Advanced Microeconomics

Professor: Eric Shi

Timekeeper: Zircon

Fall 2023

Contents

1 Choice Theory

The utility-maximization approach to choice has several characteristics that help account for its long and continuing dominance in economic analysis:

- Normative usefulness: policy-making; individuals' choices v.s. government's welfare criterion; modern democratic values.
- Positive predictions: comparative statics predictions.
- Wide scope.
- Compactness: make empirical predictions from a relatively sparse model of the choice problem, just a description of the chooser's objectives and constraints.

1.1 Preference-Based Approach

Rational choice theory starts with the idea that individuals have preferences and choose according to those. The primary task is to formalize what that means and precisely what it implies about the pattern of decisions we should observe.

1.1.1 Preference Relation

Let X be a set of possible choices. Consider a preference over the set X, as a binary relationship as follows:

$$x \succeq y \iff \text{“}x \text{ is at least as good as } y\text{”}$$

The weak preference relation $\succeq$ implies the associated strict preference relation $\succ$ and indifference relation $\sim$ as follows:

- x is *strictly preferred to* y, or $x \succ y$, if $x \succeq y$ but not $y \succeq x$.
- x is indifferent to y, or $x \sim y$, if $x \succeq y$ and $y \succeq x$.

After the primary defition, we need to make assumptions about rationality, or definition of rational preference.

The first is *completeness*, which means an agent would never be clueless faced with two choices.

Definition (Completeness) A preference relatin $\succeq$ on X is ***complete***, if for all $x, y \in X$, either $x \succeq y$, or $y \succeq x$, or both.

The second is *transitivity*, which means that an agent's weak preference cannot cycle unless among choices that are indifferent.

Definition (Transitivity) A preference relation $\succeq$ on X is ***transitive***, if whenever $x \succeq y$ and $y \succeq z$, then $x \succeq z$.

From completeness and transitivity, we have the following corollaries:

1. While the definition of transitivity involves only 3-cycles, it also extends to all n-cycles, i.e., that it implies that for any n choices $x_1, x_2, \cdots, x_n \in X$ such that $x_1 \succeq x_2$, $x_2 \succeq x_3$, ..., $x_{n-1} \succeq x_n$, we must have $x_1 \succeq x_n$. (Hint: use induction on n.)
2. Transitivity of a weak preference relation $\succeq$ implies transitity of the associated strict preference realtion $\succ$ and the indifference relation $\sim$.

Remarks:

1. The assumption of transitivity is inconsistent with certain "framing effects".
2. Preference relation is defined upon $\succeq$ for simplicity.

Definition (Rationality) A preference relation $\succeq$ on X is ***rational*** if it is both *complete* and *transitive*.

1.1.2 Choice Rule

Given preferences, we should define the agent's *choice rule* (given an "opportunity set" $B \subseteq X$) induced by preference relation $\succeq$ as

$$C_{\succeq}(B) = \{x \in B | x \succeq y, \forall y \in B\}$$

which is the set of items in B the agent likes at least as much as any of the other alternatives in B; or say, most preferred.

Remarks:

1. $C_{\succeq}(B)$ may contain more than one element.
2. $C_{\succeq}(B)$ might be empty. If B is finite and non-empty, then $C_{\succeq}(B)$ is non-empty.
   - As an example for $C_{\succeq}(B)$ to be empty, define a preference relation $x \succeq y$ iff $x \geq y$. Take $X = [0, +\infty)$ and $B = (0, 1)$. Then clearly there is no "most preferred" option.
   - To guarantee the existence of choices from infinite choice sets, we will later make some technical assumptions (such as compactness of choice sets and continuity of preferences).

Proposition Suppose $\succeq$ is complete and transitive. Then, for every finite non-empty set B, $C_{\succeq}(B) \neq \emptyset$.

Proof (By Induction) Proceed by mathematical induction on the number of elements of B.

- $|B| = 1$, say $B = \{x\}$.

  - By completeness, $x \succeq x$, so $x \in C_{\succeq}(B)$. $C_{\succeq}(B) \neq \emptyset, \forall |B| = 1$.

- Fix $n \geq 1$ and suppose that for all sets B with exactly n elements, $C_{\succeq}(B) \neq \emptyset$.
- $|B| = n + 1$.
  - Consider B_n, $|B_n| = n$. Since $C_{\succeq}(B_n) \neq \emptyset$, say $C_{\succeq}(B_n) = x^*$. Let $B = B_n \cup \{x_{n+1}\}$.
  - By completeness, we have only two (not mutually-exclusive) possibilities:
    * If $x^* \succeq x_{n+1}$, then by definition $x^* \in C_{\succeq}(B)$, so $C_{\succeq}(B) \neq \emptyset$.
    * If $x_{n+1} \succeq x^*$. Since $x^* \in C_{\succeq}(B_n)$, by definition, $x^* \succeq y, \forall y \in B_n$. By transitivity, this implies that $x_{n+1} \succeq y, \forall y \in B$. Therefore, $x_{n+1} \in C_{\succeq}(B)$, so $C_{\succeq}(B) \neq \emptyset$.
- Hence, for every set B with exact $n + 1$ elements, $C_{\succeq}(B) \neq \emptyset$. By the principle of mathematical induction, it follows that for every finite set B with any number of elements, $C_{\succeq}(B) \neq \emptyset$.

1.2 Choice-Based Approach

Much empirical work reasons in the reverse way contrary to economic theories (in preference-based approach): it looks at individuals' choices and tries to ”rationalize” those choices. That is, to figure out whether the choices are compatible with preference maximization and, if so, what they imply about those preferences. In choice-based approach, choice rule is the primitive object of the theory.

Definition (Choice Rule) Let B be the set of all nonempty subsets of X ($\mathcal{B} = 2^X \backslash \emptyset = \{B \neq \emptyset : B \subset X\}$). A ***choice rule*** is a function $C : B \to B$ with the property that for all $B \in B$, $C(B) \subseteq B$.

Remarks:

- B is the set of all nonempty **subsets** of X, which means all possible set of available choice(s) the agent is facing. And the choice rule C is a mapping from B to B, which means that the agent is choosing from a set of available choice(s) to pick his most preferred choice(s), also a subset of X.
- Here we assume that we can see the agent choose from *all* possible subsets of X, and that the agent reports *all* of his optimal choices from a given opportunity set.

By definition of choice rule, we do not impose any assumptions or restrictions on the rule or underlying preference relation. We are then interested in two questions:

- If this rule comes from maximizing some underlying preferences, what can we infer about these preferences?

- Is this choice rule consistent with the maximization of *some* complete and transitive preference relation (i.e., *rationlizable*)?

Consider the first question first. Suppose that choice rule C is consistent with the maximization of some preference relation $\succeq$, i.e., $C(\cdot) = C_{\succeq}(\cdot)$. Then, observing for some $A \subseteq X$ that $y \in A$ and $x \in C(A)$ (i.e., x is chosen when y is available) allows us to infer that $x \succeq y$. This implies that for any $B \subseteq X$ such that $x \in B$ and $y \in C(B)$, we must also have $x \in C(B)$ (Indeed, we have $x \succeq y$ and $y \succeq z$ for all $z \in B$, and so by transitivity $x \succeq z$ for all $z \in B$). By a symmetric argument, we should then have $y \in C(A)$. Thus, any rationalizable choice rule must have the following property (as a **necessary** condition):

Definition (HARP) A choice function $C : \mathcal{B} \to \mathcal{B}$ satisfies ***Houthaker's Axiom of Revealed Preference (HARP)*** if, whenever $x, y \in A \cap B$, and $x \in C(A)$ and $y \in C(B)$, we have $x \in C(B)$ and $y \in C(A)$.

In words, HARP says that if choices x, y are both available in two choice experiments, and x is chosen in one experiment and y in the other, then both x and y must be chosen in both experiments. HARP guarantees that there is no (obvious) inconsistency in the agent's choices.

It has been argued that HARP is a necessary condition for a choice rule to be rationalizable. Turns out that HARP is also sufficient for rationalizability:

Proposition Suppose $C : \mathcal{B} \to \mathcal{B}$ is nonempty-valued. Then there exists a rational (complete and transitive) preference relation $\succeq$ on X such that $C(\cdot) = C_{\succeq}(\cdot)$ if and only if C satisfies HARP.

Proof (Follow Definition)

- "Only if" part: Already argued, HARP as necessary condition for rational relation's existence.
- "If" part
  - Definition of revealed preference relation (by choice rule).
    - Suppose choice rule C satisfies HARP. Construct the "revealed preference relation" $\succeq_C$ as follows: say that $x \succeq_C y$ if and only if there exists some $A \subseteq X$ such that $y \in A$ and $x \in C(A)$.
  - $\succeq_C$ is complete.
    - Since C is nonempty-valued, pick any $x, y \in X$, we have either $x \in C(\{x, y\})$, in which case $x \succeq_C y$, or $y \in C(\{x, y\})$, in which case $y \succeq_C x$ (or both, in which case $x \sim_C y$).
  - $\succeq_C$ is transitive.

* Suppose $x \succeq_C y$ and $y \succeq_C z$, and consider $C(\{x, y, z\})$, which by hypothesis is non-empty. There are three possibilities (which are though not mutually exclusive):
    - $x \in C(\{x, y, z\})$. Then by construction of $\succeq_C$, $x \succeq_C z$.
    - $y \in C(\{x, y, z\})$. Then by HARP, since $x \succeq_C y$, we must also have $x \in C(\{x, y, z\})$, and so by case 1 we have $x \succeq_C z$.
    - $z \in C(\{x, y, z\})$. Then by HARP, since $y \succeq_C z$, we also have $y \in C(\{x, y, z\})$, and so by case 2 $x \succeq_C z$.

- $C(\cdot) = C_{\succeq_C}(\cdot)$.
  * Equivalent to show that, for all $x \in X$ and $A \in \mathcal{B}$, $x \in C(A)$ if and only if $x \in A$ and $x \succeq_C y$ for all $y \in A$.
    - "Only if" part holds by construction of $\succeq_C$.
    - "If" part: Take any $x \in C_{\succeq_C}(A)$, then $x \succeq_C y$, for all $y \in A$. Since $C(\cdot)$ is nonempty-valued, $\exists y_0 \in A$ such that $y_0 \in C(A)$. By the definition of $\succeq_C$, $\exists A_0$ such that $x, y_0 \in A_0$ and $x \in C(A_0)$. It follows that $x, y_0 \in A_0 \cap A$ and $x \in C(A_0)$ and $y_0 \in C(A)$. By HARP, $x \in C(A)$.

Remarks:

1. It is *HARP* that plays a fundamental role in the proof of "if" part, that is, to endow the revealed preference relation $\succeq_C$ with characteristics of "rationality".
2. The preceding problem develops the properties of rational choice for the case when the entire choice function $C(A)$ is observed, i.e., (i) for any given choice set, all of the agent's optimal choices are observed, not just some of them; (ii) the agent's optimal choices over all choice sets are observed. However,
   - Real data is commonly less comprehensive like that. For example, in consumer choice problems, the relevant sets A may be only budget sets that consist of affordable choices given income and prices, which is a particular subcollection of $\mathcal{B}$.
   - The first aspect of incomplete observation is virtually insolvable. For the second, to develop a theory based on more limited observations, other "axioms of revealed preference" have been developed. For example, the *weak axiom of revealed preference (WARP)* is an equivalent of HARP for choice sets restricted to budget sets described by linear prices and for choice rules restricted to be single-valued (so the conclusion of HARP can be strengthened to $x = y$). WARP is still nessary for rationalizability but proves insufficient. A stronger version, *Generalized Axiom of Revealed Preference*, has been shown both necessary and sufficient for rationalizability.

2 Consumer Theory

2.1 Setups

Choice theory assumes that a rational decision-maker selects their most preferred option from their choice set. Consumer theory can be viewed as an application of choice theory. We will focus on the *preference-based approach*, with three goals in mind:

1. Formalize the ”choice set” and ”preferences” of a consumer.
    - Set of alternatives: consumption set.
    - Choice set: budget set.
    - Preferences: utility representation of a preference relation.
2. Derive a consumer's optimal choice(s) based on the information of their choice set and the preference relation.
3. Analyze the properties of optimal choices (demand).

At the beginning, a few assumptions are needed in consumer theory:

1. *Perfect* information.
2. Consumers are *price takers*. (Prices $\mathbf{p}$ are taken as known, fixed and exogenous; No searching or bargaining for discounts)
3. Prices are *linear*. (No quantity discounts)
4. Goods are divisible. (Formally expressed by the condition $x \in \mathbb{R}^n_+$; notice that the divisibility assumption does not prevent us from applying the model to situations with discrete and indivisible goods)

For simplicity, we will view the entire positive quadrant as the consumption set. With n goods, the consumption set is given by:

$$X = \mathbb{R}^n_+ = \{\mathbf{x} \in \mathbb{R}^n : x_i \geq 0 \text{ for } i = 1, 2, \cdots, n\}$$

With n goods and income of m, given a price vector $p = (p_1, \cdots, p_n)' \geq 0, p \neq \mathbf{0}$, the budget set is given by:

$$B(\mathbf{p}, m) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{p} \cdot \mathbf{x} \leq m, \mathbf{x} \geq \mathbf{0}\}$$

Remarks: By re-interpreting the consumption goods and the budget/wealth, the derivation of the budget set can be extended to encompass other economic problems. For example, (i) consumption-leisure choice, and (ii) inter-temporal choice.

2.2 Utility Representation

The preference relation is intuitive, but difficult to work with. Having a utility representation for preferences is convenient because it turns the problem of preference maximization into a relatively familiar mathematical problem.

Definition (Utility) A preference relation $\succeq$ on X is represented by a utility function $u : X \to \mathbb{R}$ if

$$x \succeq y \iff u(x) \geq u(y)$$

Under this definition, the utility function is an *ordinal* representation. The value/utility level per se has no economic meaning and is just a convenient mathematical tool to represent the consumer's *relative rankings* of different options.

If u represents $\succeq$, then the choice rule defined upon budget set is:

$$C(B(\mathbf{p}, m); \succeq) = \left\{ \mathbf{x} \mid \mathbf{x} \text{ solves } \max_{\mathbf{x} \in B(\mathbf{p}, m)} u(\mathbf{x}) \right\}$$

Remarks: An equivalent notation of such choice rule is $C_{\succeq}(B)$.

A natural question is whether given a preference relation $\succeq$, we can always find a function $u(\cdot)$ to represent $\succeq$. In the case of preference relation defined upon a finite X, the answer is yes.

Proposition If X is finite, then any rational (complete and transitive) preference relation $\succeq$ on X can be represented by a utility function $u : X \to \{1, \cdots, n\}$, where $n = |X|$.

Proof (By induction)

- For $|X| = n = 1$, say $X = \{x\}$, we let $u(x) = 1$ and the conclusion is trivial.
- Next, suppose that preferences can be represented as described for any set with at most n elements.
- Consider a set X with $n + 1$ Elements.
  - Since $C_{\succeq}(X) \neq \emptyset$ by *Proposition 1*, the set $Y = X \backslash C_{\succeq}(X)$ has no more than n elements (by doing so we can limit our discuss and deduction back to the hypothesized case), so preferences restricted to that set can be represented by a utility function $u : Y \to \{1, 2, \cdots, n\}$.
  - We extend the domain of u to X by setting $u(x) = n+1$ for each $x \in C_{\succeq}(X)$. By construction, we have $u(x) \in \{1, \cdots, n, n+1\}$ for alll $x \in X$.

– Now we show that the constructed u represents $\succeq$, i.e., for any $x, y \in X$, $x \succeq y$ if and only if $u(x) \geq u(y)$. Suppose $x \succeq y$. There are three possibilities:

  * $x \in C_{\succeq}(X), y \in Y \iff u(x) = n+1 \geq u(y)$.
  * $x, y \in C_{\succeq}(X) \iff u(x) = u(y) = n+1$, also, $u(x) \geq u(y)$.
  * $x, y \in Y$. Then, since by construction u represents $\succeq$ on Y, $x \succeq y$ if and only if $u(x) \geq u(y)$.

Remarks: If X is infinite, things are a bit more complicated. In general, given any rational preference relation $\succeq$ on $\mathbb{R}^n_+$, it cannot always be represented by a utility function $u : \mathbb{R}^n_+ \to \mathbb{R}$.

Counter-Example (Lexicographic Preference) Consider the lexicographic preferences on the square $X = \mathbb{R}^2$ according to which $(x_1, x_2) \succ (y_1, y_2)$ if either (1) $x_1 > y_1$ or (2) $x_1 = y_1$ and $x_2 > y_2$. These preferences cannot be represented by a utility function, and this is also an example for which "indifference curves" do not exist, because the agent is never indifferent between any two choices.

Too see this, suppose by contradiction there exists a utility representation $u(x, y)$. For any $x \in \mathbb{R}_+$, consider the interval $I(x) = \big(\inf_y u(x, y), \sup_y u(x, y)\big)$. $I(x)$ is not degenerate and $I(x_1), I(x_2)$ do not overlap for $x_1 \neq x_2$. Since rational numbers are dense, there exists a rational number $r(x) \in I(x)$. Since $I(x_1), I(x_2)$ do not overlap, $r(x_1) \neq r(x_2)$ for $x_1 \neq x_2$. Consider the injective function $r : \mathbb{R}_+ \to \mathbb{Q}_+$. Since $\mathbb{R}_+$ is uncountable and $\mathbb{Q}_+$ is countable, this is impossible.

However, note that if we replace $X = \mathbb{R}^2_+$ with $X' = \mathbb{Q}^2_+$, the lexicographic preference relation would admit a utility representation. More generally, this idea is formalized into the proposition in countable set case.

Take a step forward. If a preference relation $\succeq$ is defined upon a countable set X (possibly infinite), then $\succeq$ admits a utility representation. We can always achieve this by construction, as it goes in the following proposition.

Proposition If $X \neq \emptyset$ is countable, then any rational (complete and transitive) preference relation $\succeq$ on X can be represented by a utility function $u : X \to (0, 1)$.

Proof

- First, construct a mapping of utility function $u : X \to (0, 1)$.
  – Let $(x_n)_{n=1}^{\infty}$ be an enumeration of X (that X is countable).
  – Let $u(x_1) = \frac{1}{2}$. Consider x_{n+1}, where $n \geq 1$.
    * If $x_{n+1} \sim x_i$ for some $1 \leq i \leq n$, then define $u(x_{n+1}) = u(x_i)$.

* Otherwise, define

$$M_n = \max\{\max\{u(x_i) : x_{n+1} \succ x_i, 1 \le i \le n\}, 0\}\}$$
$$m_n = \min\{\min\{u(x_i) : x_i \succ x_{n+1}, 1 \le i \le n\}, 1\}\}$$

By construction $M_n > m_n$. Define $u(x_{n+1}) = \frac{M_n+m_n}{2}$.

- Second, prove that $u(\cdot)$ is a utility representation of preference relation $\succeq$.
  - Take any $y, z \in X$. Since X is countable, $\exists i, j$ such that $y = x_i$ and $z = x_j$. Without loss of generality, suppose $i \le j$.
  - By construction, $u(x_i) = u(x_j)$ if and only if $x_i \sim x_j$. For $x_j \nsim x_i$, $u(x_j) > u(x_i)$ if and only if $x_j \succ x_i$.

Intuitively, the problem with the lexicographic preference relation is that there is sudden preference reversals. For instance, we know $(3,3) \succ (3,2)$, but $(3,2) \succ (x,3)$ for x slightly less than 3. This motivates the following additional (technical) restriction on preference relations. The "continuity" restriction on preferences is a condition starting from the simple and compelling reason that any *finite* set of observed choices that is consistent with HARP is also consistent with continuity.

Definition (Continuity) A preference relation $\succeq$ on $X \subseteq \mathbb{R}^n$ is ***continuous*** if for any sequence $\{(x^n, y^n)\}_{n=1}^{\infty}$ with $x^n \to x, y^n \to y$, and $x^n \succeq y^n$ for all n, we have $x \succeq y$.

Continuity condition implies not only that a utility representation exists, but that a *continuous* representation exists.

Proposition Any complete, transitive and continuous preference relation $\succeq$ on X on $X \subseteq \mathbb{R}^n$ can be represented by a continuous utility function $u : X \to \mathbb{R}$.

Proof

In order to have a simple, constructive proof, we prove the proposition only for the case of a monotone preference relation $\succeq$ on $X = \mathbb{R}^n_+$.

Let $e = (1, \cdots, 1)$ and consider bundles of the form $\alpha e = (\alpha, \cdots, \alpha)$ where $\alpha \ge 0$. For each $x \in \mathbb{R}^n_+$, we construct a utility number as follows: $u(x) = \max A(x)$, where $A(x) = \{\alpha \in R_+ : \alpha e \preceq x\}$. To see that the set $A(x)$ has a maximal point, note that the set is

- Nonempty, since $0 \in A(x)$ by monotonicity of $\succeq$;
- Closed, by the continuity of $\succeq$;
- Bounded, since by monotonicity of $\succeq$, $\alpha \le \max\{x_1, \cdots, x_n\}$ for each $\alpha \in A(x)$.

Now we show that we must have $u(x)e \sim x$.

1. $u(x)e \preceq x$. This is satisfied by construction of $u(x) \in A(x)$.
2. $u(x)e \succeq x$. For each $n \geq 1$, we have $u(x) + \frac{1}{n} \notin A(x)$, hence $(u(x) + \frac{1}{n})e \not\preceq x$, therefore by completeness of $\succeq$ we have $(u(x) + \frac{1}{n})e \succeq x$, which by continuity of $\succeq$ implies $\lim_{n\to\infty} (u(x) + \frac{1}{n})e = u(x)e \succeq x$.

Now it remains to show that the constructed utility function $u(\cdot)$ has

1. Ability to represent the preference relation $\succeq$, and
2. Continuity.

For representation part, note that by transitivity $x \succeq y$ if and only if $u(x)e \succeq u(y)e$ (since $u(x)e \sim x \succeq y \sim u(y)e$), and by monoticity of $\succeq$, this holds if and only if $u(x) \geq u(y)$. For continuity part, this is more subtle and not covered here.

Remarks:

1. The construction in the proof specifies the utility of any bundle x by finding the point on the 45° line on the indifference curve passing through x.
   - This specification is, of course, completely arbitrary, just for mathematical convenience.
   - To reflect this arbitrariness, utility representation of preferences is *ordinal*, i.e., only the induced preference ordering of choices is meaningful and not the exact utility numbers assigned to them.
   - In fact, if u represents $\succeq$, then $U(\cdot) = v(u(\cdot))$ also represents $\succeq$ so long as $v : \mathbb{R} \to \mathbb{R}$ is an increasing function.
2. The thought process and the introduction of the additional assumption are the most important for this part.
   - Recall that our motivation is to come up with a tractable framework to analyze the consumer's problem. One possibility is to attach a utility level to each consumption bundle (ordinal utility representation).
   - We ask whether any preference relation can be represented by a utility function. The answer is yes if the consumption set X is finite (or countable), but no if $X = \mathbb{R}^n_+$. We notice that the problem with the counter-example (e.g., the lexicographic preference relation) is that there are sudden preference reversals.
   - We then impose the restriction of "continuity" of preferences to rule out the case of sudden preference reversals and show that any rational and continuous preference relation on $X = \mathbb{R}^n_+$ can be represented by a continuous utility function.

3. The continuity of preference relation and continuity of its utility representation is not interdependent. If a preference relation can be represented by a continuous utility function, then such preference relation must be continuous. On the other hand, a continuous preference relation can be represented by a discontinuous utility function, if you like. (However, that is not convenient for mathematical issues.)

2.3 Utility Maximizing Problem

Apart from the general problem of choice theory, what makes consumer consumption decisions worthy of separate study is its particular structure that allows us to derive economically meaningful results. The structure arises because the consumer's choice sets are assumed to be defined by certain prices and the consumer's income or wealth. With this in mind, when the preference relation can be represented by a utility function $u(\cdot)$, we define the *consumer problem (CP)* as:

$$\begin{aligned} &\max_{\mathbf{x}\in\mathbb{R}^n_+} u(\mathbf{x}) \\ \text{s.t. } &\mathbf{p}\cdot\mathbf{x} \le m \\ &x_i \ge 0, \forall i = 1, 2, \cdots, n \end{aligned}$$

The idea is that, the consumer chooses a vector of goods $\mathbf{x} = (x_1, \cdots, x_n)' \ge 0$ to maximize her utility, subject to a budget constraint that given a price vector $\mathbf{p} = (p_1, \cdots, p_n)' \ge 0, \mathbf{p} \ne \mathbf{0}$, she cannot spend more than her total wealth m. In other words, the CP can be formulated as:

$$C(B(\mathbf{p}, m); \succeq) = \left\{ \mathbf{x} \mid \mathbf{x} \text{ solves } \max_{\mathbf{y}\in B(\mathbf{p},m)} u(\mathbf{y}) \right\}$$

2.3.1 Existence of Optimal Choice(s)

We are then natural to seek for the existence of an optimal choice for a utility-maximizing consumer. Luckily, the answer is yes.

Proposition Suppose that the consumer has a rational (complete and transitive) and continuous preference relation and makes rational decisions, then for $(\mathbf{p}, m) \gg \mathbf{0}$, they have (at least) one optimal choice.

Proof

- Since the consumer has a complete, transitive and continuous preference relation, then her preferences can be represented by a continuous utility function $u(\cdot)$.
- When $(\mathbf{p}, m) \gg \mathbf{0}$, the budget set $B(\mathbf{p}, m)$ is closed and bounded and hence compact.

- A continuous function on a compact set has at least one maximizer.

2.3.2 Further Assumptions

Now that existence of optimal choice is guaranteed (under rational and continuous preference relations, and $(\mathbf{p}, m) \gg \mathbf{0}$, we still need to answer:

- How to find the optimal choice(s)?
- In particular, what (additional) restrictions on the consumer's preference relation can simplify our analysis?

Locally Non-Satiation One possible simplification is to replace the inequality in the budget constraint with the equality: $\mathbf{p} \cdot \mathbf{x} = m$.

Intuitively, if a consumer thinks "more of a good is good", her preference relation should be monotone, and she would definitely exhaust all her wealth when maximizing her utility.

Definition (Monotonicity) A preference relation $\succeq$ on $X = \mathbb{R}_+^n$ is ***monotone*** if for any $\mathbf{x}, \mathbf{y} \in X$, we have:

- $\mathbf{x} \geq \mathbf{y} \Longrightarrow \mathbf{x} \succeq \mathbf{y}$;
- $\mathbf{x} \gg \mathbf{y} \Longrightarrow \mathbf{x} \succ \mathbf{y}$.

However, indeed the weaker condition of *locally non-satiated* preferences serves our purpose.

Definition (Locally Non-Satiation) A preference relation $\succeq$ on $X = \mathbb{R}_+^n$ is ***locally non-satiated*** if for any $\mathbf{x} \in X$ and $\varepsilon > 0$, there exists $\mathbf{y} \in Ball(\mathbf{x}, \varepsilon) \cap X$ such that $\mathbf{y} \succ \mathbf{x}$.

Intuitively, a preference relation is locally non-satiated if there is no *bliss* point. Moreover, a monotone preference relation must be locally non-satiated.

Proposition Suppose that the consumer has a complete, transitive, continuous and locally non-satiated preference relation and makes rational decisions, then for $(\mathbf{p}, m) \gg \mathbf{0}$, the budget constraint must hold with equality at any optimal choice $\mathbf{x}^*$, i.e., $\mathbf{p} \cdot \mathbf{x}^* = m$.

Proof

Intuitively, if the consumer does not exhaust her budget, then there must be a nearby affordable bundle which is strictly more preferred, by locally non-satiation.

- Consider any bundle $\mathbf{x} \in X = \mathbb{R}_+^n$ where $\mathbf{p} \cdot \mathbf{x} < m$. Since the preference relation $\succeq$ is locally non-satiated, for any $\varepsilon > 0$, there exists $\mathbf{y} \in Ball(\mathbf{x}, \varepsilon) \cap X$ such that $\mathbf{y} \succ \mathbf{x}$.
- We claim that $\mathbf{y} \in B(\mathbf{p}, m)$, i.e., the consumer's budget set, for $\varepsilon > 0$ small enough, so that $\mathbf{x}$ cannot be an optimal choice.
  - Let $\mathbf{y} = \mathbf{x} + \mathbf{z}$. By construction, $||\mathbf{z}|| < \varepsilon$. In particular, $|z_i| < \varepsilon$, for any $i = 1, 2, \cdots, n$. It follows that $\mathbf{p} \cdot \mathbf{y} = \mathbf{p} \cdot \mathbf{x} + \mathbf{p} \cdot \mathbf{z} \leq \mathbf{p} \cdot \mathbf{x} + n \cdot \varepsilon \cdot \max_i p_i$. As ε take appropriately small values, we finish our proof.

(Strict) Convexity Another possible simplification is the case of *unique* optimal choice. Notice that the consumer's budget set $B(\mathbf{p}, m)$ is convex. Based on this, when the utility function $u(\cdot)$ is *strictly quasi-concave*, there is a *unique* global maximizer. This translates into the following condition of strictly convex preference relations.

Definition (Convexity) A preference relation $\succeq$ on $X = \mathbb{R}_+^n$ is *convex* if for any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X$ such that $\mathbf{x} \succeq \mathbf{z}$ and $\mathbf{y} \succeq \mathbf{z}$, we have $t\mathbf{x} + (1-t)\mathbf{y} \succeq \mathbf{z}$, for any $t \in [0, 1]$. If $t\mathbf{x} + (1-t)\mathbf{y} \succ \mathbf{z}$ for any $\mathbf{x} \neq \mathbf{y}$ and $t \in (0, 1)$, then the preference relation is *strictly convex.*

Remarks:

1. Convexity can be equivalently defined as: For any $\mathbf{x}, \mathbf{y} \in X$ such that $\mathbf{y} \succeq \mathbf{x}$, we have $t\mathbf{y} + (1-t)\mathbf{x} \succeq \mathbf{x}$ for any $t \in [0, 1]$.
2. An equivalent way to describe convexity uses the indifference curves and *upper contour set* of choice bundles Upper Contour Set of $y = \{x \in X : x \succeq y\}$, graphically the area sitting upper-right above the indifference curve (included). Convexity of preferences amounts to the assumption that the upper contour set of any $y \in X$, is a *convex* set.
3. Convexity is fundamental in the standard model of competitive economics, because when consumer preferences are convex, market clearing prices exist; otherwise this may not exist. Convexity is also needed to be able to recover consumer preferences from choices from various budget sets. Convexity is often described as capturing the idea that the agent like diversity. However, whether convexity makes sense often depends on the interpretation of the goods space, in particular on the level of aggregation (e.g., over time or categories).

Proposition Suppose that the consumer has a complete, transitive, continuous and strictly convex preference relation and makes rational decisions, then for $(\mathbf{p}, m) \gg \mathbf{0}$, there is exactly one optimal choice.

Proof

The rational and continuous preference relation has guaranteed the existence of at least one optimal choice. Suppose to the contrary that the consumer has at least two optimal choices $\mathbf{x}^*$ and $\mathbf{y}^*$, and $\mathbf{x}^* \neq \mathbf{y}^*$, then by optimality $\mathbf{x}^* \sim \mathbf{y}^*$. Since the consumer's preference relation is strictly convex, construct a bundle $\mathbf{w} = \frac{1}{2}\mathbf{x}^* + \frac{1}{2}\mathbf{y}^*$, $\mathbf{w} \succ \mathbf{x}^* \sim \mathbf{y}^*$. Moreover, since the budget set $B(\mathbf{p}, m)$ is convex when $(\mathbf{p}, m) \gg \mathbf{0}$, so $\mathbf{w} \in B(\mathbf{p}, m)$. So neither $\mathbf{x}^*$ nor $\mathbf{y}^*$ can be an optimal choice, which is a contradiction.

Remarks:

Each of these (additional) properties of preferences has a corresponding property in a preference-representing utility function u.

Proposition Suppose the preference relation $\succeq$ on X can be represented by $u : X \to \mathbb{R}$. Then,

1. $\succeq$ is monotone if and only if u is non-decreasing.
2. $\succeq$ is locally non-satiated if and only if u has no local maxima in X.
3. $\succeq$ is (strictly) convex if and only if u is (strictly) quasi-concave.

2.3.3 Marshallian Demand & Indirect Utility Function

From Consumer Problem, we define the optimal value function as the *indirect utility function*:

$$v(\mathbf{p}, m) = \sup_{\mathbf{x} \in B(\mathbf{p}, m)} u(\mathbf{x})$$

Remarks: Here we rigorously use "sup" to define $v(\mathbf{p}, m)$ instead of "max", because sometimes $u(\mathbf{x})$ does not behave well.

We define the consumer's optimal choice(s) as the *Marshallian demand correspondence*:

$$\mathbf{x}^M(\mathbf{p}, m) = \{\mathbf{x} \in B(\mathbf{p}, m) : u(\mathbf{x}) = v(\mathbf{p}, m)\}$$

`$\x^M(\p,m)$ is in general a set of utility-maximizing consumption bundles.`

Assume throughout that the consumer has a rational preference relation $\succeq$ and makes rational decisions that $(\mathbf{p}, m) \gg \mathbf{0}$. Then Marshallian demand and indirect utility function must satisfy:

- Existence of optimal choice(s): If $\succeq$ is continuous, then the Marshallian demand $\mathbf{x}^M(\mathbf{p}, m) \neq \emptyset$.
- Structure of Marshallian demand: If $\succeq$ is convex, then $\mathbf{x}^M(\mathbf{p}, m)$ is a convex set. If $\succeq$ is strictly convex, then $\mathbf{x}^M(\mathbf{p}, m)$ Is a singleton.
- Homogeneity: Both $v(\mathbf{p}, m)$ and $\mathbf{x}^M(\mathbf{p}, m)$ are homogeneous of degree 0 in $(\mathbf{p}, m)$, that is, for any $t > 0$, $v(t\mathbf{p}, tm) = v(\mathbf{p}, m)$ and $\mathbf{x}^M(t\mathbf{p}, tm) = \mathbf{x}^M(\mathbf{p}, m)$.

  - Proof uses the fact that $B(\lambda p, \lambda w) = B(p, w)$, and so the consumer solves the *same* problem.

- Monotonicity of $v(\mathbf{p}, m)$: $v(\mathbf{p}, m)$ is non-increasing in $\mathbf{p}$ and non-decreasing in m. If $\succeq$ is locally non-satiated, then $v(\mathbf{p}, m)$ is strictly increasing in m.

- Walras' Law: If $\succeq$ is locally non-satiated, then $\mathbf{p} \cdot \mathbf{x} = m$, for any $\mathbf{x} \in \mathbf{x}^M(\mathbf{p}, m)$.

2.3.4 Derivation of Utility Maximization Problem

The consumer's utility maximization problem is give by:

$$\begin{aligned} &\max_{\mathbf{x} \in \mathbb{R}^n_+} u(\mathbf{x}) \\ \text{s.t. } &\mathbf{p} \cdot \mathbf{x} \le m \\ &x_i \ge 0, \forall i = 1, 2, \cdots, n \end{aligned}$$

For $(\mathbf{p}, m) \gg \mathbf{0}$, the constraint qualification is always satisfied, and we can apply the KKT necessary conditions when $u(\cdot)$ is continuously differentiable.

The Lagrangian is given by:

$$\mathcal{L}(\mathbf{x}, \lambda, \ ) = u(\mathbf{x}) + \lambda \left(m - \sum_{i=1}^{n} p_i x_i \right) + \sum_{i=1}^{n} \mu_i x_i$$

`where $\l$ is the Lagrange multiplier on the budget constraint and, for each $i$, $\mu`

$$\max_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda, \ ) = \max_{\mathbf{x}} u(\mathbf{x}) + \lambda \left(m - \sum_{i=1}^{n} p_i x_i \right) + \sum_{i=1}^{n} \mu_i x_i$$

The first-order conditions are given by:

- w.r.t. $x_i : \frac{\partial \mathcal{L}}{\partial x_i} = \frac{\partial u}{\partial x_i} - \lambda p_i + \mu_i = 0.$
- Inequality constraints: $m - \sum_{i=1}^{n} p_i x_i \ge 0, x_i \ge 0, \lambda \ge 0, \mu_i \ge 0.$
- Complementary slackness: $\lambda \left(m - \sum_{i=1}^{n} p_i x_i \right) = 0, \mu_i x_i = 0.$

Remarks: Lagrangian and First-Order Conditions

- In contrast to equality constraints, the direction of each inequality constraint affects the way we set up the Lagrangian.
  - Intuitively, we make sure the way each multiplier is non-negative and pelanize constraint violations. For instance, when the budget constraint is violated, that is, $m - \sum_{i=1}^{n} p_i x_i < 0$, the value of the Lagrangian strictly decreases.

- Despite little economic meaning, λ represents the *shadow price* of wealth, that is, *marginal utility of an additional unit of wealth*.
  - However, note that utility only has ordinal meanings. Nothing in the consumer theory developed so far suggests any basis for using the shadow price of wealth, as defined, to guide redistribution policies.
- If the preference relation is well-behaved (i.e., non-satiated and strictly convex) and the non-negativity constraints are not binding, then $\frac{\partial u}{\partial x_i} = \lambda p_i$, and we are back to the familiar "tangency conditions", that is, for all i, j:

$$MRS_{ij} = \frac{\frac{\partial u}{\partial x_i}}{\frac{\partial u}{\partial x_j}} = \frac{p_i}{p_j}$$

  The "tangency conditions" say that at the consumer's maximum, relative marginal utility of any two choices equals their relative price, that is, all possible inner "gains from trade" has been realized (thus no more inner "gain from trade").

To facilitate the derivation of the consumer's problem, we should **first check whether the preference relation is locally non-satiated and strictly convex**.

- Case 1: If *locally non-satiated* and *strictly convex*, then we proceed in two steps
  1. Directly apply the "tangency conditions": $\frac{MU_i}{p_i} = \frac{MU_j}{p_j} = \lambda$, for any i, j.
  2. Check the *non-negativity constraints* and apply the complementary slackness condition(s) if necessary.

     Note that in the previous step we first assume that the non-negativity holds and we have interior solution.
- Case 2: If neither locally non-satiated nor strictly convex, then use logic or economic intuition to tackle the problem.

2.3.5 Example

Cobb-Douglas Utility Suppose that the consumer's preference relation can be represented by the following utility function:

$$u(x_1, x_2, x_3) = (x_1 + a)(x_2 + b)(x_3 + c)$$

`where $a,b,c\ge 0$ are non-negative constants. Moreover, the consumer faces constant p`

1. First suppose that $a = b = c = 0$. Solve for the consumer's Marshallian demand correspondence $\mathbf{x}^M(x_1, x_2, x_3)$ and indirect utility function $v(p_1, p_2, p_3, m)$.

   Useful note (Cobb-Douglas Utility):
   For a Cobb-Douglas utility representation for a preference relation $\succeq$,

$$u(x_1, x_2, \cdots, x_n) = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$$

   Then the consumer would spend her income on each good according to its share, that is,

$$x_i = \frac{\alpha_i}{\sum_{j=1}^{n} \alpha_i} \cdot \frac{m}{p_i} \iff p_i x_i = \frac{\alpha_i}{\sum_{j=1}^{n} \alpha_i} \cdot m$$

   - Step 1:Apply the positive monotonic transformation for computation-convenience: $v = \ln u$.
   - Step 2: Check the locally non-satiation and convexity of preference relation for simplification of optimal solution.
     - Check that the preference relation is monotonic and hence locally non-satiated.
     - Check that the preference relation is strictly convex, that is, the utility function is strictly quasi-concave.
   - Step 3: Now that both non-satiation and convexity are satisfied, apply the "tangency conditions":

$$\frac{MU_1}{p_1} = \frac{MU_2}{p_2} = \frac{MU_3}{p_3}$$

   - Step 4: Together with the budget constraint, so we have the Marshallian demand correspondence $\mathbf{x}^M(\mathbf{p}, m)$.

$$\mathbf{x}^M(p_1, p_2, p_3, m) = \left(\frac{m}{3p_1}, \frac{m}{3p_2}, \frac{m}{3p_3} \right) \geq \mathbf{0}$$

2. Next suppose that $a, b, c > 0$. Solve for the consumer's Marshallian demand correspondence $\mathbf{x}^M(p_1, p_2, p_3, m)$ and indirect utility function $v(p_1, p_2, p_3, m)$.

The first three steps are quite similar to the previous part and thus omitted. In step 4, this time we have

$$\mathbf{x}_1^M(p_1,p_2,p_3,m) = \frac{m+p_1a+p_2b+p_3c}{3p_1} - a$$
$$\mathbf{x}_2^M(p_1,p_2,p_3,m) = \frac{m+p_1a+p_2b+p_3c}{3p_2} - b$$
$$\mathbf{x}_3^M(p_1,p_2,p_3,m) = \frac{m+p_1a+p_2b+p_3c}{3p_3} - c$$

Note that depending on the parameter values, we may or may not have $\mathbf{x}^M(\mathbf{p},m) \geq \mathbf{0}$. In other words, the non-negativity constraints may be binding and are for us to check. For simplicity, suppose that

$$2p_1a - p_2b - p_3c \geq 2p_2b - p_1a - p_3c \geq 2p_3c - p_1a - p_2b$$

The other five symmetric cases are similar. In our assumption, $p_1a \geq p_2b \geq p_3c$.

1. $m \geq 2p_1a - p_2b - p_3c$, then the non-negativity constraints are not binding, and we have the interior solution described above.
2. $2p_1a - p_2b - p_3c \geq m$, then at optimum, $x_1^* = 0$. The constraint $x_1^M \geq 0$ is binding, so the consumer only purchases goods 2 and 3, and we have

$$\mathbf{x}^M(p_1,p_2,p_3,m) = \left(0, \frac{m+p_2b+p_3c}{2p_2} - b, \frac{m+p_2b+p_3c}{2p_3} - c\right)$$

   1. $m \geq p_2b - p_3c$, then $\mathbf{x}^M$ is as above.
   2. $p_2b - p_3c > m$, then the two constraints $x_1^M \geq 0$ and $x_2^M \geq 0$ are both binding, so the consumer only purchases good 3, and

$$\mathbf{x}^M(p_1,p_2,p_3,m) = \left(0, 0, \frac{m}{p_3} - c\right)$$

Inter-Temporal Choice You have a saving $s > 0$ to spend for this year and next year. Since you are now in graduate school, you will not earn any additional income over the two years. Suppose your utility is time-separable and is given by $v(c_1,c_2) = u(c_1) + \beta u(c_2)$, where $u : \mathbb{R}_+ \to \mathbb{R}$ is a twice continuously differentiable function with $u'(c) > 0$ and $u''(c) < 0$ for any $c \in \mathbb{R}_+$ and $0 < \beta < 1$ is the discount factor. Further suppose that the going interest rate is $0 < r < 1$, which will remain constant. Let c_1^* and c_2^* be your optimal consumption choices for this year and next year.

1. Give a (necessary and sufficient) condition for $(c_1^*, c_2^*) \gg \mathbf{0}$.

The utility maximization problem is given by:

$$\begin{aligned}
&\max_{c_1,c_2} v(c_1,c_2) = u(c_1) + \beta u(c_2)\\
\text{s.t. } &c_1 + \frac{c_2}{1+r} \leq s\\
&c_1, c_2 \geq 0
\end{aligned}$$

Since $u(\cdot)$ is strictly increasing and strictly concave, $v(\cdot)$ is also strictly increasing and strictly concave (in particular, strictly quasi-concave), and we can directly apply the ”tangency condition”:

$$\frac{u'(c_1)}{1} = \frac{\beta u'(c_2)}{\frac{1}{1+r}}$$

By strictly monotonicity, $c_1 + \frac{c_2}{1+r} = s$. For $(c_1^*, c_2^*) \gg \mathbf{0}$, we need:

$$\frac{u'(s)}{u'(0)} < \beta(1+r) < \frac{u'(0)}{u'\left((1+r)s\right)}$$

1. Suppose that $(c_1^*, c_2^*) \gg \mathbf{0}$, compare c_1^* with c_2^*.

At the optimum,

$$\frac{u'(c_1)}{u'(c_2)} = \beta(1+r)$$

- If $\beta(1+r) > 1$, then $\frac{u'(c_1)}{u'(c_2)} > 1$ and $c_1^* < c_2^*$.
- If $0 < \beta(1+r) < 1$, then $\frac{u'(c_1)}{u'(c_2)} < 1$ and $c_1^* > c_2^*$.
- If $\beta(1+r) = 1$, then $\frac{u'(c_1)}{u'(c_2)} = 1$ and $c_1^* = c_2^*$.

Intuitively, $\beta(1+r)$ represents how your tradeoff between next year and this year compares with that of the market when $c_1 = c_2$.

Utility Representation: Inter-Temporal Problem Let $X = \mathbb{R}_+ \times \mathbb{N}$, where (x,t) is interpreted as receiving x yuan at time t. Consider the following six properties of preference relations on X:

- Rationality (completeness and transitivity).
- Continuity.
- There is indifference between receiving 0 yuan at time 0 and receiving 0 yuan at any other time.
- It is (strictly) better to receive any positive amount of money as soon as possible.
- Money is always desirable.
- The preference between (x,t) and $(y,t+1)$ is independent of t.

Consider the following questions:

1. Use precise mathematical language to formally define the six properties.

2. Suppose a preference relation on X can be represented by the utility function $v(x,t) = u(x)\beta^t$, where $0 < \beta < 1$ and $u(\cdot)$ is continuous, strictly increasing and $u(0) = 0$. Check whether this preference relation satisfies each of the six properties.

3. Suppose a preference relation on X satisfies all of the six properties. Show that this preference relation must admit a utility representation.

4. Use precise mathematical language to formalize the idea that "one preference is more patient than another".

5. Based on your definition in part (4), prove or disprove the following statement: A preference relation represented by $v_1(x,t) = u_1(x)\beta_1^t$ is more patient than another preference relation represented by $v_2(x,t) = u_2(x)\beta_2^t$ if $0 < \beta_2 < \beta_1$ (where $u_1(\cdot)$ and $u_2(\cdot)$ are both continuous, strictly increasing and $u_1(0) = u_2(0) = 0$).

Solution

1. Mind yourself that the time t is not continuous here, so take caution when you try to define a "limit" with regard to t.

   - Continuity: For any $t, t' \in \mathbb{N}$, and any pair of sequences $\{x(n)\}_{n=1}^{\infty}$ and $\{y(n)\}_{n=1}^{\infty}$ from $\mathbb{R}^+$ with $x(n) \to x^*$, $y(n) \to y^*$. If $(x(n), t) \succeq (y(n), t')$ for all n, we have $(x^*, t) \succeq (y^*, t')$.

   - The preference between (x,t) and $(y, t+1)$ is independent of t: For any $x, y \in \mathbb{R}^+$, and $t, t' \in \mathbb{N}$, we have $(x,t) \succeq (y, t+1) \iff (x, t') \succeq (y, t'+1)$.

2. A continuous preference relation can be represented by a discontinuous function. However, if a preference relation can be represented by a continuous function, then the preference relation must be continuous.

3. Recall the proof in our lecture of existence of utility representation for any rational and continuous preference relation.

   - Claim 1: For any pair (x,t), there is a unique number $u(x,t) \in \mathbb{R}^+$ such that $(x,t) \sim (u(x,t), 0)$.

     Proof: First, by indifference when receiving nothing, we have $(0,t) \sim (0,0)$. For any pair (x,t), we have $(x,t) \succeq (0,0)$ and $(x+1, t) \succeq (x,t)$. Then by continuity, there is y for which $(x,t) \sim (y, 0)$, and we then define $u(x,t) := y$.

   - Claim 2: The preference relation is represented by $u(x,t)$.

     By claim 1 and property that money is more desirable:

$$u(x,t) \geq u(y,t') \iff (u(x,t),0) \succeq (u(y,t'),0) \iff (x,t) \succeq (y,t')$$

4. The definition should be clearly based on preference relations and try to be somewhat math-irrelevant.

   $\succeq_1$ is more patient than $\succeq_2$ if for any (x,t) and any (y,t') with $t' > t$, $y > x$:

   $$(y,t') \succeq_2 (x,t) \Longrightarrow (y,t') \succeq_1 (x,t)$$

   The definition means that if I prefer to wait from t to t' under $\succeq_1$, then I will also prefer to wait from t to t' when I'm more patient (say under $\succeq_2$).

5. Intuitively, if the two preference relations value money differently at the baseline level (i.e., simply in terms of money), they would generate different preference over combinations of money and receiving time.

Utility Representation: Linear Case Let $\succeq$ be a rational (complete and transitive) preference relation on $X = \mathbb{R}^2_+$. Consider the following three properties:

- Additivity: If $(x_1,x_2) \succeq (y_1,y_2)$, then for any t,s such that $(x_1+t,x_2+s),(y_1+t,y_2+s) \in \mathbb{R}^2_+$, $(x_1+t,x_2+s) \succeq (y_1+t,y_2+s)$.
- Strong monotonicity: If $x_1 \geq y_1$ and $x_2 \geq y_2$, then $(x_1,x_2) \succeq (y_1,y_2)$. If in addition, $x_1 > y_1$ or $x_2 > y_2$, then $(x_1,x_2) \succ (y_1,y_2)$.
- Standard continuity: For any two sequences $\{\mathbf{x}_n\}_{n=1}^{\infty}$ and $\{\mathbf{y}_n\}_{n=1}^{\infty}$, if $\mathbf{x}_n \succeq \mathbf{y}_n$ for any n, and $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}^*$ and $\lim_{n\to\infty} \mathbf{y}_n = \mathbf{y}^*$, then $\mathbf{x}^* \succeq \mathbf{y}^*$.

Consider the following questions:

1. Show that if $\succeq$ has a linear utility representation, i.e., $u(x_1,x_2) = ax_1 + bx_2$, for some $a,b > 0$, then this preference relation satisfies the above three properties.
2. Show that these three properties are necessary for the preference relation $\succeq$ to have a linear utility representation, i.e., show that for any pair of the three properties, there is a preference relation that does not satisfy the third property.
3. Show that if $\succeq$ satisfies the three properties, then this preference relation admits a linear utility representation, i.e., there exists $a,b > 0$ such that $u(x_1,x_2) = ax_1 + bx_2$, for any $(x_1,x_2) \in \mathbb{R}^2_+$. (Hint: Think about the indifference curves/sets of this preference relation.)

Solution

1. Easy to verify. Notice that if a preference relation can be represented by a continuous function, then the preference relation must be continuous.

2. This question means the three properties are "parallel" from a preference relation to have a linear utility representation.

   - (i)(ii)$\overset{\times}{\Longrightarrow}$(iii): The lexicographic preference: $(x_1, x_2) \succeq (y_1, y_2)$ if either $x_1 > x_2$ or $x_1 = x_2$ and $y_1 > y_2$.
   - (i)(iIi)$\overset{\times}{\Longrightarrow}$(ii): $u(x_1, x_2) = x_1 - x_2$; $u(x_1, x_2) = x_1 - \frac{1}{x_2}$.
   - (iI)(iIi)$\overset{\times}{\Longrightarrow}$(i): $u(x_1, x_2) = x_1^2 + x_2^2$; $u(x_1, x_2) = x_1$.

3. Starting from possible intuitions from linear utility representation, we need to establish the following two properties of the indifference curve:

   - Property 1: The indifference curves are linear.
   - Property 2: The indifference curves are parallel, downward sloping and not thick.

   In order to establish the two properties, we first prove the following two lemmas:

   - Lemma 1: For $\mathbf{x} \neq \mathbf{y}$, if $\mathbf{x} \sim \mathbf{y}$, then for $\mathbf{z} = \frac{1}{2}\mathbf{x} + \frac{1}{2}\mathbf{y}$ and $\mathbf{z}' = 2\mathbf{y} - \mathbf{x}$, we have $\mathbf{x} \sim \mathbf{y} \sim \mathbf{z} \sim \mathbf{z}'$.
   - Lemma 2: For $\mathbf{x} \neq \mathbf{y}$, if $\mathbf{x} \sim \mathbf{y}$, then for any $w = t\mathbf{x} + (1-t)\mathbf{y}, 0 \leq t \leq 1$, we have $\mathbf{x} \sim \mathbf{y} \sim w$.

   Here we omit the technical proofs and move on with establishment of property 1. Pick any point on the horizontal axis $\mathbf{x} = (x, 0), x > 0$. By the proof of the utility representation in lecture and strong monotonicity, $\exists 0 < w < x$ such that $\mathbf{x} \sim w = we = (w, w)$. Connect $\mathbf{x}$ and w and extend it to the vertical axis. Denote the intersection of the ray $\mathbf{x}w$ with the vertical axis as $\mathbf{y} = (0, y)$. Jointly from Lemma 1 and 2 we can say the points on the line $\mathbf{xy}$ are indifferent to each other. Finally, by strong monotonicity, the indifference curves must be downward sloping and for any $x \neq x'$, we cannot have $(x, 0) \sim (x', 0)$. Moreover, if $(x, 0) \sim (w, w)$, then for any $t \geq -w$, we have $(x + t, 0) \sim (w + t, w)$, so the indifference curves are parallel.

2.4 Expenditure Minimization Problem

In order to disentangle the price effect and the income effect, we introduce the following expenditure minimization problem, holding the consumer's utility at/above a certain level.

$$
\begin{aligned}
&\min_{\mathbf{x}} \mathbf{p} \cdot \mathbf{x} \\
\text{s.t. } & u(\mathbf{x}) \geq u \\
& x_i \geq 0, \forall i = 1, 2, \cdots, n
\end{aligned}
$$

Let $F(\mathbf{p}, u) = \{\mathbf{x} \in \mathbb{R}^n_+ : u(\mathbf{x}) \geq u\}$ be the feasible set. We define the optimal (minimal) value function as the ***expenditure function***:

$$e(\mathbf{p}, u) = \inf_{\mathbf{x} \in F(\mathbf{p}, u)} \mathbf{p} \cdot \mathbf{x}$$

```
and the consumer's optimal choice(s) as the ***Hicksian demand correspondence***:
```

$$\mathbf{x}^H(\mathbf{p}, u) = \{\mathbf{x} \in F(\mathbf{p}, u) : \mathbf{p} \cdot \mathbf{x} = e(\mathbf{p}, u)\}$$

Since the consumer's utility is held at/above a certain level, when there is a price change, **the change in the Hicksian demand only captures the substitution effect**.

Similar to the analysis of the utility maximization problem, before deriving the optimal solution(s), we approach the expenditure minimization problem from two aspects:

- When can we simplify the problem (e.g., existence of solution(s), binding utility level and uniqueness of solution)?
- Properties of expenditure function and Hicksian demand correspondence.

2.4.1 Existence of Solution

Proposition Suppose $u(\cdot)$ represents a **continuous** preference relation and that $\mathbf{p} \gg \mathbf{0}$ and $F(u) \neq \emptyset$, then the expenditure minimization problem has (at least) one minimizer, that is, $\mathbf{x}^H(\mathbf{p}, u) \neq \emptyset$.

Proof

- Pick any $\mathbf{x}_0 \in F(u)$ and consider the alternative feasible set:

  $$\tilde{F} = \{\mathbf{x} \in \mathbb{R}^n_+ : u(\mathbf{x}) \geq u \text{ and } \mathbf{p} \cdot \mathbf{x} \leq \mathbf{p} \cdot \mathbf{x}_0\}$$

  It is easy to see that the expenditure minimization problem with the feasible set $\tilde{F}$ has the same solution as the original problem. Moreover, $\tilde{F}$ is closed and bounded, and hence compact.

  - Boundedness is crafted by picking $\mathbf{x}_0$ manually but without loss of generality.
  - Closedness is guaranteed by continuity of $u(\cdot)$.
    * Notice that $u(\cdot)$ need not be continuous, but we can find an alternative continuous function $v(\cdot)$ that represents the same preference relation, so $\tilde{F}$ is closed.
    * Note that the continuity of preference relation has no direct relation to the continuity of its utility representation.)

- The objective function $\mathbf{p}\cdot\mathbf{x}$ is continuous and we know that a continuous function on a compact set has at least one minimizer. That ends our proof.

2.4.2 Biding Utility Level

Proposition Suppose $u(\cdot)$ represents a **continuous** preference relation and that $\mathbf{p} \gg \mathbf{0}$, $u \geq u(\mathbf{0})$ and $F(u) \neq \emptyset$, then at any minimizer $\mathbf{x}^*$, we have $u(\mathbf{x}^*) = u$.

Proof

- By the argument in the previous proposition, we can assume without loss that $u(\cdot)$ is continuous. Suppose to the contrary that a minimizer $\mathbf{x}^*$, we have $u(\mathbf{x}^*) > u$. Since $u \geq u(\mathbf{0})$, $\mathbf{x}^* \neq \mathbf{0}$.
- By the continuity of $u(\cdot)$, we know for some $\varepsilon > 0$. $u\left((1-\varepsilon)\mathbf{x}^*\right) > u$. It follows that $(1-\varepsilon)\mathbf{x}^* \in F(u)$ and that $\mathbf{p}\cdot(1-\varepsilon)\mathbf{x}^* < \mathbf{p}\cdot\mathbf{x}^*$, which is a contradiction to the optimality of $\mathbf{x}^*$.

Remarks: The binding condition here in EMP is much weaker than that in UMP, where we do not put any additional condition on preference relation. One can understand this as the objective function in EMP is itself locally non-satiated.

2.4.3 Unique Minimizer

Proposition Suppose $u(\cdot)$ represents a **continuous** and **strictly convex** preference relation and that $\mathbf{p} \gg \mathbf{0}$ and $F(u) \neq \emptyset$, then the expenditure minimization problem has exactly one minimizer, that is, $\mathbf{x}^H(\mathbf{p}, u)$ is a singleton (single-valued).

Proof

- By the first proposition, *at least one minimizer exists*.
- Suppose to the contrary that there are two minimizers $\mathbf{x}^* \neq \mathbf{y}^*$. Then by feasibility, $u(\mathbf{x}^*) \geq u$ and $u(\mathbf{y}^*) \geq u$. Since the preference relation is strictly convex, $u(\cdot)$ is strictly quasi-concave, so for the bundle $\mathbf{w} = \frac{1}{2}\mathbf{x}^* + \frac{1}{2}\mathbf{y}^* \neq \mathbf{0}$, we have $u(\mathbf{w}) > \min\{u(\mathbf{x}^*), u(\mathbf{y}^*)\} \geq u$. By continuity, for some small $\varepsilon > 0$, $u((1-\varepsilon)\mathbf{w}) \geq u$. Moreover, $\mathbf{p}\cdot(1-\varepsilon)\mathbf{w} = \frac{1}{2}(1-\varepsilon)\mathbf{p}\cdot\mathbf{x}^* + \frac{1}{2}(1-\varepsilon)\mathbf{p}\cdot\mathbf{y}^* < \mathbf{p}\cdot\mathbf{x}^* = \mathbf{p}\cdot\mathbf{y}^*$, which is a contradiction.

2.4.4 Summary of Properties

Suppose $\mathbf{p} \gg \mathbf{0}$ and $F(u) \neq \emptyset$.

- Existence of minimizer: If $u(\cdot)$ represents a continuous preference relation and that, then the Hicksian demand $\mathbf{x}^H(\mathbf{p}, u) \neq \emptyset$.
- Structure of Hicksian demand: If $u(\cdot)$ represents a convex preference relation, then $\mathbf{x}^H(\mathbf{p}, u)$ is a convex set. If $u(\cdot)$ represents a continuous and strictly convex

preference relation, then $\mathbf{x}^H(\mathbf{p}, u)$ is a singleton.

- Homogeneity: $e(\mathbf{p}, u)$ is homogenous of degree 1 in $\mathbf{p}$, that is, for any $t > 0$, $e(t\mathbf{p}, u) = te(\mathbf{p}, u)$.
- Monotonicity of $e(\mathbf{p}, u)$: $e(\mathbf{p}, u)$ is non-decreasing in $\mathbf{p}$ and u. If $u(\cdot)$ represents a continuous preference relation, then $e(\mathbf{p}, u)$ is strictly increasing in u when $u \geq u(\mathbf{0})$.
- Binding utility level: Suppose $u(\cdot)$ represents a continuous preference relation and $u \geq u(\mathbf{0})$, then at any minimizer $\mathbf{x}^*$, $u(\mathbf{x}^*) = u$.

Since $\min \mathbf{p} \cdot \mathbf{x}$ is equivalent to $\max -\mathbf{p} \cdot \mathbf{x}$, EMP can be solved in an analogous manner to UMP. EMP and UMP share the same "tangency condition" for interior solutions:

$$\frac{MU_i}{MU_j} = \frac{p_i}{p_j}$$

2.4.5 Example

Suppose a consumer's preference relation can be represented by the following utility function:

$$u(x_1, x_2) = \ln x_1 + x_2$$

Moreover, the consumer faces constant prices $(p_1, p_2) \gg \mathbf{0}$ and has income $m > 0$.

1. Solve the consumer's utility maximization problem to derive the Marshallian demand $\mathbf{x}^M(p_1, p_2, m)$ and indirect utility function $v(p_1, p_2, m)$.

It's easy to check that the preference relation is monotonic and strictly convex. The utility miximization can then be simplified as:

$$\max_{x_1, x_2 \geq 0} u(x_1, x_2) = \ln x_1 + x_2 \quad \text{s.t.} \quad p_1 x_1 + p_2 x_2 = m$$

The "tangency condition" is: $|MRS| = \frac{MU_1}{MU_2} = \frac{1}{x_1} = \frac{p_1}{p_2}$.

Together with the binding budget constraint, we have $x_1^* = \frac{p_2}{p_1}$ and $x_2^* = \frac{m}{p_2} - 1$.

Notice the constraint $x_2 \geq 0$ may be binding, so the Marshallian demand is given by

$$\mathbf{x}^M(p_1, p_2, m) = \begin{cases} \left(\frac{p_2}{p_1}, \frac{m}{p_2} - 1\right), & \text{if } m \geq p_2 \\ \left(\frac{m}{p_1}, 0\right), & \text{if } 0 < m < p_2 \end{cases}$$

1. Solve the consumer's expenditure minimization problem to derive the Hicksian demand $\mathbf{x}^H(p_1, p_2, u)$ and expenditure function $e(p_1, p_2, u)$.

Similar to the previous part, the expenditure minimization problem can be simplified as:

$$\min_{x_1,x_2\geq 0} p_1x_1+p_2x_2 \quad \text{s.t. } u(x_1,x_2)=\ln x_1+x_2=u$$

The "tangency condition" is $|MRS|=\frac{MU_1}{MU_2}=\frac{1}{x_1}=\frac{p_1}{p_2}$.

Together with binding utility level, we have $x_1^*=\frac{p_2}{p_1}$ and $x_2^*=u-\ln\frac{p_2}{p_1}$.

Notice the constraint $x_2\geq 0$ may be binding, so the Hicksian demand is given by:

$$\mathbf{x}^H(p_1,p_2,m)=\begin{cases}\left(\frac{p_2}{p_1},u-\ln\frac{p_2}{p_1}\right), \text{ if } u\geq \ln\frac{p_2}{p_1}\\ (e^u,0), \text{ if } u<\ln\frac{p_2}{p_1}\end{cases}$$

Caution: Remember to check monotonicity and quasi-convexity beforehand!

2.5 Duality and Comparative Statics

2.5.1 Duality between UMP and EMP

Consumer's utility maximization problem (UMP):

$$\begin{aligned}&\max_{\mathbf{x}} u(\mathbf{x})\\ \text{s.t. } &\mathbf{p}\cdot\mathbf{x}\leq m\\ &x_i\geq 0, \forall i=1,2,\cdots,n\end{aligned}$$

Consumer's expenditure minimization problem (EMP):

$$\begin{aligned}&\min_{\mathbf{x}} \mathbf{p}\cdot\mathbf{x}\\ \text{s.t. } &u(\mathbf{x})\geq u\\ &x_i\geq 0, \forall i=1,2,\cdots,n\end{aligned}$$

Notice that the objective function in the UMP is precisely the constraint in the EMP and vice versa. In mathematics, the two optimization problems are called *dual problems*. By duality, Marshallian and Hicksian demand have a special relationship.

Proposition Suppose $u(\cdot)$ is a utility function that represents a continuous and locally non-satiated preference relation on $X=\mathbb{R}^n_+$, then for any $\mathbf{p}\gg\mathbf{0}$, we have:

1. For any $m\geq 0$, $\mathbf{x}^M(\mathbf{p},m)=\mathbf{x}^H(\mathbf{p},v(\mathbf{p},m))$ and $e(\mathbf{p},v(\mathbf{p},m))=m$.
2. For any $u\geq u(\mathbf{0})$, $\mathbf{x}^H(\mathbf{p},u)=\mathbf{x}^M(\mathbf{p},e(\mathbf{p},u))$ and $v(\mathbf{p},e(\mathbf{p},u))=u$.

Proof

- Fix $m > 0$ and any $\mathbf{x}_0^M \in \mathbf{x}^M(\mathbf{p}, m)$, we have:

$$e(\mathbf{p}, v(\mathbf{p}, m)) \leq \mathbf{p} \cdot \mathbf{x}_0^M = m$$

  Fix any $u \geq u(\mathbf{0})$ and any $\mathbf{x}_0^H \in \mathbf{x}^H(\mathbf{p}, u)$, we have

$$v(\mathbf{p}, e(\mathbf{p}, u)) \geq u(\mathbf{x}_0^H) = u$$

- Applying the first inequality to the wealth level $m = e(\mathbf{p}, u)$, we have:

$$e(\mathbf{p}, v(\mathbf{p}, e(\mathbf{p}, u))) \leq e(\mathbf{p}, u)$$

  On the other hand, since the preference relation is continuous, $e(\mathbf{p}, u)$ is strictly increasing in u, so from the second inequality, we have

$$e(\mathbf{p}, v(\mathbf{p}, e(\mathbf{p}, u))) \geq e(\mathbf{p}, u)$$

- Again by strict monotonicity of $e(\mathbf{p}, u)$ in u, we have $v(\mathbf{p}, e(\mathbf{p}, u)) = u$. Similarly, $e(\mathbf{p}, v(\mathbf{p}, m)) = m$.
- Finally, since the preference relation is continuous and locally non-satiated, the budget constraint must bind at the UMP and the utility level must bind at the EMP. Correspondingly,

$$\begin{aligned} \mathbf{x}^M(\mathbf{p}, m) = \{\mathbf{x} \in \mathbb{R}_+^n : \mathbf{p} \cdot \mathbf{x} = m, u(\mathbf{x}) = v(\mathbf{p}, m)\} = \mathbf{x}^H(\mathbf{p}, v(\mathbf{p}, m)) \\ \mathbf{x}^H(\mathbf{p}, u) = \{\mathbf{x} \in \mathbb{R}_+^n : u(\mathbf{x}) = u, \mathbf{p} \cdot \mathbf{x} = e(\mathbf{p}, u)\} = \mathbf{x}^M(\mathbf{p}, e(\mathbf{p}, u)) \end{aligned}$$

Remarks: Intuitively, their relationship implies that, at the "right" level of wealth and utility, the Marshallian demand correspondence is identical to the Hicksian demand correspondence.

2.5.2 Envelope Theorem

Unconstrained Optimization Let $f : X \times [\underline{\theta}, \bar{\theta}] \to \mathbb{R}$ and $V(\theta) = \sup_{\mathbf{x} \in X} f(\mathbf{x}, \theta)$.

Theorem Suppose $f(\mathbf{x}, \cdot)$ is differentiable in θ for all $\mathbf{x} \in X$. Moreover, there exists an integrable function $b : [\underline{\theta}, \bar{\theta}] \to \mathbb{R}_+$ such that $|f_\theta(\mathbf{x}, \theta)| \leq b(\theta)$ for all $\mathbf{x} \in X$ and almost all $\theta \in [\underline{\theta}, \bar{\theta}]$. Then $V(\cdot)$ is absolutely continuous and hence differentiable (a.e.). In addition, for any $\mathbf{x}^* \in \arg\max_{\mathbf{x} \in X} f(\mathbf{x}, \theta)$,

$$V'(\theta) = f_2(\mathbf{x}^*(\theta), \theta)$$

Remarks:

- The intuition is that, only the direct effect matters; the indirect effect of $\mathbf{x}^*(\theta)$ on $V(\theta)$ can be ignored, since $f(\cdot)$ does not change with $\mathbf{x}$ at the optimum.
- May get some inspiration from the simplest version.
  - Suppose x is one-dimensional and that the optimizer is unique and differentiable in θ, combined with F.O.C., we would have:

$$V'(\theta) = f_1(x^*(\theta), \theta) \cdot (x^*)'(\theta) + f_2(x^*(\theta), \theta) = f_2(x^*(\theta), \theta)$$

  - For the envelope theorem to hold, $\mathbf{x}$ need not be one-dimensional, $f(\cdot, \theta)$ need not be differentiable in $\mathbf{x}$, and the optimal $\mathbf{x}^*(\theta)$ need not be unique or differentiable in θ.

2.5.3 Constrained Optimization

Let $f, g : X \times [\underline{\theta}, \bar{\theta}] \to \mathbb{R}$ and $V(\theta) = \sup_{\mathbf{x} \in X : g(\mathbf{x}, \theta) \geq 0} f(\mathbf{x}, \theta)$. The Lagrangian is given by $\mathcal{L}(\mathbf{x}, \theta; \lambda) = f(\mathbf{x}, \theta) + \lambda g(\mathbf{x}, \theta)$.

Theorem Suppose X is compact and convex, f and g are continuous and concave in $\mathbf{x}$, $f_2(\mathbf{x}, \theta)$ and $g_2(\mathbf{x}, \theta)$ are continuous in $(\mathbf{x}, \theta)$, and there exists $\mathbf{x}_0 \in X$ such that $g(\mathbf{x}_0, \theta) > 0$ for all $\theta \in [\underline{\theta}, \bar{\theta}]$. Then $V(\cdot)$ is absolutely continuous and hence differentiable (a.e.). In addition, for any $\mathbf{x}^*(\theta) \in \arg\max_{\mathbf{x} \in X : g(\mathbf{x}, \theta) \geq 0} f(\mathbf{x}, \theta)$,

$$V'(\theta) = \mathcal{L}_2(\mathbf{x}^*(\theta), \theta; \lambda^*)$$

Roy's identity and Shepard's lemma are two direct applications of envelope theorem.

Roy's Identity ***Roy's Identity*** Suppose $u(\cdot)$ represents a locally non-satiated and strictly convex preference relation on $X = \mathbb{R}^n_+$. Then, for any $(\mathbf{p}, m) \gg \mathbf{0}$, the Marshallian demand for good i is given by:

$$x_i^M(\mathbf{p}, m) = -\frac{\partial v(\mathbf{p}, m)/\partial p_i}{\partial v(\mathbf{p}, m)/\partial m}$$

Proof

The Lagrangian of the utility maximization problem is given by:

$$\mathcal{L}(\mathbf{x}, \lambda, \mu; \mathbf{p}, m) = u(x_1, x_2, \cdots, x_n) + \lambda\left(m - \sum_{i=1}^n p_i x_i\right) + \sum_{i=1}^n \mu_i x_i$$

By the envelope theorem, we have:

$$\begin{cases} \frac{\partial v(\mathbf{p},m)}{\partial p_i} = \frac{\partial \mathcal{L}}{\partial p_i}|_{(\mathbf{x}^*,\lambda^*,\ ^*)} = -\lambda^* x_i^M(\mathbf{p},m) \\ \frac{\partial v(\mathbf{p},m)}{\partial m} = \frac{\partial \mathcal{L}}{\partial m}|_{(\mathbf{x}^*,\lambda^*,\ ^*)} = \lambda^* \end{cases} \implies x_i^M(\mathbf{p},m) = -\frac{\partial v(\mathbf{p},m)/\partial p_i}{\partial v(\mathbf{p},m)/\partial m}$$

Shepard's Lemma Shepard's Lemma Suppose $u(\cdot)$ represents a locally non-satiated and strictly convex preference relation on $X = \mathbb{R}^n_+$. Then, for any $\mathbf{p} \gg \mathbf{0}$ and $F(u) \neq \emptyset$, the Hicksian demand for good i is given by:

$$x_i^H(\mathbf{p},u) = \frac{\partial e(\mathbf{p},u)}{\partial p_i}$$

Proof

The Lagrangian of the expenditure minimization problem is given by:

$$\mathcal{L}(\mathbf{x},\lambda,\ ;\mathbf{p},u) = \sum_{i=1}^n p_i x_i - \lambda(u(\mathbf{x}) - u) - \sum_{i=1}^n \mu_i x_i$$

By the envelope theorem, we have:

$$\frac{\partial e(\mathbf{p},u)}{\partial p_i} = \frac{\partial \mathcal{L}}{\partial p_i}|_{(\mathbf{x}^*,\lambda^*,\ ^*)} = x_i^H(\mathbf{p},u)$$

Remarks: By strict convexity, both Marshallian demand and Hicksian demand are single-valued.

2.5.4 Slutsky Equation

Slutsky equation mathematically describes how Marshallian demand reacts to a price change, decomposing into income effect and substitution effect.

Theorem Suppose $u(\cdot)$ represents a continuous, locally non-satiated and strictly convex preference relation $\succeq$ on $X = \mathbb{R}^n_+$ and that $\mathbf{x}^M(\mathbf{p},m)$ and $\mathbf{x}^H(p,u)$ are both differentiable and signle-valued. Then

$$\frac{\partial x_i^M(\mathbf{p},m)}{\partial p_j} = \frac{\partial x_i^H(\mathbf{p},v(\mathbf{p},m))}{\partial p_j} - \frac{\partial x_i^M(\mathbf{p},m)}{\partial m} x_j^M(\mathbf{p},m)$$

Proof (Duality)

$$x_i^H(\mathbf{p},u) = x_i^M(\mathbf{p},e(\mathbf{p},u)),\ \text{as long as } u \geq u(\mathbf{0})$$

Take partial derivatives with respect to p_j:

$$\frac{\partial x_i^H(\mathbf{p},u)}{\partial p_j} = \frac{x_i^M(\mathbf{p},e(\mathbf{p},u))}{\partial p_j} + \frac{\partial x_i^M(\mathbf{p},e(\mathbf{p},u))}{\partial m} \cdot \frac{\partial e(\mathbf{p},u)}{\partial p_j}$$

By Shepard's lemma:

$$\frac{\partial e(\mathbf{p}, u)}{\partial p_j} = x_j^H(\mathbf{p}, u)$$

By duality,

$$\begin{cases} x_j^H(\mathbf{p}, v(\mathbf{p}, m)) = x_j^M(\mathbf{p}, m) \\ e(\mathbf{p}, v(\mathbf{p}, m)) = m \end{cases}$$

Evaluating the partial derivative equation at $u = v(\mathbf{p}, m)$, we have:

$$\frac{\partial x_i^H(\mathbf{p}, v(\mathbf{p}, m))}{\partial p_j} = \frac{x_i^M(\mathbf{p}, e(\mathbf{p}, u))}{\partial p_j} + \frac{\partial x_i^M(\mathbf{p}, m)}{\partial m} \cdot x_j^M(\mathbf{p}, m)$$

Rearrange the terms and end the proof of Slutsky Equation.

Notice the LHS represents how the Marshallian demand changes with respect to a price change; the first term on the RHS represents the substitution effect and the second term represents the income effect.

2.5.5 Comparative Statics

Definition (Comparative Statics w.r.t. p_i and m)

- Good i is a *normal* good if $x_i(\mathbf{p}, m)$ is increasing in m. It is an *inferior* good if $x_i(\mathbf{p}, m)$ is decreasing in m.
- Good i is a *regular* good if $x_i(\mathbf{p}, m)$ is increasing in p_i. It is an *Giffen* good if $x_i(\mathbf{p}, m)$ is decreasing in p_i.

Remarks: A Giffen good has to be an inferior good first.

Definition (Comparative Statics w.r.t. p_j)

- Good i is a *substitute* for good j if $x_i^H(\mathbf{p}, u)$ is increasing in p_j. It is a *complement* for good j if $x_i^H(\mathbf{p}, u)$ is decreasing in p_j.
- Good i is a *gross substitute* for good j if $x_i^M(\mathbf{p}, m)$ is increasing in p_j. It is a *gross complement* for good j if $x_i^M(\mathbf{p}, m)$ is decreasing in p_j.

Remark: Good i being a complement for good j means that, increase in p_j would shift part of original share of consumption on alternative goods but other than good i.

2.6 Consumer Welfare

Consumer surplus is a basic quantitative measure of consumer welfare. But here are some issues with it:

- What if more than one price would change at the same time?
- No equivalent or immediate interpretation in utility theory.

To address the two shortcomings, we introduce two additional measures to quantify changes in consumer welfare. That is, compensating variation and equivalent variation, on measures of "conceived wealth".

Suppose the initial price is $\mathbf{p}^0$ and $u^0 = v(\mathbf{p}, m)$, and that the final price is $\mathbf{p}'$ and $u' = v(\mathbf{p}', m)$. Consider the following two measures:

1. Compensating variation: $CV = e(\mathbf{p}^0, u^0) - e(\mathbf{p}', u^0)$.
2. Equivalent variation: $EV = e(\mathbf{p}^0, u') - e(\mathbf{p}', u')$.

Interpretations of CV and EV are:

- Notice that CV and EV have the same sign, and is positive for a price drop and negative for a price increase (though the two cases are not exhaustive).
- Next, notice that by duality,

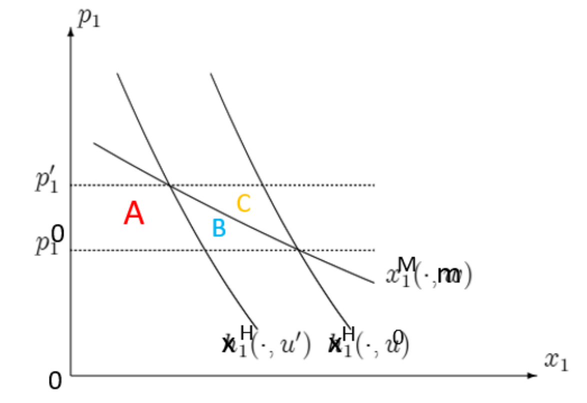
$$\begin{aligned} CV &= e(\mathbf{p}^0, u^0) - e(\mathbf{p}', u^0) = m - e(\mathbf{p}', u^0) \\ EV &= e(\mathbf{p}^0, u') - e(\mathbf{p}', u') = e(\mathbf{p}^0, u') - m \end{aligned}$$

- Intuitively, $-CV$ measures how much we need to *compensate* the consumer for them to achieve the original level of utility at the new price vector, while EV measures what is the equivalent amount of money that the consumer values this price change if the price vector were fixed at the original level.

Suppose the price of a single good i changes from p_i^0 to p_i', then

$$\begin{aligned} CV &= \int_{p_i'}^{p_i^0} \frac{\partial e(\mathbf{p}, u^0)}{\partial p_i} \mathrm{d}p_i = \int_{p_i'}^{p_i^0} x_i^H(\mathbf{p}, u^0) \mathrm{d}p_i \\ EV &= \int_{p_i'}^{p_i^0} \frac{\partial e(\mathbf{p}, u')}{\partial p_i} \mathrm{d}p_i = \int_{p_i'}^{p_i^0} x_i^H(\mathbf{p}, u') \mathrm{d}p_i \\ \Delta CS &= \int_{p_i'}^{p_i^0} x_i^M(\mathbf{p}, m) \mathrm{d}p_i \end{aligned}$$

Suppose the price of good 1 increases from p_1^0 to p_1'. Then:



$$\begin{cases} -CV = A + B + C \\ -EV = A \\ -\Delta CS = A + B \end{cases}$$

Remarks:

1. If the Marshallian demand curve is steeper than the Hicksian demand curve, it implies that the good is an inferior good. In the graph, the represented good is a normal good.

2. On any range where the good in question is either normal or inferior, then:

$$\min\{CV, EV\} \leq \Delta CS \leq \max\{CV, EV\}$$

   Notice that the two cases are not exhaustive. For example, a good may be normal good at some lower ranger of price, but reversed to inferior good at higher range of price.

2.6.1 Example: Duality and Comparative Statics

Suppose a consumer has a locally non-satiated and strictly convex preference relation on $\mathbb{R}_+^2$ that can be represented by a twice continuously differentiable utility function $u(x_1, x_2) \geq 0$. Moreover, for $(p_1, p_2) \gg \mathbf{0}$ and $u \geq 0$, the expenditure function is given by:

$$e(p_1, p_2, u) = \frac{p_1 p_2 u^2}{p_1 + p_2}$$

1. For $(p_1, p_2) \gg \mathbf{0}$ and $u > 0$, derive the Hicksian demand $\mathbf{x}^H(p_1, p_2, u)$.
2. For $(p_1, p_2, m) \gg \mathbf{0}$, derive the Marshallian demand $\mathbf{x}^M(p_1, p_2, m)$.

3. Now suppose $p_2 = 1$ and $m = 2$. Consider a price drop from $p_1^0 = 2$ to $p_1' = 1$. Calculate the compensating variation (CV), the equivalent variation (EV), and the change in consumer surplus (ΔCS) of this price change.

Solution:

1. By Shepard's Lemma,

$$\begin{aligned} x_1^H(p_1, p_2, u) &= \frac{\partial e(p_1, p_2, u)}{\partial p_1} = \frac{p_2^2 u^2}{(p_1 + p_2)^2} \\ x_2^H(p_1, p_2, u) &= \frac{\partial e(p_1, p_2, u)}{\partial p_2} = \frac{p_1^2 u^2}{(p_1 + p_2)^2} \end{aligned}$$

It follows that $\mathbf{x}^H(p_1, p_2, u) = \left(\frac{p_2^2 u^2}{(p_1+p_2)^2}, \frac{p_1^2 u^2}{(p_1+p_2)^2}\right)$

2. By duality,

$$e(p_1, p_2, v(\mathbf{p}, m)) = \frac{p_1 p_2 v(\mathbf{p}, m)^2}{p_1 + p_2} = m$$

We then have $v(\mathbf{p}, m)^2 = \frac{p_1+p_2}{p_1 p_2} m$.

Again by duality,

$$\mathbf{x}^M(p_1, p_2, m) = \mathbf{x}^H(p_1, p_2, v(\mathbf{p}, m)) = \left(\frac{p_2 m}{p_1(p_1 + p_2)}, \frac{p_1 m}{p_2(p_1 + p_2)}\right)$$

3. Apply the definition of CV, EV and ΔCS:

$$\begin{cases} CV = e(p_1^0, p_2, u^0) - e(p_1', p_2, u') = m - e(p_1', p_2, u^0) \\ EV = e(p_1^0, p_2, u') - e(p_1^0, p_2, u') = e(p_1^0, p_2, u') - m \\ \Delta CS = \int_1^2 x_1^M(p_1, p_2, m) \mathrm{d}p_1 = 2 \ln \frac{4}{3} = 0.58 \end{cases}$$

Notice that $u^0 = v(p_1^0, p_2, m) = \sqrt{3}$ and $u' = v(p_1', p_2, m) = 2$. Then the results are pinned down to be:

$$\begin{cases} CV = 2 \\ EV = \frac{2}{3} \approx 0.5 \\ \Delta CS = 2 \ln \frac{4}{3} \approx 0.58 \end{cases}$$

Aggregation Consider two consumers 1 and 2, whose preferences can be represented by the following utility functions:

$$u^1(x_1, x_2) = \begin{cases} x_1 x_2^3 & \text{if } 0 \le x_2 \le 7.7 \\ (7.7)^3 x_2 & \text{if } x_2 \ge 7.7 \end{cases}$$

$$u^2(x_1, x_2) = \begin{cases} x_1^3 x_2 & \text{if } x_1 \ge 3x_2 \\ \frac{1}{3} x_1^4 & \text{if } 0 \le x_1 \le 3x_2 \end{cases}$$

Consider the following budget sets:

- Budget set A: $p_1 = p_2 = 2$, $m = 20$.
- Budget set B: $p_1 = 3$, $p_2 = 1$, $m = 20$.

Intuitively, the failure of aggregation is due to *diverse income effects*. For instance, in the example above, the price change has a positive income effect on consumer 1, but a negative income effect on consumer 2. Aggregation is possible in the special case where all consumers have the same wealth effect, that is, $\frac{\partial \mathbf{x}^i}{\partial m^i} = \frac{\partial \mathbf{x}^j}{\partial m^j}$, for every two consumers i, j and p, m^i, m^j.

3 Profit Maximization and Rationalizability

3.1 Setups

3.1.1 Assumptions

Assumptions for firms generating profits are:

1. Perfect/complete information: no uncertainty about input/output prices, production technology, etc.
2. Perfectly competitive input and output markets: firms are price-takers in both input and output markets.
3. Input/output prices are linear, justified by perfectly competitive markets.
4. Goods are perfectly divisible.
5. The technology is exogenously given.
6. The firm's managers are perfectly controlled by the owners/shareholders.

Remarks:

1. Assumptions 1, 2 and 6 are crucial for the firm's objective of maximization:
   - Assumption 1: different risk preferences.
   - Assumption 2: consider a owner who also controls part of the input/output market.
   - Assumption 6: agency problem.
2. Firm's profit maximization object is determined by the six assumptions, instead of another assumption.

   Justification: Consider a firm jointly owned by I consumers, with consumer i's share given $\theta_i \geq 0$ (notice that $\sum_{I \in I} \theta_i = 1$), then consumer i's utility maximization problem is given by:

   $$\begin{aligned} &\max_{\mathbf{x}_i \geq \mathbf{0}} u_i(\mathbf{x}_i) \\ \text{s.t. } &\mathbf{p} \cdot \mathbf{x}_i \leq m_i + \theta_i \mathbf{p} \cdot \mathbf{y} \end{aligned}$$

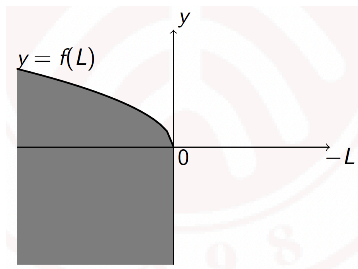
   Recall that when the preference relation is locally non-satiated, the indirect utility function $v(\mathbf{p}, m)$ is strictly increasing in m, so every consumer i would unanimously agree that a higher profit $\mathbf{p} \cdot \mathbf{y}$ is strictly more preferred.

3.1.2 Production set

A production plan is a vector $\mathbf{y} = (y_1, y_2, \cdots, y_n) \in \mathbb{R}^n$. The key observation is that y_i can be either positive or negative, with $y_i > 0$ an output and $y_i < 0$ an input.

The production set of a firm is described by $Y \subset \mathbb{R}^n$, where any $\mathbf{y} \in Y$ is a *feasible* production plan, that is, a production plan the firm can choose from.

For example, consider the example of one input and one output, suppose the production set is given by $Y = \{(-L, y) : y \leq f(L), L \geq 0\}$, where $f(\cdot)$ is the production function. The production set Y is the gray-shaded area:



Throughout our analysis, we will make the innocent technical assumptions that Y is non-empty (so as to have something to study), *closed* (to help ensure the existence of optimal production plans), and $Y \neq \mathbb{R}^n$ (so that there is some scarcity). Together with some more interesting and substantive economic properties production set might have, in conclusion basic assumptions on production sets are:

- $Y \neq \emptyset$.
- Y is closed.
- No free lunch and the possibility of inaction/shutdown: $Y \cap \mathbb{R}^n_+ = \{\mathbf{0}\}$.
- Free disposal: $\mathbf{y} \in Y \Longrightarrow \mathbf{y}' \in Y$, for any $\mathbf{y}' \leq \mathbf{y}$.

Remarks: Recall the distinction between the short run and the long run from intermediate microeconomics:

- Short run: some inputs are fixed.
- Long run: all inputs are variable.

In our discussion, we will mostly focus on the **long run** in advanced microeconomics. That is, all inputs are by default changeable.

3.1.3 Firm's Profit Maximization Problem

In the spirit of rational decision-making, the firm's problem can be framed as "choose the profit-maximizing production plan from its production set".

$$\max_{\mathbf{y}} \mathbf{p} \cdot \mathbf{y}$$
$$\text{s.t. } \mathbf{y} \in Y$$

We define the *optimal value function* as the firm's *profit function*:

$$\pi(\mathbf{p}) = \sup_{y \in Y} \mathbf{p} \cdot \mathbf{y}$$

And the firm's optimal choices as the *optimal supply correspondence*:

$$\mathbf{y}^*(\mathbf{p}) = \arg\max_{\mathbf{y} \in Y} \mathbf{p} \cdot \mathbf{y} = \{\mathbf{y} \in Y : \mathbf{p} \cdot \mathbf{y} = \pi(\mathbf{p})\}$$

Remarks:

- We have not made sufficient assumptions to ensure that a maximum profit is achieved (i.e., $y^*(\mathbf{p}) \neq \emptyset$) and so the sup in profit function cannot always be replaced with the max.

  Counter-example (constant return to scale) A firm uses labor and capital to produce its sole input. The production function is given by $f(L, K) = \sqrt{LK}$. Suppose $p = 4$ and $w = r = 1$. We can see that, if the firm chooses $L = K = t$, the profit is given by $\pi = 2t$, which is unbounded in t.

- $y^*(\mathbf{p})$, the optimal supply correspondence, is a set-valued function, which maps elements from one set, the domain of the function, to subsets of another set.

Recall our discussion in consumer theory. We will ask similar questions regarding the firm's profit maximization.

1. Suppose we observe some or all of the firm's supply decisions $y(\mathbf{p})$ (but not the production set Y), how do we know whether $y(\cdot)$ is rationalizable (i.e., consistent with profit maximization for some production set)?
2. Does the firm's profit maximization problem always have a solution?
3. Properties of the firm's profit function $\pi(\cdot)$ and the optimal supply correspondence $y^*(\cdot)$.
4. Derivation of the firm's profit maximization problem.

Note that these questions are parallel to those asked in "revealed preference" theory, with one important difference: In revealed preference theory, we observed the decision-maker's feasible sets and wanted to infer his objective function, while here we know the objective functions (profits for different prices) and want to infer the feasible set (production set).

3.2 Profit Maximization and Rationalizability

3.2.1 Rationalizability of Empirical Supply Correspondence

In practice, we do not know a firm's production set Y, but observe some of its supply choice $y(\mathbf{p})$ for $\mathbf{p} \in \mathbb{R}^n$. Hence, we define rationalizability on empirical meanings.

Definition Empirical supply correspondence $y : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is rationalized by production set Y if $y(\mathbf{p}) \subset y^*(\mathbf{p}) = \arg\max_{\mathbf{y} \in Y} \mathbf{p} \cdot \mathbf{y}$ for all $\mathbf{p} \in \mathbb{R}^n$. Empirical supply correspondence $y(\cdot)$ is rationalizable if it is rationalized by some production set.

Remarks:

- We may only observe some of the optimal choice(s) at each price $\mathbf{p}$, so in the definition it writes "$y(\mathbf{p}) \subset y^*(\mathbf{p})$".
- Practically, it is not necessary that we observe all of the firm's optimal supply decisions at all prices. We can define $y(\mathbf{p}_0) = \emptyset$ if the firm's supply decision is not observed at $\mathbf{p}_0$.
- Intuitively, an observed supply correspondence $y(\cdot)$ is rationalizable if we can find a production set Y such that $y(\cdot)$ is consistent with rational decision-making.

We are naturally interested in, what we can infer from the observations about the production set Y if the supply choices are rationalizable? Suppose at price $\mathbf{p}$ the firm chooses production plan $y((\mathbf{p})$. Here are two plausible inferences:

1. Plan $y(\mathbf{p})$ must be feasible, i.e., $y(\mathbf{p}) \in Y$.
2. Any production plan $\mathbf{y}$ other than elements in $\mathbf{y}(\mathbf{p})$ must generate no more profits than elements in $y(\mathbf{p})$ at price $\mathbf{p}$. Or equivalently, any production plan $\mathbf{y}$ that is more profitable than $\mathbf{y}(\mathbf{p})$ at price $\mathbf{p}$ cannot be feasible.

We use the first idea to construct an "inner bound" on Y that consists of all choices that the firm has actually made. We use the second idea to construct an "outer bound" on Y, which only includes plans that do not give the firm higher profits at any given price vector $\mathbf{p}$ than its observed choices at this price level.

Definition Given empirical supply correspondence $y(\cdot)$, we define the *inner bound* of the firm's production set as:

$$Y^I = \cup_{\mathbf{p} \in \mathbb{R}^n} y(\mathbf{p})$$

And the *outer bound* of the firm's production set as:

$$Y^O = \{\mathbf{y} \in \mathbb{R}^n : \mathbf{p} \cdot \mathbf{y} \leq \mathbf{p} \cdot \mathbf{y}_{\mathbf{p}}, \forall \mathbf{p} \in \mathbb{R}^n \text{ and } \mathbf{y}_{\mathbf{p}} \in y(\mathbf{p})\}$$

Intuitively, any optimal supply choice(s) must be feasible, so $Y^I \subset Y$; and any feasible production plan cannot yield a strictly higher profit at any price, so $Y \subset Y^O$.

Proposition (Rationalizable Empirical Supply Correspondence) Production set Y rationalizes empirical supply correspondence $y(\cdot)$ if and only if

$$Y^I \subset Y \subset Y^O$$

Proof

1. "Only if"
   - First consider any $\mathbf{z} \in Y^I$. By the definition of Y^I, there exists a $\mathbf{p}$ such that $\mathbf{z} \in y(\mathbf{p})$. Since $y(\cdot)$ is rationalizable, $y(\mathbf{p}) \subset y^*(\mathbf{p}) \subset Y$. It follows that $\mathbf{z} \in Y$ and $Y^I \subset Y$.
   - Next consider any $\mathbf{y} \in Y$ and $\mathbf{p} \in \mathbb{R}^n$. Since $y(\cdot)$ is rationalizable, $y(\mathbf{p}) \subset y^*(\mathbf{p})$. By the definition of $y^*(\mathbf{p})$, for any $\mathbf{y_p} \in y^*(\mathbf{p})$, $\mathbf{p} \cdot \mathbf{y} \leq \mathbf{p}\mathbf{y_p}$. It follows that $\mathbf{y} \in Y^O$ and $Y \subset Y^O$.
2. "If"
   - Fix $\mathbf{p} \in \mathbb{R}^n$ and consider any $\mathbf{y_p} \in y(\mathbf{p})$.
   - Since $y(\mathbf{p}) \subset Y^I \subset Y$, $\mathbf{y_p} \in Y$.
   - Moreover, for any $\mathbf{y} \in Y \subset Y^O$, $\mathbf{p} \cdot \mathbf{y} \leq \mathbf{p} \cdot \mathbf{y_p}$. Consequently, $\mathbf{y_p} \in y^*(\mathbf{p})$.

Remark: For the "if" part, by definition of Y^O, fix any $\mathbf{p} \in \mathbb{R}^n$, and take $\mathbf{y_p} \in y(\mathbf{p})$, $\mathbf{y_p}$ maximizes $\mathbf{p} \cdot \mathbf{y_p}$. However, with this condition we cannot simply conclude that $y(\cdot)$ is rationalizable because $\mathbf{y_p} \in Y^I$ may not in the outer bound Y^O.

The proposition means that, Y^I and Y^O carry all the information we have about the production set based on rational decision-making. The proposition immediately implies the following two corollaries:

Corollary (Weak Axiom of Profit Maximization, WAPM) Let $P = \{\mathbf{p} \in \mathbb{R}^n : y(\mathbf{p}) \neq \emptyset\}$. Empirical supply correspondence $y(\cdot)$ is rationalizable if and only if $Y^I \subset Y^O$, that is, $\mathbf{p} \cdot \mathbf{y_p} \geq \mathbf{p} \cdot \mathbf{y_{p'}}$, for any $\mathbf{p} \in P$ and $\mathbf{y_p} \in y(\mathbf{p})$, $\mathbf{y_{p'}} \in y(\mathbf{p'})$.

One simple consequence of this characterization is that when checking rationalizability we can restrict attention to supply functions rather than correspondences. (Simply put, compared with the preceding proposition, the production set Y is "left out" here.)

WAPM directly implies "law of supply":

Corollary (Law of Supply) Suppose empirical supply correspondence $y(\cdot)$ is rationalizable and let $P = \{\mathbf{p} \in \mathbb{R}^n : y(\mathbf{p}) \neq \emptyset\}$. Then for any $\mathbf{p} \in P$ and $\mathbf{y_p} \in y(\mathbf{p})$, $\mathbf{y_{p'}} \in y(\mathbf{p'})$,

$$(\mathbf{p} - \mathbf{p'}) \cdot (\mathbf{y_p} - \mathbf{y_{p'}}) \geq 0$$

Proof

Since $y(\cdot)$ is rationalizable, by WAPM, we have

$$\mathbf{p} \cdot \mathbf{y_p} \geq \mathbf{p} \cdot \mathbf{y_{p'}}$$

Switching the role of $\mathbf{p}$ and $\mathbf{p}'$, we have

$$\mathbf{p}' \cdot \mathbf{y_{p'}} \geq \mathbf{p}' \cdot \mathbf{y_p}$$

Adding the two equations above, we get the "law of supply".

In particular, if there is a single output and $y(\cdot)$ is single-valued, then

$$(p - p')(y(p) - y(p')) \geq 0$$

In other words, any rationalizable supply function must be **(weakly) upward sloping**.

Corollary Let $P = \{\mathbf{p} \in \mathbb{R}^n : y(\mathbf{p}) \neq \emptyset\}$. Empirical supply correspondence $y(\cdot)$ is rationalizable if and only if:

1. Any selection $\hat{y} : P \to \mathbb{R}^n$ is rationalizable.
2. For any two selections $\hat{y}$ and $\tilde{y}$ and any $\mathbf{p} \in P$, $\mathbf{p} \cdot \hat{y}(\mathbf{p}) = \mathbf{p} \cdot \tilde{y}(\mathbf{p})$. (Or equivalently, $\pi(\mathbf{p})$ is single-valued for each $\mathbf{p} \in P$.)

Remarks:

- The first statement of this corollary is equivalent to WAPM applied to $\mathbf{p}' \neq \mathbf{p}$,
- The second statement of this corollary is equivalent to WAPM applied to $\mathbf{p}' = \mathbf{p}$.
- Thus, when given a supply correspondence, we only need to check that
  1. Each selection from it is a rationalizable supply function, and
  2. The profit function $\pi(\mathbf{p}) = \mathbf{p} \cdot \mathbf{y}$ is single-valued at any given $\mathbf{p} \in P$; or equivalently speaking, $\pi(\mathbf{p})$ does not depend on which selection is chosen.

  Since checking the second condition is trivial, we can focus on rationalizability of supply functions (single-valued correspondence).

Verifying rationalizability by checking all the WAPM inequalities is difficult when the set of observations is large. Fortunately, it turns out that when we have a continuum of observations, rationalizability can be verified much more easily using differential conditions. Specifically, we now suppose that we observe the firm's supply choices on an open convex set P of prices (e.g., P could be the set of all strictly positive price vectors).

Proposition (Rationalizability: Differentiable Case) Consider an empirical supply correspondence $y(\cdot)$ whose domain $P = \{\mathbf{p} \in \mathbb{R}^n : y(\mathbf{p}) \neq \emptyset\}$ is an open convex set. Suppose that $\pi(\mathbf{p}) = \mathbf{p} \cdot y(\mathbf{p})$ is a differentiable function on $\mathbf{p} \in P$. Then $y(\cdot)$ is rationalizable if and only if:

1. (Hotelling's Lemma) $\nabla\pi(\mathbf{p}) = \mathbf{y_p}$, for any $\mathbf{p} \in P$ and $\mathbf{y_p} \in y(\mathbf{p})$.
2. $\pi(\cdot)$ is a convex function.

Proof

1. "If"
    - Fix $\mathbf{q} \in P$ and take any $\mathbf{y_p} \in y(\mathbf{q})$. Consider the "difference function": $G(\mathbf{q};\mathbf{p}) = \mathbf{p} \cdot \mathbf{y_q} - \pi(\mathbf{p})$.
    - It suffices to show that, $G(\mathbf{q};\cdot)$ is maximized at $\mathbf{q} = \mathbf{p}$.
    - Since $\pi(\cdot)$ is a convex function, then $G(\mathbf{q};\cdot)$ is a concave function. Since $G(\mathbf{q};\cdot)$ is differentiable in $\mathbf{p}$, the first-order condition is both necessary and sufficient.
    - The F.O.C.: $\mathbf{y_q} - \nabla\pi(\mathbf{p})|_{\mathbf{p}=\mathbf{q}} = 0$, which is precisely the Hotelling's Lemma.
2. "Only if"
    - The proof above also shows WAPM implies the Hotelling's lemma. Indeed, it is just an application of envelope formula. It remains to show $\pi(\cdot)$ is a convex function.
    - By rationalizability, there exists Y such that $\pi(\mathbf{p}) = \sup_{\mathbf{y}\in Y} \mathbf{p} \cdot \mathbf{y}$.
    - Take any $\mathbf{p}, \mathbf{q} \in P$ and $t \in (0,1)$. If $\pi(\mathbf{p}) = +\infty$ or $\pi(\mathbf{q}) = +\infty$, then clearly $\pi(t\mathbf{p} + (1-t)\mathbf{q}) \leq t\pi(\mathbf{p}) + (1-t)\pi(\mathbf{q})$. Otherwise,

$$\begin{aligned}\pi(t\mathbf{p} + (1-t)\mathbf{q}) =& \sup_{\mathbf{y}\in Y}(t\mathbf{p} + (1-t)\mathbf{q}) \cdot \mathbf{y} \\ \leq& t \cdot \sup_{\mathbf{y}\in Y} \mathbf{p} \cdot \mathbf{y} + (1-t) \cdot \sup_{\mathbf{y}\in Y} \mathbf{q} \cdot \mathbf{y} \\ =& t\pi(\mathbf{p}) + (1-t)\pi(\mathbf{q})\end{aligned}$$

Proposition (Rationalizability: General Case) Consider an empirical supply function $y : P \to \mathbb{R}^n$, where $P \subset \mathbb{R}^n$ is a convex set. Then $y(\cdot)$ is rationalizable if and only if:

1. (Producer Surplus Formula): $\pi(\mathbf{p}) = \mathbf{p} \cdot y(\mathbf{p})$ satisfies, for any smooth path $\rho : [0,1] \to P$, with $\rho(0) = \mathbf{p}$ and $\rho(1) = \mathbf{p}'$,

$$\pi(\mathbf{p}') - \pi(\mathbf{p}) = \int_0^1 y(\rho(t)) \cdot \rho'(t)\mathrm{d}t$$

2. (Law of Supply): For $\mathbf{p}, \mathbf{p}' \in P$,

$$(\mathbf{p}-\mathbf{p}')\cdot(y(\mathbf{p})-y(\mathbf{p}'))\geq 0$$

Proof:

- "Only if" part

  We have already shown the "law of supply", so it suffices to show the "producer surplus formula".

  Let $\phi(t)=\pi(\rho(t))$. Consider the "difference function"

$$\begin{aligned}\delta(\theta;t):=&\rho(t)\cdot y(\rho(\theta))-\pi(\rho(t))\\ =&\rho(t)\cdot y(\rho(\theta))-\phi(t)\end{aligned}$$

  By rationalizability of $y(\cdot)$,

$$\begin{aligned}&\delta(\theta;t)\leq 0=\delta(\theta;\theta)\\ \Longrightarrow&\left.\frac{\partial\delta(\theta;t)}{\partial t}\right|_{t=\theta}=0\\ \Longleftrightarrow&\left.\frac{\partial\delta(\theta;t)}{\partial t}\right|_{t=\theta}=y(\rho(\theta))\cdot\rho'(\theta)-\phi'(\theta)=0\\ \Longleftrightarrow&\ \phi'(\theta)=y(\rho(\theta))\cdot\rho'(\theta)\\ \Longrightarrow&\phi(1)-\phi(0)=\int_0^1 y(\rho(t))\cdot\rho'(t)\mathrm{d}t\\ \Longrightarrow&\pi(\mathbf{p}')-\pi(\mathbf{p})=\int_0^1 y(\rho(t))\cdot\rho'(t)\mathrm{d}t\end{aligned}$$

  Remark: In fact in order to derive the integral result from the derivatives, we have to prove the continuity of $\phi(\cdot)$. It can be shown that $\phi(\cdot)$ is Lipschitz continuous and hence absolutely continuous. The proof is rather technical and omitted here.

- "If" part

  In order to show the rationalizability of $y(\cdot)$, it suffices to show WAPM. Take a straight line for math convenience.

$$\rho(t)=\mathbf{p}+t(\mathbf{p}'-\mathbf{p})$$

  This way works because the path integral is not path-dependent. We aim to prove

$$\pi(\mathbf{p}') - \mathbf{p}' \cdot y(\mathbf{p}) \geq 0$$

Making tweaks to the difference in profit:

$$\begin{aligned}
\pi(\mathbf{p}') - \mathbf{p}' \cdot y(\mathbf{p}) &= \pi(\mathbf{p}') - \pi(\mathbf{p}) + \pi(\mathbf{p}) - \mathbf{p}' \cdot y(\mathbf{p}) \\
&= (\pi(\mathbf{p}') - \pi(\mathbf{p})) - (\mathbf{p}' - \mathbf{p}) \cdot y(\mathbf{p}) \\
&= \int_0^1 y(\rho(t)) \cdot \rho'(t) \mathrm{d}t - ((\mathbf{p}' - \mathbf{p}) \cdot y(\mathbf{p})) \\
&= \int_0^1 y(\mathbf{p} + t(\mathbf{p}' - \mathbf{p})) \cdot \rho'(t) \mathrm{d}t - ((\mathbf{p}' - \mathbf{p}) \cdot y(\mathbf{p})) \\
&= \int_0^1 (\mathbf{p}' - \mathbf{p}) \cdot \left[y\left(\mathbf{p} + t(\mathbf{p}' - \mathbf{p})\right) - y(\mathbf{p}) \right] \mathrm{d}t
\end{aligned}$$

Let

$$\begin{cases} \mathbf{q}' = \mathbf{p} + t(\mathbf{p}' - \mathbf{p}) \\ \mathbf{q} = \mathbf{p} \end{cases}$$

One direct observation is that $\mathbf{q}' - \mathbf{q} = t(\mathbf{p}' - \mathbf{p})$. Therefore, we can simplify the preceding integral as:

$$\int_0^1 (\mathbf{p}' - \mathbf{p}) \cdot \left[y\left(\mathbf{p} + t(\mathbf{p}' - \mathbf{p})\right) - y(\mathbf{p}) \right] \mathrm{d}t = \int_0^1 \frac{1}{t} (\mathbf{q}' - \mathbf{q}) \cdot \left(y(\mathbf{q}') - y(\mathbf{q}) \right) \mathrm{d}t$$

Since the law of supply holds, $(\mathbf{q}' - \mathbf{q}) \cdot (y(\mathbf{q}') - y(\mathbf{q})) \geq 0$. Therefore,

$$\int_0^1 \frac{1}{t} (\mathbf{q}' - \mathbf{q}) \cdot \left(y(\mathbf{q}') - y(\mathbf{q}) \right) \mathrm{d}t \geq 0$$

3.2.2 Example: Rationalizability without Observing the Outputs

Consider a price-taking firm with m inputs and n outputs. Suppose we can observe the firm's input choices $\mathbf{z}$ for different input prices $\mathbf{w} \in \mathbb{R}_+^m$, but cannot observe its output choices or output prices (and may not even know the number of outputs n). We DO know that output prices, whatever they are, do not change between the observations.

1. Suppose $m = 2$, and that at input prices $\mathbf{w}^1 = (1, 1)$ the firm chooses the input vector $\mathbf{z}^1 = (10, 15)$ and at input prices $\mathbf{w}^2 = (2, 3)$ it chooses the input vector

$\mathbf{z}^2 = (13, 14)$. Is this pair of observations rationalizable (i.e., consistent with profit maximization for some production set and output prices)?

2. In general, give a necessary and sufficient condition for two input price-demand observations $\mathbf{z}^1, \mathbf{w}^1 \in \mathbb{R}^m_+$ and $\mathbf{z}^2, \mathbf{w}^2 \in \mathbb{R}^m$ to be rationalizable. Prove both necessity and sufficiency.

3. Now suppose instead that we have the following observations:
   - At prices $\mathbf{w}^1 = (1, 1)$, the firm chooses the input vector $\mathbf{z}^1 = (10, 15)$;
   - At prices $\mathbf{w}^2 = (2, 3)$, the firm chooses the input vector $\mathbf{z}^2 = (13, 13)$;
   - At prices $\mathbf{w}^3 = (4, 1)$, the firm chooses the input vector $\mathbf{z}^3 = (8, 9)$.

   Are these three observations jointly rationalizable?

1. Suppose instead we know the output price $\mathbf{p}$ and output vectors $\mathbf{y}^1$, $\mathbf{y}^2$, then by WAPM,

$$\begin{cases} (\mathbf{w}^1, \mathbf{p}) \cdot (-\mathbf{z}^1, \mathbf{y}^1) \geq (\mathbf{w}^1, \mathbf{p}) \cdot (-\mathbf{z}^2, y^2) \\ (\mathbf{w}^2, \mathbf{p}) \cdot (-\mathbf{z}^2, \mathbf{y}^2) \geq (\mathbf{w}^2, \mathbf{p}) \cdot (-\mathbf{z}^1, y^1) \end{cases} \implies \begin{cases} \mathbf{p} \cdot \mathbf{y}^1 - 25 \geq \mathbf{p} \cdot \mathbf{y}^2 - 27 \\ \mathbf{p} \cdot \mathbf{y}^2 - 68 \geq \mathbf{p} \cdot \mathbf{y}^1 - 75 \end{cases} \implies -2 \leq \mathbf{p} \cdot \mathbf{y}^1 - \mathbf{p} \cdot$$

   which apparently leads to a contradiction.

2. In the same way, rationalizability requires that

$$\begin{cases} (\mathbf{w}^1, \mathbf{p}) \cdot (-\mathbf{z}^1, \mathbf{y}^1) \geq (\mathbf{w}^1, \mathbf{p}) \cdot (-\mathbf{z}^2, y^2) \\ (\mathbf{w}^2, \mathbf{p}) \cdot (-\mathbf{z}^2, \mathbf{y}^2) \geq (\mathbf{w}^2, \mathbf{p}) \cdot (-\mathbf{z}^1, y^1) \end{cases} \implies \mathbf{w}^1 \cdot (\mathbf{z}^1 - \mathbf{z}^2) \leq \mathbf{p} \cdot (\mathbf{y}^1 - \mathbf{y}^2) \leq \mathbf{w}^2(\mathbf{z}^1 - \mathbf{z}^2)$$
$$\implies (\mathbf{w}^1 - \mathbf{w}^2) \cdot (\mathbf{z}^1 - \mathbf{z}^2) \leq 0$$

   So $(\mathbf{w}^1 - \mathbf{w}^2) \cdot (\mathbf{z}^1 - \mathbf{z}^2) \leq 0$ is a necessary condition. Then prove it is also sufficient.

   Take any output price $\mathbf{p}$ and output vectors $\mathbf{y}^1, \mathbf{y}^2$ that satisfies $\mathbf{w}^1 \cdot (\mathbf{z}^1 - \mathbf{z}^2) \leq \mathbf{p} \cdot (\mathbf{y}^1 - \mathbf{y}^2) \leq \mathbf{w}^2(\mathbf{z}^1 - \mathbf{z}^2)$. It is trivial that output price $\mathbf{p}$ and production set $Y = \{(-z^1, \mathbf{y}^1), (-\mathbf{z}^2, \mathbf{y}^2)\}$ rationalize the pair of observations.

3. Again, when try to rationalize those choices, use necessary condition $\begin{cases} (\mathbf{w}^1, \mathbf{p}) \cdot (-\mathbf{z}^1, \mathbf{y}^1) \geq (\mathbf{w}^1, \mathbf{p}) \cdot (-\mathbf{z}^2, y^2) \\ (\mathbf{w}^2, \mathbf{p}) \cdot (-\mathbf{z}^2, \mathbf{y}^2) \geq (\mathbf{w}^2, \mathbf{p}) \cdot (-\mathbf{z}^1, y^1) \end{cases}$ for pairwise checks. Then we can get

$$\begin{aligned} -1 &\le \mathbf{p} \cdot (\mathbf{y}^1 - \mathbf{y}^2) \le 0 \\ 22 &\le \mathbf{p} \cdot (\mathbf{y}^2 - \mathbf{y}^3) \le 24 \\ -14 &\le \mathbf{p} \cdot (\mathbf{y}^3 - \mathbf{y}^1) \le -8 \end{aligned}$$

Even though at first glance there is no apparent contradiction, if we try to take the sum of the first two inequalities:

$$\begin{cases} -1 \le \mathbf{p} \cdot (\mathbf{y}^1 - \mathbf{y}^2) \le 0 \\ 22 \le \mathbf{p} \cdot (\mathbf{y}^2 - \mathbf{y}^3) \le 24 \end{cases} \implies 21 \le \mathbf{p} \cdot (\mathbf{y}^1 - \mathbf{y}^3) \le 24 \iff -24 \le \mathbf{p} \cdot (\mathbf{y}^3 - \mathbf{y}^1) \le -21$$

which contradicts with the third equation. Thus the choices cannot be rationalized.

3.3 Profit Maximization Problem

Questions to analyze on firm's profit maximization problem is much of the same to those in consumer's utility maximization problem:

1. Suppose we observe some or all of the firm's supply decisions $y(\mathbf{p})$ (but not the production set Y), how do we know whether $y(\cdot)$ is rationalizable (i.e., consistent with profit maximization for some production set)?
2. Does the firm's profit maximization problem always have a solution?
3. Properties of the firm's profit function$\pi(\cdot)$ and optimal supply correspondence $y^*(\cdot)$.
4. Derivation of the firm's profit maximization problem.

3.3.1 Returns to Scale

Recall the earlier counter-example on the non-existence of optimal supply correspondence. Intuitively, in some cases the firm can simply replicate its existing production plan infinitely many times and earns infinitely increasing profits. This motivates the following definitions.

Definition (Returns to Scale)

The production set Y exhibits:

- Non-increasing returns to scale if $\mathbf{y} \in Y \implies t\mathbf{y} \in Y, \forall t \in [0, 1]$.
- Non-decreasing returns to scale if $\mathbf{y} \in Y \implies t\mathbf{y} \in Y, \forall t \in [1, +\infty)$.

- Constant returns to scale if $\mathbf{y} \in Y \Longrightarrow t\mathbf{y} \in Y, \forall t \geq 0$.

If a firm shows non-decreasing returns to scale, then intuitively there is no scarcity in its production ability. In the special case, there are two possibilities: earning nothing or earning everything.

Proposition If the production set $Y \neq \emptyset$ exhibits non-decreasing returns to scale and the possibility of inaction/shutdown, then for any $\mathbf{p} \in \mathbb{R}^n$, $\pi(\mathbf{p}) = 0$ or ∞.

Proof First fix any $\mathbf{p} \in \mathbb{R}^n$. By the possibility of inaction/shutdown, $\mathbf{0} \in Y$, so $\pi(\mathbf{p}) \geq \mathbf{p} \cdot \mathbf{0} = 0$. Now suppose instead that $0 < \pi(\mathbf{p}) < +\infty$, then there must exist $\mathbf{y}_0 \in Y$ such that $\mathbf{p} \cdot \mathbf{y}_0 > \pi(\mathbf{p}) - \varepsilon > 0$, for any $\varepsilon > 0$ small enough. By non-decreasing returns to scale, $t\mathbf{y} \in Y$, for any $t \geq 1$. Notice that $\mathbf{p} \cdot (t\mathbf{y}_0) = t(\mathbf{p} \cdot \mathbf{y}_0) > t\pi(\mathbf{p}) - t\varepsilon > \pi(\mathbf{p})$, for any $t > 1$ and $\varepsilon > 0$ small, which is a contradiction. It follows that $\pi(p) = 0$ or $+\infty$.

3.3.2 Properties of Profit Function and Supply Correspondence

Suppose the production set Y is closed and satisfies the free disposal property. Let $\pi(\cdot)$ be the profit function and $y^*(\cdot)$ the associated optimal supply correspondence. Then for $\mathbf{p} \gg \mathbf{0}$,

- $\pi(\cdot)$ is homogenous of degree 1.
- $\pi(\cdot)$ is a convex function.
- $y^*(\cdot)$ is homogenous of degree 0.
- If Y is a convex set, then $y^*(\mathbf{p})$ is a convex set for all $\mathbf{p} \gg \mathbf{0}$. If Y is a strictly convex set, then $y^*(\mathbf{p})$ is either empty or single-valued.
- (Hotelling's Lemma) If $y^*(\mathbf{p})$ is single-valued, then $\pi(\cdot)$ is differentiable at $\mathbf{p}$ and $\nabla\pi(\mathbf{p}) = y^*(\mathbf{p})$, that is, $\frac{\partial \pi(\mathbf{p}_i)}{\partial p_i} = y_i^*(\mathbf{p})$, for $i = 1, 2, \cdots, n$.

3.3.3 Derivation of Profit Maximization Problem

The firm's profit maximization problem:

$$
\begin{gathered}
\max_{\mathbf{y}} \mathbf{p} \cdot \mathbf{y} \\
\text{s.t. } T(\mathbf{y}) \leq 0
\end{gathered}
$$

where $T(\cdot)$ is the transformation function. (One convenient way to represent prod

The Lagrangian is given by:

$$
\mathcal{L}(\mathbf{y}, \lambda) = \mathbf{p} \cdot \mathbf{y} - \lambda T(\mathbf{y}) = \sum_{i=1}^{n} p_i y_i - \lambda T(\mathbf{y})
$$

The F.O.C.s are:

$$\lambda \cdot \nabla T(\mathbf{y}) = \mathbf{p}$$

For most discussions in the course, we will focus on the case of a single output:

$$\max_{z} p \cdot f(\mathbf{z}) - \sum_{i=1}^{m} w_i z_i$$
$$\text{s.t. } \mathbf{z} \geq \mathbf{0}$$

In this special case, the Lagrangian is given by:

$$\mathcal{L}(\mathbf{z}, \mu_i) = p \cdot f(\mathbf{z}) - \sum_{i=1}^{m} w_i z_i + \sum_{i=1}^{m} \mu_i z_i$$

The F.O.C.s are:

$$\begin{cases} p \cdot \frac{\partial f(\mathbf{z})}{\partial z_i} - w_i + \mu_i = 0 \\ \mu_i z_i = 0, \mu_i \geq 0 \end{cases} \quad , \text{ for all} i = 1, 2, \cdots, m$$

Equivalently, the F.O.C.s can be written as $p \cdot \frac{\partial f(\mathbf{z})}{\partial z_i} \leq w_i$, with equality if $z_i > 0$.

Interpretation of the first-order conditions is straightforward and intuitive. The LHS represents the firm's marginal value/benefit of using additional unit of input i at the optimum, while the RHS represents the marginal cost. The F.O.C.s require that the marginal value/benefit cannot exceed the marginal cost, with equality if the input is ever used at the optimum.

Systematic procedures to solve the firm's profit-maximization problem:

- In the single-output and two-input case, check **returns to scale** and whether the production technology is strictly convex and monotonic.
- If decreasing returns to scale, monotonic and strictly convex, apply the ”tangency conditions”:

$$p \cdot MP_i = w_i$$

- Check whether the inputs are non-negative and whether the profit is non-negative and whether the profit is non-negative and bounded.

3.3.4 Example: Cobb-Douglas Production Technology

A firm uses two inputs: labor (L) and capital (K) to produce a single output (Y). The production function is given by:

$$f(L,K) = L^{\frac{1}{4}}K^{\frac{1}{2}}$$

Suppose that the input and output prices are $(w,r,p) \gg \mathbf{0}$. Solve the firm's profit maximization problem to derive the profit function $\pi(w,r,p)$ and optimal supply correspondence $y^*(w,r,p)$.

- Step 1: The production function is Cobb-Douglas, and hence strictly convex and monotonic. Moreover, if $f(L,K) \geq y$, then $f(tL,tK) = t^{\frac{3}{4}}f(L,K) \geq t(L,K) \geq ty$, for any $0 \leq t \leq 1$, so decreasing returns to scale.

- Step 2: F.O.C.s given by

$$\begin{cases} [L] : p \cdot \frac{\partial f(L,K)}{\partial L} = w \\ [K] : p \cdot \frac{\partial f(L,K)}{\partial K} = r \end{cases} \implies \begin{cases} L^* = \frac{p^4}{64w^2r^2} \\ K^* = \frac{p^4}{32wr^3} \end{cases}$$

- Check non-negativity:

  Clearly, $L^*, K^* > 0$. It follows that

$$\begin{cases} y^*(w,r,p) = \left(-\frac{p^4}{64w^2r^2}, -\frac{p^4}{32wr^3}, \frac{p^3}{16wr^2}\right) \\ \pi(w,r,p) = \frac{p^4}{64wr^2} > 0 \end{cases}$$

4 Cost Minimization Problem

Given that we already have a direct approach to solve for the firm's profit maximization problem, it is still useful to take a detour and analyze the cost minimization problem for the following two reasons:

- The cost minimization problem is more well-behaved than the profit maximization problem and has close connection with the expenditure minimization problem in consumer theory (e.g., recall the existence of solutions).
- As we shall see later, the indirect approach is more insightful if the firm has monopoly power in the output market (but is still a price-taker in the input market).

Let $Z(y) = \{\mathbf{z} \in \mathbb{R}^n_+ : f(\mathbf{z}) \geq y\}$ be the firm's feasible set. We define the optimal (minimal) value function as the *cost function*:

$$c(\mathbf{w}, y) = \inf_{\mathbf{z} \in Z(y)} \mathbf{w} \cdot \mathbf{z}$$

And the firm's optimal factor choice(s) as the *conditional factor demand correspondence*:

$$\mathbf{z}(\mathbf{w}, y) = \{\mathbf{z} \in Z(y) : \mathbf{w} \cdot \mathbf{z} = c(\mathbf{w}, y)\}$$

Notice that $\min f(\mathbf{x})$ is equivalent to $\max(-f(\mathbf{x}))$, so the cost minimization problem can be viewed as the profit maximization problem on the restricted production set $Y_y = \{(-\mathbf{z}, y) : \mathbf{z} \in \mathbb{R}^n_+, y \leq f(\mathbf{z})\}$.

Proposition Consider a conditional factor demand function $\mathbf{z} : W \times \mathbb{R} \to \mathbb{R}^m$ for a fixed output y on an open convex set $W \subset \mathbb{R}^m_+$ such that $c(\mathbf{w}, y) = \mathbf{w} \cdot \mathbf{z}(\mathbf{w}, y)$ is differentiable in $\mathbf{w}$. Then $\mathbf{z}$ is rationalizable by some production function if and only if,

1. (Shepard's lemma) $\nabla_{\mathbf{w}} c(\mathbf{w}, y) = \mathbf{z}(\mathbf{w}, y)$.
2. $c(\cdot, y)$ is concave in $\mathbf{w}$.

Remarks: The same result holds for the expenditure function $e(\mathbf{p}, u)$.

Suppose the production function $f(\cdot)$ is continuous and the production set Y satisfies the free disposal property. Then for $\mathbf{w} \gg \mathbf{0}$, cost function and conditional demand have the following properties:

- Existence of conditional factor demand: If $Z(y) \neq \emptyset$, then the conditional factor demand correspondence $\mathbf{z}(\mathbf{w}, y) \neq \emptyset$.
- Structure of conditional factor demand: If the production technology is convex (i.e., the upper contour set $\{\mathbf{z} \geq \mathbf{0} : f(\mathbf{z}) \geq y\}$ is a convex set for any $y \geq 0$),

then $\mathbf{z}(\mathbf{w}, y)$ is a convex set. If the production technology is strictly convex and $Z(y) \neq \emptyset$, then $\mathbf{z}(\mathbf{w}, y)$ is single-valued.

- Homogeneity: $c(\mathbf{w}, y)$ is homogeneous of degree 1 in $\mathbf{w}$, $\mathbf{z}(\mathbf{w}, y)$ is homogeneous of degree 0 in $\mathbf{w}$. **If the production function $f(\cdot)$ exhibits constant returns to scale, then $c(\mathbf{w}, y)$ and $\mathbf{z}(\mathbf{w}, y)$ are homogeneous of degree 1 in y.**
- Monotonicity: $c(\mathbf{w}, y)$ is non-decreasing in $\mathbf{w}$ and is strictly increasing in y for $y \geq 0$.
- Binding production level: For $y > 0$ and $Z(y) \neq \emptyset$, at any minimizer $\mathbf{z}^*$, $f(\mathbf{z}^*) = y$.
- **Convexity: If $f(\cdot)$ is a concave function, then $c(\mathbf{w}, \cdot)$ is a convex function of y.**
- Shepard's lemma: If $\mathbf{z}(\mathbf{w}, y)$ is single-valued, then $c(\mathbf{w}, y)$ is differentiable with respect to w_i and $\dfrac{\partial c(\mathbf{w}, y)}{\partial w_i} = z_i(\mathbf{w}, y)$.

Remarks: Notice that there is no counterpart in EMP in terms of the bolded properties. The essential reason is that, in consumer theory, we only care about utility representation, whose ordinal meaning matters; however in producer theory, the production function has cardinal meanings, and that makes the difference.

Derivation of Cost Minimization Problem

The cost minimization problem is given by:

$$\begin{aligned} &\min_{\mathbf{z}} \mathbf{w} \cdot \mathbf{z} \\ \text{s.t. } & f(\mathbf{z}) \geq y \\ & z_i \geq 0, \forall i = 1, 2, \cdots, m \end{aligned}$$

Notice that the CMP is almost identical to EMP in consumer theory. The Lagrangian is given by:

$$\mathcal{L}(\mathbf{z}; \lambda, \mu) = \mathbf{w} \cdot \mathbf{z} - \lambda(f(\mathbf{z}) - y) - \sum_{i=1}^{n} \mu_i z_i$$

The F.O.C.s are given by:

- w.r.t. z_i: $w_i - \lambda \dfrac{\partial f(\mathbf{z})}{\partial z_i} - \mu_i = 0$.
- Inequality constraints: $f(\mathbf{z}) \geq y$, $z_i \geq 0$, $\lambda \geq 0$, $\mu_i \geq 0$.
- Complementary slackness: $\lambda(f(\mathbf{z}) - y) = 0$, $\mu_i z_i = 0$.

For $y \geq 0$, we have binding production level (i.e., $f(\mathbf{z}) = y$), so the F.O.C.s can be alternatively framed as

$$\lambda \frac{\partial f(z)}{\partial z_i} \leq w_i, \text{ with equality if } z_i > 0$$

The economic intuition of $\lambda \frac{\partial f(z)}{\partial z_i} \leq w_i$ is that, the left-hand side corresponds to the marginal benefit, or shadow value of additional one unit of input z_i, and the right-hand side means the marginal cost of such investment. The F.O.C. states that at the optimum, the marginal benefit of inputs cannot exceed their marginal cost; or more precisely, for those deployed inputs, their marginal benefit just equals marginal cost, while for inputs that are not ever invested, their marginal benefit must be no more than the marginal cost, otherwise the firm can cut down its cost further, meaning the current solution has not reached the optimum.

Remarks: From the Lagrangian of CMP, by the envelope theorem we can see:

$$\frac{\partial c(\mathbf{w}, y)}{\partial y} = \lambda^*$$

Specifically, λ^* **measures the firm's marginal cost of producing an additional unit of output (valued by the firm instead of the market).** Hence, another interpretation for the F.O.C. is that, the firm's internal valuation of additional input cannot exceed that of the market, otherwise the firm can just invest more in and earn more out. Or equivalently, the F.O.C. says that at the optimum, the marginal cost of production using factor i, which is $\frac{w_i}{MP_i}$, should be weakly greater than the total marginal cost, with equality if the factor is employed at the optimum.

Example (Perfect Substitutes Production Technology) A firm uses two inputs: labor (L) and capital (K) to produce a single output (Y). The production function is given by

$$f(L, K) = 2L + K$$

Suppose the input prices are $(w, r) \gg \mathbf{0}$. Solve the firm's cost minimization problem to derive the cost function $c(w, r, y)$ and the conditional factor demand correspondence $(L^*(w, r, y),\ K^*(w, r, y))$.

Solution The production technology is monotonic and weakly convex, but not strictly convex. The marginal cost of production using L is given by $\frac{w}{MP_L}$, and the marginal cost of production using K is given by $\frac{r}{MP_K}$. There are three cases for discussion:

- If $0 < w < 2r$, $\frac{w}{MP_L} < \frac{r}{MP_K}$, $(L^*, K^*) = \left(\frac{y}{2}, 0\right)$ and $c(w, r, y) = w\frac{y}{2}$.
- If $w > 2r > 0$, $\frac{w}{MP_L} > \frac{r}{MP_K}$, $(L^*, K^*) = (0, y)$ and $c(w, r, y) = ry$.

- If $w = 2r > 0$, $\dfrac{w}{MP_L} = \dfrac{r}{MP_K}$, $(L^*, K^*) = \left(\frac{y-k}{2}, k\right)$ (any $0 \leq k \leq y$) and $c(w, r, y) = w\frac{y}{2}$.

With the cost function, we can restate the firm's profit maximization problem as:

$$\max_{y \geq 0} py - c(\mathbf{w}, y)$$

which is an indirect approach.

Clearly, the F.O.C. is given by

$$p \leq \frac{\partial c(\mathbf{w}, y)}{\partial y}, \text{ with equality if } y > 0$$

If we relax the assumption of perfect competition and instead assume that the firm is a monopolist in the output market and a price-taker in the input market, then the firm no longer takes the output price as given. We can restate the firm's profit maximization problem as:

$$\max_{y \geq 0} p(y)\, y - c(\mathbf{w}, y)$$

The F.O.C. is given by

$$p'(y)\, y + p(y) \geq \frac{\partial c(\mathbf{w}, y)}{\partial y}, \text{ with equality if } y > 0$$

The interpretation is similar to the perfectly competitive case. We see the indirect approach is more useful in this case and the cost minimization problem is central to the analysis.

5 Comparative Statics Analysis

A central question in economics is how an endogenous variable changes with an exogenous variable. For instance, in the indirect approach of single-product profit maximization:

$$\max_{y\geq 0} py - c(\mathbf{w}, y)$$

Holding the input prices $\mathbf{w}$ as given, we are interested in how the firm's optimal supply (correspondence) y^* changes with the output price p.

More generally, let $F : X \times \Theta \to \mathbb{R}$, where $X, \Theta \subset \mathbb{R}$, we are interested in how the maximizer

$$x^*(\theta) = \arg\max_{x\in X} F(x, \theta)$$

would change with the parameter θ.

5.1 Univariate Comparative Statics

The classical approach is to apply the implicit function theorem. We need at least two sets of assumptions:

- $F(\cdot, \cdot)$ is twice continuously differentiable.
- The maximizer $x^*(\theta)$ is unique and is characterized by the first-order condition.

A set of sufficient conditions for the second assumption is:

- The choice set X is convex.
- $F(\cdot, \theta)$ is strictly concave in x.
- The solution is interior.

Under the two sets of assumptions, the first-order condition is:

$$F_x(x(\theta), \theta) = 0$$

Differentiating on both sides with respect to θ, we have:

$$\begin{aligned} F_{xx}(x(\theta), \theta) \cdot x'(\theta) + F_{x\theta}(x(\theta, \theta)) &= 0 \\ \Longrightarrow x'(\theta) &= -\frac{F_{x\theta}(x(\theta), \theta)}{F_{xx}(x(\theta), \theta)} \end{aligned}$$

If $F\left(\cdot, \theta\right)$ is strictly concave in x, then $F_{xx}\left(\cdot, \cdot\right) < 0$ (assuming no reflection point), so $x\left(\cdot\right)$ is strictly increasing in θ if $F_{x\theta}\left(\cdot, \cdot\right) > 0$.

- Advantage

- Explicit expression for $x'(\theta)$, so quantitive analysis is possible.

- Disadvantage
  - Technical: Strong assumptions and tedious calculation in some examples.
  - Substantive: The requirement of strict concavity in x counters our intuition (recall from consumer theory that the set of maximizers is unaffected by any positive monotonic transformation).

Counter-Example Use the first-order approach to determine whether (and when) the firm's supply curve is (weakly) upward sloping.

Taking $\mathbf{w}$ as given, the profit maximization problem is given by

$$\max_{y \geq 0} py - c(y)$$

Notice that the choice set $Y = [0, +\infty)$ is convex, and when $c(\cdot)$ is strictly convex, the objective function is strictly concave. Assuming interiority, the first-order condition is given by:

$$p = c'(y(p))$$

Differentiating on both sides with respect to p, we get $y'(p) = \dfrac{1}{c''(y(p))}$, so $c(\cdot)$ being strictly convex is a sufficient condition (but notice that the only possible corner is $y = 0$). However, the requirement of $c(\cdot)$ being strictly convex is not always a reasonable assumption. Moreover, this assumption is not necessary. Recall that we have *Law of Supply*, which states that the firm's supply curve is weakly upward sloping without other assumptions.

In conclusion, we argued that the differentiability and concavity of $F(\cdot, \theta)$ are not important for comparative statics analysis. By comparison, $F_{x\theta}(\cdot, \cdot) > 0$ captures the *complementarity* between x and θ. Naturally, we would like to ask whether $F_{x\theta}(\cdot, \cdot) > 0$ alone is sufficient for $x^*(\theta)$ to be increasing in θ. Our goal in this part is to formulate a discrete analogue of this condition and show that it is sufficient for $x^*(\theta)$ to be increasing in θ.

For the general case (assumptions of $F(\cdot, \cdot)$ being continuously twice differentiable is too strong an assumption), we introduce the following definition of (strict) increasing differences.

Definition (Increasing Differences) Let $F : X \times \Theta \to \mathbb{R}$, where $X, \Theta \subset \mathbb{R}$. We say that $F(\cdot, \cdot)$ has *increasing differences* in (x, θ) if for any $x, x' \in X$ and $\theta, \theta' \in \Theta$ such that $x' > x$ and $\theta' > \theta$, we have

$$F(x', \theta') - F(x, \theta') \geq F(x', \theta) - F(x, \theta)$$

If the inequalities are strict for all such $x, x' \in X$ and $\theta, \theta' \in \Theta$, $F(\cdot, \cdot)$ has *strictly increasing differences* in (x, θ).

Intuitively, ”increasing differences” captures the case where the *marginal value* of x is higher at a higher value of θ. In other words, there is *complementarity* between x and θ.

Increasing Differences for Smooth Functions Suppose $X = [\underline{x}, \bar{x}]$ and $\Theta = \left[\underline{\theta}, \bar{\theta}\right]$, where $X, \Theta \subset \mathbb{R}$.

1. If $F(\cdot, \cdot)$ is continuously differentiable in both x and θ, show that $F(\cdot, \cdot)$ has increasing differences in (x, θ) if and only if either of the following two conditions holds:
   - $F_x(x, \cdot)$ is non-decreasing in θ, for all x.
   - $F_\theta(\cdot, \theta)$ is non-decreasing in x, for all θ.
2. If $F(\cdot, \cdot)$ is twice continuous differentiable in both x and θ, show that $F(\cdot, \cdot)$ has increasing differences in (x, θ) if and only if $F_{x\theta}(\cdot, \cdot) \geq 0$, for all (x, θ).

There is one technical question we need to address. When $F(\cdot, \theta)$ is not strictly concave in x, the maximizer $x^*(\theta)$ need not be unique. We shall introduce two natural ways of comparison of sets.

Definition (Comparison of Sets) For any two sets A and B, we say that:

- $A \leq B$ *in the strong set order* if for any $a \in A$ and $b \in B$, we have $\min\{a, b\} \in A$ and $\max\{a, b\} \in B$.
- $A \leq B$ *pointwise* if for any $a \in A$ and $b \in B$, we have $a \leq b$.

Univariate Topkis's Theorem Let $F : X \times \Theta \to \mathbb{R}$, where $X, \Theta \subset \mathbb{R}$, and $x^*(\theta) = \arg\max_{x \in X} F(x, \theta)$. Then for any $\theta' > \theta$,

1. If $F(\cdot, \cdot)$ has increasing differences in (x, θ), then $x^*(\theta) \leq x^*(\theta')$ in the strong set order.
2. If $F(\cdot, \cdot)$ has strictly increasing differences in (x, θ), then $x^*(\theta) \leq x^*(\theta')$ pointwise.

Proof Take any $x \in x^*(\theta)$ and $x' \in x^*(\theta')$. Suppose $x > x'$, then by revealed preference,

$$\begin{aligned} F(x, \theta) &\geq F(x', \theta) \\ F(x', \theta') &\geq F(x, \theta') \end{aligned}$$

By increasing differences,

$$F(x, \theta) - F(x', \theta) \leq F(x, \theta') - F(x', \theta')$$

Jointly we have:

$$
\begin{aligned}
&0 \le F(x,\theta) - F(x',\theta) \le F(x,\theta') - F(x',\theta') \le 0 \\
\Longrightarrow &\begin{cases} F(x,\theta) = F(x',\theta) \\ F(x,\theta') = F(x',\theta') \end{cases} \Longrightarrow \begin{cases} x \in x^*(\theta') \\ x' \in x^*(\theta) \end{cases}
\end{aligned}
$$

If $F(\cdot,\cdot)$ has strictly increasing differences in (x,θ), then the combined inequality cannot hold, so it must be that $x^*(\theta) \le x^*(\theta')$ pointwise.

In particular, if $F(\cdot,\cdot)$ is twice continuously differentiable in both x and θ and $x^*(\theta)$ is single-valued, then $F_{x\theta}(\cdot,\cdot) > 0$ is sufficient for $x^*(\theta)$ to be (weakly) increasing in θ.

Example: Law of Supply Again consider the firm's profit maximization problem. How would the firm's supply (correspondence) change with the output price p?

The objective function is $F(y,p) = py - c(\mathbf{w},y)$. $Y = [0,+\infty)$ and $P = [0,+\infty)$ are both intervals in $\mathbb{R}$. Moreover, $F_p(y,p) = y$, which is strictly increasing in y, so $F(\cdot,\cdot)$ has increasing differences.

Example: Tax on Monopoly Consider a monopolist which faces a downward sloping demand curve $Q^D(p)$. Now suppose the government levies a unit tax t on the firm. How would the "before-tax" price p received by the firm change with the unit tax t?

The objective function is $F(p,t) = (p-t)\,Q^D(p) - c\left(Q^D(p)\right)$. $P = [0,+\infty)$ and $T = [0,+\infty)$ are both intervals in $\mathbb{R}$. Moreover, $F_t(p,t) = -Q^D(p)$, which is strictly increasing in p (since the demand curve is downward sloping), so $F(\cdot,\cdot)$ has strictly increasing differences in (p,t). Consequently, the firm's before-tax price p increases with t pointwise.

Remarks: Notice that whether $x^*(\theta)$ increases/decreases with the parameter θ is an *ordinal* property, while (strictly) increasing differences is still a *cardinal* property. Indeed, we know from the discussion on consumer theory that $\max_x F(x,\theta)$ and $\max_x G(x,\theta)$ have the same set of maximizers if $G = \varphi \circ F$, for $\varphi(\cdot)$ strictly increasing. Nevertheless, G has (strictly) increasing differences in (x,θ) does not necessarily mean F has (strictly) increasing differences in (x,θ). In other words, the requirement of (strictly) increasing differences is still too strong for monotone comparative statics. For our purpose, if we can find a positive monotonic transformation $G = \varphi \circ F$, which has (strictly) increasing differences in (x,θ), then we know $x^*(\theta)$ increases with θ in the strong set order (pointwise).

Example: Tax on Monopoly Consider the effects of an increase in the market size on monopoly quantity (and monopoly price). Each consumer in the market has an identical inverse function given by $p^D(q)$. Suppose the number of consumers N is exogenously given, and that the firm's cost function is $c(Q)$, where $Q = Nq$ is the total quantity sold (i.e., the number of consumers times per unit purchase). Discuss how the firm's cost function $c(\cdot)$ would affect the optimal *per-consumer* quantity $q^*(N)$ as the number of consumers N increases.

Solution The objective function is

$$F(q,N) = N \cdot p^D(q) \cdot q - c(Nq)$$

Notice that it is hard to check ID of $F(\cdot,\cdot)$ (you can try it out yourself and turns out to be blocked by some terms that would rely on unassumed information). Consider $G(q,N) = \dfrac{F(q,N)}{N}$. Then if $G(\cdot,\cdot)$ is twice continuously differentiable (which can be relaxed), we have

$$\begin{aligned} G(q,N) &= \frac{\pi(q,N)}{N} = p^D(q) \cdot q - \frac{c(Nq)}{N} \\ G_N(q,N) &= -\frac{q \cdot c'(Nq) \cdot N - c(Nq)}{N^2} \\ G_{Nq}(q,N) &= -qc''(Nq) \end{aligned}$$

If $c(\cdot)$ is concave, then $G_{Nq}(q,N) \geq 0$ and $G(q,N)$ has increasing differences in (q,N), so $q^*(\cdot)$ weakly increases with N. If $c(\cdot)$ is convex, then $G_{Nq}(q,N) \leq 0$ and $G(q,N)$ has increasing differences in $(q,-N)$, so $q^*(\cdot)$ weakly decreases with N.

5.2 Multivariate Comparative Statics

Consider a two-variable maximization problem: $(x_1^*(\theta) .x_2^*(\theta)) = \arg\max_{(x_1,x_2)\in X\subset\mathbb{R}^2} F(x_1,x_2,\theta)$. If we know that $F(\cdot,\cdot)$ has (strictly) differences in (x_1,θ), we cannot conclude that $x_1^*(\theta)$ is weakly increasing in θ unless x_2^* is independent of θ. Intuitively, as θ varies, there is the direct effect of θ on x_1, but there is also the effect of θ on x_2 and consequently the indirect effect of x_2 on x_1.

Recall from univariate Topkis's theorem that $A \leq B$ in the strong set order if for any $a \in A$ and $b \in B$, we have $\min\{a,b\} \in A$ and $\max\{a,b\} \in B$. In the multivariate case, we still need to address two technical questions:

- For two vectors $\mathbf{x}$ and $\mathbf{y}$, how should we define $\min\{\mathbf{x},\mathbf{y}\}$ and $\max\{\mathbf{x},\mathbf{y}\}$.
- Under the definition, there is no guarantee that $\min\{\mathbf{x},\mathbf{y}\}$ or $\max\{\mathbf{x},\mathbf{y}\}$ is in the choice set X.

For the first question, we introduce the definition of meet and *join*, the first of which sets the greatest lower bound of the two vectors, and the second of which sets the smallest upper bound of the two vectors.

Definition (Meet & Join) For $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$, define the operations *meet* and *join*, as follows:

$$\mathbf{x} \wedge y = (\min \{x_1, y_1\}, \min \{x_1, y_1\}, \cdots, \min \{x_1, y_1\})$$
$$\mathbf{x} \vee y = (\max \{x_1, y_1\}, \max \{x_1, y_1\}, \cdots, \max \{x_1, y_1\})$$

where $\mathbf{x} \wedge \mathbf{y}$ is called the greatest lower bound of $\mathbf{x}$ and $\mathbf{y}$, and $\mathbf{x} \vee \mathbf{y}$ is called the smallest upper bound of $\mathbf{x}$ and $\mathbf{y}$.

For the second concern, we clarify the boundary of what we are discussing on by defining a structure called *sublattice*.

Definition (Sublattice) A set $X \subset \mathbb{R}^n$ is a *sublattice* if for any $\mathbf{x}, \mathbf{y} \in X$, both $\mathbf{x} \wedge \mathbf{y} \in X$ and $\mathbf{x} \vee \mathbf{y} \in X$.

Remarks: $\mathbb{R}^n$ itself is a *lattice*, so any set in $\mathbb{R}^n$ is called a *sublattice*.

Exercise Check whether the following sets are sublattices.

- $X = X_1 \times X_2 \times \cdots \times X_n$, where $X_i \subset \mathbb{R}$, for $i = 1, 2, \cdots, n$.
- $X = \{\mathbf{x} \in \mathbb{R}^n_+ : \mathbf{p} \cdot \mathbf{x} \leq m\}$, where $\mathbf{p} \gg \mathbf{0}$ is the price vector and $m > 0$ is the income.
- $X = \left\{(x_1, x_2) \in \mathbb{R}^2 : x_2 \leq g(x_1), \text{ for } g(\cdot) \text{ strictly increasing}\right\}$.

Not hard to find the second set is not a sublattice, while the first and the last are sublattices.

Similar to properties of complimentarity of two variables, in the multivariate case we introduce the definition of *supermodularity*.

Definition (Supermodularity) Let $F : X \to \mathbb{R}$, where $X \subset \mathbb{R}^n$ is a sublattice. We say that *supermodular* if for any $\mathbf{x}, \mathbf{y} \in X$, we have

$$F(\mathbf{x} \wedge \mathbf{y}) + F(\mathbf{x} \vee \mathbf{y}) \geq F(\mathbf{x}) + F(\mathbf{y})$$

Remarks:

- Note when $X = X_1 \times X_2 \subset \mathbb{R}^2$, supermodularity is equivalent to increasing differences in (x_1, x_2).

  Equivalently, you may resort to the 2-D graph to help understand its meaning:

$$F(\mathbf{x} \vee \mathbf{y}) - F(\mathbf{x}) \geq F(\mathbf{x} \wedge \mathbf{y}) - F(\mathbf{y})$$
$$F(\mathbf{x} \vee \mathbf{y}) - F(\mathbf{y}) \geq F(\mathbf{x} \wedge \mathbf{y}) - F(\mathbf{x})$$

- More generally, when $X = X_1 \times X_2 \times \cdots \times X_n \subset \mathbb{R}^n$, **supermodularity is equivalent to increasing differences in** $\left(x_i, x_j\right)$ **for all pairs of** $i \neq j$**.**

Putting all together, we have the following multivariate Topkis's Theorem.

Multi-Variable Topkis's Theorem Let $F : X \times \Theta \to \mathbb{R}$, where $X \subset \mathbb{R}^n$ is a sublattice and $\Theta \subset \mathbb{R}$. Consider $\mathbf{x}^*(\theta) = \arg\max_{\mathbf{x} \in X} F(\mathbf{x}, \theta)$. If F is supermodular, then for any $\theta' > \theta$ and $\mathbf{x} \in \mathbf{x}^*(\theta)$ and $\mathbf{x}' \in \mathbf{x}^*(\theta')$, we have

$$\begin{aligned} \mathbf{x} \wedge \mathbf{x}' &\in \mathbf{x}^*(\theta) \\ \mathbf{x} \vee \mathbf{x}' &\in \mathbf{x}^*(\theta') \end{aligned}$$

Proof Since X is a sublattice, $\mathbf{x} \wedge \mathbf{x}' \in X$ and $\mathbf{x} \vee \mathbf{x}' \in X$. By revealed preference,

$$\begin{cases} F(\mathbf{x}, \theta) \geq F(\mathbf{x} \wedge \mathbf{x}', \theta) \\ F(\mathbf{x}', \theta') \geq F(\mathbf{x} \vee \mathbf{x}', \theta') \end{cases} \Longrightarrow F(\mathbf{x}, \theta) + F(\mathbf{x}, \theta') \geq F(\mathbf{x} \wedge \mathbf{x}', \theta) + F(\mathbf{x} \vee \mathbf{x}', \theta')$$

Since F is supermodular,

$$\begin{aligned} F(\mathbf{x}, \theta) + F(\mathbf{x}', \theta) &\leq F((\mathbf{x}, \theta) \wedge (\mathbf{x}', \theta')) + F((\mathbf{x}, \theta) \vee (\mathbf{x}', \theta')) \\ &= F(\mathbf{x} \wedge \mathbf{x}', \theta) + F(\mathbf{x} \vee \mathbf{x}', \theta') \end{aligned}$$

Jointly, it must be that

$$F(\mathbf{x}, \theta) + F(\mathbf{x}, \theta') = F(\mathbf{x} \wedge \mathbf{x}', \theta) + F(\mathbf{x} \vee \mathbf{x}', \theta')$$

Consequently, we have

$$\begin{aligned} F(\mathbf{x} \wedge \mathbf{x}', \theta) &= F(\mathbf{x}, \theta) \\ F(\mathbf{x} \vee \mathbf{x}', \theta') &= F(\mathbf{x}, \theta') \end{aligned}$$

which indicates that $\mathbf{x} \wedge \mathbf{x}' \in \mathbf{x}^*(\theta)$ and $\mathbf{x} \vee \mathbf{x}' \in \mathbf{x}^*(\theta')$.

Example Suppose a firm uses two inputs (L and K) to produce a single output y, and the production function is given by

$$y = f(L, K)$$

where $f(\cdot, \cdot)$ is twice continuously differentiable.

1. Discuss how the optimal demand for capital $K^*(w, r, p)$ would be affected by an increase in the rental rate of the capital r.

2. Discuss how the optimal demand for labor $L^*(w, r, p)$ would be affected by an increase in the rental rate of the capital r.

Solution Fixing w and p, the objective function is given by

$$\pi(L, K, r) = pf(L, K) - wL - rK$$

Immediately, we have $\pi_{Lr} = 0$ and $\pi_{Kr} = -1$. Then since $\pi()$ is supermodular in $(-K, r)$, so $K^*(w, r, p)$ weakly decreases with r. Next consider the indirect path of r through K to L.

- If $f_{LK} > 0$, $\pi()$ is supermodular in $(-L, -K, r)$, so $L^*(w, r, p)$ weakly decreases with r.
- If $f_{LK} < 0$, $\pi()$ is supermodular in $(L, -K, r)$, so $L^*(w, r, p)$ weakly increases with r.

Remarks: Similar to increasing differences, supermodularity is a *cardinal* property, which is again too strong. Indeed, the weaker requirement of *quasi-supermodularity* serves our purpose.

Definition (Quasi-Supermodularity) Let $F : X \to \mathbb{R}$, where $X \subset \mathbb{R}^n$ is a sublattice. We say that F is *quasi-supermodular* if for any $\mathbf{x}, \mathbf{y} \in X$, we have

$$\begin{aligned} F(\mathbf{x}) \geq F(\mathbf{x} \wedge \mathbf{y}) &\Longrightarrow F(\mathbf{x} \vee \mathbf{y}) \geq F(\mathbf{y}) \\ \text{and } F(\mathbf{x}) > F(\mathbf{x} \wedge \mathbf{y}) &\Longrightarrow F(\mathbf{x} \vee \mathbf{y}) > F(\mathbf{y}) \end{aligned}$$

Under such extension, the analysis can be further extended beyond the case of $X \subset \mathbb{R}^n$ and X forming a lattice structure.

6 Uncertainty

Example (Motivating Expected Utility) Consider an investor who must decide how much of their initial wealth w to put into a risky asset. The risky asset can have any of the positive or negative rates of return r_i with probabilities p_i, $i = 1, 2, \cdots, n$. Suppose the investor is an expected utility maximizer and their utility for x amount of money for sure can be represented by a twice continuously differentiable, strictly increasing and strictly concave utility function $u(x)$. Let a^* be the investor's optimal amount of money to put in the risky asset. Give a necessary and sufficient condition for the investor to have strict incentives to invest in the risky asset, that is, $a^* > 0$ and is strictly preferred to $a = 0$.

Solution Since the investor is an expected utility maximizer, their optimization problem is given by:

$$\max_{0 \leq a \leq w} U(a) = \sum_{i=1}^{n} p_i \cdot u(w + ar_i)$$

A sufficient condition is $U'(0) = u'(w) \cdot \sum_{i=1}^{n} p_i r_i > 0$, that is, the expected rate of return $\sum_{i=1}^{n} p_i r_i > 0$.

Next suppose $\sum_{i=1}^{n} p_i r_i \leq 0$. Since $u(\cdot)$ is strictly concave, $U''(a) = \sum_{i=1}^{n} p_i \cdot r_i^2 \cdot u''(w + ar_i) \leq 0$, that is, $U(\cdot)$ is also strictly concave. Consequently, $U'(a) \leq U'(0) \leq 0$. It follows that the condition is also necessary.

The expected utility representation is convenient, but rather ad-hoc. Similar to the certainty benchmark, we want to understand what restrictions on the agent's preference relation induce an expected utility representation.

6.1 Basic Setups

6.1.1 Model Primitives

For simplicity, let $X = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\}$ denote a finite set of *sure* outcomes (i.e., no uncertainty).

- A *simple lottery* $\mathbf{p} = p_1 \circ \mathbf{x}_1 + p_2 \circ \mathbf{x}_2 + \cdots + p_n \circ \mathbf{x}_n$ is a probability distribution over the sure outcomes. When there is no confusion, we will also use $\mathbf{p} = \{p_1, p_2, \cdots, p_n\}$ to denote a simple lottery.
- A compound lottery is a probability distribution over a finite number of simple lotteries. For instance, let $\mathbf{p}^j$ $(j = 1, 2, \cdots, k)$ be k simple lotteries, then $\sum_{j=1}^{k} \alpha_j \mathbf{p}^j$ is a compound lottery.

To facilitate our analysis, we introduce two assumptions (while the first one can be relaxed):

1. The probabilities are **objective**, that is, regarded as objective/exogenously given facts (in contrast to subjective evaluation) by the decision-maker.
2. The decision-maker is a **consequentialist**, that is, only care about the outcomes.

   Under this assumption, the compound lottery $\sum_{j=1}^{k} \alpha_j \mathbf{p}^j$ is identical to the simple lottery it induces, that is,

$$\sum_{j=1}^{k} \alpha_j \mathbf{p}^j \iff \left(\sum_{j=1}^{k} \alpha_j \mathbf{p}_1^j, \sum_{j=1}^{k} \alpha_j \mathbf{p}_2^j, \cdots, \sum_{j=1}^{k} \alpha_j \mathbf{p}_n^j \right)$$

Given the two assumptions, we can restrict to the space of simple lotteries $\mathcal{P} = \Delta(X)$.

$$\Delta(X) = \{(p_1, \cdots, p_n) : p_i \geq 0 \ \ \forall i, p_1 + \cdots + p_n = 1\}$$

6.1.2 Assumptions on Preference Relations

Consider the agent's preference relation $\succeq$ over $\mathcal{P}$. To ensure a utility representation $U : \mathcal{P} \to \mathbb{R}$, we maintain the three basic assumptions as in the certainty benchmark.

- Completeness: For any $\mathbf{p}^1, \mathbf{p}^2 \in \mathcal{P}$, either $\mathbf{p}^1 \succeq \mathbf{p}^2$ or $\mathbf{p}^2 \succeq \mathbf{p}^1$ (or both).
- Transitivity: For any $\mathbf{p}^1, \mathbf{p}^2, \mathbf{p}^3 \in \mathcal{P}$, if $\mathbf{p}^1 \succeq \mathbf{p}^2$ and $\mathbf{p}^2 \succeq \mathbf{p}^3$, then $\mathbf{p}^1 \succeq \mathbf{p}^3$.
- Continuity: For any two sequences $(\mathbf{p}^i)_{i=1}^n, (\mathbf{q}^i)_{i=1}^n \in \mathcal{P}$, if $\mathbf{p}^I \succeq \mathbf{q}^I$ $(\forall i = 1, 2, \cdots, n)$ and $\lim_{n\to\infty} \mathbf{p}^i = \mathbf{p}^*$, $\lim_{n\to\infty} \mathbf{q}^i = \mathbf{q}^*$, then $\mathbf{p}^* \succeq \mathbf{q}^*$.

Remarks: There is an alternative definition of "continuity". Let $\succeq$ be a complete, transitive and continuous preference relation on $\mathcal{P}$, then for any $\mathbf{p}^1, \mathbf{p}^2, \mathbf{p}^3 \in \mathcal{P}$ such that $\mathbf{p}^1 \succeq \mathbf{p}^2 \succeq \mathbf{p}^3$, there exists $t \in [0, 1]$ such that $\mathbf{p}^2 \sim t\mathbf{p}^1 + (1-t)\,\mathbf{p}^3$.

To ensure an expected utility representation, we introduce an additional restriction.

Definition (Independence) A preference relation $\succeq$ on $\mathcal{P}$ satisfies *independence* if for any $\mathbf{p}^1, \mathbf{p}^2, \mathbf{p}^3 \in \mathcal{P}$ and $t \in (0, 1)$, we have

$$\mathbf{p}^1 \succeq \mathbf{p}^2 \iff t\mathbf{p}^1 + (1-t)\,\mathbf{p}^3 \succeq t\mathbf{p}^2 + (1-t)\,\mathbf{p}^3$$

Intuitively, this assumption requires that the relative ranking of two lotteries will not be affected by the mixture with a third lottery. Notice that both $t\mathbf{p}^1 + (1-t)\,\mathbf{p}^3$ and $t\mathbf{p}^2 + (1-t)\,\mathbf{p}^3$ are *compound* lotteries, which should be understood as the *simple lotteries* then induce.

Moreover, if the preference relation $\succeq$ on $\mathcal{P}$ satisfies independence, then for any $\mathbf{p}^1, \mathbf{p}^2, \mathbf{p}^3 \in \mathcal{P}$ and $t \in (0, 1)$, we have

- $\mathbf{p}^1 \succ \mathbf{p}^2 \iff t\mathbf{p}^1 + (1-t)\,\mathbf{p}^3 \succ t\mathbf{p}^2 + (1-t)\,\mathbf{p}^3$
- $\mathbf{p}^1 \sim \mathbf{p}^2 \iff t\mathbf{p}^1 + (1-t)\,\mathbf{p}^3 \sim t\mathbf{p}^2 + (1-t)\,\mathbf{p}^3$
- If $\mathbf{p}^1 \succeq \mathbf{p}^2$ and $\mathbf{p}^3 \succeq \mathbf{p}^4$, then $t\mathbf{p}^1 + (1-t)\,\mathbf{p}^3 \succeq t\mathbf{p}^2 + (1-t)\,\mathbf{p}^4$.

6.2 Expected Utility Representation

Before starting the representation theorem, we first formalize what we mean by an expected utility function.

Definition A utility function $U : \mathcal{P} \to \mathbb{R}$ has an *expected utility form* (or a von Neumann-Morgenstern expected utility function) if there is an assignment of numbers $(u_1, u_2, \cdots, u_n)$ to each of the n sure outcomes $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$ such that for any $\mathbf{p} \in \mathcal{P}$, we have $U(\mathbf{p}) = \sum_{i=1}^{n} p_i u_i$.

Remarks: $\mathbf{x}_i$ can be represented by $\mathbf{e}_i = (0, \cdots, 0, 1, 0, \cdots, 0)$ and we have $U(\mathbf{x}_i) = U(\mathbf{e}_i) = u_i$.

Notice that an expected utility function satisfies $U(\mathbf{p}) = \sum_{i=1}^{n} p_i U(\mathbf{e}_i)$, that is, linear in the probabilities. The following result shows that this observation holds more generally.

Proposition A linear function $U : \mathcal{P} \to \mathbb{R}$ has an expected utility form if and only if it is *linear*, that is, for any k (simple) lotteries $\mathbf{p}^j \in \mathcal{P}, (j = 1, 2, \cdots, k)$ and associated probabilities $\alpha_j \geq 0$ with $\sum_{j=1}^{k} \alpha_j = 1$, we have

$$U\left(\sum_{j=1}^{k} \alpha_j \mathbf{p}^j\right) = \sum_{j=1}^{k} \alpha_j U(p_j)$$

Proof

- "If" part
  - For any $\mathbf{p} \in \mathcal{P}$, by definition we have $\mathbf{p} = \sum_{i=1}^{n} p_i \mathbf{e}_i$.
  - By linearity, $U(\mathbf{p}) = \sum_{i=1}^{n} p_i U(\mathbf{e}_i)$.
  - Letting $u_i = U(\mathbf{e}_i)$ yields the desired result.
- "Only if" part
  - Since the decision-maker is a consequentialist, the compound lottery $\sum_{j=1}^{k} \alpha_j \circ \mathbf{p}^j$ is identical to the simple lottery it induces, that is

$$\alpha_1 \circ \mathbf{p}^1 + \alpha_2 \circ \mathbf{p}^2 + \cdots + \alpha_k \circ \mathbf{p}^k \iff \left(\sum_{j=1}^{k} \alpha_j p_1^j, \sum_{j=1}^{k} \alpha_j p_2^j, \cdots, \sum_{j=1}^{k} \alpha_j p_n^j\right)$$

- Since the utility function has the expected utility form, we have

$$U\left(\sum_{j=1}^{k}\alpha_j p_1^j, \sum_{j=1}^{k}\alpha_j p_2^j, \cdots, \sum_{j=1}^{k}\alpha_j p_n^j\right) = \sum_{i=1}^{n}\left(\sum_{j=1}^{k}\alpha_j p_i^j\right) u_i$$

- On the other hand, $U\left(\mathbf{p}^j\right) = \sum_{i=1}^{n} p_i u_i$, so

$$\sum_{j=1}^{k}\alpha_j U\left(\mathbf{p}^j\right) = \sum_{j=1}^{k}\alpha_j\left(\sum_{i=1}^{n} p_i^j u_i\right)$$

- Jointly, we have proved the result:

$$\begin{aligned} U\left(\sum_{j=1}^{k}\alpha_j \mathbf{p}^j\right) &= U\left(\sum_{j=1}^{k}\alpha_j p_1^j, \sum_{j=1}^{k}\alpha_j p_2^j, \cdots, \sum_{j=1}^{k}\alpha_j p_n^j\right) \\ &= \sum_{i=1}^{n}\left(\sum_{j=1}^{k}\alpha_j p_i^j\right) u_i \\ &= \sum_{j=1}^{k}\alpha_j U\left(\mathbf{p}^j\right) \end{aligned}$$

Theorem A complete and transitive preference relation $\succeq$ on $\mathcal{P}$ satisfies continuity and independence if and only if it admits an expected utility representation $U : \mathcal{P} \to \mathbb{R}$.

Proof

- "If" part
  - It is easy to verify that, if U is an expected utility representation of $\succeq$, then $\succeq$ must satisfy continuity and independence.
- "Only if" part
  - Let $\succeq$ be given and assume it satisfies continuity and independence. We will construct an expected utility function $U(p) = \sum_{i=1}^{n} p_i u_i$ that represents $\succeq$ as follows:
    * First, with a slight abuse of notation, interpret the consequences $x_1, \cdots, x_n$ as degenerate lotteries, so that x_i puts probability 1 on x_i and probability 0 on all other consequences. Then any lottery p can be viewed as a mixture of those degenerate lotteries: $p = \sum_i p_i x_i$.
    * Label the consequences so that (when viewed as degenerate lotteries) x_1 is the best and x_n is the worst, i.e., $x_1 \succeq x_i \succeq x_n$ for all i.

* By continuity, for each consequence x_i, there is some $\lambda_i \in [0,1]$ such that $x_i \sim \lambda_i x_1 + (1-\lambda_i) x_n$.
* By independence (iterated n times for each consequence), for any lottery $p = \sum_i p_i x_i$,

$$p \sim \sum_i p_i \left(\lambda_i x_1 + (1-\lambda_i) x_n\right) = \left(\sum_i p_i \lambda_i\right) x_1 + \left(1 - \sum_i p_i \lambda_i\right) x_n$$

  · If $x_1 \sim x_n$ then by independence, for all p we have

$$p \sim \left(\sum_i p_i \lambda_i\right) x_1 + \left(1 - \sum_i p_i \lambda_i\right) x_1 = x_1$$

  Hence letting $u_i = 0$ for all i gives an EU representation of $\succeq$.

  · If $x_1 \succ x_n$, then for $1 \geq \lambda \geq \lambda' \geq 0$, letting $\delta = \lambda - \lambda' \in [0,1]$ and $\hat{p} = \frac{\lambda'}{1-\delta} x_1 + \frac{1-\lambda}{1-\delta} x_n$ and using the Independence Axiom,

$$\begin{aligned} \lambda x_1 + (1-\lambda) x_n &= \delta x_1 + (1-\delta)\hat{p} \\ &\succ \delta x_1 + (1-\delta)\hat{p} \\ &= \lambda' x_1 + (1-\lambda') x_n \end{aligned}$$

  Thus, letting $u_i = \lambda_i$ yields an EU representation of $\succeq$.

6.3 Measures of Risk

6.3.1 Risk Attitudes

Let the set of sure outcomes $X = \mathbb{R}$ (note: the technical issue since $\mathbb{R}$ is infinite). In this setup, a lottery is fully characterized by the distribution function F. Be expected utility $U(F) = \int_X u(x)\,\mathrm{d}F(x)$, where $u(x)$ is the utility for receiving x amount for sure. For simplicity, suppose $u(\cdot)$ is strictly increasing and continuous and that $U(F) = \mathbf{e}[F]u(x) < +\infty$.

A decision-maker is strictly *risk-averse* if for any non-degenerate lottery F, the sure outcome $\delta_{\mathrm{E}_F} \succ F$, that is,

$$U\left(\delta_{\mathrm{E}_F}\right) = u\left(\int_X x \mathrm{d}F(x)\right) > U(F) = \int_X u(x)\,\mathrm{d}F(x)$$

In other words, a decision-maker is (strictly) risk-averse if and only if $u(\cdot)$ is (strictly) concave. The marginal utility of additional unit of money is decreasing for a risk-averse individual.

Remarks: "Risk-loving/seeking" and "risk-neutral" attitudes can be similarly defined.

Definition (Certainty Equivalent) The *certainty equivalent* $c(F,u)$ of a money lottery is characterized by

$$u(c(F,u)) = U(F)$$

Intuitively, $c(F,u)$ is the sure outcome money equivalent of a risk lottery.

Remarks: Certainty equivalent depends on initial wealth. For simplicity, we assume $w_0 = 0$, i.e., the agent does not have initial wealth; then $c(F,u) \leq \mathbf{e}F$ if and only if the agent is risk-averse.

6.3.2 Measure of Risk

We know that an agent is (strictly) risk-averse if and only if $u(\cdot)$ is (strictly) concave. Can we use $u(\cdot)$ to quantify/compare different levels of risk aversion (recall the cardinal interpretation of vNM expected utility)? Therefore, we introduce the measures of risk aversion.

Consider a risk-averse agent who has an initial wealth w and faces a (small) fair gamble.

- Scenario 1: The gamble $\tilde{\varepsilon}$ is measured in monetary unit. Then consider $u(w-a) = U(w+\tilde{\varepsilon})$.
- Scenario 2: The gamble $\tilde{\delta}$ is measured in percentage of the initial wealth. Then consider $u((1-r)w) = U\left(\left(1+\tilde{\delta}\right)w\right)$.

Since the agent is risk-averse, $a, r > 0$. Intuitively, a measures how much money the agent is willing to give up to avoid the gamble; r measures the percentage of initial wealth the agent is willing to give up to avoid the gamble. If the agent were risk-neutral, $a = r = 0$, so fixing w and the gamble, both a and r measure the agent's level of risk aversion.

When $\tilde{\varepsilon}$ is small, a is small, so by Taylor expansion and expected utility:

$$\begin{aligned}
u(w-a) &\approx u(w) - u'(w)\cdot a \\
U(w+\tilde{\varepsilon}) &= \mathbf{e}[\varepsilon]u(w+\varepsilon) = \int_\varepsilon u(w+\varepsilon)\,\mathrm{d}\varepsilon \\
&\approx \int_\varepsilon \left[u(w) + u'(w)\cdot\varepsilon + \frac{1}{2}u''(w)\cdot\varepsilon^2\right]\mathrm{d}\varepsilon \\
&= u(w)\int_\varepsilon 1\mathrm{d}\varepsilon + u'(w)\int_\varepsilon \varepsilon\mathrm{d}\varepsilon + \frac{1}{2}u''(w)\int_\varepsilon \varepsilon^2\mathrm{d}\varepsilon \\
&= u(w) + \frac{1}{2}u''(w)\cdot\mathrm{Var}[\tilde{\varepsilon}]
\end{aligned}$$

It follows that

$$\begin{aligned} u(w) - u'(w) \cdot a &= u(w) + \frac{1}{2} u''(w) \cdot \mathrm{Var}[\tilde{\varepsilon}] \\ \Longrightarrow a &\approx -\frac{1}{2} \cdot \frac{u''(w)}{u'(w)} \cdot \mathrm{Var}[\tilde{\varepsilon}] \end{aligned}$$

Similarly, we use Taylor expansion and expected utility to compute r:

$$u(w(1-r)) = U\left(w\left(1+\tilde{\delta}\right)\right) = \mathbf{e}[G(\delta)]u(w(1+\delta))$$

$$\begin{aligned} u(w(1-r)) &\approx u(w) - u'(w) \cdot (wr) \\ U\left(w + w\tilde{\delta}\right) &= \mathbf{e}[\delta]u(w + w\delta) = \int_\delta u(w + w\delta)\,\mathrm{d}G(\delta) \\ &\approx \int_\delta \left[u(w) + u'(w) \cdot (w\delta) + \frac{1}{2} u''(w) \cdot (w\delta)^2\right] \mathrm{d}\varepsilon \\ &= u(w) \int_\delta 1\mathrm{d}G(\delta) + u'(w) \int_\delta w\delta \mathrm{d}G(\delta) + \frac{1}{2} u''(w) \int_\delta (w\delta)^2\, \mathrm{d}G(\delta) \\ &= u(w) + \frac{1}{2} u''(w) \cdot w^2 \mathrm{Var}\left[\tilde{\delta}\right] \end{aligned}$$

It follows that

$$\begin{aligned} u((1-r)w) &= U\left(\left(1+\tilde{\delta}\right)w\right) \\ \Longrightarrow r &\approx -\frac{wu''(w)}{2u'(w)} \mathrm{Var}\left[\tilde{\delta}\right] \end{aligned}$$

In conclusion,

- $A(x,u) = -\dfrac{u''(x)}{u'(x)}$, coefficient of absolute risk aversion.
- $R(x,u) = -\dfrac{xu''(x)}{u'(x)}$, coefficient of relative risk aversion.

Proposition (Equivalence of Different Measures of Risk Aversion) The following definitions of u being "more risk averse" than v are equivalent:

1. For any lottery F and sure outcome δ_X. if $F \succeq_u \delta_X$, then $F \succeq_v \delta_X$.
2. For any lottery F, $c(F,u) \leq c(F,v)$.
3. The function u is "more concave" than v, that is, there exists some increasing and concave function g such that $u = g \circ v$.

4. $r(x) = \dfrac{u'(x)}{v'(x)}$ is non-increasing in x.

5. For any x, $A(x,u) = -\dfrac{u''(x)}{u'(x)} \geq -\dfrac{v''(x)}{v'(x)} = A(x,v)$.

Proof

- (1) $\Longleftrightarrow$ (2)
  - (1) $\Longrightarrow$ (2): By definition of certainty equivalent, $\delta_{c(F,u)} \sim_u F$. By condition (1), we have $\delta_{c(F,v)} \sim_v F \succeq_v \delta_{c(F,u)}$. Recall that $v(\cdot)$ is strictly, so we have $c(F,u) \leq c(F,v)$.
  - (2) $\Longrightarrow$ (1): If $\delta_{c(F,u)} \sim F \succeq_u \delta_x$, then by strict monotonicity of u, it follows that $c(F,u) \geq x$. So $c(F,v) \geq c(F,u) \geq x \Longrightarrow c(F,v) \geq x$, and we have $c(F,v) \sim F \succeq_v \delta_x$.
- (2) $\Longleftrightarrow$ (3)
  - Since u and v are strictly increasing, v^{-1} is well defined and $g = u \circ v^{-1}$ is strictly increasing.
  - Again by strict monotonicity of u, we have $u(c(F,u)) \leq u(c(F,v))$ for all lotteries F.

$$\begin{aligned} u(c(F,u)) &= \int_X u(x)\,\mathrm{d}F(x) = \int_X g(v(x))\,\mathrm{d}F(x) \\ u(c(F,v)) &= g(v(c(F,v))) = g\left(\int_X v(x)\,\mathrm{d}F(x)\right) \end{aligned}$$

  - By Jensen's inequality, $\int_X g(v(x))\,\mathrm{d}F(x) \leq g\left(\int_X v(x)\,\mathrm{d}F(x)\right)$ iff g is concave.
- (3) $\Longleftrightarrow$ (4)
  - Since u and v are strictly increasing, v^{-1} is well defined and $g = u \circ v^{-1}$ is strictly increasing.
  - When u and v are both differentiable, $u'(x) = g'(v(x)) \cdot v'(x) \Longrightarrow \frac{u'(x)}{v'(x)} = g'(v(x))$.
  - Since v is strictly increasing, $\frac{u'(x)}{v'(x)}$ is non-increasing in x iff g' is non-increasing, that is, g is concave.
- (4) $\Longleftrightarrow$ (5)
  - When $r(x) > 0$, $r(x)$ is non-increasing in x iff $\ln r(x) = \ln u'(x) - \ln v'(x)$ is non-increasing in x.

– When u and v are twice differentiable,

$$(\ln r(x))' = \frac{u''(x)}{u'(x)} - \frac{v''(x)}{v'(x)}$$

– It follows that $r(x) = \frac{u'(x)}{v'(x)}$ is non-increasing in x iff $A(x,u) \geq A(x,v)$.

Example Consider an investor who must decide how much of their initial wealth w to put into a risky asset. The risky asset can have any of the positive or negative rates of return r_i with probabilities p_i, $i = 1, 2, \cdots, n$. Suppose the investor is an expected utility maximizer and their utility for x amount of money for sure can be represented by a twice continuously differentiable, strictly increasing and strictly concave utility function $u(x)$. Let a^* be the investor's optimal amount of money to put in the risky asset. If $\sum_{i=1}^{n} p_i r_i > 0$,

1. Give a sufficient condition for $a^* < w$.
2. Suppose $a^* < w$ and that $A(x,u)$ is strictly decreasing in x, how would $a^*(w)$ change with w?

6.4 Comparison of Risky Prospects

The earlier proposition shows we can use the coefficient of absolute risk aversion $A(x,u)$ to compare the risk attitudes of two individuals, as well as the risk attitudes of the same individual at two wealth levels. But how about the comparisons of two lotteries? In particular, when can we unambiguously say that $F \succeq G$ for any (risk-averse) decision-maker?

When comparing monetary lotteries, an agent would consider

- (Expected) payoffs.
- Variations in payoffs.

Intuitively, $F \succeq G$ Unambiguously if one of the following two cases:

1. F yields a higher return than G with a higher probability.
2. F and G have the same expected return, but there is less variation/risk in F.

6.4.1 First-Order Stochastic Dominance

Proposition (First-Order Stochastic Dominance) The following two conditions are equivalent:

1. For any non-decreasing $u : \mathbb{R} \to \mathbb{R}$,

$$\int_X u(t)\,\mathrm{d}F(t) \geq \int_X u(t)\,\mathrm{d}G(t)$$

2. $F(x) \leq G(x)$, for any x.

In this case, we say F **first-order stochastically dominates** G.

Proof Assume that u is continuously differentiable, and $u'(\cdot) > 0$. Let $X = [\underline{x}, \bar{x}]$.

$$\begin{aligned}
&\int_{\underline{x}}^{\bar{x}} u(t)\,\mathrm{d}F(t) - \int_{\underline{x}}^{\bar{x}} u(t)\,\mathrm{d}G(t) \\
&= \left[u(t)\,F(t) \Big|_{\underline{x}}^{\bar{x}} - \int_{\underline{x}}^{\bar{x}} u'(t)\,F(t)\,\mathrm{d}t \right] - \left[u(t)\,G(t) \Big|_{\underline{x}}^{\bar{x}} - \int_{\underline{x}}^{\bar{x}} u'(t)\,G(t)\,\mathrm{d}t \right] \\
&= \left[(u(\underline{x}) - u(\bar{x})) - \int_{\underline{x}}^{\bar{x}} u'(t)\,F(t)\,\mathrm{d}t \right] - \left[(u(\underline{x}) - u(\bar{x})) - \int_{\underline{x}}^{\bar{x}} u'(t)\,G(t)\,\mathrm{d}t \right] \\
&= \int_{\underline{x}}^{\bar{x}} u'(t)\,(G(t) - F(t))\,\mathrm{d}t
\end{aligned}$$

It follows that $\int_{\underline{x}}^{\bar{x}} u(t)\,\mathrm{d}F(t) \geq \int_{\underline{x}}^{\bar{x}} u(t)\,\mathrm{d}G(t)$ iff $G(x) \geq F(x)$ a.e.

Notice that the condition (1) says that any decision-maker who prefers more money would prefer lottery F over G, no matter whether the agent's risk attitude.

6.4.2 Second-Order Stochastic Dominance

Definition (Mean-Preserving Spread) Let X, Y be two random variables with distribution functions F and G, respectively. Then G is a *mean-preserving spread* of F if $Y = X + \tilde{\varepsilon}$, where $\mathbf{e}\tilde{\varepsilon}|X = 0$.

Intuitively, the two lotteries F and G yield the same expected return, but G involves more risk.

Proposition (Second-Order Stochastic Dominance) Suppose $\mathbf{e}F = \mathbf{e}G$, then the following three conditions are equivalent:

1. G is a mean-preserving spread of F.

2. For any (non-decreasing) concave $u : \mathbb{R} \to \mathbb{R}$,

$$\int_X u(t)\,\mathrm{d}F(t) \geq \int_X u(t)\,\mathrm{d}G(t)$$

3. For any x,

$$\int_{-\infty}^{x} F(t)\,\mathrm{d}t \leq \int_{-\infty}^{x} G(t)\,\mathrm{d}t$$

In this case, we say F **second-order stochastically dominates** G.

Proof For simplicity, suppose $X = [\underline{x}, \bar{x}]$ and that u is twice continuously differentiable.

- (1) $\iff$ (2)
  - By our earlier calculation, $\int_{\underline{x}}^{\bar{x}} u(t)\,\mathrm{d}F(t) - \int_{\underline{x}}^{\bar{x}} u(t)\,\mathrm{d}G(t) = \int_{\underline{x}}^{\bar{x}} u'(t)\,(G(t) - F(t))\,\mathrm{d}t$.
  - Again by integration by parts and noticing $\mathbf{e}F = \mathbf{e}G$, we have:

$$\int_{\underline{x}}^{\bar{x}} u(t)\,\mathrm{d}F(t) - \int_{\underline{x}}^{\bar{x}} u(t)\,\mathrm{d}G(t) = \int_{\underline{x}}^{\bar{x}} u''(x)\left[\int_{\underline{x}}^{x} G(t)\,\mathrm{d}t - \int_{\underline{x}}^{x} F(t)\,\mathrm{d}t\right]\mathrm{d}x$$

  - It follows that $\int_{\underline{x}}^{\bar{x}} u(t)\,\mathrm{d}F(t) \geq \int_{\underline{x}}^{\bar{x}} u(t)\,\mathrm{d}G(t)$ iff $\int_{\underline{x}}^{x} G(t)\,\mathrm{d}t \geq \int_{\underline{x}}^{x} F(t)\,\mathrm{d}t$ a.e..
- (1) $\iff$ (3)
  - We will only show (3) $\implies$ (1).
  - By the law of iterated expectation,

$$\mathbf{e}[G]u(Y) = \mathbf{e}[F]\mathbf{e}[G]u(Y)\,|X$$

  - Since $u(\cdot)$ is concave, by Jensen's inequality

$$\mathbf{e}[G]u(Y) = \mathbf{e}[F]\mathbf{e}[G]u(Y)\,|X \leq \mathbf{e}[F]u(\mathbf{e}[G]Y|X) = \mathbf{e}[F]u(Y)$$

Remarks:

- Condition (2) says that any risk-averse decision-maker would prefer lottery F over G. (For $\mathbf{e}F = \mathbf{e}G$, we actually do not need $u(\cdot)$ to be non-decreasing.)
- If $\mathbf{e}F \neq \mathbf{e}G$, then conditions (2) and (3) are still equivalent, but we actually need $u(\cdot)$ to be non-decreasing and concave.

6.4.3 Likelihood Ratio Dominance

Likelihood Ratio Dominance Let F and G be two distribution function with common support $[\underline{x}, \bar{x}]$. Suppose the density functions exist and are given by f and g, respectively. We say F dominates G in the likelihood ratio order if $\frac{f(x)}{g(x)}$ is non-decreasing in x.

Intuitively, compared with G, F puts higher probabilities on higher returns.

Proposition (Relationships Between LR and FOSD) If F dominates G in the likelihood ratio order, then F first-order stochastically dominates G.

Proof Since $F(\cdot)$ and $G(\cdot)$ are both non-decreasing, in order to show that $F(x) \leq G(x)$, it suffices to show that $\dfrac{F(x)}{1-F(x)} \leq \dfrac{G(x)}{1-G(x)}, \forall x \in (\underline{x}, \bar{x})$. We have

$$
\begin{aligned}
LHS &= \frac{\int_{\underline{x}}^{x} f(t)\,\mathrm{d}t}{\int_{x}^{\bar{x}} f(t)\,\mathrm{d}t} \\
&= \frac{\int_{\underline{x}}^{x} \frac{f(t)}{g(t)} \cdot g(t)\,\mathrm{d}t}{\int_{x}^{\bar{x}} \frac{f(t)}{g(t)} \cdot g(t)\,\mathrm{d}t}
\end{aligned}
$$

Since $\frac{f(x)}{g(x)}$ is non-increasing,

$$
\begin{aligned}
\forall t \in [\underline{x}, x], \frac{f(t)}{g(t)} &\leq \frac{f(x)}{g(x)} \\
\forall t \in [x, \bar{x}], \frac{f(t)}{g(t)} &\geq \frac{f(x)}{g(x)}
\end{aligned}
$$

Therefore, we have

$$
LHS = \frac{\int_{\underline{x}}^{x} \frac{f(t)}{g(t)} \cdot g(t)\,\mathrm{d}t}{\int_{x}^{\bar{x}} \frac{f(t)}{g(t)} \cdot g(t)\,\mathrm{d}t} \leq \frac{\frac{f(x)}{g(x)} \cdot \int_{\underline{x}}^{x} g(t)\,\mathrm{d}t}{\frac{f(x)}{g(x)} \cdot \int_{x}^{\bar{x}} g(t)\,\mathrm{d}t} = \frac{\int_{\underline{x}}^{x} g(t)\,\mathrm{d}t}{\int_{x}^{\bar{x}} g(t)\,\mathrm{d}t} = \frac{G(x)}{1-G(x)}
$$

In other words, compared with FOSD, the likelihood ratio order is a stronger requirement on higher returns.

6.5 Comparative Statics Under Risk

Our earlier discussion on comparative statics didn't specifically consider uncertainty. Whether and how would the comparative statics results carry over to the setting of choices under uncertainty?

6.5.1 Supermodularity of Expected Utility

Lemma (Supermodularity of Expected Utility) Let $X \subset \mathbb{R}^n$ be a sublattice and $T \subset \mathbb{R}$. Suppose $u : X \times T \to \mathbb{R}$ is supermodular in $\mathbf{x}$. Then for any distribution function F on T, the function $U : X \to \mathbb{R}$ defined by

$$U(\mathbf{x}) = \int_t u(\mathbf{x}, t)\, \mathrm{d}F(t)$$

is supermodular in $\mathbf{x}$.

Proof Take any $\mathbf{x}, \mathbf{x}' \in X$, we have

$$\begin{aligned} U(\mathbf{x} \wedge \mathbf{x}') + U(\mathbf{x} \vee \mathbf{x}') &= \int_t [u(\mathbf{x} \wedge \mathbf{x}', t) + u(\mathbf{x} \vee \mathbf{x}', t)]\, \mathrm{d}F(t) \\ &\geq \int_t [u(\mathbf{x}, t) + u(\mathbf{x}', t)]\, \mathrm{d}F(t) \\ &= U(\mathbf{x}) + U(\mathbf{x}') \end{aligned}$$

where the inequality follows from the supermodularity of $u(\mathbf{x}, t)$ in $\mathbf{x}$. Consequently, $U(\cdot)$ is supermodular in $\mathbf{x}$.

In other words, expected utility preserves supermodularity.

6.5.2 FOSD and Increasing Difference

Lemma (FOSD and Increasing Differences) Let $X \subset \mathbb{R}^n$ be a sublattice and $T, \Theta \subset \mathbb{R}$. Suppose $u : X \times T \to \mathbb{R}$ has increasing differences in $(\mathbf{x}, t)$, and $\{F_\theta\}_{\theta \in \Theta}$ is a family of distribution function on T such that $F_\theta \geq_{FOSD} F_{\theta'}$ if $\theta > \theta'$. Then the function $U : X \times \Theta \to \mathbb{R}$ defined by

$$U(\mathbf{x}, \theta) = \int_t u(\mathbf{x}, t)\, \mathrm{d}F_\theta(t)$$

has increasing differences in $(\mathbf{x}, \theta)$.

Proof Take any $\mathbf{x} > \mathbf{x}'$ and $\theta > \theta'$, we have

$$\begin{aligned} U(\mathbf{x}, \theta) - U(\mathbf{x}', \theta) &= \int_t [u(\mathbf{x}, t) - u(\mathbf{x}', t)]\, \mathrm{d}F_\theta(t) \\ &\geq \int_t [u(\mathbf{x}, t) - u(\mathbf{x}', t)]\, \mathrm{d}F_{\theta'}(t) \\ &= U(\mathbf{x}, \theta') - U(\mathbf{x}', \theta') \end{aligned}$$

where the inequality holds because $F_\theta \geq_{FOSD} F_{\theta'}$ and $\delta(t) := u(\mathbf{x}, t) - u(\mathbf{x}', t)$ is non-decreasing in t (by increasing differences of $u(\cdot, \cdot)$ in $(\mathbf{x}, t)$). Consequently, $U(x, \theta)$ has increasing differences in $(\mathbf{x}, \theta)$.

Intuitively, θ can be interpreted as a signal of t. In other words, FOSD preserves increasing differences.

6.5.3 Comparative Statics Under Risk

Proposition (Comparative Statics Under Risk) Let $X \subset \mathbb{R}^n$ be a sublattice and $T, \Theta \subset \mathbb{R}$. Suppose $u : X \times T \to \mathbb{R}$ is supermodular in $\mathbf{x}$ and has increasing differences in $(\mathbf{x}, t)$. Further suppose $\{F_\theta\}_{\theta \in \Theta}$ is a family of distribution function on T such that $F_\theta \geq_{FOSD} F_{\theta'}$ if $\theta > \theta'$. Then for the function $U : X \times \Theta \to \mathbb{R}$ defined by

$$U(\mathbf{x}, \theta) = \int_t u(\mathbf{x}, t) \, \mathrm{d}F_\theta(t)$$

The (set of) maximizer(s) $\arg\max_{\mathbf{x} \in X} U(\mathbf{x}, \theta)$ is non-decreasing in θ (in the strong set order).

Proof

- Fixing any θ, by the first lemma, $U(\mathbf{x}, \theta)$ is supermodular in $\mathbf{x}$.
- By the second lemma, $U(\mathbf{x}, \theta)$ has increasing differences in $(\mathbf{x}, \theta)$.
- Since $X \times \Theta$ forms a product set, $U(\mathbf{x}, \theta)$ is supermodular in $(\mathbf{x}, \theta)$.
- By the multivariate Topkis' theorem, the (set of) maximizers $\arg\max_{\mathbf{x} \in X} U(\mathbf{x}, \theta)$ is non-decreasing in θ (in the strong set order).

Example Suppose a monopolist faces an uncertain demand and must make a production decision prior to learning the realized demand for its product. The monopolist does learn some information about demand prior to choosing its output. Formally, suppose the inverse demand function $p(q, t) = \hat{p}(q) + t$, where $\hat{p}(q)$ is the estimated inverse demand, and t is a random noise. The ex-post profit of the firm is therefore

$$\pi(q, t) = p(q, t) \cdot q - c(q)$$

where $c(q)$ is the monopolist's cost function. The monopolist observes a signal θ that is informative about the parameter t. Specifically, suppose the distribution of t conditional on θ is $F_\theta(\cdot)$ and $F_\theta \geq_{FOSD} F_{\theta'}$ for $\theta > \theta'$. Determine how would the monopolist's optimal output $q^*(\theta)$ change with the observed signal θ.

Solution The monopolist solves the following maximization problem:

$$\max_{q \geq 0} \Pi(q, \theta) = \int_t \pi(q, t) \, \mathrm{d}F_\theta(t)$$

Notice that $\pi(q,t) = p(q,t) \cdot q - c(q)$ and that $\frac{\partial \pi(q,t)}{\partial t} = q$, which is itself increasing in q.

It follows that $\pi(q,t)$ has increasing differences in (q,t). Since $F_\theta \geq_{FOSD} F_{\theta'}$ for $\theta > \theta'$, $\Pi(q,\theta)$ has increasing differences in (q,θ) and $q^*(\theta)$ is non-decreasing in θ in the strong set order.

Example Consider an agent with an initial wealth w, who faces a random loss $\tilde{l} \in [0,w]$. To counter the potential loss, the agent may purchase any fraction of insurance $y \in [0,1]$. Specifically, if the agent purchases y unit of insurance, then they pay a total price of yp upfront, and get paid yl if they incur a loss of l. Suppose the agent is an expected utility maximizer and has a Bernoulli utility function $u(x)$ for amount of money for sure. Further suppose $u(\cdot)$ is twice continuously differentiable with $u'(\cdot) > 0$ and $u''(\cdot) < 0$.

1. First suppose $p \leq \mathbf{e}\tilde{l}$. Show the optimal amount of insurance $y^* = 1$.
2. Next suppose $p > \mathbf{e}\tilde{l}$. Show the optimal amount of insurance $y^* < 1$.
3. Let $A(x,u)$ be the coefficient of absolute risk aversion. Write down its expression and show that if $A(x,u) = c_0$ (a constant), then the optimal amount of insurance y^* is independent of w.
4. Now suppose $A(x,u)$ strictly decreases with x. Show the optimal amount of insurance y^* is non-increasing in the agent's initial wealth w.

Solution

Since the agent is an expected utility maximizer, their choice problem is given by

$$\max_{y \in [0,1]} U(y) = \mathbf{e}u(w - yp - (1-y)l)$$

Simple calculation shows

$$\begin{aligned} U'(y) &= \mathbf{e}(l-p) \cdot u'(w - l + (l-p)y) \\ U''(y) &= \mathbf{e}(l-p)^2 \cdot u''(w - l + (l-p)y) \leq 0 \end{aligned}$$

It follows that $U'(y) \geq U(1) = u'(w-p)\left(\mathbf{e}\tilde{l} - p\right) \geq 0$, with strict inequality as long as $\tilde{l}$ is not degenerate or $p < \mathbf{e}\tilde{l}$. Consequently, $y^* = 1$.

When $p > \mathbf{e}\tilde{l}$, $U'(1) = u'(w-p)\left(\mathbf{e}\tilde{l} - p\right) < 0$. Since $U''(y) \leq 0$, with strict inequality as long as $\tilde{l}$ is not degenerate, we have $y^* < 1$.

By definition, $A(x,u) = -\dfrac{u''(x)}{u'(x)}$.

Take any $w_1 < w_2$, and let

$$
\begin{aligned}
v_1(x) &= u(w_1 + x) \\
v_2(x) &= u(w_2 + x)
\end{aligned}
$$

Then at initial wealth $w_i \, (i = 1, 2)$, the agent's maximization problem is given by

$$
\max_{y \in [0,1]} V_i(y) = \mathbf{e} v_i(-l + (l - p) y)
$$

From the previous two parts, it suffices to focus on the case of $p > \mathbf{e}\tilde{l}$.

By our earlier characterization, $A(x, u) = c_0$ implies $\dfrac{v_1'(x)}{v_2'(x)} = c_1$, which is also a constant.

Simple calculation shows

$$
\begin{aligned}
V_i'(y) &= \mathbf{e}(l - p) v_i'(-l + (l - p) y) \\
V_i''(y) &= \mathbf{e}(l - p)^2 v_i''(-l + (l - p) y) \leq 0
\end{aligned}
$$

If $V_i'(0) \leq 0$, then $V_2'(0) \leq 0$, so $y_1^* = y_2^* = 0$. Otherwise, if $V_1'(y) = 0$, then $V_2'(y) = 0$. so $y_1^* = y_2^* = y^*$.

Suppose by contradiction that $y_2^* > y_1^* \geq 0$. By our earlier characterization, $A(x, u)$ strictly decreases with x implies $\dfrac{v_1'(x)}{v_2'(x)}$ is non-increasing in x.

Consequently, we have

$$
\begin{aligned}
V_1'(y_2^*) &= \int_0^p v_1'(-l + (l - p) y_2^*)(l - p) \, \mathrm{d}F(l) \\
&\quad + \int_p^w v_1'(-l + (l - p) y_2^*)(l - p) \, \mathrm{d}F(l) \\
&= \int_0^p \frac{v_1'(-l + (l - p) y_2^*)}{v_2'(-l + (l - p) y_2^*)} v_2'(-l + (l - p) y_2^*)(l - p) \, \mathrm{d}F(l) \\
&\quad + \int_p^w \frac{v_1'(-l + (l - p) y_2^*)}{v_2'(-l + (l - p) y_2^*)} v_2'(-l + (l - p) y_2^*)(l - p) \, \mathrm{d}F(l) \\
&\geq \frac{v_1'(-p)}{v_2'(-p)} \int_0^w v_2'(-l + (l - p) y_2^*)(l - p) \, \mathrm{d}F(l) \\
&= 0
\end{aligned}
$$

For non-degenerate $\tilde{l}$, we have $V_1'(y_1^*) \geq V_1'(y_2^*) \geq 0$, which is a contradiction.

7 General Equilibrium

- Central idea: Markets clear.
- Partial v.s. General Equilibrium.
  - Partial equilibrium: Suppose the prices of all other goods are held fixed, analyze the market equilibrium in a single market.
  - General equilibrium: All markets clear simultaneously.
- When is the partial equilibrium approach valid?
  - The prices of the goods in all other markets are held fixed, that is, the effect on the prices of all other goods can be neglected.
  - No wealth effect of the market under consideration.

Example A large number of towns n, each with a price-taking identical firm that produces a single consumption good with the production function $y = f(l)$, where $f'(l) > 0$ and $f''(l) < 0$. The single consumption good is traded in the national market. Suppose L units of inelastic labor supply in total. Workers can freely move across towns to seek the highest possible wage (complete information). Normalize $p_c = 1$ and let w_i denote the equilibrium wage in town i. Now suppose that town 1 levies a small tax $t > 0$ on firm 1 for each unit of labor hired. Which group of individual(s) will bear the tax burden?

Solution

- Equilibrium without tax
  - Whether we adopt the partial or general equilibrium approach, the competitive equilibrium without the labor tax is identical.
  - Since workers can move freely and $f'(l) > 0$ and $f''(l) < 0$, we must have

$$\begin{cases} l^* = \frac{1}{n} \cdot L \\ w_1 = w_2 = \cdots = w_n = w_0 \\ w_0 = f'(l^*) = f'\left(\frac{l}{n}\right) \end{cases}$$

- Partial equilibrium in town 1 with tax
  - Assuming the wage rate in other markets are unaffected, we must have $w_1(t) = w_0 + t$.
  - At the new equilibrium, $f'(l_1(t)) = w_1(t) = w_0 + t$.
  - Hence we can see that firm 1 bears all the tax burden.

- General equilibrium with town 1 levied with tax t
  - Let $l_1(t)$ denote the equilibrium amount of labor in town 1 with a t unit tax, and $l_{-1}(t)$ denote the equilibrium amount of labor in any other town when town 1 imposes a t unit tax and $w(t)$ the equilibrium wage rate received by the worker. When all labor markets clear, we should have:

$$\begin{cases} l_1(t) + (n-1)\, l_{-1}(t) = L \\ f'(l_1(t)) = w_1(t) = w(t) + t \\ f'(l_{-1}(t)) = w(t) \end{cases}$$
$$\Longrightarrow \begin{cases} f'(L - (n-1)\, l_{-1}(t)) = w(t) + t \\ f'(l_{-1}(t)) = w(t) \end{cases}$$
$$\Longrightarrow \begin{cases} -(n-1)\, l'_{-1}(t) \cdot f''(L - (n-1)\, l_{-1}(t)) = w'(t) + 1 \\ l_{-1}(t)\, f''(l_{-1}(t)) = w'(t) \end{cases}$$

  - Set $t \to 0$, so we have $l_{-1}(0) = \frac{1}{n} \cdot L$. From the two equations above we obtain $w'(0) = -\frac{1}{n}$.
  - Let $\Pi(w(t))$ denote the total profit of the firms when the wage rate is $w(t)$ and $\pi(w)$ the profit of a representative firm. Naturally, $\Pi(w(t)) = \pi(w(t) + t) + (n-1)\, \pi(w(t))$, and

$$\frac{\partial \Pi(w(t))}{\partial t} = (w'(t) + 1)\, \pi'(w(t) + t) + (n-1)\, w'(t)\, \pi'(w(t))$$
$$\Longrightarrow \left.\frac{\partial \Pi(w(t))}{\partial t}\right|_{t \to 0} = (w'(0) + 1)\, \pi'(w(0)) + (n-1)\, w'(0)\, \pi'(w(0)) = 0$$

  - From quantitative analysis we can see that, the firms as a whole do not bear the tax burden and the workers bear all the tax burden.
  - Intuitively, it must be the workers that bear all the tax burden. Even though the labor supply for any firm is perfectly elastic, the **total** labor supply is **perfectly inelastic**. The impact on the wages in other towns, even though small, is not negligible. Indeed, the labor supply for any given firm is perfectly elastic, so we cannot ignore the general equilibrium effect.

7.1 Pure Exchange Economy

7.1.1 Introduction

- Three central activities in a market: Consumption, production and trade.

- Pure exchange economy: Each agent has an endowment that they can trade with, and there is no production.

Consider a perfectly competitive economy with two agents ($i = A, B$) and two goods ($j = 1, 2$). Suppose the agents' preference relations and initial endowments are given by

$$u^A(x_1, x_2) = x_1^2 x_2 \text{ with } e^A = (1, 2)$$
$$u^B(x_1, x_2) = x_1 x_2^2 \text{ with } e^B = (2, 1)$$

Can we come up with a price vector (p_1, p_2) that clear both markets?
Idea: Each agent's "income" is endogenously determined by the equilibrium price vector.

Suppose there is an equilibrium price vector (p_1, p_2), then agent A's "income" is $m^A = p_1 + 2p_2$ and agent B's "income" is $m^B = 2p_1 + p_2$. From this we can pin down the optimal individual demands

$$\mathbf{x}^A = \left(\frac{2(p_1 + 2p_2)}{3p_1}, \frac{1(p_1 + 2p_2)}{3p_2}\right)$$
$$\mathbf{x}^B = \left(\frac{1(2p_1 + p_2)}{3p_1}, \frac{2(2p_1 + p_2)}{3p_2}\right)$$

In equilibrium, market demand equals market endowment for both goods:

$$\frac{2(p_1 + 2p_2)}{3p_1} + \frac{1(2p_1 + p_2)}{3p_1} = 3$$
$$\frac{1(p_1 + 2p_2)}{3p_2} + \frac{2(2p_1 + p_2)}{3p_2} = 3$$

From either equation, we have $\dfrac{p_1^*}{p_2^*} = 1$. (Otherwise if there comes a contradiction, the equilibrium cannot stand.)

Four observations in this two-agent, two-good pure exchange economy.

1. There is an equilibrium price vector that clears both markets.
2. Only the relative price matters in equilibrium.
3. When one market clears, the other market clears miraculously.
4. The equilibrium allocation is Pareto efficient.

7.1.2 Basic Setups

Model Preliminaries

- $I = \{1, 2, \cdots, n\}$ agents and $J = \{1, 2, \cdots, m\}$ goods.
- $\succeq_i$: preference relation of agent i.
- $\mathbf{e}^i = \{e_1^i, e_2^I, \cdots, e_m^i\}$ initial endowment of agent i (property rights are well-defined and no co-ownership).
- Denote the economy $\mathcal{E} = \{\succeq_i, \mathbf{e}^i\}_{I \in \mathcal{J}}$.

Assumptions on Market Structure

- Perfect/complete information.
- Perfectly competitive markets.
  - Agents are price-takers.
  - Prices are linear.
- Goods are perfectly divisible.

Assumptions on Agents' Preferences and Endowments For any agent $i \in \mathcal{J}$:

- (A1) The preference relation $\succeq^i$ is rational (complete and transitive) and continuous.
- (A2) The preference relation $\succeq^i$ is monotonic.
- (A3) The preference relation $\succeq^i$ is (weakly) convex.
- (A4) $e_j^i > 0$, for all $j \in \mathcal{J}$.

Notice that for the first assumption, the existence of $u^i(\cdot)$ is guaranteed for any $i \in I$, so we can alternatively denote the economy as $\mathcal{E} = \left(u^i(\cdot), \mathbf{e}^i\right)_{i \in \mathcal{J}}$.

7.1.3 Walrasian Equilibrium

Definition **Definition (Walrasian Equilibrium)** A **Walrasian Equilibrium** for a perfectly competitive pure exchange economy $\mathcal{E} = \left(u^i(\cdot), \mathbf{e}^i\right)_{i \in \mathcal{J}}$ is a price vector $\mathbf{p} \geq \mathbf{0}$ and an allocation $\left(\mathbf{x}^i\right)_{i \in \mathcal{J}}$ such that:

1. Utility maximization: Agent $i \in \mathcal{J}$ solves:

$$\max_{\mathbf{x}^i \geq \mathbf{0}} u^i\left(\mathbf{x}^i\right) \text{ s.t. } \mathbf{p} \cdot \mathbf{x}^i \leq \mathbf{p} \cdot \mathbf{e}^i$$

2. Market clearing (for all goods):

$$\sum_{i \in \mathcal{I}} \mathbf{x}^i (\mathbf{p}) = \sum_{i \in \mathcal{I}} \mathbf{e}^i$$

Under Walrasian equilibrium, all markets clear at the same time. A Walrasian equilibrium specifies both the equilibrium price vector and the allocation.

For simplicity, suppose for any $i \in \mathcal{I}$, $\mathbf{x}^i (\mathbf{p})$ is always unique. This is guaranteed if we strengthen (A3) to assume each $\succeq_i$ is strictly convex and $\mathbf{p} \gg \mathbf{0}$.

The (aggregate) **excess demand** for good j is given by

$$z_i (\mathbf{p}) = \sum_{i \in \mathcal{I}} x_j^i (\mathbf{p}) - \sum_{i \in \mathcal{I}} e_j^i$$

With this formulation, a Walrasian equilibrium can be equivalently formulated as

$$z_j (\mathbf{p}) = 0, \forall j \in \mathcal{J}$$

Proposition (Properties of Excess Demand) Let $\mathcal{E}$ be a perfectly competitive pure exchange economy that satisfies (A1) and (A2), then the aggregate excess demand function $\mathbf{z} (\mathbf{p})$ satisfies:

1. Homogeneity of degree 0: $z_j (\mathbf{p})$ is homogeneous of degree 0 in $\mathbf{p}$, for any $j \in \mathcal{J}$.
2. Walras' law: $\mathbf{p} \cdot \mathbf{z} (\mathbf{p}) = 0$.

Proof

- Homogeneous of degree 0
  - $\mathbf{x}^i (\mathbf{p})$ solves $\max_{\mathbf{x}^i \geq \mathbf{0}} u^i (\mathbf{x}^i)$ s.t. $\mathbf{p} \cdot \mathbf{x}^i \leq \mathbf{p} \cdot \mathbf{e}^i$.
  - For any $t > 0$, the solution must be the same as: $\max_{\mathbf{x}^i \geq \mathbf{0}} u^i (\mathbf{x}^i)$ s.t. $(t\mathbf{p}) \cdot \mathbf{x}^i \leq (t\mathbf{p}) \cdot \mathbf{e}^i$. Hence, $\mathbf{x}^i (\mathbf{p}) = \mathbf{x}^i (t\mathbf{p})$.
  - For any $t > 0$, $z_j (t\mathbf{p}) = \sum_{i=1}^n x_j^i (t\mathbf{p}) - \sum_{i=1}^n e_j^i = \sum_{i=1}^n x_j^i (\mathbf{p}) - \sum_{i=1}^n e_j^i = z_j (\mathbf{p})$.
- Walras' law
  - Budget balance for any agent $i \in \mathcal{I}$, $\mathbf{p} \cdot \mathbf{x}^i (\mathbf{p}) = \mathbf{p} \cdot \mathbf{e}^i$. (This must hold because monotonicity of preference relation, which is stronger than locally non-satiation).
  - It follows that

$$\begin{aligned}\mathbf{p}\cdot\mathbf{z}(\mathbf{p}) &= \mathbf{p}\cdot\left(\sum_{i=1}^{n}\mathbf{x}^i(\mathbf{p})-\sum_{i=1}^{n}\mathbf{e}^i\right)\\ &= \sum_{i=1}^{n}\left(\mathbf{p}\cdot\mathbf{x}^i(\mathbf{p})-\mathbf{p}\cdot\mathbf{e}^i\right)\\ &= 0\end{aligned}$$

Existence ***Theorem (Existence of Walrasian Equilibrium)*** Let $\mathcal{E} = \left(u^i(\cdot), \mathbf{e}^i\right)_{i\in\mathcal{I}}$ be a perfectly competitive pure exchange economy that satisfies (A1) through (A4), then a Walrasian equilibrium exists.

Proof Focus on the case of $m = 2$ and $\succeq^i$ being strictly convex for any $i \in \mathcal{I}$. Given (A1) and the strict convexity of $\succeq^i$, $\mathbf{x}^i(p_1, p_2)$ is always unique for any $\mathbf{p} \gg \mathbf{0}$. When $m = 2$, Walras' law says that

$$p_1 z_1(p_1, p_2) + p_2 z_2(p_1, p_2) = 0$$

then a Walrasian equilibrium is given by $z_1(p_1, p_2) = 0$ or $z_2(p_1, p_2) = 0$. By (A2) and (A3), the only possibility is $p_1, p_2 > 0$ in equilibrium, so we can normalize $p_1 = 1$ or $p_2 = 1$.

By theorem of the maximum, given (A1) to (A3), $x_i^1(p_1, 1)$ is continuous for any $i \in \mathcal{I}$. Hence, $z_1(p_1, 1)$ is continuous on $(0, +\infty)$. Given monotonicity and strict convexity, $u^i(\cdot)$ is strongly monotone. Therefore, $\lim_{p_1\to 0+} x_1^i(p_1, 1) = +\infty$ and $\lim_{p_1\to 0+} z_1(p_1, 1) = +\infty$. Symmetrically, we have $\lim_{p_2\to 0+} x_2^i(1, p_2) = +\infty$ and $\lim_{p_2\to 0+} z_2(1, p_2) = +\infty$. By continuity, $\exists 0 < \underline{p} < \bar{p}$ such that $z_1\left(\underline{p}, 1\right) > 0 > z_1(\bar{p}, 1)$. By the intermediate value theorem, $\exists p_1^* \in \left(\underline{p}, \bar{p}\right)$ such that $z_1(p_1^*, 1) = 0$.

7.2 Allocation

- First theorem of welfare economics: Any Walrasian equilibrium is Pareto efficient.
- Second theorem of welfare economics: By adjusting the initial endowments. any Pareto efficient allocation can be supported by a Walrasian equilibrium.

7.2.1 Definition

Definition (Feasible Allocation) For a pure exchange economy $\mathcal{E}$, an allocation $\left(\mathbf{x}^i\right)_{i=1}^n \geq \mathbf{0}$ is **feasible** if for any good j, $\sum_{i=1}^n x_j^i \leq \sum_{i=1}^n e_j^i$.

That is, the total allocation for any good cannot be more than the total initial endowment.

Definition (Pareto Efficient Allocation) Given an economy $\mathcal{E}$, a feasible allocation $\left(\mathbf{x}^i\right)_{i=1}^n$ is (strongly) **Pareto efficient** if there is no other feasible allocation $\left(\mathbf{w}^i\right)_{i=1}^n$ such that $\mathbf{w}^i \succeq_i \mathbf{x}^i$ for all $i \in \mathcal{I}$, with $\mathbf{w}^k \succ_k \mathbf{x}^k$ for some $k \in \mathcal{I}$.

Under Pareto efficient allocation, there is no way to make any agent strictly better off without harming some other agent(s).

Definition (Core Allocation) A feasible allocation $\left(\mathbf{x}^i\right)_{i=1}^n$ is in the **core** of $\mathcal{E}$ if there is no group $\mathcal{I}_0 \in I$ and an alternative allocation $\left(\mathbf{w}^i\right)_{i=1}^n$ such that:

1. For any $j \in \mathcal{J}$, $\sum_{i\in\mathcal{I}_0} w_j^i \leq \sum_{i\in\mathcal{I}_0} e_j^i$.
2. $\mathbf{w}^i \succeq_i \mathbf{x}^i$ for all $i \in \mathcal{I}_0$, with $\mathbf{w}^k \succ_k \mathbf{x}^k$ for some $k \in \mathcal{I}_0$.

Without any group (not necessarily the grand coalition), there is no way to make any agent strictly better off without harming some other agent(s).

7.2.2 Strengthened First Theorem of Welfare Economics

- S_C: The set of Pareto efficient allocations.
- S_C: The set of core allocations.
- S_W: The set of Walrasian equilibrium allocations.

Theorem (Strengthened First Theorem of Welfare Economics) Let $\mathcal{E}$ be a perfectly competitive pure exchange economy satisfying (A1) and (A2), then $S_W \subseteq S_C \subseteq S_P$.

- $S_W \subseteq S_P$: First theorem of welfare economics.
- $S_W \subseteq S_C$: Not only efficient, but also fair, which is a justification for the price machanism.
- Very minimal requirement on the agents' preference relations, but perfect competition, complete information, no externality and complete markets are indispensable.

Proof By definition, $S_C \subseteq S_P$, so it suffices to show $S_W \subseteq S_C$. Suppose not, then there exists a Walrasian equilibrium $\left(\mathbf{p}; \left(\mathbf{x}^i\right)_{i\in\mathcal{I}}\right)$, a group $\mathcal{I}_0 \in \mathcal{I}$ and an alternative allocation $\left(\mathbf{w}^i\right)_{i\in\mathcal{I}}$ such that:

$$\begin{cases} \mathbf{w}^i \succeq \mathbf{x}^i, \text{ for all } i \in \mathcal{I}_0 \\ \mathbf{w}^k \succeq \mathbf{x}^k, \text{ for some } k \in \mathcal{K}_0 \subseteq \mathcal{I}_0 \end{cases}$$

By direct revealed preference,

$$\begin{cases} \mathbf{p}\cdot\mathbf{w}^k > \mathbf{p}\cdot\mathbf{x}^k, \text{ for all } k\in\mathcal{K}_0 \\ \mathbf{p}\cdot\mathbf{w}^i \geq \mathbf{p}\cdot\mathbf{x}^i, \text{ for all } i\in\mathcal{I}_0\backslash\mathcal{K}_0 \end{cases}$$

Summing over $i\in\mathcal{I}_0$, we have

$$\begin{aligned} &\mathbf{p}\cdot\sum_{i\in\mathcal{I}_0}\mathbf{w}^i \geq \mathbf{p}\cdot\sum_{i\in\mathcal{I}_0}\mathbf{x}^i = \mathbf{p}\cdot\sum_{i\in\mathcal{I}_0}\mathbf{e}^i \\ \Longrightarrow &\sum_{i\in\mathcal{I}_0}\mathbf{w}^i \not\leq \sum_{i\in\mathcal{I}_0}\mathbf{e}^i \end{aligned}$$

Theorem (Second Theorem of Welfare Economics) Let $\mathcal{E}$ be a perfectly competitive pure exchange economy satisfying (A1) through (A3). If $\mathbf{x}^i \gg \mathbf{0}$, for all i, is Pareto efficient, then there exists an initial endowment $\left(\mathbf{e}^i\right)_{i\in\mathcal{I}}$ and a price vector $\mathbf{p}\geq\mathbf{0}$ such that $\left(\mathbf{p};\left(\mathbf{x}^i\right)_{i\in\mathcal{I}}\right)$ is a Walrasian equilibrium given these endowments.

However, the second theorem of welfare economics is less exciting, because the government cannot

1. Arbitrarily redistribute initial endowment(s).
2. Know each agent's preference in the economy.

7.3 GE with Production

7.3.1 Model Preliminaries

Recall that in producer theory, we used production sets and ownership shares to describe the firms and their production technologies.

Introduce K firms ($\mathcal{K}=\{1,2,\cdots,K\}$), with respective production set Y^k. Let $\alpha^{ki}\geq 0$ be agent i's ownership share of firm k. The production economy can then be described as

$$\mathcal{E}=\left(\left(u^i,\mathbf{e}^i,\left(\alpha^{ki}\right)_{k\in\mathcal{K}}\right)_{i\in\mathcal{I}},\left(Y^k\right)_{k\in\mathcal{K}}\right)$$

Definition (Walrasian Equilibrium with Production) A **Walrasian Equilibrium** for a perfectly competitive production economy $\mathcal{E}=\left(\left(u^i,\mathbf{e}^i,\left(\alpha^{ki}\right)_{k\in\mathcal{K}}\right)_{i\in\mathcal{I}},\left(Y^k\right)_{k\in\mathcal{K}}\right)$ is a vector $\left(\mathbf{p},\left(\mathbf{x}^i\right)_{i\in\mathcal{I}},\left(\mathbf{y}^k\right)_{k\in\mathcal{K}}\right)$ such that:

1. Profit maximization: Firm k solves

$$\max_{y^k\in Y^k}\mathbf{p}\cdot\mathbf{y}^k$$

2. Utility maximization: Agent i solves

$$\max_{\mathbf{x}^i \geq \mathbf{0}} u^i\left(\mathbf{x}^i\right) \text{ s.t. } \mathbf{p}\cdot\mathbf{x}^i \leq \mathbf{p}\cdot\mathbf{e}^i + \sum_{k\in\mathcal{K}} \alpha^{ik}\mathbf{p}\cdot\mathbf{y}^k$$

3. Market clearing (for all goods):

$$\sum_{i\in\mathcal{I}} \mathbf{x}^i\left(\mathbf{p}\right) = \sum_{i\in\mathcal{I}} \mathbf{e}^i + \sum_{k\in\mathcal{K}} \mathbf{y}^k\left(\mathbf{p}\right)$$

That is, in Walrasian equilibrium

- All markets clear at the same time.
- All consumers and firms optimize.

Assumptions on production technology:

- (A5): For all firms $k \in \mathcal{K}$, $Y^k \neq \emptyset$ is closed and convex.
- (A6): For all firms $k \in \mathcal{K}$, shut-down and free-disposal properties, that is, $\mathbf{0} \in Y^k$, and $\mathbf{y} \in Y^k$ implies $\mathbf{y}' \in Y^k$, for all $\mathbf{y}' \leq \mathbf{y}$.
- (A7): Irreversibility: Let $Y = \cup_{k\in\mathcal{K}} Y^k$, then $Y \cap (-Y) = \{\mathbf{0}\}$.

Remarks:

- $Y^k \neq \emptyset$ is innocuous, otherwise since firm k can produce nothing, we can just discard Y^k.
- Irreversibility just says that the production cannot be completely reversed.

7.3.2 Existence of Walrasian Equilibrium with Production

Theorem (Existence of Walrasian Equilibrium with Production) Let $\mathcal{E} = \left(\left(u^i, \mathbf{e}^i, \left(\alpha^{ki}\right)_{k\in\mathcal{K}}\right)_{i\in\mathcal{I}}, \left(Y^k\right)_{k\in\mathcal{K}}\right)$ be a perfectly competitive production economy satisfying (A1) through (A7), then a Walrasian equilibrium $\left(\mathbf{p}, \left(\mathbf{x}^i\right)_{i\in\mathcal{I}}, \left(\mathbf{y}^k\right)_{k\in\mathcal{K}}\right)$ exists.

Remarks:

- Generalize the existence result on pure exchange economies.
- The first and second theorem of welfare economics also generalize to production economics.

Recall that when a production technology is constant returns to scale, the firm's profit must be 0 or $+\infty$. For market clearing, it must be that each firm is making 0 profit.

For simplicity, consider two goods. Transform b units of good 1 into (at most) c units of good 2, that is, linear transformation. The production set can be depicted as

$$Y = \left\{ (y_1, y_2) \,| y_1 \leq 0, y_2 \geq 0, y_2 \leq -\frac{c}{b} y_1 \right\}$$

Alternatively, we can summarize the production technology with the vector $a = \left(-1, \frac{c}{b}\right)$.

If a linear activity production is ever used in a Walrasian equilibrium (i.e., $\lambda^* > 0$), then $(p_1^*, p_2^*) \cdot (a_1, a_2) = 0$.

Example Consider a perfectly competitive economy $\mathcal{E}$ with two goods (1 and 2) and two agents (A and B). The agents' utility functions and initial endowments are as follows:

$$u^A (x_1, x_2) = x_1 x_2, \mathbf{e}^A = (1, 0)$$
$$u^B (x_1, x_2) = x_1 x_2^2, \mathbf{e}^B = (0, 1)$$

First suppose the economy is pure exchange with no production.

1. Derive the set of Pareto efficient allocations.
2. Derive the set of core allocations.
3. Derive a Walrasian equilibrium $\left((p_1, p_2); \left(x_1^A, x_2^A, x_1^B, x_2^B\right)\right)$.

Next we introduce production: Suppose there is a perfectly competitive firm which can transform one unit of good 1 into (at most) one unit of good 2, that is, $\mathbf{a} = (-1, 1)$.

1. Derive a Walrasian equilibrium with production $\left((p_1, p_2,); \left(x_1^A, x_2^A, x_2^B, x_2^B\right), \lambda\right)$.

Solution Since both agents' preference relations are monotonic, any Pareto efficient allocation must satisfy

$$\mathbf{x}^A + \mathbf{x}^B = \mathbf{e}^A + \mathbf{e}^B$$

Clearly, $((1,1),(0,0))$ and $((0,0),(1,1))$ are Pareto efficient. (**Remarks:** $((1,0),(0,1))$ or $((0,1),(1,0))$ are not Pareto efficient, though we cannot make any agent better off in material without making the other worse off in material meaning. However, material meaning is not equivalent to utility meaning.)

Otherwise, due to no gain from trade, we must have

$$|MRS^A| = |MRS^B|$$

Therefore, the set of efficient allocations is given by

$$S_P = \left\{ \left(x_1^A, x_2^A, x_1^B, x_2^B\right) \geq \mathbf{0} : x_1^A x_2^B = 2 x_2^A x_1^B, x_1^A + x_1^B = x_2^A + x_2^B = 1 \right\}$$

When there are only two agents, the only subgroups besides the grand coalition are $\{1\}$ and $\{2\}$. In other words, the only additional assumption is that **no agent can receive a worse allocation than their endowment**, which is also termed individual rationality. Since the initial endowment is the worst possible allocation for each agent, the set of core allocations $S_C = S_P$.

Similar to Example 2 from the last lecture, suppose there is an equilibrium price vector (p_1, p_2), then agent A's "income" $m^A = p_1$ and agent B's "income" $m^B = p_2$. The optimal individual demands

$$\begin{aligned}\left(x_1^A, x_2^A\right) &= \left(\frac{1}{2}, \frac{p_1}{2p_2}\right)\\ \left(x_1^B, x_2^B\right) &= \left(\frac{p_2}{3p_1}, \frac{2}{3}\right)\end{aligned}$$

In equilibrium, market demand equals market endowment for each good.

$$\begin{cases}\frac{1}{2} + \frac{p_2}{3p_1} = 1\\ \frac{p_1}{2p_2} + \frac{2}{3} = 1\end{cases} \Longrightarrow \frac{p_1^*}{p_2^*} = \frac{2}{3}$$

Hence, a Walrasian equilibrium is $\left(\mathbf{p} = (2, 3), \mathbf{x}^A = \left(\frac{1}{2}, \frac{1}{3}\right), \mathbf{x}^B = \left(\frac{1}{2}, \frac{2}{3}\right)\right)$. (**Remarks:** Here "a" is emphasized because the relative prices do not matter in Walrasian equilibrium.)

Suppose the production technology is used in equilibrium, then we must have $(p_1, p_2) \cdot \mathbf{a} = \mathbf{0}$, that is, $p_1 = p_2$, $\dfrac{p_1}{p_2} = 1$.

In equilibrium, market demand equals market supply for each good.

$$\begin{cases}\frac{1}{2} + \frac{1}{3} & = 1 - \lambda\\ \frac{1}{2} + \frac{2}{3} & = 1 + \lambda\end{cases} \Longrightarrow \lambda^* = \frac{1}{6}$$

A Walrasian equilibrium is given by $\left(\mathbf{p}^* = (1, 1), \mathbf{x}^A = \left(\frac{1}{2}, \frac{1}{2}\right), \mathbf{x}^B = \left(\frac{1}{3}, \frac{2}{3}\right), \lambda^* = \frac{1}{6}\right)$.