f(Y \mid Z')\right|.$$

Theorem 8.4. *Suppose that the set of contracts X is acyclic and that all firms' preferences are fully substitutable and satisfy the Laws of Aggregate Demand and Supply. Then, for each firm, the difference between the number of contracts that firm buys and the number of contracts that firm sells is invariant across stable allocations.*

Proof. Consider any stable allocation A associated with the fixed point $\left(X^S, X^B\right)$ and the seller-optimal stable allocation $\hat{A}$ associated with the fixed point $\left(\hat{X}^S, \hat{X}^B\right)$. Consider an arbitrary firm f. We have that

$$\begin{aligned}\left|C_S^f\left(\hat{X}^S \mid \hat{X}^B\right)\right| - \left|C_B^f\left(\hat{X}^B \mid \hat{X}^S\right)\right| &\geq \left|C_S^f\left(X^S \mid \hat{X}^B\right)\right| - \left|C_B^f\left(\hat{X}^B \mid X^S\right)\right| \\ &\geq \left|C_S^f\left(X^S \mid X^B\right)\right| - \left|C_B^f\left(X^B \mid X^S\right)\right|\end{aligned}$$

where the first inequality follows by the Law of Aggregate Supply, as $\hat{X}^S \supseteq X^S$, and the second follows by the law of aggregate demand, as $\hat{X}^B \subseteq X^B$. Hence, the difference between the number of contracts f sells and the number f buys is weakly greater under $\left(\hat{X}^S, \hat{X}^B\right)$ than $\left(X^S, X^B\right)$. However, the sum across all firms of the difference between the number of contracts bought and the number of contracts sold equals 0 . It follows that, for each firm,

$$\left|C_S^f\left(\hat{X}^S \mid \hat{X}^B\right)\right| - \left|C_B^f\left(\hat{X}^B \mid \hat{X}^S\right)\right| = \left|C_S^f\left(X^S \mid X^B\right)\right| - \left|C_B^f\left(X^B \mid X^S\right)\right|.$$

□

Furthermore, the Laws of Aggregate Demand and Supply are the weakest possible conditions that ensure this additional structure on the set of stable allocations.

Theorem 8.5. *Suppose that the set of contracts X is acyclic and exhaustive. If the preferences of some firm f fail to satisfy either the Law of Aggregate Demand or the Law of Aggregate Supply but are fully substitutable, then there exist fully substitutable preferences for the other firms satisfying the Laws of Aggregate Demand and Supply such that there exist two stable allocations across which the difference between the number of contracts f buys and the number of contracts f sells varies.*

Proof. Since the contractual set X is acyclic, there is an ordering $\triangleleft$ on firms such that $x_S \triangleleft x_B$ for all $x \in X$. By symmetry, we need only consider the following three possible cases.

Case 1: There exist contracts $x, y, z \in X$ and sets $Y, Z \subseteq X$ such that $Y_B = \{f\}$ and $Z_S = \{f\}$ such that

$$x, y \in C_S^f(Z \mid Y) \text{ and } x, y \notin C_S^f(\{z\} \cup Z \mid Y).$$

Let every firm $\tilde{f} \in F - \{x_B, y_B, z_B, f\}$ have preferences such that

$$C^{\tilde{f}}(W) = (W \cap (Y \cup Z))|_{\tilde{f}}.$$

There are two subcases to consider: a) We have either $f \triangleleft x_B \triangleleft z_B$ or $f \triangleleft y_B \triangleleft z_B$; assume the former (the latter is symmetric). By exhaustivity of X, there exist contracts $\hat{y}, \hat{z} \in X$ such that

$\hat{y}_B = y_B, \hat{z}_B = z_B$, and $\hat{y}_S = \hat{z}_S = x_B$. Let x_B, y_B, z_B have preferences such that

$$C^{x_B}(W) = \begin{cases} (W \cap (\{x\} \cup Y \cup Z \cup \{\hat{y}\}))|_{x_B} & \{\hat{y}, \hat{z}\} \subseteq W \\ (W \cap (\{x\} \cup Y \cup Z \cup \{\hat{y}, \hat{z}\}))|_{x_B} & \text{otherwise} \end{cases}$$
$$C^{y_B}(W) = \begin{cases} (W \cap (Y \cup Z \cup \{y\}))|_{y_B} & \{y, \hat{y}\} \subseteq W \\ (W \cap (Y \cup Z \cup \{y, \hat{y}\}))|_{y_B} & \text{otherwise} \end{cases}$$
$$C^{z_B}(W) = \begin{cases} (W \cap (Y \cup Z \cup \{\hat{z}\}))|_{z_B} & \{\hat{z}, z\} \subseteq W \\ (W \cap (Y \cup Z \cup \{\hat{z}, z\}))|_{z_B} & \text{otherwise} \end{cases}.$$

There are at least two stable allocations: $C^f(Z \cup Y) \cup \{\hat{z}\}$ and $C^f(\{z\} \cup Z \cup Y) \cup \{\hat{y}\}$. By the cross-side compelementarity of the preferences of f, $\left|C_B^f(Y \mid Z)\right| \geq \left|C_B^f(Y \mid \{z\} \cup Z)\right|$ and the same-side subsitutability of the preferences of f, $\left|C_S^f(Z \mid Y)\right| < \left|C_S^f(\{z\} \cup Z \mid Y)\right|$, and hence the conclusion of the preceding theorem fails for f.

b) We have $f \lhd z_B \lhd x_B, y_B$. By exhaustivity of X, there exists a contract $\hat{x} \in X$ such that $\hat{x}_S = z_B$ and $\hat{x}_B = x_B$. Let x_B, y_B, z_B have preferences such that

$$C^{x_B}(W) = \begin{cases} (W \cap (Y \cup Z \cup \{x\}))|_{x_B} & \{x, \hat{x}\} \subseteq W \\ (W \cap (Y \cup Z \cup \{x, \hat{x}\}))|_{x_B} & \text{otherwise} \end{cases}$$
$$C^{y_B}(W) = (W \cap (Y \cup Z \cup \{y\}))|_{y_B}$$
$$C^{z_B}(W) = \begin{cases} (W \cap (Y \cup Z \cup \{\hat{x}, z\}))|_{z_B} & \{\hat{x}, z\} \subseteq W \\ (W \cap (Y \cup Z))|_{z_B} & \text{otherwise} \end{cases}.$$

There are at least two stable allocations: $C^f(Z \cup Y)$ and $C^f(\{z\} \cup Z \cup Y) \cup \{\hat{x}\}$, and hence (again by the full substitutability of the preferences of f as above) the conclusion of the preceding theorem fails for f.

Case 2: There exist contracts $x, y, z \in X$ and sets $Y, Z \subseteq X$ such that $Y_B = \{f\}$ and $Z_S = \{f\}$ such that

$$x \in C_S^f(Z \mid Y), y \notin C_B^f(Y \mid Z) \text{ and } x \notin C_S^f(\{z\} \cup Z \mid Y), y \in C_B^f(Y \mid \{z\} \cup Z)$$

By exhaustivity of X, there exists a contract $\hat{y}$ such that $\hat{y}_S = y_S$ and $\hat{y}_B = z_B$. Let x_B, y_B, z_B have preferences such that

$$C^{x_B}(W) = (W \cap (\{x\} \cup Y \cup Z))|_{x_B}$$
$$C^{y_B}(W) = \begin{cases} (W \cap (Y \cup Z \cup \{y\}))|_{y_B} & \{y, \hat{y}\} \subseteq W \\ (W \cap (Y \cup Z \cup \{y, \hat{y}\}))|_{y_B} & \text{otherwise} \end{cases}$$
$$C^{z_B}(W) = \begin{cases} (W \cap (Y \cup Z \cup \{\hat{y}\}))|_{z_B} & \{\hat{y}, z\} \subseteq W \\ (W \cap (Y \cup Z \cup \{\hat{y}, z\}))|_{z_B} & \text{otherwise} \end{cases}.$$

Finally, let every firm $\tilde{f} \in F - \{x_B, y_B, z_B, f\}$ have preferences such that

$$C^{\tilde{f}}(W) = (W \cap (Y \cup Z))|_{\tilde{f}}.$$

There are at least two stable allocations: $C^f(Z \cup Y) \cup \{\hat{y}\}$ and $C^f(\{z\} \cup Z \cup Y)$, and hence (again by the full substitutability of the preferences of f as above) the conclusion of the preceding theorem fails for f.

Case 3: There exist contracts $x, y, z \in X$ and sets $Y, Z \subseteq X$ such that $Y_B = \{f\}$ and $Z_S = \{f\}$ such that

$$x, y \notin C_B^f(Y \mid Z) \text{ and } x, y \in C_B^f(Y \mid \{z\} \cup Z).$$

By exhaustivity of X, there exists a contract $\hat{y}$ such that $\hat{y}_S = y_S$ and $\hat{y}_B = z_B$. The same preferences constructed in the previous case again render at least two allocations stable: $C^f(Z \cup Y) \cup \{\hat{y}\}$ and $C^f(\{z\} \cup Z \cup Y)$, and hence the conclusion of the preceding theorem fails for f. □

3. Incentive compatibility and weak Pareto

THEOREM 8.6. *Suppose that the set of contracts X is acyclic and that firms' preferences are fully substitutable and satisfy the Laws of Aggregate Demand and Supply. If additionally, for all $g \in G \subseteq F$, the preferences of g exhibit unit demand, then any mechanism that selects the buyer-optimal stable allocation is (group) strategy-proof for G.*

This follows Hatfield and Kojima (2009) which we have proven already for the one-to-one case. It follows isomorphically.

PROPOSITION 8.1. *Suppose that the set of contracts X is acyclic and that firms' preferences are fully substitutable and satisfy the Laws of Aggregate Demand and Supply. If additionally, for all $g \in G \subseteq F$, the preferences of g exhibit unit demand, then there does not exist an individually rational allocation that every member of G strictly prefers to the buyer-optimal stable allocation.*

PROOF. Suppose by way of contradiction that there is an individually rational allocation $\{y_d\}_{d\in D}$ that each doctor strictly prefers to the doctor-optimal stable allocation. For each $d \in D$, let $\succsim'_d$ be preferences that declare y_d as the unique acceptable contract. It is easy to show that $\{y_d\}_{d\in D}$ is the unique stable allocation when each d declares $\succsim'_d$. But this contradicts the incentive compatibility theorem. □

Part 3

Matching with transfers

CHAPTER 9

Coalitional games, assignment games and the core

1. Coalitional games with transferable utility

Definition 9.1. *A coalitional game with transferable utility (TU game) is a pair $(N;v)$ where*

- *$N = \{1, 2, \ldots, n\}$ is a finite set of players. A subset of N is called a coalition. The collection of all the coalitions is denoted by 2^N.*
- *$v : 2^N \to \mathbb{R}$ is a function associating every coalition S with a real number $v(S)$, satisfying $v(\emptyset) = 0$. This function is called the coalitional function of the game.*

We call the real number $v(S)$ the *worth* of the coalition S.

Definition 9.2. *A coalitional game $(N;v)$ is superadditive if for any pair of disjoint coalitions S and T,*

$$v(S) + v(T) \leq v(S \cup T)$$

Definition 9.3. *A coalitional game $(N;v)$ is monotonic if for any pair of coalitions S and T, such that $S \subseteq T$,*

$$v(S) \leq v(T).$$

Definition 9.4. *Let $\mathcal{U}$ be a family of coalitional games (over any set of players). A solution concept (over $\mathcal{U}$) is a function φ associating every game $(N;v) \in \mathcal{U}$ with a subset $\varphi(N;v)$ of $\mathbb{R}^N$. A solution concept is called a point solution iffor every coalitional game $(N;v) \in \mathcal{U}$, the set $\varphi(N;v)$ contains only one element.*

Definition 9.5. *A coalitional structure is a partition $\mathcal{B}$ of the set of players N.*

Definition 9.6. *Let $(N;v)$ be a game, and let $\mathcal{B}$ be a coalitional structure. A vector $x \in \mathbb{R}^N$ is called efficient* [8] *for the coalitional structure $\mathcal{B}$ if for every coalition $S \in \mathcal{B}$,*

$$x(S) = v(S).$$

A vector x is called individually rational if for every player $i \in N$,

$$x_i \geq v(i).$$

Definition 9.7. *Let $(N;v)$ be a coalitional game, and let $\mathcal{B}$ be a coalitional structure. An imputation for the coalitional structure $\mathcal{B}$ is a vector $x \in \mathbb{R}^N$ that is efficient for the coalitional structure $\mathcal{B}$, and individually rational. The set of all imputations for the coalitional structure $\mathcal{B}$ is denoted by $X(\mathcal{B};v)$.*

2. The core

Definition 9.8. *Let $(N;v)$ be a coalitional game. An imputation $x \in X(N;v)$ is coalitionally rational if for every coalition $S \subseteq N$*

$$x(S) \geq v(S).$$

Definition 9.9. *The core of a coalitional game $(N;v)$, denoted by $\mathcal{C}(N;v)$, is the collection of all coalitionally rational imputations,*

$$\mathcal{C}(N;v) := \{x \in X(N;v) : x(S) \geq v(S), \quad \forall S \subseteq N\}.$$

Theorem 9.1. *The core of a coalitional game is the intersection of a finite number of half-spaces, and is therefore a convex set. In addition, the core is a compact set.*

PROOF. For each coalition S the set $\left\{x \in \mathbb{R}^N : x(S) \geq v(S)\right\}$ is a closed half-space. The core is the intersection of $2^n - 1$ half-spaces $\left\{x \in \mathbb{R}^N : x(S) \geq v(S)\right\}$, for all $\emptyset \neq S \subseteq N$, and the half-spaces $\left\{x \in \mathbb{R}^N : x(N) \leq v(N)\right\}$.

Every half-space is convex, and the intersection of convex sets is convex. The core is therefore a convex set. Every half-space is closed, and the intersection of closed sets is closed. The core is therefore a closed set. Finally, since the core is a subset of $X(N;v)$, it is bounded. A closed and bounded set is compact. □

DEFINITION 9.10. *A collection of coalitions $\mathcal{D}$ is a balanced collection if there exists a vector of positive numbers $(\delta_S)_{S\in\mathcal{D}}$ such that*

$$\sum_{\{S\in\mathcal{D}:i\in S\}} \delta_S = 1, \quad \forall i \in N$$

Denote by $\mathcal{P}(N) := \{S \subseteq N, S \neq \emptyset\}$ the collection of nonempty coalitions. Denote by P the collection of all weights weakly balancing $\mathcal{P}(N)$,

$$P := \left\{\delta = (\delta_S)_{S\in\mathcal{P}(N)} : \delta_S \geq 0 \quad \forall S \in \mathcal{P}(N), \sum_{S\in\mathcal{P}(N)} \delta_S \chi^S = \chi^N\right\}$$

This set is a polytope in the space $\mathbb{R}^{2^n-1}$, and is nonempty: it contains, for example, the vector δ in which $\delta_{\{i\}} = 1$ for every $i \in N$, and $\delta_S = 0$ for every S containing at least two players.

THEOREM 9.2 (The Bondareva-Shapley Theorem). *A necessary and sufficient condition for the nonemptiness of the core of a coalitional game $(N;v)$ is that*

$$v(N) \geq \sum_{S\in\mathcal{P}(N)} \delta_S v(S), \quad \forall \delta = (\delta_S)_{S\in\mathcal{P}(N)} \in P$$

We can prove this using linear programming.

PROOF. **Step 1**: Consider the following linear program with the variables $(\delta_S)_{S\in\mathcal{P}(N)}$.

$$\begin{array}{ll} \text{Compute:} & Z_P := \max \sum_{S\in\mathcal{P}(N)} \delta_S v(S), \\ \text{subject to:} & \sum_{\{S:i\in S\}} \delta_S = 1, \quad \forall i \in N, \\ & \delta_S \geq 0, \quad \forall S \in \mathcal{P}(N). \end{array}$$

The set of feasible solutions of this linear program is the set P, the collection of all weights weakly balancing $\mathcal{P}(N)$. Because P is compact and nonempty, Z_P is finite.

Step 2: The dual problem is the following problem with the variables $(x_i)_{i\in N}$ (verify!).

$$\begin{array}{ll} \text{Compute:} & Z_D := \min x(N), \\ \text{subject to:} & x(S) \geq v(S), \quad \forall S \in \mathcal{P}(N). \end{array}$$

As already shown Z_P is finite, and therefore the duality theorem implies that Z_D is also finite, and equals Z_P.

Step 3: Let x be a vector in the core. Then $x(S) \geq v(S)$ for every coalition S, and therefore x satisfies all the constraints of the dual problem. The value of the objective function at x is $x(N) = v(N)$; hence $Z_D \leq v(N)$.

Step 4: Let x be a feasible solution of the dual problem at which the minimum is attained, i.e., $x(N) = Z_D$. Since x satisfies the constraints of the dual problem, it is coalitionally rational. We show that $x(N) = v(N)$. Since $Z_D \leq v(N)$, it follows that $x(N) = \sum_{i\in N} x_i = Z_D \leq v(N)$. For $S = N$, the constraint $x(S) \geq v(S)$ is $x(N) \geq v(N)$, so that we deduce that $x(N) = v(N)$. It follows that x is in the core, and therefore the core is not empty.

Step 5: $Z_P \leq v(N)$ if and only if $\sum_{\{S\in\mathcal{P}(N):i\in S\}} \delta_S v(S) \leq v(N)$ for every feasible solution $\delta = (\delta_S)_{S\in\mathcal{P}(N)}$, i.e., if and only if the Bondareva-Shapley Theorem holds. □

3. Assignment games

We can model assignment games as coalitional games. Consider producers $N = \{1, 2, \ldots, n\}$ who trade $L = \{1, 2, \ldots, l\}$ commodities. A vector of commodities (often called a *bundle*) is denoted by $x = (x_j)_{j=1}^{L} \in \mathbb{R}_{+}^{L}$. The bundle of producer i will be denoted by x_i, and the quantity of commodity j in this bundle will be denoted by $x_{i,j}$. Each producer has a "production technology" represented by a production function $u_i : \mathbb{R}_{+}^{L} \to \mathbb{R}$: if $x_i \in \mathbb{R}_{+}^{L}$ is the bundle of commodities owned by producer i, then that producer can produce the sum of money $u_i(x_i)$.

Every producer i has an initial endowment that is a bundle $a_i \in \mathbb{R}_{+}^{L}$ of goods, and the producers can trade goods between each other.

If the coalition S is formed, the total bundle of goods available to the coalition is $a(S) := \sum_{i \in S} a_i \in \mathbb{R}_{+}^{L}$. The coalition can allocate to each of its members a bundle $x_i \in \mathbb{R}_{+}^{L}$, subject to the constraint

$$x(S) = \sum_{i \in S} x_i = \sum_{i \in S} a_i = a(S).$$

Definition 9.11. *An allocation for a coalition S is a collection of bundles of commodities $(x_i)_{i \in S}$, where $x_i \in \mathbb{R}_{+}^{L}$ for every producer $i \in N$, satisfying $x(S) = a(S)$.*

Theorem 9.3. *For every coalition S, the set X^S is compact.*

Easy to prove: X^S is bounded because the total quantity of commodities is bounded; X^S is closed because every half-space is closed.

Every market can be associated with a coalitional game, in which the set of players is the set of producers $N = \{1, 2, \ldots, n\}$, and the worth of each nonempty coalition $S \subseteq N$ is

$$v(S) = \max\left\{\sum_{i \in S} u_i(x_i) : x = (x_i)_{i \in S} \in X^S\right\}$$

Theorem 9.4. *If for every $i \in N$ the production function u_i is continuous, the maximum of*

$$v(S) = \max\left\{\sum_{i \in S} u_i(x_i) : x = (x_i)_{i \in S} \in X^S\right\}$$

is attained for every coalition S.

Proof. Since all production functions $(u_i)_{i \in N}$ are continuous, the function $\sum_{i \in S} u_i$, as the sum of a finite number of continuous functions, is also a continuous function. Since the maximum of a continuous function over a compact set is always attained, and since the set X^S is compact we deduce that the maximum is attained. □

Theorem 9.5 (Shapley and Shubik, 1969). *The core of a market game is nonempty.*

We prove this using the Bondareva-Shapley Theorem.

Proof. The proof of this theorem relies on the Bondareva-Shapley Theorem; we will prove that every market game is a balanced game. For every coalition S, choose an imputation $x^S = \left(x_i^S\right)_{i \in S} \in \mathbb{R}_{+}^{S \times L}$ at which the maximum of $v(S)$ is attained. Then:

- $x_i^S \in \mathbb{R}_{+}^{L}$ for every player i,
- $x^S(S) = \sum_{i \in S} x_i^S = a(S)$, and
- $\sum_{i \in S} u_i\left(x_i^S\right) = v(S)$.

Let $\delta = (\delta_S)_{S \in \mathcal{P}(N)} \in P$ be a vector that weakly balances $\mathcal{P}(N)$. We need to show that

$$v(N) \geq \sum_{S \in \mathcal{P}(N)} \delta_S v(S).$$

For every $i \in N$ denote

$$z_i := \sum_{\{S \in \mathcal{P}(N) : i \in S\}} \delta_S x_i^S \in \mathbb{R}_{+}^{L}.$$

We first show that $z = (z_i)_{i\in N}$ is a feasible bundle, i.e., that $z(N) = a(N)$. Note that by the definition of z_i,

$$z(N) = \sum_{i\in N} z_i = \sum_{i\in N} \sum_{\{S\in\mathcal{P}(N):i\in S\}} \delta_S x_i^S$$

By changing the order of summation,

$$z(N) = \sum_{S\in\mathcal{P}(N)} \sum_{i\in S} \delta_S x_i^S = \sum_{S\in\mathcal{P}(N)} \left(\delta_S \sum_{i\in S} x_i^S\right) = \sum_{S\in\mathcal{P}(N)} \delta_S x^S(S).$$

Since $x^S(S) = a(S)$, by changing the order of summation, one has

$$z(N) = \sum_{S\in\mathcal{P}(N)} \delta_S a(S) = \sum_{S\in\mathcal{P}(N)} \left(\delta_S \sum_{i\in S} a_i\right) = \sum_{i\in N} \left(a_i \sum_{\{S\in\mathcal{P}(N):i\in S\}} \delta_S\right).$$

Since δ is a vector of balancing weights, $\sum_{\{S\in\mathcal{P}(N):i\in S\}} \delta_S = 1$ for every player $i \in N$, and therefore

$$z(N) = \sum_{i\in N} a_i = a(N)$$

that is, z is indeed a feasible bundle. By this, and from the definition of the function v we deduce that

$$v(N) \geq \sum_{i\in N} u_i(z_i)$$

Then we substitute for z_i:

$$\begin{aligned}
v(N) &\geq \sum_{i\in N} u_i(z_i) \\
&= \sum_{i\in N} u_i \left(\sum_{\{S\in\mathcal{P}(N):i\in S\}} \delta_S x_i^S\right) \\
&\geq \sum_{i\in N} \sum_{\{S\in\mathcal{P}(N):i\in S\}} \delta_S u_i\left(x_i^S\right) \\
&= \sum_{S\in\mathcal{P}(N)} \sum_{i\in S} \delta_S u_i\left(x_i^S\right) \\
&= \sum_{S\in\mathcal{P}(N)} \left(\delta_S \sum_{i\in S} u_i\left(x_i^S\right)\right) \\
&= \sum_{S\in\mathcal{P}(N)} \delta_S v(S),
\end{aligned}$$

where we use the concavity of the functions $(u_i)_{i\in N}$, and that $v(S) = \sum_{i\in N} u_i\left(x_i^S\right)$. It follows that the game is balanced, and therefore the core is nonempty. □

4. More on linear programming

4.1. Linear Programming. Linear programming problems constitute the special cases of constrained maximization problems for which both the constraints and the objective function are linear in the variables $(x_1, \ldots, x_N)$. A general linear programming problem is typically written in the form

$$\begin{aligned}
\text{Max}_{(x_1,\ldots,x_N)\geq 0} \qquad & f_1x_1 + \cdots + f_Nx_N \\
\text{s.t. } & a_{11}x_1 + \cdots + a_{1N}x_N \leq c_1 \\
& \vdots \\
& a_{K1}x_1 + \cdots + a_{KN}x_N \leq c_K,
\end{aligned}$$

or, in matrix notation,

$$\begin{aligned} \mathrm{Max}_{x\in\mathbb{R}^{\mathrm{N}}_{+}} \quad & f\cdot x \\ \text{s.t. } & Ax \le c, \end{aligned}$$

where A is the $K \times N$ matrix with generic entry a_{kn}, and $f \in \mathbb{R}^N, x \in \mathbb{R}^N$, and $c \in \mathbb{R}^K$ are (column) vectors.

A most interesting fact about the linear programming problem (M.M.1) is that with it we can associate another linear programming problem, called the dual problem, that has the form of a minimization problem with K variables (one for each constraint of the original, or primal, problem) and N constraints (one for each variable of the primal problem):

$$\begin{aligned} &\mathrm{Min}_{(\lambda_1,\ldots,\lambda_K)\ge 0}\, c_1\lambda_1 + \cdots + c_K\lambda_K \\ &\text{s.t. } a_{11}\lambda_1 + \cdots + a_{K1}\lambda_k \ge f_1 \\ &\vdots \\ &a_{1N}\lambda_1 + \cdots + a_{KN}\lambda_K \ge f_N, \end{aligned}$$

or, in matrix notation,

$$\begin{aligned} \mathrm{Max}_{\lambda\in\mathbf{R}}\, k \quad & c\cdot i \\ \text{s.t. } & A^\top\lambda \ge f, \end{aligned}$$

where $\lambda \in \mathbb{P}^K$ is a column vector.

Suppose that $x \in \mathbb{R}^N_+$ and $\lambda \in \mathbb{R}^K_+$ satisfy, respectively, the constraints of the primal and the dual problems. Then

$$f\cdot x \le \left(A^\top\lambda\right)\cdot x = \lambda\cdot(Ax) \le \lambda\cdot c = c\cdot\lambda.$$

Thus, the solution value to the primal problem can be no larger than the solution value to the dual problem. The duality theorem of linear programming, now to be stated, says that these values are actually equal. The key for an understanding of this fact is that, as the notation suggests, the dual variables $(\lambda_1,\ldots,\lambda_K)$ have the interpretation of Lagrange multipliers.

Theorem 9.6. *Suppose that the primal problem attains a maximum value $v \in \mathbb{R}$. Then v is also the minimum value attained by the dual problem.*

Proof. Let $\bar{x} \in \mathbb{R}^N$ be a maximizer vector. Denote by $\bar\lambda = \left(\bar\lambda_1,\ldots,\bar\lambda_K\right) \ge 0$ the Lagrange multipliers associated with this problem. Formally, we regard $\bar\lambda$ as a column vector. Then, applying we have

$$A^\top\bar\lambda = f \quad \text{ and } \quad \bar\lambda\cdot(c - A\bar{x}) = 0$$

Hence, $\bar\lambda$ satisfies the constraints of the dual problem (since $A^\top\bar\lambda \ge f$) and

$$c\cdot\bar\lambda = \bar\lambda\cdot c = \bar\lambda\cdot A\bar{x} = \left(A^{\mathrm{T}}\bar\lambda\right)\cdot\bar{x} = f\cdot\bar{x}$$

We know that $c\cdot\lambda \ge f\cdot\bar{x}$ for all $\lambda \in \mathbb{R}^K_+$ such that $A^\top\lambda \ge f$. Therefore $c\cdot\bar\lambda \le c\cdot\lambda$ if $A^\top\lambda \ge f$. So $\bar\lambda$ solves the dual problem and therefore the value of the dual problem, $c\cdot\bar\lambda$, equals $f\cdot\bar{x}$, the value of the primal problem.

$\square$

Part 4

Object allocation

CHAPTER 10

House allocation and exchange

There are a finite set of agents A and finite set of objects S. Define an assignment analogously to a one-to-one matching. Only the agents are strategic actors in this model so we say "an object is assigned to an agent." We don't care for object preferences, but they do have priority.

Consider an initial assignment or endowment of agents to objects $\mu : A \cup S \to A \cup S \cup \{\emptyset\}$. Each agent has strict preferences $\succ_a$ over all objects $S \cup \{\emptyset\}$, including her own assignment. Agent a prefers the assignment μ to assignment ν if and only if she prefers $\mu(a)$ to $\nu(a)$. There is a fixed or random priority list π over all agents A.

1. House allocation

All agents and goods are initially unassigned: $\mu(a) = \emptyset$ for all $a \in A$. A (direct assignment) mechanism ϕ takes in the initial assignment, priority list, and the preferences of the agents and produces a new assignment.

Definition 10.1. *A assignment is individually rational if no agent strictly prefers his endowment to his assignment.*

We define the normatively appealing properties with respect to the agents.

Definition 10.2. *A assignment is Pareto-efficient if there is no other assignment that makes all agents weakly better off and at least one agent strictly better off.*

Definition 10.3. *A mechanism is (group) strategy-proof no (group of) agent(s) can misrepresent their preferences and obtain a strictly better assignment under the mechanism.*

Definition 10.4 (Serial Dictatorship). *The Serial Dictatorship mechanism is defined as follows:*

- *If there is no priority list, define a priority list π over agents.*
- *At Step t, pick an agent withe t^{th} priority according to π and assign to them their most preferred remaining object (or leave them unassigned). Remove that object.*

The algorithm stops after the agent with the lowest priority has been assigned her object.

Theorem 10.1. *Serial dictatorship is individually rational, Pareto-efficient, and group-strategyproof. Any Pareto-efficient assignment can be achieved by a serial dictatorship.*

2. House exchange

Let us assume that $|A| = |S|$ and every agent is initially assigned an object: $\mu(a) \neq \emptyset$ for all $a \in A$. Individual rationality is now less of a given. In order to improve the assignment, agents are going to have to get together and trade their houses.

Develop notion of the core—the set of undominated assignments.

Definition 10.5. *An assignment μ is in the core if there is no set of agents $C \subseteq A$ and analogously assignment v such that*

- *for any $a \in C, v(a)$ is the initially assigned object of some $b \in C$ (agents in C only trade objects among themselves).*
- *$v(a) \succeq_a \mu(a)$ for all $a \in C$ and $v(a) \succ_a \mu(a)$ for some $a \in C$ (no agents in C are worse off and at least one agent is better off).*

Any assignment in the core is individually rational and Pareto-efficient.

THEOREM 10.2 (Shapley and Shubik, 1974). *The core of any house exchange is non-empty.*

Like serial dictatorship for house allocation, we define a mechanism that facilitates results.

DEFINITION 10.6 (Top Trading Cycles). *Define the Top Trading Cycles (TTC) algorithm as follows: In Step* $t \geq 1$

- *For each agent, we draw an arrow starting from that agent to the agent who owns her most preferred object among the objects that have not been removed. If an agent's most preferred object is already hers, then the arrow goes to hersel (we have a self-cycle).*
- *There must be at least one cycle: if we start from an agent the cycle and we follow the arrows, we eventually visit that agent again.*
- *Each agent in a cycle is assigned the object owned by the individual she is pointing to. Agents in a cycle and the object they are assigned to are removed from the problem.*

The algorithm stops when all agents have been removed or there are no acceptable objects left for any agent who has not yet been removed.

THEOREM 10.3 (Roth and Postlewaite, 1977). *The core of any house exchange is unique.*

We can prove both of the above theorems in one go.

PROOF. Consider a sequence of disjoint set of agents assigned in k steps of the TTC algorithm: $A_1, A_2, \ldots, A_k$.

If an assignment is in the core, then all agents in A_1 must get the object they are assigned by the TTC algorithm. Those who are not could block. Then all agents in A_2 can only be made better off by getting an object assigned in the first step, which is impossible because the agents from A_1 would block. So the best that agents in A_2 can do in a core outcome is to get the object assigned to them by the TTC. Repeat this for the remaining $3, \ldots, k$ steps of TTC to grant that TTC is the unique core assignment. $\square$

THEOREM 10.4 (Ma, 1994). *For house exchange, a mechanism is individually rational, Pareto-efficient, and strategyproof if and only if it is TTC.*

CHAPTER 11

School choice

1. Algorithms

There are a disjoint sets of schools S with quota q_s and students A. Assume that $\mu(a) = \emptyset$ for $a \in A$. Each school has quota q_s and its priority list π_s.

A notion of fairness is important:

Definition 11.1. *An assignment eliminates justified envy if the corresponding matching is stable.*

We present some commonplace algorithms and discuss their properties.

1.1. Boston algorithm.

Definition 11.2 (Boston Algorithm). *The Boston Algorithm is defined as follows: Step $t = 1$:*

- *Each student "proposes" to her first-choice school.*
- *Each school immediately accepts the highest-priority proposing students up to its quota and rejects all other students.*

Step $t \geq 2$:

- *Each student rejected in Step $t - 1$ applies to her next highest choice.*
- *Each school immediately accepts the highest-priority students proposing in Step t up to its remaining capacity (= its quota minus the sum of admitted students from previous steps).*

The algorithm terminates when no more proposals are made by the students.

Proposition 11.1. *The Boston mechanism is neither stable nor strategyproof.*

1.2. Deferred acceptance.

Definition 11.3 (Deferred Acceptance). *Consider finite sets H and D with agents $h \in H$ and $d \in D$. The (doctor-proposing) deferred acceptance algorithm is as follows:*

1. $\forall d \in D : d$ proposes to its most preferred firm to whom she hasn't already proposed.
2. $\forall h \in H : h$ tentatively accepts its best proposal and rejects the rest.
3. $\forall d \in D$: If d is rejected, she proposes to the next h on its list.
Repeat steps 2 and 3 until there are no more proposals. Return the matching and call it μ.

We already covered that DA is stable and optimal for the proposing side.

1.3. Top trading cycles.

Definition 11.4 (Top Trading Cycles). *Define the Top Trading Cycles (TTC) algorithm as follows: In Step $t \geq 1$*

- *For each agent, we draw an arrow starting from that agent to the agent who owns her most preferred object among the objects that have not been removed. If an agent's most preferred object is already hers, then the arrow goes to hersel (we have a self-cycle).*
- *There must be at least one cycle: if we start from an agent the cycle and we follow the arrows, we eventually visit that agent again.*
- *Each agent in a cycle is assigned the object owned by the individual she is pointing to. Agents in a cycle and the object they are assigned to are removed from the problem.*

The algorithm stops when all agents have been removed or there are no acceptable objects left for any agent who has not yet been removed.

PROPOSITION 11.2. *TTC is Pareto efficient and strategyproof.*

With these algorithms in mind, let's assess some facts.

PROPOSITION 11.3. *Student-optimal DA algorithm is not Pareto-efficient.*

PROOF. Example from Roth (1982). Consider three students i_1, i_2, i_3, and three schools s_1, s_2, s_3, each of which has only one seat. The priorities of schools and the preferences of students are as follows:

$$\begin{array}{llll} s_1 : i_1 - i_3 - i_2 & i_1 : s_2 & s_1 & s_3 \\ s_2 : i_2 - i_1 - i_3 & i_2 : s_1 & s_2 & s_3 \\ s_3 : i_2 - i_1 - i_3 & i_3 : s_1 & s_2 & s_3. \end{array}$$

Let us interpret the school priorities as school preferences and consider the associated college admissions problem. In this case there is only one stable matching:

$$\begin{pmatrix} i_1 & i_2 & i_3 \\ s_1 & s_2 & s_3 \end{pmatrix}.$$

But this matching is Pareto-dominated by

$$\begin{pmatrix} i_1 & i_2 & i_3 \\ s_2 & s_1 & s_3 \end{pmatrix}.$$

Here agents i_1 and i_2 have the highest priorities for schools s_1 and s_2, respectively. So there is no way student i_1 can be assigned a school that is worse than school s_1 and hence she shall be assigned either s_2 or s_1. Similarly there is no way student i_2 can be assigned a school that is worse than school s_2 and hence she shall be assigned either s_1 or s_2. Thus students i_1 and i_2 should share schools s_1 and s_2 among themselves. Stability forces them to share these schools in a Pareto-inefficient way: This is because if students i_1 and i_2 are assigned schools s_2 and s_1 respectively, then we have a situation where student i_3 prefers school s_1 to her assignment s_3 and she has a higher priority for school s_2 than student i_2 does.

We see complete elimination of justified envy may conflict with Pareto efficiency. □

PROPOSITION 11.4. *Top Trading Cycles algorithm does not eliminate justified envy.*

PROPOSITION 11.5. *For general school priorities, it may not be possible to eliminate justified envy in a Pareto-efficient assignment.*

The last of which disappears if school priorities are sufficiently correlated (Ergin, 2002).

THEOREM 11.1 (Kesten, 2010). *There is no Pareto-efficient and strategyproof mechanism that Pareto dominates the student-optimal DA algorithm.*

THEOREM 11.2 (Kesten, 2010). *There is no Pareto-efficient and strategyproof mechanism that selects the Pareto-efficient matching which eliminates justified envy whenever it exists.*

THEOREM 11.3 (Kesten, 2010). *Given any set of schools S and capacity vector $q = (q_s)_{s \in S}$, there always exist a set of students A and a school choice problem $\left(S, (\succ_a)_{a \in A}\right)$ for which student-optimal DA algorithm assigns each student at either his or her worst choice or his or her next worst choice.*

To absolve the inefficiency, Ergin and Erdil (2006) propose running DA and then TTC.

CHAPTER 12

Complex constraints

We often deal with complex constraints in matching markets. Scarf's lemma is useful to find stable matches especially in the presence of complementarities.

1. Scarf's lemma

Let $\mathcal{A}$ be an $n \times m$ non-negative matrix with at least one non-zero entry in each row and $q \in \mathbb{R}_+^n$. Associated with each row $i \in \{1, .., n\}$ of $\mathcal{A}$ is a strict order $\succ_i$ over the set of columns j for which $\mathcal{A}_{i,j} > 0$. We can associate each row with an agent and each column to the characteristic vector of a coalition of agents. So $\mathcal{A}_{ij} = 1$ means that agent i is in the j^{th} coalition. Then, $\succ_i$ can be interpreted as agent i 's preference ordering over all the columns/coalitions of $\mathcal{A}$ of which i is a member. This can be easiest seen by an example. Consider

$$\mathcal{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Here, the rows represent agents and the columns represent coalitions.

- Agent 1 (Row 1): The non-zero entries are in the 1st and 3rd columns, so we can define a strict order, say, $1 \succ_1 3$ (i.e., agent 1 prefers the 1st coalition over the 3rd).
- Agent 2 (Row 2): The non-zero entries are in the 2nd and 3rd columns, so we can define a strict order, say, $2 \succ_2 3$ (i.e., agent 2 prefers the 2nd coalition over the 3rd).
- Agent 3 (Row 3): The non-zero entries are in the 1st and 2nd columns, so we can define a strict order, say, $2 \succ_3 1$ (i.e., agent 3 prefers the 2nd coalition over the 1st).

Now consider the system $\left\{x \in \mathbb{R}_+^m : \mathcal{A}x \leq q\right\}$. The positive elements of a feasible x are the feasible coalitions that can be formed. We now present a definition of dominance.

Definition 12.1. *A vector $x \geq 0$ satisfying $\mathcal{A}x \leq q$ dominates column k of $\mathcal{A}$ if there exists a row i such that $\sum_{j=1}^{n} \mathcal{A}_{ij}x_j = q_i$ and for all column $l \in \{1, .., m\}$ such that $\mathcal{A}_{i,l} > 0$ and $x_l > 0, l \succeq_i k$.*

We illustrate the above with an example.

Example 2. Consider an instance of the stable matching problem that consists of two hospitals (h_1, h_2), each with capacity 1, and two single doctors (d_1, d_2). This is the setting of Gale and Shapley 1962. The preferences are as follows:

$$d_1 \succ_{h_1} d_2; d_1 \succ_{h_2} d_2; h_2 \succ_{d_1} h_1; h_2 \succ_{d_2} h_1.$$

We can describe the set of feasible matchings as the solution to a system of inequalities. Introduce $x_{(d_i,h_j)} \in \{0, 1\}$ for $i \in \{1, 2\}$ and $j \in \{1, 2\}$ where $x_{(d_i,h_j)} = 1$ if and only if d_i is assigned to h_j and zero otherwise. In the 4×4 matrix $\mathcal{A}$ each row corresponds to an agent (a hospital or a doctor), and each column corresponds to a doctor-hospital pair. In this example $q = \mathbf{1}$.

$$\begin{array}{c} \\ h_1 \\ h_2 \\ d_1 \\ d_2 \end{array}\left(\begin{array}{cccc} (d_1,h_1) & (d_1,h_2) & (d_2,h_1) & (d_2,h_2) \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right) \cdot x \leq \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} ; \text{ order } : \begin{array}{c} \mathrm{col}_1 \succ_{h_1} \mathrm{col}_3 \\ \mathrm{col}_2 \succ_{h_2} \mathrm{col}_4 \\ \mathrm{col}_2 \succ_{d_1} \mathrm{col}_1 \\ \mathrm{col}_3 \succ_{d_2} \mathrm{col}_4 \,. \end{array}$$

An entry $\mathcal{A}_{ij}$ of the matrix $\mathcal{A}$ is 1 if and only if the agent corresponding to row i is a member of the coalition corresponding to column j. Otherwise, $\mathcal{A}_{ij} = 0$. We model capacity constraints with $\mathcal{A}x \leq q$ and respecting that each doctor can be assigned to at most one hospital. For each row i of $\mathcal{A}$, the strict order on the set of columns j for which $\mathcal{A}_{ij} \neq 0$ is the same as the preference ordering of agent i.

Now leveraging our definition of dominance: Every integer solution to $\mathcal{A}x \leq 1$ corresponds to a matching and vice versa. Notice $x = (1,0,0,1)^T$ corresponds to the matching $(d_1,h_1);(d_2,h_2)$. It is not stable because it is blocked by (d_1,h_2). The solution $x = (0,1,1,0)^T$ is a dominating solution and corresponds to a stable matching.

Theorem 12.1 (Scarf's Lemma). *Let $\mathcal{A}$ be an $n \times m$ non-negative matrix and $q \in \mathbb{R}^n_+$. Then, there exists an extreme point of $\left\{x \in \mathbb{R}^m_+ : \mathcal{A}x \leq q\right\}$ that dominates every column of $\mathcal{A}$.*

Proof. The proof is by reduction to the existence of a Nash equilibrium in a two-person game. By scaling we can assume that q is the vector of all 1's, denoted $\mathbf{1}$. Let $\mathcal{C}$ be a matrix defined as follows:

- If $\mathcal{A}_{ij} = 0$, then $\mathcal{C}_{ij} = 0$.
- If $\mathcal{A}_{ij} > 0$ and j is ranked t-th in the preference list of i, then $\mathcal{C}_{ij} = -N^t$, for $N \geq n$.

The j^{th} columns of $\mathcal{A}$ and $\mathcal{C}$ will be denoted $\mathcal{A}^j$ and $\mathcal{C}^j$ respectively.

We associate a 2-person game with the pair $(\mathcal{A},\mathcal{C})$. The payoff matrix for ROW will be $\mathcal{A}$. The payoff matrix for COLUMN will be $\mathcal{C}$. Let (x^*,y^*) be an equilibrium pair of mixed strategies for this game where x^* is a mixed strategy for ROW player (payoff matrix $\mathcal{A}$) and y^* is the mixed strategy for COLUMN (payoff matrix $\mathcal{C}$).

Let R^* be ROW's expected payoff and C^* be the expected payoff to COLUMN. Clearly $\mathcal{A}x^* \leq R^* 1$. We show that $\frac{x^*}{R^*}$ is a dominating solution.

Suppose the columns of $\mathcal{A}$ are sorted so that $x^*_i > 0$ for $1 \leq i \leq k$ and $x^*_i = 0$ for $k+1 \leq i \leq n$. COLUMN's expected payoff when playing each column $1,\ldots,k$ is exactly C^*, and expected payoff when playing any other column is at most C^*. If COLUMN plays one of the first k columns uniformly at random, her expected payoff is also C^*, i.e.,

$$\frac{1}{k} y^* \cdot (\mathcal{C}^1 + \ldots + \mathcal{C}^k) = \sum_i y^*_i \frac{(\mathcal{C}_{i1} + \mathcal{C}_{i2} + \ldots + \mathcal{C}_{ik})}{k} = C^*$$

Choose any column j. We show that x^* dominates j. As x^* is a best response to y^* it must be that

$$\sum_i y^*_i \mathcal{C}_{ij} \leq C^* = \sum_i y^*_i \frac{(\mathcal{C}_{i1} + \mathcal{C}_{i2} + \ldots + \mathcal{C}_{ik})}{k}.$$

$$\Rightarrow \exists y^*_r > 0 \text{ s.t. } \mathcal{C}_{rj} \leq \frac{(\mathcal{C}_{r1} + \mathcal{C}_{r2} + \ldots + \mathcal{C}_{rk})}{k}.$$

Hence,

$$y^*_r > 0 \Rightarrow \mathcal{A}_{r1} x^*_1 + \ldots + \mathcal{A}_{rk} x^*_k = R^* > 0.$$

Therefore, at least one of $\mathcal{A}_{r1},\ldots,\mathcal{A}_{rk}$ is non-zero. Hence, at least one of $\mathcal{C}_{r1},\ldots,\mathcal{C}_{rk}$ is non-zero and $\mathcal{C}_{ij} \neq 0 \Rightarrow \mathcal{A}_{ij} \neq 0$.

Assume that among the columns $1,\ldots,k$, that $\mathcal{A}_{r1} \neq 0$ and column 1 is the least preferred. Let ℓ be the rank of column 1 in that preference ordering. Recall $\mathcal{C}_{r1} = -N^\ell$. Hence,

$$\frac{(\mathcal{C}_{r1} + \mathcal{C}_{r2} + \ldots + \mathcal{C}_{rk})}{k} \leq \frac{\mathcal{C}_{r1}}{k} = \frac{-N^l}{k} < -N^{l-1}.$$

By definition, $N > n \geq k$. Therefore, $\mathcal{C}_{rj} < -N^{\ell-1}$. This shows first, that $\mathcal{C}_{rj} < 0$ and thus $\mathcal{A}_{rj} \neq 0$, and j is in the preference list of row i; second, the ranking of j is below column 1 . In other words, row i does not prefer j to any of the columns 1 to k. $\square$

2. Matching with couples

Let H be the set of hospitals, D^1 the set of single doctors, and D^2 the set of couples. Each couple $c \in D^2$ is denoted by $c = (f, m)$. For each couple $c \in D^2$ we denote by f_c and m_c the first and second member of c. The set of all doctors, D is given by $D^1 \cup \{m_c \mid c \in D^2\} \cup \{f_c \mid c \in D^2\}$.

Each single doctor $s \in D^1$ has a strict preference relation $\succ_s$ over $H \cup \{\emptyset\}$ where $\emptyset$ denotes the outside option for each doctor. Each couple $c \in D^2$ has a strict preference relation $\succ_c$ over $H \cup \{\emptyset\} \times H \cup \{\emptyset\}$, i.e., over pairs of hospitals including the outside option. The need for ordered pairs arises because couples will have preferences over which member is assigned to which hospital. While a couple consists of two agents they should be thought of as a single agent with preferences over ordered pairs of slots. This is the source of complementarities.

Each hospital $h \in H$ has a fixed capacity $k_h > 0$. A hospital's preferences are characterized by a choice function: $Ch_h(\cdot) : 2^D \to 2^D$. While a choice function can be associated with every strict preference ordering over subsets of D, the converse is not true. The choice function is sufficient to recover a partial order over subsets of D and is responsive.

Theorem 12.2. *In the presence of couples, a stable matching need not exist.*

Proof. Consider the following example taken from Klaus and Klijn, 2005. Suppose two hospitals h_1 and h_2 and three doctors $\{d_1, d_2, d_3\}$. Doctors $\{d_1, d_2\}$ are a couple while d_3 is a single doctor. The capacity of each hospital is 1 and the priority ordering of h_1 is

$$d_1 \succ_{h_1} \succ d_3 \succ_{h_1} \emptyset \succ_{h_1} d_2.$$

The priority ordering for hospital h_2 is

$$d_3 \succ_{h_2} \succ d_2 \succ_{h_2} \emptyset \succ_{h_2} d_1.$$

The preference ordering of the couple is $(h_1, h_2) \succ_{(d_1,d_2)} \emptyset$, while that of doctor d_3 is $h_1 \succ h_2 \succ \emptyset$. $\square$

To deploy Scarf's lemma, we can make the following construction. For each single doctor d and hospital h, that are mutually acceptable, let $x_{(d,h)} = 1$ if d is assigned to h and 0 otherwise. Similarly, for each couple $c \in D^2$ and distinct $h, h' \in H$, such that (h, h') is acceptable to c and the first and second member of c are acceptable to h and h', respectively, let $x_{(c,h,h')} = 1$ if the first member of c is assigned to h and the second is assigned to h'. Let $x_{(c,h,h')} = 0$ otherwise. Finally, for a couple c and a hospital h that are mutually acceptable, let $x_{(c,h,h)} = 1$ if both members of the couple are assigned to hospital $h \in H$ and 0 otherwise. Every $0 - 1$ solution to the following system is a feasible matching and vice versa.

$$\sum_{d \in D^1} x_{(d,h)} + \sum_{c \in D^2} \sum_{h' \neq h} x_{(c,h,h')} + \sum_{c \in D^2} \sum_{h' \neq h} x_{(c,h',h)} + \sum_{c \in D^2} 2x_{(c,h,h)} \leq k_h \quad \forall h \in H$$

$$\sum_{h \in H} x_{(d,h)} \leq 1 \forall d \in D^1$$

$$\sum_{h,h' \in H} x_{(c,h,h')} \leq 1 \forall c \in D^2$$

Let $\mathcal{A}$ be the matrix whose entries are the coefficients of the system. Each agent (single doctor, couple and hospital) is represented by a single row. Each column/variable corresponds to a coalition of agents (an assignment of a single doctor to a hospital or a couple to a pair of hospital slots, that are mutually acceptable).

To apply Scarf's lemma we need each of the rows to have an ordering over the columns that are in the support of that row. For the rows associated with a single doctor and a couple, we can use their preference ordering over the hospitals (and pairs of hospitals in the case of couples).

Notice hospital preferences can invoke complementarities. If one is content to model hospital preferences in this way, we can immediately invoke Scarf's lemma to deduce the existence of a fractional stable matching.

3. Matching with sizes

In practical problems (e.g. resettlement, schools) we often deal with multiple dimensional knapsack constraints. Let L be the set of localities and F be the set of families. For each $\ell \in L$ and $f \in F$ let $v_{1,f}$ be a cardinal value $v_{\ell,f}$ that increases with the probability of member of f finding employment in ℓ.

For each $f \in F$, let a^1_f, a^2_f, a^3_f be the number of children, adults, and elderly in the family f. For each $l \in L$ let $c^1_\ell, c^2_e ll, c^3_\ell$ be the limit on the number of children, adults, and elderly that locality ℓ can absorb.

Use choice function $C_\ell(K)$ to designate the choice function of locality ℓ. Locality ℓ will select a subset of families in K that satisfies its capacity constraints on children, adults and the elderly and maximizes the sum of $v_{\ell,f}$ s chosen.

We can write this as a linear programming problem. Define $z_{\ell f} := 1$ if family f is selected by locality ℓ and zero otherwise. Then $C_\ell(K)$ solves

$$\begin{array}{ll} \max & \sum_{f\in K} v_{\ell f} z_{\ell f} \\ \text{s.t.} & \sum_{f\in K} a^s_f z_{\ell f} \le c^s_\ell, \text{ for } s \in \{1,2,3\} \\ & z_{\ell f} \in \{0,1\} \end{array}$$

Proposition 12.1. *A pairwise stable matching need not exist and even if it did, it need not be Pareto optimal.*

If one follows the previous sections, each column of the matrix $\mathcal{A}$ corresponds to a family and location pair. Therefore, domination would only capture pair-wise stability. To capture group stability, we need the columns to correspond to all potential blocking coalitions. The resulting $\mathcal{A}$ matrix is dense which means the rounding step cannot guarantee small violations of the soft capacity constraints.

Part 5

Random allocation mechanisms

CHAPTER 13

Lotteries and tie-breaking

Recall we can use the Birkhoff-von Neumann Theorem, to say the ex-ante allocation is equivalent to a probability distribution over ex post mappings of agents to goods.

Definition 13.1. *Let x and y be two random allocations. Assume that the components of x_i and y_i are ordered by $\succ_i$. Then x_i stochastically dominates y_i if*

$$\forall k \in \{1, \ldots, n\} : \quad \sum_{j=1}^{k} x_{ij} \geq \sum_{j=1}^{k} y_{ij}$$

Definition 13.2. *Random allocation x is said to be envy free if for any two agents i and j, agent i 's allocation stochastically dominates agent j' s, i.e., $x_i \succeq_i^{sd} x_j$.x is ordinally efficient if x is not stochastically dominated by another random allocation x'.*

Definition 13.3. *Let x be the random allocation produced by a mechanism when agent i reports her preference relation truthfully and the rest report whatever they wish. Let x' be the random allocation produced when agent i misreports her preference relation and the rest do not change. We will say that the mechanism is strategyproof if $x_i \succeq_i^{sd} x'_i$.*

Definition 13.4. *Let x and y be two ex post allocations, i.e., mappings of agents to goods. We will say that y dominates x if for each agent $i, y_i \succeq_i x_i$ and for at least one agent $j, y_j \succ_j x_j$; we will denote this as $y \succ x$. We will say that x is ex post Pareto efficient if there is no mapping of agents to goods, y, such that $y \succ x$.*

1. Random priority

Reintroduce serial dictatorship mechanism: All n good are declared available. Pick an ordering π of the n agents and in the i^{th} iteration, let the agent $\pi(i)$ pick her most preferred good among the currently available goods; declare the chosen good unavailable.

Proposition 13.1. *Serial dictatorship is strategyproof and the allocation produced by it is Pareto optimal.*

Proof. Strategyproofness is straightforward: In each iteration, the active agent has the opportunity of obtaining the best available good, according to her preference list. Therefore, misreporting preferences can only lead to a suboptimal allocation.

We next show Pareto optimality. Let μ be the allocation produced by the priority mechanism. For contradiction, assume that μ' is an allocation that dominates μ. Let i be the first index such that $\mu'(\pi(i)) \succ_{\pi(i)} \mu(\pi(i))$. Clearly, for $j < i$, agent $\pi(j)$ is assigned the same good under μ and μ'. Therefore, in the i^{th} iteration, agent $\pi(i)$ has available good $\mu'(\pi(i))$. Since $\pi(i)$ picks the best available good, $\mu(\pi(i)) \succeq_{\pi(i)} \mu'(\pi(i))$, leading to a contradiction. ☐

Now introduce random serial dictatorship. It iterates over all $n!$ orderings of the n agents. For each ordering π, it runs serial dictatorship and when an agent chooses a good, it assigns a $\frac{1}{n!}$ share of the good to the agent. Clearly, at the end of the process, each agent is assigned a total of one unit of probability shares over all goods.

Proposition 13.2. *Random serial dictatorship is strategyproof and ex post Pareto optimal.*

PROOF. In each of the $n!$ iterations, truth-revealing is the best strategy of an agent. Therefore, it is strategyproof.

We next argue that the random allocation output can be decomposed into a convex combination of perfect matchings of agents to goods such that the allocation made by each perfect matching is Pareto optimal. This is obvious if we choose the n ! perfect matchings corresponding to the $n!$ orderings of agents. Clearly, two different orderings may yield the same perfect matching, therefore the convex combination may be over fewer than $n!$ perfect matchings. Therefore, it is ex post Pareto optimal. □

EXAMPLE 3. Let $n = 4, A = \{1, 2, 3, 4\}$ and $G = \{a, b, c, d\}$. Let the preferences of the agents be as follows.

$$\begin{array}{lcccc} 1: & a & b & c & d \\ 2: & a & b & c & d \\ 3: & b & a & d & c \\ 4: & b & a & d & c \end{array}$$

Then RP will return the following random allocation.

Agent	a	b	c	d
1	5/12	1/12	5/12	1/12
2	5/12	1/12	5/12	1/12
3	1/12	5/12	1/12	5/12
4	1/12	5/12	1/12	5/12

However, it is stochastically dominated by the following random allocation.

Agent	a	b	c	d
1	1/2	0	1/2	0
2	1/2	0	1/2	0
3	0	1/2	0	1/2
4	0	1/2	0	1/2

2. Probabilistic serial

Now define the probabilistic serial (PS) mechanism as follows: As before, we will view each good as one unit of probability shares. Think of probability as a cereal which can be consumed at the rate of one unit per hour. In one hour, PS will assign one unit of probability share to each agent using the following continuous process. Initially, each agent is allocated probability from her most preferred cereal; clearly, if m agents are simultaneously being allocated good j, then j is getting depleted at rate m units per hour. As soon as an agent's preferred cereal is fully allocated, she moves on to her most preferred cereal that is still available. Because each agent has a total order, each agent will get a unit of cereal.

PROPOSITION 13.3. *The allocation computed by PS is envy-free.*

PROOF. At each point in the algorithm, each agent is obtaining probability share of her most favorite good. Therefore, at any time in the algorithm, agent i cannot prefer agent j' s current allocation to her own. Hence PS is envy-free. □

PROPOSITION 13.4. *The random allocation computed by PS is ordinally efficient.*

PROOF. During the run of PS on the given instance, let $t_0 = 0, t_1, \ldots, t_m = 1$ be the times at which some agent exhausts the good she is currently being allocated. By induction on k, we will prove that at time t_k, the partial allocation computed by PS is ordinally efficient among all allocations which give t_k amount of probability shares to each agent.

The induction basis, for $k = 0$, is obvious since the empty allocation is vacuously ordinally efficient. Let A^k denote the allocation at time t_k and let A_i^k denote the allocation made to agent i under A^k. Assume the induction hypothesis, namely that the assertion holds for k, i.e., A^k is ordinally efficient.

For the induction step, we need to show that A^{k+1} is ordinally efficient. Suppose not and let it be stochastically dominated by random allocation P. Let $\alpha = t_{k+1} - t_k$. For each agent i, remove α amount of the least desirable probability shares from P_i to obtain P'_i. Since $P_i \succeq_i^{sd} A_i^{k+1}$, by the property stated above, $P'_i \succeq_i^{sd} A_i^k$. By the induction hypothesis, $A_i^k \succeq_i^{sd} P'_i$ as well. Therefore, $P'_i = A_i^k$. In the time period between t_k and t_{k+1}, each agent obtains α units of probability shares of her most preferred good remaining. Therefore, $A_i^{k+1} \succeq_i^{sd} P_i$, leading to a contradiction. $\square$

Proposition 13.5. *PS is not strategyproof.*

Proof. Proof by counterexample. Let $n = 3, A = \{1, 2, 3\}$ and $G = \{a, b, c\}$. Let the preferences of the agents be as follows.

$$\begin{array}{llll} 1: & a & b & c \\ 2: & a & c & b \\ 3: & b & a & c \end{array}$$

PS will return the following allocation.

Agent	a	b	c
1	1/2	1/4	1/4
2	1/2	0	1/2
3	0	3/4	1/4

However, if Agent 3 lies and reports her preference list as $a \succ b \succ c$, then PS will return the following allocation.

Agent	a	b	c
1	1/3	1/2	1/6
2	1/3	0	4/6
3	1/3	1/2	1/6

Therefore by lying, agent 3 obtains 5/6 units of her favorite two goods, instead of only 3/4 units.

$\square$

CHAPTER 14

Pseudomarkets

We follow Pycia (2021). Pseudomarkets are allocation mechanism where participants provide a representation of their cardinal utilities, from which a resulting allocation is set to be in a Walrasian equilibrium. The key is that participants have budgets of fictitious (token) money; it is fictitious in that no actual money transfers are made. The assumptions follow directly from the paper:

- Finite one-sided matching market with agents $i, j, k \in I = \{1, \ldots, |I|\}$.
- Indivisible objects or goods $x, y, z \in X = \{1, \ldots, |X|\}$.
- Each object x is represented by a number of identical copies $|x| \in \mathbb{N}$. If agents have outside options, we treat them as objects in X; this implies that $\sum_{x \in X} |x| \geq |I|$.
- Agents demand a probability distribution over objects; let $q_i^x \in [0,1]$ be the probability that agent i obtains a copy of object x. Agent i 's random assignment $q_i = \left(q_i^1, \ldots, q_i^{|X|}\right)$ is a probability distribution on X; we denote the set of such probability distributions by $\Delta(X)$ and we also refer to them as individual assignments.
- The utility from random assignment q_i is $u_i\left(q_i\right) = v_i \cdot q_i = \sum_{x \in X} v_i^x \cdot q_i^x$ where $v_i = \left(v_i^x\right)_{x \in X} \in [0, \infty)^{|X|}$ is the vector of agent i 's von Neumann-Morgenstein valuations for objects $x \in X$.
- The set of economy-wide random assignments is $\Delta(X)^I$.

Definition 14.1. *An economy-wide assignment $q = \left(q_i^x\right)_{i \in I, x \in X}$ is feasible if $\sum_{i \in I} q_i^x \leq |x|$ for every object x.*

Theorem 14.1 (Birkhoff-von Neumann Theorem). *A feasible random assignment can be expressed as a lottery over feasible deterministic assignments.*

Definition 14.2 (Equilibria). *A feasible economy-wide assignment q^* and a vector $p^* \in \mathbb{R}_+^X$ constitute an equilibrium for a constraint vector $w^* \in \mathbb{R}_+^I$ if $q^* = \left(q_i^*\right)_{i \in I}$ satisfies $p^* \cdot q_i^* \leq w_i^*$ for all $i \in I$ and $u_i\left(q_i\right) > u_i\left(q_i^*\right) \Longrightarrow p^* \cdot q_i > w_i^*$ for all $\left(q_i\right)_{i \in I} \in \Delta(X)^{\bar{I}}$.*

Theorem 14.2 (Existence). *For any budget vector $w^* \in \mathbb{R}_+^I$, there is an equilibrium (q^*, p^*) satisfying market clearing.*

Proof. For simplicity say the price vector $p^* \in \mathbb{R}_+^X$ has only nonnegative prices. Let $w = \max\left\{w_1, \ldots, w_{|I|}\right\}$ and let $t(p) = \left(\max\left\{0, p_x\right\}\right)_{x \in X}$ be a projection of the vector $p \in \mathbb{R}^X$ onto $\mathbb{R}_+^X$. For sufficiently large $M > 0$, the price vector adjustment function

$$f(p) = t(p) + \left(\sum_{i \in I} d_i(t(p)) - (|x|)_{x \in X}\right)$$

maps the compact Cartesian-product space $\times_{x \in X}[-|x|, M]$ into itself because, if price p_x of good x is higher than $w|I|$, then the sum $\sum_{i \in I} d_i(t(p))_x$ of agents' demands for x is lower than $|x|$ and thus $f(p)_x \leq p_x$.

The properties of demand correspondences d_i established above imply that f is upper hemi-continuous and takes values that are non-empty, convex, and compact. Thus, by Kakutani's fixed-point theorem, there exists a fixed point $\hat{p} \in f(\hat{p})$. The fixed point property of $\hat{p}$ implies that for each x

$$\left(|x| + \left(\hat{p}_x - t_x(\hat{p})\right)\right)_{x \in X} \in \sum_{i \in I} d_i(t(\hat{p})),$$

and hence there is $q^* \in \times_{i\in I} d_i\,(p^*)$ for which

$$(|x| + (\hat{p}_x - t_x(\hat{p})))_{x\in X} = \sum_{i\in I} q_i^*.$$

By construction, q^* and $p^* = t(\hat{p})$ are in equilibrium provided q^* is feasible, that is $0 \leq \sum_{i\in I} q_i^{*x} \leq |x|$ for each $x \in X$. The first inequality is from $q_i^* \in d_i\,(p^*)$ and the second inequality from $t_x(\hat{p}) \geq \hat{p}_x$. The market-clearing property is satisfied because, for positive p_x^*, we have $\hat{p}_x - t_x(\hat{p}) = 0$ and hence $\sum_{i\in I} q_i^{*x} = |x| + (\hat{p}_x - t_x(\hat{p})) = |x|$. $\square$

Part 6

Auction design

CHAPTER 15

Auctions

1. Basics

We will stick to the independent private value (IPV) model. Be explicit that common values have strategic conditioning that erase some of these results.

We have n risk-neutral bidders that compete for a single unit. Bidder i values the unit at v_i, which is private information to her, but it is common knowledge that each v_i is independently drawn from the same continuous distribution $F(v)$ on $[\underline{V}, \bar{V}]$ (so $F(\underline{V}) = 0, F(\bar{V}) = 1$) with density $f(v)$. We call a bidder's value her type.

We often use the uniform distribution categorized by

$$\left(F(v) = \frac{v - \underline{V}}{\bar{V} - \underline{V}}, f(v) = \frac{1}{\bar{V} - \underline{V}}\right).$$

Note it corresponds to linear demand, in that the proportion of buyers with valuations above any price p (that is $\frac{\bar{V}-p}{\bar{V}-\underline{V}}$) is linear in p. The expected k^{th} highest value among n values independently drawn from the uniform distribution on $[\underline{V}, \bar{V}]$ is $\underline{V} + \left(\frac{n+1-k}{n+1}\right)(\bar{V} - \underline{V})$

There are four main categories of single-unit auctions.

1.1. First price sealed bid auction. Every player independently chooses a bid without seeing others' bids. The object is sold to the bidder who makes the highest bid. The winner pays her bid.

Proposition 15.1. *Under the uniform distribution, the optimal bid in the first price sealed bid auction is*

$$b(v_i) = \underline{V} + \left(\frac{n-1}{n}\right)(v_i - \underline{V}).$$

Proof. Consider player i chooses to mimic having value $\widetilde{v}$ instead of bidding her value; that is she bids $b(\widetilde{v}) = \widetilde{b}$. She then beats any other bidder j with the probability that $v_j < \widetilde{v}$, which equals $F(\widetilde{v})$ in equilibrium. So mimicking $\widetilde{v}$ would beat all the other $(n-1)$ bidders with probability $(F(\widetilde{v}))^{n-1}$ and yields expected surplus to player i of

$$S(v_i, \widetilde{v}) = (v_i - b(\widetilde{v}))(F(\widetilde{v}))^{n-1}.$$

The optimal bid is then equivalent to choosing the best $\widetilde{v}$ to mimic, which we can do by looking at the first-order condition

$$\frac{\partial S(v_i, \widetilde{v})}{\partial \widetilde{v}} = -b'(\widetilde{v})(F(\widetilde{v}))^{n-1} + (v_i - b(\widetilde{v}))(n-1)(F(\widetilde{v}))^{n-2} f(\widetilde{v}).$$

For the bidding function $b(v)$ to be a (symmetric) equilibrium, i's best-response to all others bidding according to this function must be to do likewise, i.e. her optimal choice of $\widetilde{b}$ is $b(v_i)$ and of $\widetilde{v}$ is v_i. So

$$\frac{\partial S}{\partial \widetilde{v}}(v_i, \widetilde{v}) = 0 \text{ at } \widetilde{v} = v_i$$

$$\Rightarrow \quad b'(v_i) = (v_i - b(v_i))(n-1)\frac{f(v_i)}{F(v_i)}.$$

Using the boundary condition

$$b(\underline{V}) = \underline{V}$$

(it is obvious type $\underline{V}$ will not bid more than $\underline{V}$, and we assume the auctioneer will not accept lower bids than $\underline{V}$), we can solve the differential equation for the equilibrium. Apply F and f from the uniform distribution and get

$$b'(v_i) = (v_i - b(v_i))(n-1)\left(\frac{1}{v_i - \underline{V}}\right)$$

which is solved by

$$b(v_i) = \underline{V} + \left(\frac{n-1}{n}\right)(v_i - \underline{V}).$$

□

Proposition 15.2. *Under the uniform distribution, the expected revenue in the first price sealed bid auction is* $\underline{V} + \left(\frac{n-1}{n+1}\right)(\bar{V} - \underline{V})$.

Proof. The seller's expected revenue is

$$E\left\{\max_{i=1,..,n} b(v_i)\right\} = E\left\{\underline{V} + \left(\frac{n-1}{n}\right)\left(\max_{i=1,..,n} v_i - \underline{V}\right)\right\}.$$

Using the uniform distribution we get

$$E\left\{\max_{i=1,..,n} v_i\right\} = \underline{V} + \left(\frac{n}{n+1}\right)(\bar{V} - \underline{V})$$

so the expected revenue is $\underline{V} + \left(\frac{n-1}{n+1}\right)(\bar{V} - \underline{V})$. □

1.2. Dutch auction. The auctioneer starts at a very high price, say $\bar{V}$, and then lowers the price continuously. The first bidder who calls out that she will accept the current price wins the object at the current price.

Proposition 15.3. *The Dutch auction is strategically equivalent to the first price sealed bid auction. The players' bidding functions and expected payoff of the seller are exactly the same.*

1.3. Second price sealed bid auction. Every player independently chooses a bid without seeing the other players' bids, and the object is sold to the bidder who makes the highest bid. The winner pays the second highest bid.

Proposition 15.4. *In the second price sealed bid auction, it is optimal for a bidder to bid her true value.*

Proof. Consider bidder i bids some $v_i - x$. If the highest bid is w, then:

(1) If $v_i - x > w$, agent i will win the auction and pay w, just as if agent i bid v_i.
(2) If $w > v_i$, agent i will lose the auction and get nothing, just as if agent i bid v_i.
(3) But if $v_i > w > v_i - x$, bidding $v_i - x$ causes agent i to lose the auction and get nothing, whereas if agent i had bid v_i, agent i would have won the auction and paid w for a net surplus of $v_i - w$. So agent i never gains and might lose if agent i bids $v_i - x$.

Consider agent i bidding $v_i + x$. If the highest bid other than agent i's is w, then:

(1) If $v_i > w$, agent i will win and pay w, just as if agent i bid v_i.
(2) If $w > v_i + x$, agent i will lose and pay nothing, just as if agent i bid v_i.
(3) But if $v_i + x > w > v_i$, having bid $v_i + x$ causes agent i to "win" an auction agent i otherwise would have lost, and agent i has to pay $w > v_i$, so agent i gets negative surplus. So bidding $v_i + x$ may hurt agent i compared with bidding v_i, but it never helps agent i.

□

Proposition 15.5. *Under the uniform distribution, the expected revenue in the second price sealed bid auction is* $\underline{V} + \left(\frac{n-1}{n+1}\right)(\bar{V} - \underline{V})$.

Proof. Since everyone is bidding her true value, the seller's expected revenue is the expected second highest value of the n values. This is $\underline{V} + \left(\frac{n-1}{n+1}\right)(\bar{V} - \underline{V})$. □

1.4. English auction. The price rises until one bidder remains.

PROPOSITION 15.6. *The English auction is strategically equivalent to the second price sealed bid auction. The players' bidding functions and expected payoff of the seller are exactly the same.*

The next-to-last person will drop out when her value is reached, so the person with the highest value will win at a price equal to the value of the second-highest bidder.

2. Revenue equivalence

Let $S_i(v_i)$ be the expected surplus that bidder i will obtain in equilibrium from participating in the mechanism as a function of her type, v_i. Let $P_i(v_i)$ be her probability of receiving the object in the equilibrium. Then we have

$$S_i(v) \geq S_i(\widetilde{v}) + (v - \widetilde{v})P_i(\widetilde{v}).$$

Since in equilibrium type v must not want to mimic type $v + dv$, we have

$$S_i(v) \geq S_i(v + dv) + (-dv)P_i(v + dv)$$

and since $v + dv$ must not want to mimic type v we have

$$S_i(v + dv) \geq S_i(v) + (dv)P_i(v).$$

Redistributing yields

$$P_i(v + dv) \geq \frac{S_i(v + dv) - S_i(v)}{dv} \geq P_i(v)$$

and taking the limit as $dv \to 0$ we obtain

$$\frac{dS_i}{dv} = P_i(v).$$

Integrating up,

$$S_i(v) = S_i(\underline{V}) + \int_{x=\underline{V}}^{v} P_i(x)dx.$$

Now consider any two mechanisms which have the same $S_i(\underline{V})$ and the same $P_i(v)$ function for all v and for every player i. They have the same $S_i(v)$ function. So any given type of player i makes the same expected payment in each of the two mechanisms (since $S_i(v) = vP_i(v) - E$ (payment by type v of player i), since the bidder is risk-neutral). So also i 's expected payment averaged across her different possible types, v, is the same for both mechanisms. Since this is true for all bidders, i, the mechanisms yield the same expected revenue for the auctioneer. This is the Revenue Equivalence Theorem.

THEOREM 15.1 (Revenue Equivalence Theorem). *For any two auctions and any two equilibria of those auctions: (i) if the good is allocated identically across these two equilibria, and (ii) the same expected payment is made by the lowest type v=0 in in both of these two equilibria, then every type of every bidder makes the same expected payment in these two equilibria.*