Lecture Note
Microeconomic Theory I

Duk Gyoo Kim

Current version: March 16, 2023

Abstract

This lecture note is prepared for the first sequence of graduate-level microeconomic theory. Some sections may come with complete sentences, but I may merely list items for the other sections. During this semester, this note will always be incomplete. I never intend this note to replace the existing textbooks, so please refer to the textbooks when this note confuses you rather than helps.

0 Some thoughts

Is microeconomic theory dead? Yes and no. Yes in the sense that some theories are outdated and never give any new insights, and that tremendous amount of empirical data awaits researchers to analyze it. No in the sense that

- it helps you to train yourself to think like economists,
- some classic analyses can give profound insights about new environments, and
- more practically, all other economists assume that you have already learned topics typically covered in microeconomic theory.

The first sequence of graduate-level microeconomic theory tends to be math heavy: This is the same for every grad program you can think of. Given that it is common knowledge that only a few students would need advanced microeconomic theory for their research, there must be some reasons why economists use math. We use mathematics as a tool

- to describe model and assumptions precisely,
- to make analysis rigorous,

- to obtain results that may not be available through verbal arguments, and
- to reduce unnecessary debates.

1 Consumer Theory

There are four building blocks for consumer theory: consumption set, budget set, preference relation, and behavioral assumption.

1.1 Consumption Set and Budget Constraint

Assume that there are $L \in \mathbb{N}^+$ goods. A consumption set is a set of all possible consumption bundles $\equiv X = \{(x_1, x_2, \ldots, x_L)\}$.

- We typically consider $X = \mathbb{R}^L_+$, but it could be more specific.
- We assume $X \subseteq \mathbb{R}^L_+$ is closed and convex.

A budget set (or budget constraint), $B(p,m)$, is a set of all affordable consumption bundles given prices $p = (p_1, p_2, \ldots, p_L)$ and income $m \geq 0$, which is

$$B(p,m) = \{x \in X | p \cdot x \leq m\}$$

I am sure that you had drawn a budget constraint with two goods.

1.2 Preference

Preference is a binary relation on the consumption set X.

- $x \succeq y$: "$x \in X$ is at least as good as $y \in X$." or "x is weakly preferred to y."
- $x \succ y$ iff $x \succeq y$ and $y \not\succeq x$: "x is strictly preferred to y."
- $x \sim y$ iff $x \succeq y$ and $y \succeq x$: "x is indifferent to y."

We can derive subsets of X from the preference relation. For a given bundle $y \in X$,

- Upper contour set: $P(y) = \{x \in X | x \succeq y\}$.
- Strictly upper contour set: $P_s(y) = \{x \in X | x \succ y\}$.
- The lower contour set and strictly lower contour set are denoted by $L(y)$ and $L_s(y)$.

- Indifferent set (or indifferent curve): $I(y) = \{x \in X | x \sim y\}$.

We hope the preference relation to have some desirable properties.

1. Completeness: $\forall x, y \in X$, $x \succeq y$ or $y \succeq x$.

2. Transitivity: $\forall x, y, z \in X$, if $x \succeq y$ and $y \succeq z$, then $x \succeq z$.

We call the preference relation is rational if it is complete and transitive. In most situations, we assume an individual's preference is rational, but you can find some counterexamples. (If you ever said, "they are like apples and oranges, so I really can't tell which one I prefer," or "I am on a diet. I prefer (diet) coke over salads, salads over hamburgers, but when coke and hamburgers are present, then I lose my control and go with hamburgers!" then you are not considered rational.) Some textbooks include reflectivity ("For any $x \in X$, $x \sim x$.") as another desired property of the preference relation, but completeness and transitivity implies reflectivity.

Some more desirable properties are:

3. Monotonicity: $\forall x, y \in X$, if $x \geq y$, $x \succeq y$, while if $x \gg y$, then $x \succ y$.

   Strict monotonicity: $\forall x, y \in X$ with $x \neq y$, if $x \geq y$, then $x \succ y$.

4. Continuity: $\forall y \in X$, $P(y)$ and $L(y)$ are closed. (Or, $P_s(y)$ and $L_s(y)$ are open.)

5. Non-satiation: $\forall x \in X$, $\exists y \in X$ such that $y \succ x$. (There is no bliss point.)

6. Local non-satiation: $\forall x \in X$ and $\forall \epsilon > 0$, $\exists y \in X$ with $|y - x| < \epsilon$ such that $y \succ x$. (The indifference curve is a line, not a band.)

7. Convexity: If $x \succeq y$, then $tx + (1-t)y \succeq y$ $\forall t \in [0,1]$.

   Strict convexity: If $x \succeq y$ and $x \neq y$, then $tx + (1-t)y \succ y$ $\forall t \in (0,1)$.

Remark: If $\succeq$ is a locally non-satiated upper semicontinuous preference, then there is a monotonic preference $\succeq^*$ that generates the same demand. Thus, local non-satiation is no more general than monotinicity from the point of view of competitive demand theory. See Kim Border's lecture note.

Digression: A set $X \subseteq \mathbb{R}^n$ is open if for any $x \in X$, you can draw a tiny "epsilon-ball" around x and the epsilon ball is inside of X (that is, any element in the epsilon ball is an element of X). A set X is closed if the complement of the set is open. One way to check the closedness is to consider a sequence x^n approaches x. If every x^n is in the set X, but $x = \lim_{n\to\infty} x^n$ is not in the set X, then it is not closed.

- -

It is known that a lexicographic preference does not hold continuity.

Exercise: Assume $L = 2$. A lexicographic preference $\succeq$ is defined as follows: $x \succeq y$ if and only if either $x_1 \geq y_1$ or $x_1 = y_1$ and $x_2 \geq y_2$. Show that this preference does not hold continuity. [Hint: compare a sequence $(1+1/n, 1)$ and $(1,2)$.]

1.3 Utility Function

Definition 1. *A function $u : X \to \mathbb{R}$ represents the preference relation $\succeq$ if and only if*

$$\forall x, y \in X, u(x) \geq u(y) \Leftrightarrow x \succeq y.$$

Some famous utility functions:

1. Cobb-Douglas: $u(x) = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_L^{\alpha_L}$ with $\alpha_l > 0$ for all l. It is continuous, strictly monotone, and strictly convex in $\mathbb{R}^L_{++}$.

2. Linear: $u(x) = \sum_{l=1}^{L} \alpha_l x_l$. It is continuous, strictly monotone, and convex in $\mathbb{R}^L_+$.

3. Leontief: $u(X) = \min\{\alpha_1 x_1, \ldots, \alpha_L x_L\}$. It is continuous, monotone, and convex in $\mathbb{R}^L_+$.

4. Homogeneous function of degree n: For every $a > 0$, $u(ax) = a^n u(x)$. A homogeneous function of degree 1 is called a homothetic function. Cobb-Douglas function with $\sum_{l=1}^{L} \alpha_l = 1$ is a homothetic function.

Note that we can compare any two numbers, and numbers are transitive. This means, if a preference is represented by a utility function (that maps consumption bundles to numbers), then it must be complete and transitive. However, the converse is not true: Not every rational preference can be represented by a utility function.

Theorem 1 (Existence of a Utility Function)**.** *Suppose that preference relation $\succeq$ is complete, transitive, continuous, and strictly monotonic. Then, there exists a continuous utility function $u : \mathbb{R}^L_+ \to \mathbb{R}$ representing $\succeq$.*

Proof: Let $e = (1, 1 \ldots, 1) \in \mathbb{R}^L_+$. Given any vector $x \in \mathbb{R}^L_+$, let $u(x)$ is defined such that $x \sim u(x)e$, that is, $(x_1, x_2, \ldots, x_L) \sim (u(x), u(x), \ldots, u(x))$. We first show that $u(x)$ exists and is unique. Then, we show that $u(\cdot)$ represents $\succeq$. Finally, we show that $u(\cdot)$ is continuous.
Existence: Let $B = \{t \in \mathbb{R} | te \succeq x\}$ and $W = \{t \in \mathbb{R} | x \succeq te\}$. Because of the strict monotonicity of $\succeq$, both B and W are nonempty, and $B \cup W = \mathbb{R}$. Also, the continuity of $\succeq$ implies both B and W are closed. Since the real line is continuous, there exists $t_x \in \mathbb{R}$ such that $t_x e \sim x$.

Uniqueness: If $t_1 e \sim x$ and $t_2 e \sim x$, then $t_1 e \sim t_2 e$ by the transitivity of $\succeq$. By strict monotonicity, $t_1 = t_2$. (If not, say $t_1 > t_2$, it violates the strict monotonicity.)
Representativeness: Consider any $x^1, x^2 \in \mathbb{R}^L_+$. Without loss of generality, suppose $x^1 \succeq x^2$. $u(x^1)e \sim x^1$ and $x^2 \sim u(x^2)e$ imply $u(x^1)e \succeq u(x^2)e$, which in turn implies $u(x^1) \geq u(x^2)$.
Continuity: Consider a sequence $\{x^m\}$ with $x^m \to x$. Suppose for the sake of contradiction that $u(x^m) \not\to u(x)$. Consider the case $u' := \lim_{m\to\infty} u(x^m) > u(x)$. Then, by monotonicity, $u'e \succ u(x)e$. Let $\hat{u} := \frac{u'+u(x)}{2}$. By monotonicity, $\hat{u}e \succ u(x)e$. Since $u(x^m) \to u' > \hat{u}$, there exists $M \in \mathbb{N}_+$ such that for all $m > M$, $u(x^m) > \hat{u}$. For all such m, $x^m \sim u(x^m)e \succ \hat{u}e$. By the continuity of $\succeq$, this would imply $x \succeq \hat{u}e$, which in turn implies $u(x)e \sim x \succeq \hat{u}e$, which is a contradiction. The opposite case ($u' < u(x)$) is analogous. $\square$

We have learned that the "well-behaved" preference relation can be represented by a utility function. Recall that lexicographic preference is not continuous, which means that it cannot be represented by a utility function. (Try to cook up one function, you will soon realize a problem.)

The utility function is not unique: A monotonic transformation of one utility function is another utility function representing preference relation $\succeq$.

Proposition 1 (Invariance of Utility Function to Positive Monotonic Transforms). *If $u(x)$ represents some preference $\succeq$ and $f : \mathbb{R} \to \mathbb{R}$ is strctly increasing, then $v(x) = f(u(x))$ represents the same preference.*

Proof: For all $x, y \in \mathbb{R}^L_+$, $u(x) \geq u(y)$ if and only if $v(x) = f(u(x)) \geq f(u(y)) = v(y)$. $\square$

The utility function inherits the properties of the preference relation that it represents. For example, $\succeq$ is strictly monotonic if and only if $u(\cdot)$ is strictly increasing. $\succeq$ is (strictly) convex if and only if $u(\cdot)$ is (strictly) quasiconcave.

- -

Digression: A function $f : X \to \mathbb{R}$ defined on a convex set $X \subset \mathbb{R}^L$ is quasiconcave if for all $x, y \in X$ and $\lambda \in [0,1]$, $f(\lambda x + (1-\lambda)y) \geq \min\{f(x), f(y)\}$. It is strictly quasiconcave if for all $x, y \in X$ and $\lambda \in (0,1)$, $f(\lambda x + (1-\lambda)y) > \min\{f(x), f(y)\}$. As long as you rely on the Lagrangian method, you have quasiconcavity in mind. Why are we bothered by "quasi"concavity when concavity is easy to understand? Because quasiconcavity is an ordinal property (like the utility function), while concavity is a cardinal property. A monotone transformation of a quasiconcave function is still quasiconcave, but a monotone transformation of a concave function is not necessarily concave. Think about $f(x) = \sqrt{x}$, and a monotone transformation $v(f) = f^4$. $v(f(x)) = x^2$ is not concave.

It is also known that $u(x)$ representing a convex preference relation $\succeq$ if and only if $u(x)$ is quasiconcave. (Check JR.)

- -

Although we have never said anything about the diffentiability of the utility function, we often assume that the utility function is smooth enough to be differentiable (twice). If the utility function is differentiable, we call the first-order partial derivative of $u(x)$ with respect to x_i the marginal utility of good i,

$$MU_i(x) = \frac{\partial u(x)}{\partial x_i}.$$

A ratio of two marginal utilities of goods i and j is called the marginal rate of substitution between i and j,

$$MRS_{ij}(x) = \frac{\partial u(x)/\partial x_i}{\partial u(x)/\partial x_j},$$

which measures the (absolute value of) gradient of indifference hyperplane at consumption bundle x. (For two goods, it captures the slope of indifference curve at x.) A good feature of the MRS is that it is invariant to the monotone transform of the utility function. If $v(x) = f(u(x))$, and $f(\cdot)$ is monotonic, then

$$\frac{\partial v(x)/\partial x_1}{\partial v(x)/\partial x_2} = \frac{f'(x)\partial u(x)/\partial x_1}{f'(x)\partial u(x)/\partial x_2} = \frac{\partial u(x)/\partial x_1}{\partial u(x)/\partial x_2}$$

(Thus the MRS is an ordinal, not cardinal, concept.)

It is known that if the utility function is quasiconcave and twice continuously differentiable, its Hessian matrix $H(x)$ satisfies $y'H(x)y \le 0$ for any vector y such that $\nabla u(x) \cdot y = 0$. (The Hessian matrix of a twice continuously differentiable concave function is negative semi-definite. So you can say that $H(x)$ of a quasiconcave utility function is negative semi-definite in the subspace $\{y \in \mathbb{R}^L | \nabla u(x) y = 0\}$.)

- -

Digression: "A twice continuously differentiable concave function's Hessian is negative semi-definite." In this digression, we learn some necessary steps to understand the quoted sentence. A $N \times N$ matrix M is negative semi-definite if for all $z \in \mathbb{R}^N$, $z'Mz \le 0$. A twice continuously differentiable function on a convex set $X \subset \mathbb{R}^N$ is concave if for all $x, y \in X$ and for all $\lambda \in [0,1]$, $f(\lambda x + (1-\lambda)y) \ge \lambda f(x) + (1-\lambda) f(y)$. Equivalently, a function is concave if and only if

$$f(x+z) \le f(x) + \nabla f(x) \cdot z \quad \forall x \in X \text{ and } x+z \in X.$$

This is because

$$
\begin{aligned}
&f(\lambda(x+z)+(1-\lambda)x) \geq \lambda f(x+z)+(1-\lambda)f(x)\\
\Leftrightarrow &f(x+\lambda z) \geq f(x)+\lambda(f(x+z)-f(x))\\
\Leftrightarrow &f(x)+\frac{f(x+\lambda z)-f(x)}{\lambda} \geq f(x+z)
\end{aligned}
$$

$\lim_{\lambda\to 0}\frac{f(x+\lambda z)-f(x)}{\lambda} = \nabla f(x)\cdot z$. Taking a second-order Taylor expansion of the function $\phi(\alpha) = f(x+\alpha z)$ around $\alpha = 0$, we have

$$
f(x+\alpha z) = f(x)+\nabla f(x)\cdot(\alpha z)+\frac{\alpha^2}{2}z'D^2 f(x+\beta z)z,
$$

where $\beta \in [0,\alpha]$. Since f is concave, $\frac{\alpha^2}{2}z'D^2f(x+\beta z)z = f(x+\alpha z)-f(x)-\nabla f(x)\cdot(\alpha z) \leq 0$. Since α, hence β, can be arbitrarily small, this gives the conclusion $z'D^2f(x)z \leq 0$. For more details, check MWG pages 930–940.

- -

1.4 Utility Maximization and Marshallian Demand

Let's review the Lagrangian method first.

- -

Lagrangian Method

Let f and h_i, $i = 1,\ldots,m$ be concave $\mathbb{C}^1$ functions defined on the open and convex set $U \subset \mathbb{R}^n$. Consider the following maximization problem:

$$
\max_{x\in U} f(x) \quad \text{subject to } x \in D = \{z \in U | h_i(z) \geq 0, i = 1,\ldots,m\}
$$

(What if f and h_i are convex? Redefine $-f$ and $-h_i$ as f and h_i.) We typically call $f(x)$ the objective function, and h_is the constraints. Set up the Lagrangian for this problem as follows:

$$
L(x,\lambda) = f(x)+\sum_{i=1}^{m}\lambda_i h_i(x),
$$

where $\lambda = (\lambda_1,\ldots,\lambda_m)$. We call a pair $(x^*,\lambda^*) \in U\times\mathbb{R}^m_+$ *saddle point* of L if it satisfies the following conditions:

$$
\nabla f(x^*)+\lambda_i^*\nabla h_i(x^*) = 0, \forall i = 1,\ldots,m \tag{1}
$$

$$
h_i(x^*) \geq 0 \text{ and } \lambda_i^* h_i(x^*) = 0, \forall i = 1,\ldots,m \tag{2}
$$

Here (1) implies that (x^*,λ^*) maximizes L, and (2) implies that given x^*, λ^* minimizes L. The term, saddle point, is derived from these two conditions. Condition (2) is called complementary slackness: If $h_i(x^*) = 0$ (that is, if the constraint i is binding,) then λ_i^* can be slack, and if $h_i(x^*) > 0$, then λ_i^* should be zero. Checking this complementary slackness condition is important to figure out "corner solutions."

Theorem 2 (Kuhn-Tucker)**.** *$x^* \in U$ solves the above problem if and only if there is $\lambda^* \in \mathbb{R}_+^m$ such that (x^*, λ^*) is a saddle point of L.*

Remark: If functions f and h_i are quasi-concave, then x^* maximizes f over D provided at least one of the following conditions holds: (a) $\nabla f(x^*) \neq 0$; (b) f is concave. (It is called Arrow–Enthoven theorem.)

Exercise: $\max \sqrt{x+1} - x$ subject to $x \geq 0$ and $\frac{x}{3} + \frac{1}{2} \geq 0$. (Try to use the Lagrangian.)

- -

A fundamental idea in the consumer theory is that a rational consumer chooses the most preferred consumption bundle from the set of affordable alternatives. A consumer's decision is then described as a utility maximization problem: For $p \gg 0$ and $m > 0$,

$$\max_{x \in \mathbb{R}_+^L} u(x) \quad \text{s.t. } x \in B(p,m)$$

- By the Weierstrass' extreme value theorem, at least one solution exists. ($B(p,m)$ is compact (closed and bounded) and $u(x)$ is continuous.)
- If $u(x)$ is strictly quasiconcave, then the solution is unique. [Prove this claim.]
- If $u(x)$ is locally non-satiated, then the budget constraint is binding at the optimum, that is, $p \cdot x = m$. [Prove this claim.]

Marshallian demand function is

$$x(p,m) = \arg\max_{x \in \mathbb{R}_+^L} u(x) \quad \text{s.t. } x \in B(p,m).$$

- By calling it "function," we already assume that the solution of the maximization problem is unique. Have in mind that we mostly deal with strictly quasiconcave utility functions.
- By Berge's Maximum Theorem, $x(p,m)$ is continuous. ($u(x)$ is continuous, and $B(p,m)$ is continuous and convex.)
- Note that $B(p,m) = B(tp,tm)$ for $t > 0$, which implies that $x(tp,tm) = x(p,m)$. Thus the Marshallian demand function is homogeneous of degree zero.

For notational convenience, let $\lambda = (\lambda_0, \lambda_1, \ldots, \lambda_L)$. Set up the Lagrangian function as

$$L(x,\lambda) = u(x) + \lambda_0[m - p \cdot x] + \sum_{l=1}^{L} \lambda_l x_l,$$

where $\lambda_0 \geq 0$ is the Lagrangian multiplier associated with the budget constraint, and $\lambda_l \geq 0$ is the multiplier associated with the nonnegativty condition of good i. (You cannot consume negative amount of goods.)

- According to the Kuhn-Tucker theorem, if $x^* = x(p,m)$, then there exists $\lambda^* \in \mathbb{R}_+^{L+1}$ such that
$$\frac{\partial u(x^*)}{\partial x_l} - \lambda^* p_l + \lambda_l^* = 0 \quad \forall l = 1, \ldots, L$$
and
$$\lambda_0^*[m - p \cdot x^*] = 0 \text{ and } \lambda_l^* x_l^* = 0 \quad \forall l = 1, \ldots, L$$
It implies
$$\frac{\partial u(x^*)}{\partial x_l} \leq \lambda^* p_l \text{ with equality if } x_l^* > 0.$$

- If $x^* \gg 0$, then we call it an interior solution. For all $l,k = 1,2,\ldots,L$,
$$\frac{\partial u(x^*)}{\partial x_l} = \lambda^* p_l \text{ and } \frac{\partial u(x^*)}{\partial x_k} = \lambda^* p_k, \text{ implying } \frac{\partial u(x^*)/\partial x_l}{\partial u(x^*)/\partial x_k} = \frac{p_l}{p_k}.$$
That is, $MRS_{lk}(x)$ equals the price ratio p_l/p_k at x^*. (This doesn't hold when x^* is not interior.)

- λ^* measures the change in utility from a marginal increase in m. Assuming $x(p,m) \gg 0$, we have
$$\sum_{l=1}^{L} \frac{\partial(u(x(p,m))}{\partial x_l} \frac{\partial x_l(p,m)}{\partial m} = \sum_{l=1}^{L} \lambda^* p_l \frac{\partial x_l(p,m)}{\partial m} = \lambda^*,$$
where the last equality is derived from $p \cdot x(p,m) = m$. (Taking a derivative with respect to m, you get $p \cdot \nabla x(p,m) = 1$, or $\sum_{l=1}^{L} p_l \frac{\partial x_l(p,m)}{\partial m} = 1$.)

- -

Worked-out Example: Marshallian Demand for Cobb-Douglas

$u(x) = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_L^{\alpha_L}$. The consumer solves

$$\max_{x_1,\ldots,x_L} x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_L^{\alpha_L} \quad \text{s.t. } p_1 x_1 + p_2 x_2 + \cdots + p_L x_L \leq m.$$

The first order conditions:

$$\frac{\alpha_l}{x_l}u(x) = \lambda p_l, \quad \forall l = 1,2,\ldots,L$$

Thus, at the optimum, $p_k x_k = \frac{\alpha_k}{\alpha_1} p_1 l_1$ for all $k = 1,\ldots,L$. Substituting them into the budget constraint, we have

$$\frac{\alpha_1}{\alpha_1}p_1x_1 + \frac{\alpha_2}{\alpha_1}p_1x_1 + \cdots + \frac{\alpha_L}{\alpha_1}p_1x_1 = \frac{\alpha_1 + \alpha_2 + \cdots + \alpha_L}{\alpha_1}p_1x_1 = m \Rightarrow x_1 = \frac{\alpha_1}{\alpha_1 + \cdots + \alpha_L}\frac{m}{p_1}$$

Similarly, we can find $x_l = \frac{\alpha_l}{\sum_{l=1}^L \alpha_l}\frac{m}{p_l}$. That is, with a Cobb-Douglas utility function, the Marshallian demand for good l is

[relative weight of good l] × [the largest quantity of good l affordable with m and p_l].

If $\sum_{l=1}^L \alpha_l = 1$, (that is, if the function is homothetic) $x_l = \alpha_l \frac{m}{p_l}$.

The indirect utility function $v : \mathbb{R}^L_+ \times \mathbb{R}_+ \to \mathbb{R}$ is given by

$$v(p,m) = \max_{x \in \mathbb{R}^L_+} u(x) \text{ s.t. } p \cdot x \leq m.$$

Hence, $v(p,m) = u(x(p,m))$. (The Marshallian demand is the maximizing argument of the utility function subject to the budget constraint, and the indirect utility function is the maximized value of the utility function subject to the budget constraint.)

Proposition 2 (Properties of the Indirect Utility Function). *If $u(x)$ is continuous and locally non-satiated on $\mathbb{R}^L_+$ and $(p,m) \gg 0$, then the indirect utility function is*

1. *Homogeneous of degree zero, that is $v(tp,tm) = v(p,m)$,*
2. *Nonincreasing in p and strictly increasing in m,*
3. *Quasiconvex in p and m,*
4. *Continuous in p and m.*

proof(sketch): (1) $v(tp,tm) = u(x(tp,tm)) = u(x(p,m)) = v(p,m)$. (2) For $p' > p$, $B(p',m) \subseteq B(p,m)$. Thus, $v(p,m) = \max_{x \in B(p,m)} u(x) \geq \max_{x \in B(p',m)} u(x) = v(p',m)$. Local non-satiation condition implies $\lambda^* > 0$. Thus $v(p,m') > v(p,m)$ if $m' > m$. (4) follows from the Berge's Theorem of Maximum. For (3), consider (p,m) and (p',m') such that $v(p,m) \leq \bar{v}$ and $v(p',m') \leq \bar{v}$. Let $(p'',m'') = (tp + (1-t)p', tm + (1-t)m')$ for $t \in (0,1)$. We need to show $v(p'',m'') \leq \bar{v}$. We need to show that $B(p'',m'') \subseteq B(p,m) \cup B(p',m')$, but I leave it to you. Then $v(p'',m'') = \max_{x \in B(p'',m'')} u(x) \leq \max_{x \in B(p,m) \cup B(p',m')} u(x) \leq \bar{v}$.

1.5 Expenditure Minimization and Hicksian Demand

Now we study the dual problem of the utility maximization, the expenditure minimization problem: Given $p \gg 0$ and utility level $u \in \mathbb{R}$,

$$\min_{x \in \mathbb{R}^L_+} p \cdot x \quad \text{s.t.} \quad u(x) \geq u$$

The solution of this expenditure minimization problem, $h(p,u)$, is the **Hicksian demand**. The **expenditure function** is defined as $e(p,u) := p \cdot h(p,u)$.

Proposition 3 (Properties of the Expenditure Function)**.** *If $u(x)$ is continuous and locally non-satiated on $\mathbb{R}^L_+$, then $e(p,u)$ for $p \gg 0$ is*

(1) homogeneous of degree 1 in p;

(2) strictly increasing in u and nondecreasing in p;

(3) continuous in p and u; and

(4) concave in p.

In addition, if $u(x)$ is strictly quasiconcave, we have

(5) (Shephard's lemma) $h_l(p,u) = \frac{\partial e(p,u)}{\partial p_l}$.

proof(sketch): (1) First, note that $h(p,u)$ is homogeneous of degree 0. (check by yourself: the minimizing argument is invariant to the scale of the objective function.) Thus, $e(tp,u) = tp \cdot h(tp,u) = tp \cdot h(p,u) = te(p,u)$. (2, 3) check by yourself. (4) Given utility level u, let $p'' = \lambda p + (1-\lambda)p'$, we need to show $e(p'',u) \geq \lambda e(p,u) + (1-\lambda)e(p',u)$.

$$e(p'',u) = p'' \cdot h(p'',u) = \lambda p \cdot h(p'',u) + (1-\lambda)p' \cdot h(p'',u) \geq \lambda p \cdot h(p,u) + (1-\lambda)p' \cdot h(p',u),$$

where the last inequality follows since $p \cdot h(p'',u) \geq p \cdot h(p,u)$ and $p' \cdot h(p'',u) \geq p' \cdot h(p',u)$. The Sheppard's lemma can be proven in two ways. For any price vector $p \gg 0$ consider p' such that $p'_l > p_l$ and $p'_k = p_k$ for all $k \neq l$. Since $e(p',u) = p' \cdot h(p',u) \leq p' \cdot h(p,u) = p \cdot h(p,u) + (p'_l - p_l)h_l(p,u)$, $e(p',u) - e(p,u) \leq (p'_l - p_l)h_l(p,u)$. Similarly, since $e(p,u) = p \cdot h(p,u) \leq p \cdot h(p',u)$, $e(p',u) - e(p,u) \geq (p'_l - p_l)h_l(p',u)$. Combining the two inequalities above, we have

$$h_l(p',u) \leq \frac{e(p',u) - e(p,u)}{p'_l - p_l} \leq h_l(p,u).$$

By continuity of $h_l(p,u)$, $\lim_{p'\to p}\frac{e(p',u)-e(p,u)}{p'_l-p_l}=\frac{\partial e(p,u)}{\partial p_l}=h_l(p,u)$. Another way of showing (5) is to use the envelope theorem. For $\lambda=(\lambda_0,\lambda_1,\dots,\lambda_L)\geq 0$,

$$e(p,u)=\min_x L(x,\lambda)=\min_x p\cdot x+\lambda_0(u-u(x))-\sum_{l=1}^{L}\lambda_l x_l$$

By the envelope theorem,

$$\frac{\partial e(p,u)}{\partial p_l}=\left.\frac{\partial L(x,\lambda)}{\partial p_l}\right|_{x=h(p,u)}=h_l(p,u).$$

- -

The Envelope Theorem

(For more details and proofs, refer to the mathematical appendix in one of the textbooks.)

Consider the maximization problem, $\max_{x\in\mathbb{R}^L_+} f(x,\theta)$ s.t. $h(x,\theta)\geq 0$, where x is a vector of choice variables, and θ is a m-by-1 vector of parameters. Suppose the solution to this constrained maximization problem exists and unique. Denote the unique maximizing argument as $x(\theta)$ and the value function as $V(\theta)=f(x(\theta),\theta)$. If $f(x,\theta)$, $h(x,\theta)$, and $x(\theta)$ are continuously differentiable in θ, then it is easy to describe $\frac{\partial V(\theta)}{\partial\theta_j}$. Let $L(x,\theta,\lambda)$ be the problem's associated Lagrangian function, and let $x(\theta),\lambda(\theta)$ solve the Kuhn-Tucker conditions. The Envelope theorem states that for every θ,

$$\frac{\partial V(\theta)}{\partial\theta_j}=\left.\frac{\partial L(x,\theta,\lambda)}{\partial\theta_j}\right|_{(x,\lambda)=(x(\theta),\lambda(\theta))}\qquad j=1,\dots,m$$

Although $\frac{\partial V(\theta)}{\partial\theta_j}=\sum_{i=1}^{L}\frac{\partial f(x(\theta),\theta)}{\partial x_i}\frac{\partial x_i(\theta)}{\partial\theta_j}+\frac{\partial f(x(\theta),\theta)}{\partial\theta_j}$ looks complicated, the Envelope theorem says that you can simply take the derivative of the Lagrangian with respect to θ_j and plug the solution $x(\theta),\lambda(\theta)$ to it.

- -

1.6 Some Important Identity

Theorem 3. *Suppose that the utility function is continuous and strictly increasing. Then, for all $p\gg 0$, m, and u, (1) $e(p,v(p,m))=m$, and (2) $v(p,e(p,u))=u$.*

Proof(sketch): (1) Note that, by definition, $e(p,v(p,m))\leq m$ since m is large enough to achieve $v(p,m)$. Suppose for the sake of contradiction that $e(p,v(p,m))<m$. Since $e(p,v(p,m))$ is the expenditure to achieve $v(p,m)$, $m-e(p,v(p,m))>0$ is the extra money to buy some more consumption goods even after achieving $v(p,m)$. By strict monotonicity, it will result in a higher utility, which is a contradiction to the claim that $v(p,m)$ is the maximized value that can be attained with m.

Prove (2) by yourself. (The same logic applies. Note that $v(p,e(p,u)) \geq u$. Suppose $v(p,e(p,u)) > u$, then it means $e(p,u)$ is not actually minimized expenditure.)

This identity condition leads to the duality between Marshallian and Hicksian demand functions.

Theorem 4. *Suppose that the utility function is continuous and strictly increasing. Then, for all $p \gg 0$, m, and u, (1) $x(p,m) = h(p,v(p,m))$, and (2) $h(p,u) = x(p,e(p,u))$.*

Proof: (1) Note that $u(x(p,m)) = v(p,m)$ by definition, and $e(p,v(p,m)) = m = p \cdot x(p,m)$ by Theorem 3. Thus,

$$x(p,m) = \arg\min_{x \in \mathbb{R}^L_+} p \cdot x \text{ s.t. } u(x) \geq v(p,m),$$

which implies $x(p,m) = h(p,v(p,m))$. (2) Note that $e(p,u) = p \cdot h(p,u)$ by definition, and $v(p,e(p,u)) = u = u(h(p,u))$ by Theorem 3. Thus,

$$h(p,u) = \arg\max_{x \in \mathbb{R}^L_+} u(x) \text{ s.t. } p \cdot x \leq e(p,u),$$

which implies $h(p,u) = x(p,e(p,u))$. □

Using these identity conditions, we can derive one application.

Proposition 4 (Roy's Identity)**.** *If $x(p,m)$ is the Marshallian demand function, then*

$$x_l(p,m) = -\frac{\partial v(p,m)/\partial p_l}{\partial v(p,m)/\partial m},$$

provided that the RHS is well defined and $p_l > 0$ and $m > 0$.

Proof: Recall the identity, $u = v(p,e(p,u))$. Differentiate this with p_l and evaluate at $u = v(p,m)$ to get

$$\begin{aligned}
0 &= \frac{\partial v(p,m)}{\partial p_l} + \frac{\partial v(p,m)}{\partial m}\frac{\partial e(p,v(p,m))}{\partial p_l} \\
&= \frac{\partial v(p,m)}{\partial p_l} + \frac{\partial v(p,m)}{\partial m} h_l(p,v(p,m)) && \text{(by Shephard's lemma)} \\
&= \frac{\partial v(p,m)}{\partial p_l} + \frac{\partial v(p,m)}{\partial m} x_l(p,m) && \text{(by duality)}
\end{aligned}$$

Just like you can prove the Shephard's lemma using the Envelope theorem, you can prove the Roy's identity using the Envelope theorem as well. Try at home. □

1.7 Properties of Demand Functions

By the maximum theorem, we know $h(p,u)$ is continuous. Let's assume $h(p,u)$ is differentiable from now on.

Proposition 5 (Properties of Hicksian demand)**.** *Consider the following substitution matrix:* $D_p h(p,u) := \left(\frac{\partial h_l}{\partial p_k}\right)_{l,k} = \begin{pmatrix} \frac{\partial h_1}{\partial p_1} & \frac{\partial h_1}{\partial p_2} & \cdots & \frac{\partial h_l}{\partial p_L} \\ \frac{\partial h_2}{\partial p_1} & \frac{\partial h_2}{\partial p_2} & \cdots & \frac{\partial h_2}{\partial p_L} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_L}{\partial p_1} & \frac{\partial h_L}{\partial p_2} & \cdots & \frac{\partial h_L}{\partial p_L} \end{pmatrix}$.

1. $D_p(p,u) = D_p^2 e(p,u)$. *(That is, for all* l,k, $\frac{\partial h_l(p,u)}{\partial p_k} = \frac{\partial^2 e(p,u)}{\partial p_l \partial p_k}$.*)*
2. $D_p h(p,u)$ *is negative semidefinite.*
3. $D_p h(p,u)$ *is symmetric.*
4. $D_p h(p,u)p = 0$. *(That is,* $\sum_{k=1}^{L} \frac{\partial h_l(p,u)}{\partial p_k} p_k = 0$ *for all* $l = 1,\dots,L$.*)*

Proof(sketch): (1) directly follows from $h_l(p,u) = \frac{\partial e(p,u)}{\partial p_l}$. (2) is because $e(p,u)$ is concave in p and twice differentiable in p, and (3) follows from the symmetry of second derivatives. For (4), note $h(tp,u) = h(p,u)$. Differentiating both sides with t yields the result. □

From Proposition 5, we can derive some nice features of Hicksian demand.

Corollary 1. *For each good* l*, it holds that*

1. $\frac{\partial h_l}{\partial p_l} \le 0$*: (compensated) own price effect is nonpositive.*
2. $\frac{\partial h_l}{\partial p_k} = \frac{\partial h_k}{\partial p_l}$ *for all* $k \neq l$*: cross price effects are symmetric.*
3. $\frac{\partial h_l}{\partial p_k} \ge 0$ *for some* $k \neq l$*: each good has at least one substitute.*

Proof: (1) Let e_l is a vector whose lth entity is 1 and 0 otherwise. Since $D_p h(p,u)$ is negative semidefinite, $e_l' D_p h(p,u) e_l = \frac{\partial h_l}{\partial p_l} \le 0$. (2) directly follows from the symmetry of $D_p h(p,u)$. For (3), suppose $\frac{\partial h_l}{\partial p_k} < 0$ for all $k \neq l$. Then $D_p h(p,u)p = 0$ is violated. □

Theorem 5 (Slutsky Equation)**.**

$$\frac{\partial x_l(p,m)}{\partial p_k} = \frac{\partial h_l(p,v(p,m)}{\partial p_k} - \frac{\partial x_l(p,m)}{\partial m} x_k(p,m)$$

Proof: Let $x^* := x(p,m)$ and $u^* := u(x^*)$. By the identity shown in Theorem 4,

$$h_l(p,u^*) = x_l(p,e(p,u^*)).$$

Differentiating this with p_k yields

$$\frac{\partial h_l(p,u^*)}{\partial p_k} = \frac{\partial x_l(p,e(p,u^*))}{\partial p_k} + \frac{\partial x_l(p,e(p,u^*))}{\partial m}\frac{\partial e(p,u^*)}{\partial p_k}.$$

By Shephard's lemma and the identity, $\frac{\partial e(p,u^*)}{\partial p_k} = h_k(p,u^*) = h_k(p,v(p,m)) = x_k(p,m)$. Plugging this into the equation above yields

$$\frac{\partial h_l(p,u^*)}{\partial p_k} = \frac{\partial x_l(p,e(p,u^*))}{\partial p_k} + \frac{\partial x_l(p,e(p,u^*))}{\partial m}x_k(p,m),$$

which leads the result by rearrangement. □

The Slutsky equation nicely demonstrates that a price change involves two effects:

- $\frac{\partial h_l(p,v(p,m))}{\partial p_k}$: *substitution effect* which measures the change in demand due to change in relative prices.

- $\frac{\partial x_l(p,m)}{\partial m}x_k(p,m)$: *income effect* which measures the change in demand due to the change in the 'purchasing' power.

We examined the properties of substitution matrix of Hicksian demand in Proposition 5. Let's think about the Marshallian version of it.

Corollary 2 (Symmetric and Negative Semidefinite Slutsky Matrix)**.** *The Slutsky matrix is defined as*

$$S(p,m) := \left(\frac{\partial x_l(p,m)}{\partial p_k} + \frac{\partial x_l(p,m)}{\partial m}x_k(p,m)\right)_{l,k}.$$

$S(p,m)$ is symmetric and negative semidefinite.

Proof: Letting $u^* = v(p,m)$,

$$\left(\frac{\partial x_l(p,m)}{\partial p_k} + \frac{\partial x_l(p,m)}{\partial m}x_k(p,m)\right)_{l,k} = \left(\frac{\partial h_l(p,u^*)}{\partial p_k}\right)_{l,k},$$

which is symmetric and negative semidefinite by Proposition 5. □

1.8 Money Metric (Indirect) Utility Functions

Recall that the utility functions are ordinal, so a statement like "the utility is 10" is not informative at all for welfare comparison. It would be desirable if the atilt function can be translated to a certain comparable form. Having that motivating in mind, we will construct other functions from the expenditure function, which will be useful for the welfare analysis.

The **money metric utility function** is defined by

$$m(p,x) := e(p,u(x))$$

- The money metric measures how much money the consumer would need to achieve the same utility level as he could achieve with the bundle x.
- Since $m(p,x)$ is the expenditure function, it has all properties of $e(p,x)$: $m(p,x)$ is monotonic, homogeneous of degree one, and concave in p.
- For a fixed p, $m(p,x)$ is monotonic transformation of the utility function. (Check by yourself: If $u(x') > u(x)$, then $m(p,x') > m(p,x)$.) Therefore, $m(p,x)$ by itself is a utility function.

The **money metric indirect utility function** is defined by

$$\mu(p;q,m) := e(p,v(q,m))$$

- This measures the amount of money the consumer would need at price p to achieve the same utility as he could achieve under the price q and income m.
- For a fixed p, $\mu(p;q,m)$ is a monotone transform of the indirect utility function.
- Both money metric and money metric indirect utility functions are invariant to the monotonic transformation of the underlying utility function.

Exercise: Suppose $f(u(x))$ is a monotonic transformation of $u(x)$. Verify that $e(p,u(x)) = e(p,f(u(x))$.

Later we will use this money metric indirect utility function for welfare evaluation of price changes.

2 Topics in Consumer Theory

2.1 Integrability

If the demand function $x(p,m) \in \mathbb{C}^1$ is derived from "well-behaved" rational preference, then it satisfies

(P1) Homogeneity: $x(tp,tm) = x(p,m)$.

(P2) Budget balancedness (or Walras' Law): $p \cdot x(p,m) = m$.

(P3) Symmetry: $S(p,m)$ is symmetric.

(P4) Negative Semidefinite: $S(p,m)$ is negative semidefinite.

The integrability problem is about the possibility of recovering the preference. (Why does it matter? The policy suggestions must be based on better understanding about affected citizens' preferences. Weren't you criticizing that microeconomic theory is bad because it all starts from "let's assume"?) If we observe a demand function $x(p,m)$ that satisfies properties (P1)–(P4) above, then can we find preference that generated $x(p,m)$? The answer is 'yes' according to Antonelli (1886) and Hurwicz and Uzawa (1971). We loosely follow Hurwicz and Uzawa (1971) with the two steps.

- Recover $e(p,u)$ from $x(p,m)$.
- Recover preferences from $e(p,u)$.

Let's work on the example first.

- -

Example: Suppose $L = 3$, and the consumer's demand function is

$$x_l(p,m) = \alpha_l \frac{m}{p_l}, \quad l = 1,2,3,$$

where $\alpha_l \in [0,1]$ such that $\alpha_1+\alpha_2+\alpha_3 = 1$. One can easily check that this demand function satisfies the integrability conditions (P1)–(P4). By Shepard's lemma,

$$\frac{\partial e(p,u)}{\partial p_l} = \alpha_l \frac{m}{p_l} = \alpha_l \frac{e(p,u)}{p_l}, \quad l = 1,2,3.$$

This is a differential equation with respect to $e(p,u)$, which can be rewritten as

$$\frac{\partial \ln e(p,u)}{\partial p_l} = \frac{\alpha_l}{p_l},$$

which implies that for some monotone increasing function $c(\cdot)$,

$$\ln e(p,u) = \alpha_1 \ln p_1 + \alpha_2 \ln p_2 + \alpha_3 \ln p_3 + c(u)$$

Thus, we have

$$e(p,u) = e^{c(u)} p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3}$$

A monotone transformation on $e^{c(u)}$ does not affect the partial differential equations. By replacing $e^{c(u)}$ with u, we obtain

$$e(p,u) = u p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3}.$$

Now we have recovered the expenditure function from the observed demand function.

The next step is to obtain the utility function from $e(p,u)$. Note from the duality that $e(p, v(p,m)) = m$. Let's normalize $m = 1$, and wrote $v(p,1)$ as $v(p)$ for simplicity.

$$e(p, v(p)) = v(p) p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3} = 1.$$

Then,

$$u(x) = \min_{p | p \cdot x = 1} v(p).$$

Why? Since $p \cdot x = 1$, a consumer can always attain $u(x)$ with a given price p because x is affordable. This means $v(p,1)$, the maximized utility value of the utility subject to $p \cdot x = 1$, must be greater than or equal to $u(x)$. Thus, the minimum value of $v(p)$ will be $u(x)$.

From $v(p) p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3} = 1$, $v(p) = p_1^{-\alpha_1} p_2^{-\alpha_2} p_3^{-\alpha_3}$. Set up the Lagrangian for the minimization problem of $v(p)$ subject to $p_1 x_1 + p_2 x_2 + p_3 x_3 = 1$. You can verify that $u(x) = \kappa x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3}$, where $\kappa = \frac{1}{\alpha_1^{\alpha_1} \alpha_2^{\alpha_2} \alpha_3^{\alpha_3}}$. Try at home.

The example above can be neatly solved, but in general, it is easier said than done. First, we need to How to recover $e(p,u)$ from $x(p,m)$. For a fixed u^0 and (p^0, m^0), by Shephard's Lemma and Duality between Hicksian and Marshallian demand functions, we can consider the following system of partial differential equations:

$$\frac{\partial e(p,u^0)}{\partial p_1} = h_1(p,u^0) = x_1(p,e(p,u^0))$$
$$\frac{\partial e(p,u^0)}{\partial p_2} = h_2(p,u^0) = x_2(p,e(p,u^0))$$
$$\vdots$$
$$\frac{\partial e(p,u^0)}{\partial p_L} = h_L(p,u^0) = x_L(p,e(p,u^0))$$

with initial condition, $e(p^0,u^0) = m^0$. It is known (by Frobenius' theorem) that this system of PDEs has a solution if and only if the $L \times L$ derivative matrix of the functions on the the right hand side is symmetric, that is,

$$\left(\frac{\partial x_l(p,e(p,u^0)}{\partial p_k} + \frac{\partial x_l(p,e(p,u^0))}{\partial m}\frac{\partial e(p,u^0)}{\partial p_k}\right)_{l,k} := S(p,m)$$

is symmetric. Do you see the famous form in the big parentheses? It is the Slutsky equation. Thus, the necessary and sufficient condition for the recovery of an underlying expenditure function is the symmetry and negative semidefiniteness of the Slutsky matrix. (By (P3), it holds.)

The following proposition tells us how to recover preference from $e(p,u)$.

Proposition 6. *Given $e(p,u)$, define $V_u := \{x \in \mathbb{R}^L_+ | p \cdot x \geq e(p,u), \forall p \gg 0\}$. Then,*

$$e(p,u) = \min_{x \in \mathbb{R}^L_+} p \cdot x \quad s.t. \quad x \in V_u.$$

That is, V_u is an upper contour set of utility $u \in \mathbb{R}$.

Proof: From the definition of V_u, $e(p,u) \leq \min\{p \cdot x | x \in V_u\}$. We need to prove that $e(p,u) \geq \min\{p \cdot x | x \in V_u\}$. For any p and p', the concavity of $e(p,u)$ in p implies that

$$e(p',u) \leq e(p,u) + \nabla_p e(p,u) \cdot (p' - p),$$

where $\nabla_p e(p,u) = (\frac{\partial e}{\partial p_1} \cdots \frac{\partial e}{\partial p_L}$. Since $e(p,u)$ is homogeneous of degree one in p, $e(p,u) = \nabla_p e(p,u) \cdot p$ by Euler's homogeneous function theorem. Thus, $e(p',u) \leq \nabla_p e(p,u) \cdot p'$, which means $\nabla_p e(p,u) \in V_u$. It follows that $\min\{p \cdot x | x \in V_u\} \leq p \cdot \nabla_p e(p,u) = e(p,u)$. □

- -

Digression: Euler's homogeneous function theorem. If f is a continuously differentiable function of n variables that is homogeneous of degree k, then

$$kf(x_1,x_2,\ldots,x_n)=\sum_{i=1}^{n}x_i\frac{\partial f}{\partial x_i}(x_1,x_2,\ldots,x_n).$$

Its proof is simple. For simplicity, we set $x=(x_1,x_2,\ldots,x_n)$. Differentiate both sides of $f(ax)=a^k f(x)$ with respect to a. Take the limit of the result when a approaches 1.

- -

One minor remark: not every system of PDEs has a solution. It means that we sometimes cannot recover the expenditure function from the demand function. Before attempting to recover, try to check whether (P1)–(P4), especially the symmetry, hold.

2.2 Revealed Preference

While integrability was about recovering the preferences from the demand functions, revealed preference theory is about recovering the preferences from the observed choices. If you buys one consumption bundle instead of another affordable bundle, then the first bundle is considered to be revealed preferred to the second. The observation conveys important information about your tastes. In this section, we make assumptions about the observed choices, instead of assumptions about the (unobserved) preferences.

Let (p,x) denote a price-quantity data such that the bundle x is chosen under the price p. For example, x^0 is a choice bundle under the price p^0, and x^t is a choice bundle under the price p^t.

Definition 2 (Weak Axiom of Revealed Preference (WARP)). *The observations satisfy WARP if $p^0\cdot x^1\le p^0\cdot x^0$ and $x^0\ne x^1$, then we must have $p^1\cdot x^0>p^1\cdot x^1$.*

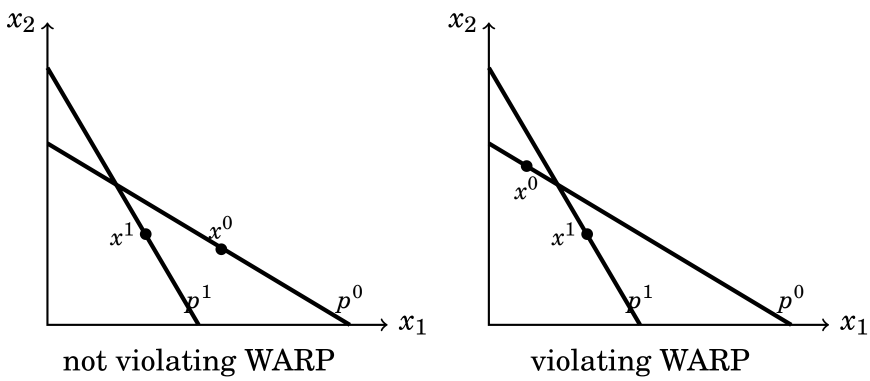

Figure 1: WARP

That is, WARP holds if whenever x^0 is revealed preferred to x^1 (because x^0 is chosen when x^1 is affordable), x^1 is never revealed preferred to x^0. Figure 1 illustrates the choice bundles that are not violating WARP (left), and violating WARP (right).

Whether the axiom is satisfied is closely related to whether there exists a utility function that would yield the observed choices as the outcome of the utility-maximizing process, or rationalizes, the data. (It is worth pointing out that if $x(p,m)$ is a Marshallian demand function with the prices p and income m, then $x(p,m)$ must satisfy WARP. Try to see this at home.)

Let $x(p,m)$ denote the choice made by a consumer facing prices p and income m. We refer to $x(p,m)$ as a choice function. (Although it looks like a demand function, we did not mention anything about preferences.) If the observed choices satisfy WARP and budget balancedness ($p \cdot x(p,m) = m$ for $p \gg 0$), then the choice function inherits many properties of the Marshallian demand function.

Proposition 7. *If $x(p,m)$ satisfies WARP and budget balancedness, then it is homogeneous of degree zero in (p,m), and the associated Slutsky matrix is negative semidefinite.*

Proof: For homogeneity of degree zero, we want to show $x(p,m) = x(tp,tm)$. Suppose x is chosen when prices are p and income is m, and x' is chosen when prices are $p' = tp$ and income is $m' = tm$. Note x is affordable under p' and m' and x' is affordable under p and m. Due to budget balancedness, $p' \cdot x' = tm = tp \cdot x$. Substituting tp with p', we get $p' \cdot x' = p' \cdot x$. Substituting p' with tp, we get $p \cdot x' = p \cdot x$. If $x \neq x'$, then it must violate WARP. Thus, $x(p,m) = x(tp,tm)$.

Now we prove the negative semidefinteness. Fix $p \gg 0$, $m > 0$, and let $x := x(p,m)$. For any other price vector p' and $x' := x(p', p' \cdot x)$, WARP implies that

$$p \cdot x \leq p \cdot x' \tag{3}$$

with a strict inequality if $x \neq x'$. (That is, if $x \neq x'$, then WARP implies that x' should not be affordable.) By budget balancedness,

$$p' \cdot x = p' \cdot x' \tag{4}$$

Subtracting 3 from 4, we have for all prices p',

$$(p' - p) \cdot x \geq (p' - p) \cdot x' \tag{5}$$

Since 5 holds for all prices p', let $p' = p + tz$, where $t > 0$ and arbitrary $z \in \mathbb{R}^L$. 5 becomes

$$tz \cdot x \geq tz \cdot x(p + tz, (p + tz) \cdot x) \tag{6}$$

For any arbitrary z, we can choose $\bar{t} > 0$ such that $p + tz \gg 0$ for all $t \in [0, \bar{t}]$. Since the inequality holds when $t=0$, the function $f(t):[0,\bar{t})\to\mathbb{R}$ defined by

$$f(t)\equiv x(p+tz,(p+tz)\cdot x),$$

that is, the right-hand side of equation 6, is maximized on $[0,\bar{t})$ at $t=0$. We thus must have $f'(0)\leq 0$. Taking the derivative of $f(t)$ and evaluating at $t=0$ yields

$$f'(0)=\sum_l\sum_k z_l\left[\frac{\partial x_l(p,m)}{\partial p_k}+x_k(p,m)\frac{\partial x_l(p,m)}{\partial m}\right]z_k=z\cdot S(p,m)z\leq 0.$$

Thus, $S(p,m)$ is negative semidefinite. □

Regarding the integrability, we have shown that WARP+budget balancedness satisfies (P1) homogeneity, (P2) budget balancedness, and (P4) Negative semidefiniteness. It means, if we show (P3) symmetry, then the underlying preference can be recovered from the observed choice function.

If there are only two goods, the symmetry is known to be satisfied. (You can check at home that homogeneity and budget balancedness imply $\frac{\partial^2 e(p,u)}{\partial p_1\partial p_2}=\frac{\partial^2 e(p,u)}{\partial p_2\partial p_1}$. Intuitively, there are no transitive cycles with two goods.) However, if there are three or more goods, the symmetry may not hold.

Definition 3 (Strong Axiom of Revealed Preference (SARP))**.** *The observations satisfy SARP if, for every sequence of distinct price-choice bundles* (p^0,x^0), (p^1,x^1), $\ldots$, (p^k,x^k), *where* x^{m-1} *is revealed preferred to* x^m, $m=1,2,\ldots,k$, *it is not the case that* x^k *is revealed preferred to* x^0. *That is,* $p^k\cdot x^0>p^k\cdot x^n$.

If $x(p,m)$ satisfies SARP, it is known that it must also have the symmetric Slutsky matrix. Thus, SARP is essentially equivalent to the existence of a utility function that rationalizes the data.

So far, we have focused on the situation where the choice *function* is observable; in effect, we have assumed that we had an infinitely large collection of price and quantity data. Because real-world data sets will never contain more than a finite number of sample points, we need to work on revealed preference when the number of observations is finite.

Definition 4 (Generalized Axiom of Revealed Preference (GARP))**.** *The observations satisfy GARP if, for every sequence of distinct price-choice bundles* (p^0,x^0), (p^1,x^1), $\ldots$, (p^k,x^k), *where* x^{m-1} *is revealed preferred to* x^m, $m=1,2,\ldots,k$, *it is not the case that* x^k *is revealed preferred to* x^0. *That is,* $p^k\cdot x^0\geq p^k\cdot x^n$.

To that end, Afriat (1967) introduced the GARP, a slightly weaker requirement than SARP, and proved an analogue of the integrability theorem. According to Afriat's theorem, a finite set of observed price and quantity data satisfy GARP if and only if there exists a continuous, increasing, and concave utility function that rationalises the data.

Remarks: Three textbooks have different approaches on the revealed preference theory. MWG introduces SARP to say that the choice function is essentially the demand function. JR (which I follow for this section) introduces WARP and provides a rationale for considering SARP. V focuses on GARP in detail. Not surprisingly, many observed data violate GARP, but this should not the end of the story. We can tell the smallest "adjustment" that make the individual choice data consistent with GARP, and call that adjustment as the quantitative deviation from rationality; see Choi et al. (2014) for example. Rationality of social preferences (that is, examining whether an individual's prosocial decision is rationally derived) can be evaluated with using GARP; see Andreoni and Miller (2002) for example.

2.3 Welfare Evaluation of Price Changes

Suppose that you are at Region 0 and get a job offer from Region 1. The proposed change is from (p^0, m^0) to (p^1, m^1). How to evaluate your welfare changes?

One way to measure the welfare change involved in moving from (p^0, m^o) to (p^1, m^1) is to calculate $v(p^1, m^1) - v(p^0, m^0)$. However, this measure is subject to the monotone transformation. (Check the discussion in subsection 1.8.) It may be used to merely check the sign of the welfare change, but not good for the cost-benefit analysis.

Instead, we can use the money metric indirect utility function.

$$\mu(q; p^1, m^1) - \mu(q; p^0, m^0) = e(q, u^1) - e(q, u^0) \quad \text{for some } q,$$

where $u^0 = v(p^0, m^0)$ and $u^1 = v(p^1, m^1)$.

- This measures the difference between the utilities $v(p^0, m^0)$ and $v(p^1, m^1)$ in monetary terms using q as the base price.

- *Equivalent Variation (EV)*: Setting $q = p^0$,

  $$EV := \mu(p^0; p^1, m^1) - \mu(p^0; p^0, m^0) = \mu(p^0; p^1, m^1) - m^0,$$

  which measures what income change at the current prices would be equivalent to the proposed change in terms of its impact on utility.

  In other words, $EV + m^0 = \mu(p^0; p^1, m^1) = e(p^0, v(p^1, m^1))$ means that you need to spend $e(p^0, v(p^1, m^1))$ in Region 1 to achieve welfare level that could have achieved in Region 1, $v(p^1, m^1)$. If EV is positive, then it means you need more than your current income, m^0.

- *Compensating Variation (CV)*: Setting $q = p^1$,

$$CV := \mu(p^1; p^1, m^1) - \mu(p^1; p^0, m^0) = m^1 - \mu(p^1; p^0, m^0),$$

  which measures what income change at the new prices would be necessary to compensate the consumer for the proposed change.

Suppose that $p_1^0 > p_1^1$ while all other prices and income levels are the same, $p_{-1}^0 = p_{-1}^1 = p_{-1}$ and $m^0 = m^1 = m$. Then, EV and CV can be alternatively expressed as

$$EV = e(p^0, u^1) - e(p^1, u^1) = \int_{p_1^1}^{p_1^0} h_1(p_1, p_{-1}, u^1) dp_1$$

$$CV = e(p^0, u^0) - e(p^1, u^0) = \int_{p_1^1}^{p_1^0} h_1(p_1, p_{-1}, u^0) dp_1$$

If consumption good 1 is a normal good, that is, $\frac{\partial x_1(p,m)}{\partial m} > 0$, then $EV > CV$.

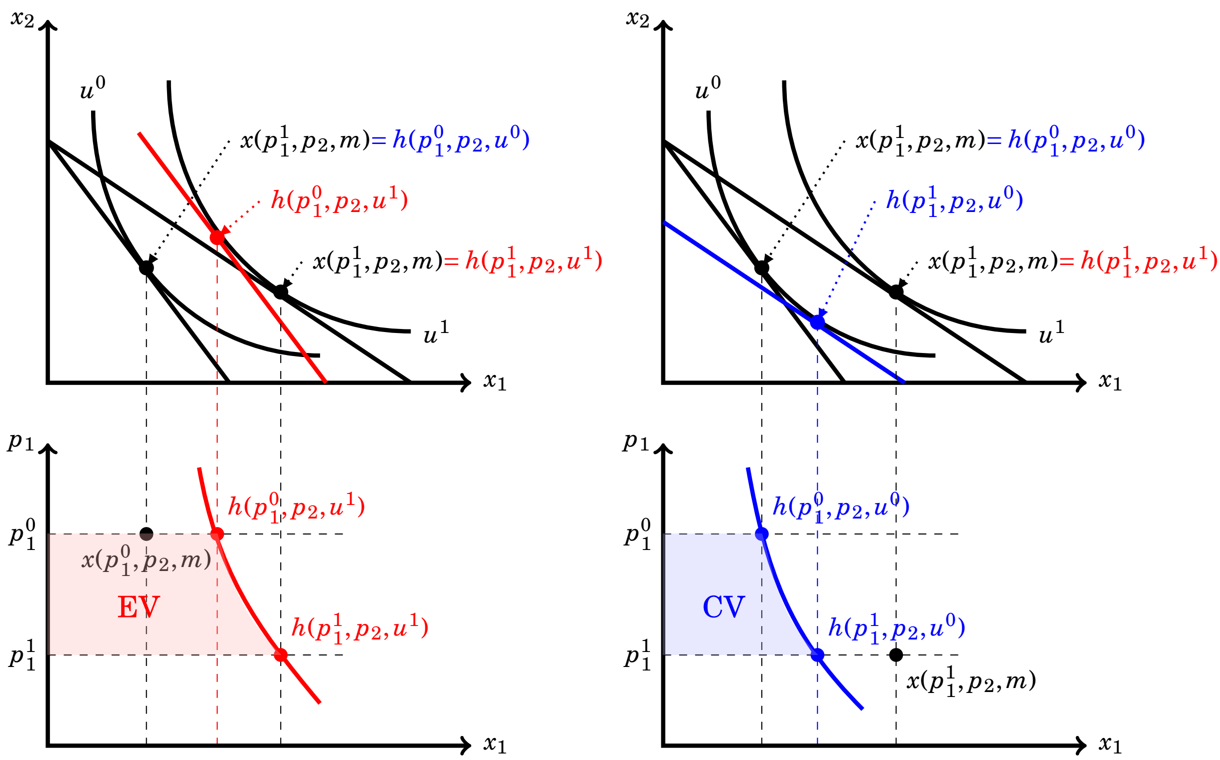

Figure 2: EV and CV

Price of good 1 decreases from p_1^0 to p_1^1, while p_2 and m remain constant.

(Do you see that the change in 'consumer surplus' lie in between CV and EV? Check the area connecting p_1^0–$x(p_1^0, p_2, m)$–$x(p_1^1, p_2, m)$–p_1^1. That would be an easier approximation when the price change is small.)

Remark: Figure 2 figure shows that if the income effect on the consumption of good 1 is positive (negative), EV is greater (smaller) than CV. It means that EV=CV when there is no income effect on the consumption of good 1. Later it will be useful to know that a quasilinear utility function with respect to the numeraire good (good 0) has zero income effect.

A **quasilinear utility function** has the following form

$$U(x_1, x_2, \dots, x_L) = x_1 + \phi(x_2, \dots, x_L),$$

where $\phi(\cdot)$ is increasing concave. You will see the demand for good 1 is determined from the budget constraint, after the demands for other goods are figured out.

Exercise: Assume $L = 2$. A utility function is given as follows: $U(x_1, x_2) = \ln(x_1) + x_2$. Good 2 is the numeraire, so the prices are normalized to $p_2 = 1$. The income is m. Derive $x_1(p_1; m)$, and show that $\frac{\partial x_1(p_1;m)}{\partial m} = 0$.

2.4 Aggregate Demand

Suppose that there are I consumers with Marshallian demand functions $x^i(p, m)$ for consumer $i = 1, \dots, I$. Given prices $p \in \mathbb{R}_+^L$ and wealth levels $(m^1, \dots, m^I)$, aggregate demand is written as

$$X(p, m^1, \dots, m^I) = \sum_{i=1}^{I} x^i(p, m^i).$$

Is the aggregate demand representative? That is, when would the aggregate demand be the same as the demand of a single representative consumer with aggregate wealth $m = \sum_{i=1}^{I} m^i$?

One such case is where the aggregate demand can be expressed as a function of prices and aggregate wealth:

$$X(p, m^1, \dots, m^I) = x(p, m).$$

- One requirement here is that for all wealth levels $(m^1, \dots, m^I)$ and its differential change $(dm^1, \dots, dm^I)$ satisfying $\sum_{i=1}^{I} dm^i = 0$,

$$\sum_{i=1}^{I} \frac{\partial x_l^i(p, m^i)}{\partial m^i} dm^i = 0$$

  for every l. In words, reallocation of wealth (taking out some consumers' money and distributing out to others) should not change the aggregate demand.

- This will hold if

$$\frac{\partial x_l^i(p,m^i)}{\partial m^i} = \frac{\partial x_l^j(p,m^j)}{\partial m^j}$$

for all i, j. That is, the wealth effect must be the same across consumers.

Proposition 8 (Gorman Form Indirect Utility Function)**.** *If each consumer's indirect utility function has the form of $v^i(p,m^i) = a^i(p)+b(p)m^i$, a representative consumer exists, that is, $X(p,m^1,\cdot,m^I) = x(p,m)$, where $m = \sum_{i=1}^I m^i$.*

Proof: By Roy's identity, consumer i's demand for good l takes the form

$$x_l^i(p,m^i) = -\frac{\frac{\partial v^i(p,m^i)}{\partial p_l}}{\frac{\partial v^i(p,m^i)}{m^i}} = -\frac{\frac{\partial a^i(p)}{\partial p_l}}{b(p)} - \frac{\frac{\partial b(p)}{\partial p_l}}{b(p)} m^i.$$

Let's denote $\frac{\frac{\partial a^i(p)}{\partial p_l}}{b(p)} = a_l^i(p)$ and $\frac{\frac{\partial b(p)}{\partial p_l}}{b(p)} = b_l(p)$. The aggregate demand for good l takes the form

$$x_l(p,m) = -\left[\sum_{i=1}^I a_l^i(p) + b_l(p)m\right].$$

This can be generated by a representative consumer whose indirect utility function is given by

$$v(p,m) = \sum_{i=1}^I a^i(p) + b(p)m.$$

To verify this, apply Roy's identity to $V(p,m)$. □

Remarks

- We only state the sufficiency, but it is actually necessary and sufficient condition. Check Deaton and Muellbauer (1980) for the proof of necessity (If a representative consumer exists, then every consumer has the Gorman form indirect utility function.)

- Two utility functions whose indirect utility function is of German form:

  1. Homothetic utility function: $u^i(x) = g(h(x))$, where g is a strictly increasing function (that is, a monotone transformation) and h is homogeneous of degree 1. $\to v^i(p,m^i) = b(p)m^i$. (That is, $a^i(p) = 0$.)
  2. Quasi-linear utility function: $u^i(x) = x_i + w^i(x_2,\ldots,x_L) \to v^i(p,m^i) = a^i(p) + m^i$. (That is, $b(p) = 1$.)

3 Choice under Uncertainty

Remarks: Although this section is typically called choice under "uncertainty," we deal with a specific type of uncertainty called "risk." Risk applies to situations where we do not know the outcome of a given situation, but can accurately measure the odds. Imagine a situation where your roll a dice: you don't know the outcome of the dice roll, but you know that each outcome is equally likely. Uncertainty is a broader concept; it also applies to situations where we cannot know all the information we need to set accurate probabilities. For those who are more interested in, search for terms like "Knightian uncertainty" (named after Frank Knight who first distinguished risk from uncertainty) and "ambiguity."

3.1 Expected Utility Theory

Imagine a decision maker who faces a choice among a number of risky alternatives. Each risky alternative results in one of outcomes 1 to N, following some probability distribution over those outcomes. We call such a risky alternative a *lottery*.

Definition 5. *A* ***simple lottery*** *L is a list $L = (p_1, \dots, p_N)$ with $p_n > 0$ for all $n \in \{1, \dots, N\}$ and $\sum_{n=1}^{N} p_n = 1$ where p_n is the probability of outcome n occurring.*

Example: Suppose that the set of outcomes is $\{\$0, \$10, \$100\}$, then $L = (0.2, 0.7, 0.1)$ means that the decision maker can obtain \$0, \$10, and \$100 with probabilities 0.2, 0.7, and 0.1, respectively.

Let $\mathscr{L}$ denote the set of all simple lotteries. In case of n outcomes, $\mathscr{L}$ can be represented by the $n-1$-dimensional unit simplex. For example, if there are three outcomes, then the following two dimensional unit simplex represents all simple lotteries. The three-dimensional figure (right) in Figure 3 shows $(p_1, p_2, p_3) \in [0,1]^3$ such that $p_1 + p_2 + p_3 = 1$. It is simpler to draw in two dimensions (left).

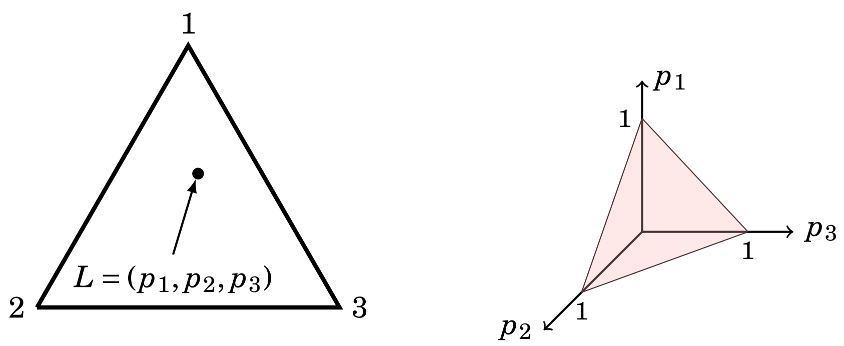

Figure 3: Representation of the set of all lotteries
Any point in the triangle represents a simple lottery.

Definition 6. *Given K simple lotteries $L_k = (p_1^k, \ldots, p_N^k)$, $k = 1, \ldots, K$, a* ***compound lottery*** *is a list $(L_1, \ldots, L_K; \alpha_1, \ldots, \alpha_K)$, where α_k is the probability of lottery k occurring, with $\sum_k \alpha_k = 1$.*

That is, a compound lottery is a lottery whose outcomes are lotteries.

Example: Suppose that you flip a coin to choose one of two similarly attractive jobs. If the coin lands head, you get job 1 whose payoff is \$40K with probability 0.3 and \$100K with probability 0.7. If the coin lands tail, you get job 2 whose payoff is \$65 for sure. A compound lottery $(L_1, L_2; 1/2, 1/2)$ where $L_1 = (0.3, 0, 0.7)$ and $L_2 = (0, 1, 0)$ when the outcomes are $(40K, 65K, 100K)$.

Let $\succeq$ denote the decision maker's preference relation on $\mathscr{L}$. We will axiomatically represent the preference relation on the lotteries using the *expected utility*. Here comes the overview first.

- We assume completeness, transitivity, reduction of compound lotteries, continuity, independence, and the existence of best and worst lotteries.
- We introduce the expected utility form, especially von Neumann-Morgenstern (v.N-M) expected utility function.
- The above assumptions are necessary and sufficient conditions for $\succeq$ on $\mathscr{L}$ to be represented by a v.N-M utility function.
- Each assumption may not hold in reality. We will discuss some of them.

In addition to completeness and transitivity, we assume:

- **Reduction of compound lotteries**: For any $L, L' \in \mathscr{L}$,

$$(L, L'; \alpha, 1-\alpha) \sim \alpha L + (1-\alpha)L' = (\alpha p_1 + (1-\alpha)p_1', \ldots, \alpha p_N + (1-\alpha)p_N').$$

For example, the compound lottery considered in the previous example is equally preferred to a simple lottery $L = (0.15, 0.5, 0.35)$.

- **Continuity**: For any three lotteries $L, L', L'' \in \mathscr{L}$ satisfying $L \succ L' \succ L''$, there exists $\alpha \in (0,1)$ such that
$$\alpha L + (1-\alpha)L'' \sim L'.$$

It means that there are no "jumps" in preferences.

- **Independence**: For any three lotteries $L, L', L'' \in \mathscr{L}$ and $\alpha \in (0,1)$, we have

$$L \succeq L' \text{ if and only if } \alpha L + (1-\alpha)L'' \succeq \alpha L' + (1-\alpha)L''.$$

In words, mixing each of two lotteries with the third lottery, then the preference order of two mixtures does not depend on the third one. This assumption is conceptually related to the "independence of irrelevant alternatives."

The independence axiom implies that the indifference curves (or in general, $N-1$-dimensional indifference hyperplanes) are straight and parallel on the unit simplex representing $\mathscr{L}$. Figure 4 illustrates two situations where indifference curves violate the independence axiom. If the indifference curve is not a straight line, it means $L \sim L'$ but $L \nsim \alpha L + (1-\alpha)L' \nsim L'$ for some $\alpha \in (0,1)$. If the indifference curves are not parallel, then mixing two indifferent lotteries (L and L') with another lottery (L'') can violate the independence axiom.

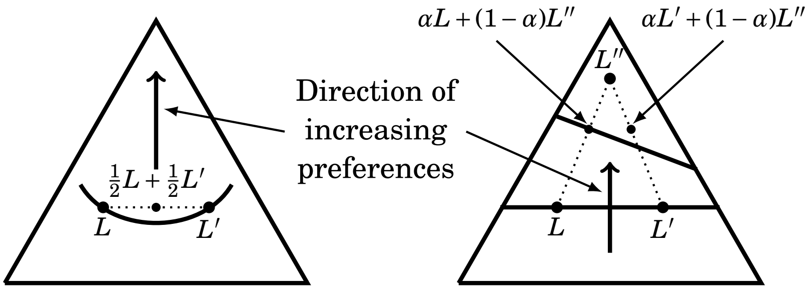

Figure 4: Violations of Independence Axiom

- **Best and worst lotteries**: There exist two lotteries $\overline{L}$ and $\underline{L}$ such that for all $L \in \mathscr{L}$, $\overline{L} \succeq L \succeq \underline{L}$.

Definition 7. *The utility function $U : \mathscr{L} \to \mathbb{R}$ is called a von Neumann-Morgenstem (v. N-M) expected utility function if for every $L \in \mathscr{L}$,*

$$U(L) = p_1 u_1 + p_2 u_2 + \cdots + p_N u_N,$$

where u_n is the utility assigned to outcome n. u_k is often called Bernoulli utility function in case each outcome is a monetary value.

Theorem 6 (Expected Utility Theorem)**.** *The above assumptions are necessary and sufficient for the preference $\succeq$ on $\mathscr{L}$ to admit a utility representation of the expected utility form.*

Proof: The necessity part is straightforward. (Check by yourself if the v.N-M expected utility function satisfies all the axioms above.) We prove the sufficiency part.

First, by the independence and the existence of the best and worst lotteries, for any $\alpha, \beta \in [0,1]$,

$$\beta\overline{L} + (1-\beta)\underline{L} \succ \alpha\overline{L} + (1-\alpha)\underline{L} \quad \text{if and only if} \quad \beta > \alpha. \tag{7}$$

By the continuity, for any $L \in \mathscr{L}$, there exists an $\alpha_L \in [0,1]$ such that $\alpha\overline{L}+(1-\alpha)\underline{L} \sim L$. Also, such α_L is unique due to (7).

Consider a function $U : \mathscr{L} \to \mathbb{R}$ that assigns α_L to each $L \in \mathscr{L}$. Then, such $U(\cdot)$ represents $\succeq$ since for any two lotteries $L, L' \in \mathscr{L}$, we have

$$L \succ L' \text{ if and only if } \alpha_L\overline{L} + (1-\alpha_L)\underline{L} \succeq \alpha_{L'}\overline{L} + (1-\alpha_{L'})\underline{L},$$

which implies that $L \succeq L'$ if and only if $\alpha_L \geq \alpha_{L'}$.

Next, we show that $U(\cdot)$ is linear, that is, for any $L, L' \in \mathscr{L}$ and $\beta \in [0,1]$,

$$U(\beta L + (1-\beta)L') = \beta U(L) + (1-\beta)U(L'). \tag{8}$$

Note that $L \sim U(L)\overline{L} + (1-U(L))\underline{L}$ and $L' \sim U(L')\overline{L} + (1-U(L'))\underline{L}$. Applying the independence axiom (twice), we obtain

$$\begin{aligned} \beta L + (1-\beta)L' &\sim \beta[U(L)\overline{L} + (1-U(L))\underline{L}] + (1-\beta)L' \\ &\sim \beta[U(L)\overline{L} + (1-U(L))\underline{L}] + (1-\beta)[U(L')\overline{L} + (1-U(L'))\underline{L}]. \end{aligned}$$

By the reduction of compound lottery, $\beta[U(L)\overline{L}+(1-U(L))\underline{L}]+(1-\beta)[U(L')\overline{L}+(1-U(L'))\underline{L}]$ is equivalent to

$$[\beta U(L) + (1-\beta)U(L')]\overline{L} + [1-\beta U(L)) - (1-\beta)U(L')]\underline{L}.$$

Thus, (8) follows from the definition of $U(\cdot)$. We are almost done. Consider N simple lotteries L_n, $n = 1, \ldots, N$, where L_n has $p_n = 1$. (That is, outcome n occurs for sure.) Denote $U(L_n) = u_n$ for all n. Then any lottery $L = (p_1, \ldots, p_N)$ can be described by a linear combination of such simple lotteries, $p_1L_1 + \cdots + p_NL_N$. Therefore, $U(L) = U(p_1L_1 + \cdots + p_NL_N) = p_1U(L_1) + \cdots + p_NU(L_N) = p_1u_1 + \cdots + p_Nu_N$. □

The following example, called Allais Paradox, casts doubt on the independence axiom of the expected utility theory.

Example (Allais Paradox): There are three monetary outcomes, \$2.5 million, \$0.5 million, and \$0. A decision maker is asked to make two decisions. The first one is to choose between

$$L_1 = (0,1,0) \quad \text{and} \quad L_1' = (0.10, 0.89, 0.01).$$

The second one is to choose between

$$L_2 = (0, 0.11, 0.89) \quad \text{and} \quad L_2' = (0.10, 0, 0.90).$$

People often exhibit $L_1 \succ L_1'$ and $L_2' \succ L_2$. However, the expected utility theory cannot explain this seemingly natural decisions. If the expected utility theory holds true, then $L_1 \succ L_1'$ and $L_2' \succ L_2$ imply that

$$U(0.5M) > 0.1U(2.5M) + 0.89U(0.5M) + 0.01U(0) \Rightarrow 0.11U(0.5M) > 0.1U(2.5M) + 0.01U(0)$$
$$0.1U(2.5M) + 0.9U(0) > 0.11U(0.5M) + 0.89U(0) \Rightarrow 0.1U(2.5M) + 0.01U(0) > 0.11U(0.5M)$$

, which are contratictory. This paradox suggests that people may want to avoid regret.

The following example, called Ellsberg Paradox, casts doubt on whether people can always assign probabilities to the uncertain events.

Example (Ellsberg Paradox): A decision maker is told that among 300 balls contained in an urn, 100 balls are red and 200 are either blue or green. The decision maker is asked to make two decisions. The first one is to choose between

- Gamble A: You receive $1000 if the ball is red.
- Gamble B: You receive $1000 if the ball is blue.

The second one is to choose between

- Gamble C: You receive $1000 if the ball is not red.
- Gamble D: You receive $1000 if the ball is not blue.

People often exhibit that $A \succ B$ and $C \succ D$. $A \succ B$ implies that $p(R) > p(B)$, while $C \succ D$ implies that $1 - p(R) > 1 - p(B)$, contradicting to $p(R) > p(B)$. This paradox suggests that people exhibit ambiguity aversion.

3.2 Risk Aversion

If a lottery yields the monetary outcomes, it can be represented by a cumulative distribution function $F : M \to [0,1]$. We can extend the expected utility theory to obtain the v.N-M utility function

$$U(F) = \int u(x)dF(x)$$

where $u(x)$ is a Bernoulli utility assigned to x amount of money. (The outcome of the lottery can be discrete, and in that case $U(F) = \sum_i u(x_i)p(x_i)$, where i is the outcome index, and $p(x_i)$ is the probability that the outcome x_i occurs.)

Definition 8 (Risk Aversion)**.** *A decision maker is **risk averse** if for any lottery F,*

$$\int u(x)dF(x) \leq u\left(\int xdF(x)\right).$$

Two ways to appreciate the worth of lottery $F(\cdot)$ is to look at the *certainty equivalent*, denoted $c(F,u)$ and defined by

$$u(c(F,u)) = \int u(x)dF(x),$$

and the *probability premium*, denoted $\pi(x,\varepsilon,u)$ and defined by

$$u(x) = \left(\frac{1}{2} + \pi(x,\varepsilon,u)\right)u(x+\varepsilon) + \left(\frac{1}{2} - \pi(x,\varepsilon,u)\right)u(x-\varepsilon).$$

In words, the certainty equivalent is the amount of money for which the individual is indifferent between the gamble $F(\cdot)$ and the certain amount $c(F,u)$. The probability premium is the excessive in winning probability over fair odds that makes the individual indifferent between the certain outcome x and a gamble between $x+\varepsilon$ and $x-\varepsilon$.

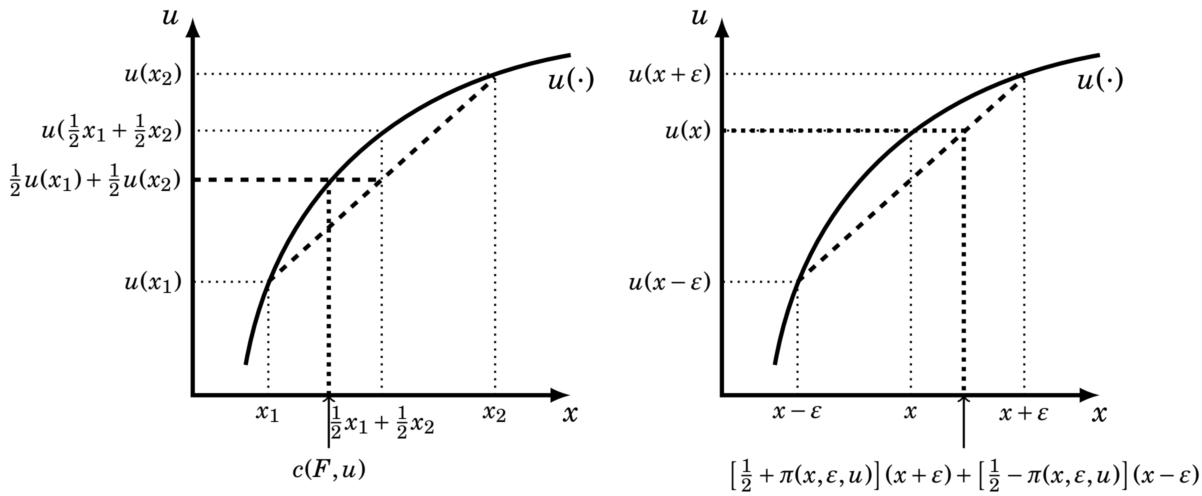

Figure 5: The Certainty Equivalent and the Probability Premium

Proposition 9. *The following three statements are equivalent.*

1. *The decision maker is risk averse.*
2. $u(\cdot)$ *is concave.*
3. $c(F,u) \leq \int xdF(x)$.
4. $\pi(x,\varepsilon,u) \geq 0$ *for all* x,ε.

Proof: (1 ⇔ 2): By Jensen's inequality, $u(\cdot)$ is concave if and only if $\int u(x)dF(x) \leq u\left(\int xdF(x)\right)$.
(1 ⇔ 3): $u(c(F,u)) = \int u(x)dF(x) \leq u\left(\int xdF(x)\right) \Leftrightarrow c(F,u) \leq \int xdF(x)$.
(2 ⇔ 4): (For rotational simplicity, I write $\pi(x,\varepsilon,u)$ as π.) $\frac{1}{2}u(x-\varepsilon)+\frac{1}{2}u(x+\varepsilon) \leq u(\frac{1}{2}(x-\varepsilon)+\frac{1}{2}(x+\varepsilon)) = u(x) = \left(\frac{1}{2}+\pi\right)u(x+\varepsilon)+\left(\frac{1}{2}-\pi\right)u(x-\varepsilon)$, where the first inequality is due to Jensen's inequality, and the last equality is by definition. Canceling common terms, we have $0 \leq \pi(u(x+\varepsilon)+u(x-\varepsilon))$, leading to $0 \leq \pi$. □

Example (Demand for a risky asset): Suppose that there are two assets, a *safe* asset with 1 dollar return per dollar and a *risky* one with a random return of z dollars per dollar. Assume that $\int zdF(z) > 1$, meaning that a risk is actuarially favorable. Letting α and β denote the amounts invested in the risky and safe assets, the utility maximization problem of the investor with wealth w is

$$\max_{\alpha,\beta\geq 0}\int u(\alpha z+\beta)dF(z) \text{ subject to } \alpha+\beta = w \qquad \text{or} \qquad \max_{\alpha\in[0,w]}\int u(\alpha z+w-\alpha)dF(z)$$

The Kuhn-Tucker first-order condition at the optimum α^* is

$$\phi(\alpha^*) := \int u'(w+\alpha^*(z-1))(z-1)dF(z)\begin{cases}\leq 0 & \text{if } \alpha^* < w\\ \geq 0 & \text{if } \alpha^* > 0\end{cases}$$

Since $\phi(0) = u'(w)\int(z-1)dF(z) > 0$, $\alpha^* = 0$ cannot be optimal. This implies that if a risky asset is actuarially favorable, an investor will always accept some amount of risk, regardless of how risk averse she is.

3.3 Measuring Risk Aversion

Definition 9. *The Arrow-Pratt coefficient of absolute risk aversion is*

$$r_A(x) = -\frac{u''(x)}{u'(x)},$$

at wealth level x.

Intuitively, the more concave the utility function, the more risk averse the consumer. Thus, risk aversion could be measured by the "curvature" of the function, $u''(x)$. The problem is that $u''(x)$ is not invariant to the monotone transformation of the utility function. That's why we normalize the second derivative by dividing by the first, we get a reasonable measure.

- -

Digression: To give a bit more rationale, consider a simple gamble that the consumer with an initial wealth level w gets x with probability p and y with probability $1-p$. The consumer's acceptance set, $A(w)$ is the set of all pairs (x, y) that the consumer would accept at an initial wealth level w. Let $y(x)$ denote the smallest y that consumer is willing to take the gamble given x. See Figure 6.

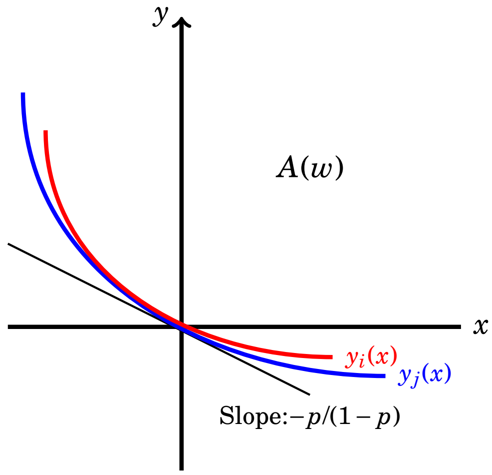

Figure 6: The Acceptance Set and Risk Aversion

The acceptance set is the upper right region of the curve. Consumer i is locally more risk averse than consumer j if i's acceptance set is contained in j's set in a neighborhood of (0,0).

For a pair of outcomes $(x, y(x))$, the consumer is indifferent between taking the gamble and not taking it.

$$pu(w+x)+(1-p)u(w+y(x)) = u(w)$$

Differentiating this identity with respect to x and evaluating this derivative at $x = 0$, we have

$$pu'(w)+(1-p)u'(w)y'(0) = 0.$$

Thus the slope at $(0,0)$ is $-\frac{p}{1-p}$. Given the same slope, consumer i is said to be locally more risk averse than consumer j if i's acceptance set is contained in the j's set, that is, if consumer i's acceptance boundary $y_i(x)$ is "more curved" than that of j. We can check the curvature of the acceptance set by examining $y''(x)$ around $(0,0)$. Twice differentiating the identity with respect to x and evaluating this second derivative at $x = 0$, we have

$$pu''(w)+(1-p)u''(w)y'(0)y'(0)+(1-p)u'(w)y''(0) = 0.$$

Using the fact that $y'(0) = -\frac{p}{1-p}$, we have

$$y''(0) = \frac{p}{(1-p)^2}\left[-\frac{u''(w)}{u'(w)}\right],$$

which is proportional to $r_A(w)$.

- -

Given two utility functions $u_1(\cdot)$ and $u_2(\cdot)$, we say that $u_2(\cdot)$ is *more risk averse* than $u_1(\cdot)$ if $r_A(x,u_2) \geq r_A(x,u_1)$ for every x.

Proposition 10. *Let $u_1(\cdot)$ and $u_2(\cdot)$ be two differentiable, increasing, and concave utility functions of wealth. The following are equivalent:*

1. $r_A(x,u_2) \geq r_A(x,u_1)$ *for every* x.
2. $u_2(x) = G(u_1(x))$ *for some increasing concave function* $G(\cdot)$.
3. $c(F,u_2) \leq c(F,u_1)$ *for any* $F(\cdot)$.
4. $\pi(x,\varepsilon,u_2) \geq \pi(x,\varepsilon,u_1)$ *for any* $F(\cdot)$.

Proof: (1) implies (2). Define $G(u_1)$ implicitly by $u_2(w) = G(u_1(w))$. Differentiate this twice to find $u_2'(w) = G'(u_1(w))u_1'(w)$ and $u_2''(w) = G''(u_1(w))u_1'(w)^2 + G'(u_1(w))u_1''(w)$. Since $u_1'(w) > 0$ and $u_2'(w) > 0$, the first equation establishes $G'(\cdot) > 0$. Dividing the second equation and rearranging it to get $\frac{G''(u_1)}{G'(u_1)}u_1'(w) = \frac{u_2''(w)}{u_2'(w)} - \frac{u_1''(w)}{u_1'(w)} = -r_A(x,u_2) + r_A(x,u_1) \leq 0$. Since $G'(u_1)$ and $u'(w)$ are positive, $G''(u_1) \leq 0$.
(2) implies (3). $u_2(c(F,u_2)) = \int u_2(x)dF(x) = \int G(u_1(x))dF(x) \leq \int u_1(x)dF(x) = u_1(c(F,u_1))$. Thus, $c(F,u_2) \leq c(F,u_1)$.
(3) implies (4). (For notational simplicity, I will write $\pi(x,\varepsilon,u_i)$ and π_i for $i = 1,2$.) Denote by $F(\cdot)$ the distribution function that puts probability $\frac{1}{2} - \pi_2$ on $x-\varepsilon$ and $\frac{1}{2}+\pi_2$ on $x+\varepsilon$, so that $c(F,u_2) = x$. By (3), $c(F,u_1) \geq x$, or $u_1(c(F,u_2)) \geq u_1(x)$. Since $u_1(c(F,u_1)) = (\frac{1}{2} - \pi_2)u_1(x-\varepsilon) + (\frac{1}{2}+\pi_2)u_1(x+\varepsilon) = \frac{1}{2}(u_1(x-\varepsilon)+u_1(x+\varepsilon)) + \pi_2(u_1(x+\varepsilon) - u_1(x-\varepsilon))$ and $u_1(x) = (\frac{1}{2}-\pi_1)u_1(x-\varepsilon) + (\frac{1}{2}+\pi_1)u_1(x+\varepsilon) = \frac{1}{2}(u_1(x-\varepsilon)+u_1(x+\varepsilon)) + \pi_1(u_1(x+\varepsilon)-u_1(x-\varepsilon))$, $\pi_2 \geq \pi_1$.
(4) implies (1). Since $\pi(x,0,u_1) = \pi(x,0,u_2) = 0$, (4) implies that $\left.\frac{\partial\pi(x,\varepsilon,u_2)}{\partial\varepsilon}\right|_{\varepsilon=0} \geq \left.\frac{\partial\pi(x,\varepsilon,u_1)}{\partial\varepsilon}\right|_{\varepsilon=0}$. Twice differentiating $u(x) = \left(\frac{1}{2}+\pi(x,\varepsilon,u)\right)u(x+\varepsilon) + \left(\frac{1}{2}-\pi(x,\varepsilon,u)\right)u(x-\varepsilon)$ with respect to ε and evaluating it at $\varepsilon = 0$, we have $u''(x) + 4\left.\frac{\partial\pi(x,\varepsilon,u)}{\partial\varepsilon}\right|_{\varepsilon=0} u'(x) = 0$, which means $r_A(x,u) = 4\left.\frac{\partial\pi(x,\varepsilon,u)}{\partial\varepsilon}\right|_{\varepsilon=0}$. Thus, $\left.\frac{\partial\pi(x,\varepsilon,u_2)}{\partial\varepsilon}\right|_{\varepsilon=0} \geq \left.\frac{\partial\pi(x,\varepsilon,u_1)}{\partial\varepsilon}\right|_{\varepsilon=0}$ implies $r_A(x,u_2) \geq r_A(x,u_1)$. □

The risk preference might depend on the wealth level, x. We say that the Bernoulli utility function $u(\cdot)$ exhibits decreasing/ constant/ increasing absolute risk aversion (DARA/ CARA/ IARA) if $r_A(x,u)$ is decreasing/ constant/ increasing in x.

The CARA utility function takes the following form:

$$u(x) = -e^{-\alpha x},\ a > 0 \quad \to r_A(x,u) = a, \forall x.$$

Example (Demand for a risky asset, continued): Suppose that the risk averse investor has a DARA Bernoulli utility function. Would he invest more as he holds more wealth? Consider two wealth levels w_1 and w_2, $w_2 > w_1$. Define $u_1(x) := u(w_1 + x)$ and $u_2(x) := u(w_2 + x)$. Because of the DARA property, $u_1(x) = \psi(u_2(x))$ for some concave function $\psi(\cdot)$. Then, the utility maximization problem of the investor with wealth w_i is

$$\max_{0 \le \alpha \le w_i} \int u(w_i + \alpha(z-1))dF(z) = \int u_i(\alpha(z-1)dF(z).$$

The first order condition is

$$\phi_1(\alpha_1^*) := \int (z-1)u_1'(\alpha_1^*(z-1))dF(z) = 0$$

for the investor with w_1, and

$$\int (z-1)u_2'(\alpha_2^*(z-1))dF(z) = 0$$

for the investor with w_2. (We should have considered the corner solution case with $\alpha_i^* = w_i$ to be more rigorous, but we assume interior solutions. The second order condition for maximum is satisfied; check it by yourself.)

Since $u_1(\cdot)$ is concave, or $u_1'(\cdot)$ is decreasing, $\phi_1(\cdot)$ is decreasing. This means that $\alpha_2^* > \alpha_1^*$ (i.e., the richer, the more risky assets) if $\phi_1(\alpha_2^*) < 0$, which holds because

$$\begin{aligned}
&\phi_1(\alpha_2^*) = \int (z-1)u_1'(\alpha_2^*(z-1))dF(z) = \int (z-1)\psi'(u_2(\alpha_2^*(z-1)))u_2'(\alpha_2^*(z-1))dF(z) \\
&= \int_{-\infty}^{1} (z-1)\psi'(u_2(\alpha_2^*(z-1)))u_2'(\alpha_2^*(z-1))dF(z) + \int_{1}^{\infty} (z-1)\psi'(u_2(\alpha_2^*(z-1)))u_2'(\alpha_2^*(z-1))dF(z) \\
&< \int_{-\infty}^{1} (z-1)\psi'(u_2(0))u_2'(\alpha_2^*(z-1))dF(z) + \int_{1}^{\infty} (z-1)\psi'(u_2(0))u_2'(\alpha_2^*(z-1))dF(z) \\
&= \psi'(u_2(0)) \int (z-1)u_2'(\alpha_2^*(z-1))dF(z) = 0,
\end{aligned}$$

where the inequality above is due to $\psi'(u_2(-)) > \psi'(u_2(0)) > \psi'(u_2(+))$. Note that the range of z is decomposed to $(-\infty, 1]$ and $[1, \infty)$ so that $(z-1)$ is negative in the first range and positive in the second range.

Thus, the demand of risky asset is increasing in wealth, that is, the risky asset is a normal good if the investor has a DARA utility function. Also, the risky asset is an inferior good if the investor has an IARA utility function. (But in practice, you don't see an investor with IARA utility function. Can you imagine someone who is willing to play a \$10 gamble with \$1M wealth but find it too risky with \$1B wealth?)

Another useful measure of risk aversion is the coefficient of *relative risk aversion* at x defined as $r_R(x,u) := -x\frac{u''(x)}{u'(x)} = x r_A(x,u)$. The constant relative risk aversion (CRRA) utility takes the power function form:

$$u(x) = \begin{cases} \frac{x^{1-\eta}}{1-\eta} & \text{if } \eta \geq 0, \eta \neq 1 \\ \ln(x) & \text{if } \eta = 1 \end{cases}$$

The RRA of this utility function is η, and only the power form (and its affine transformation) has a constant RRA. ($u(x) = \ln x = \lim_{\eta \to 1} \frac{x^{1-\eta}}{1-\eta}$.)

Since $r_R(x) = x r_A(x)$, a consumer who has a DRRA utility function must exhibit DARA.

3.4 Comparison between Two Lotteries

We have learned how to compare risk attitudes. Let us learn how to compare between lotteries. To avoid technical subtleties of dealing with $-\infty$ and ∞, assume that the lottery has the support of $[a, b]$, with $a, b \in \mathbb{R}$. For notational simplicity, by lottery F, I mean a lottery of which outcome is drawn from a cumulative probability distribution $F : [a, b] \to [0, 1]$.

Definition 10. *Lottery F first-order stochastically dominates (FOSD) lottery G if, for every nondecreasing function $u : \mathbb{R} \to \mathbb{R}$,*

$$\int u(x) dF(x) \geq \int u(x) dG(x)$$

Proposition 11. *F FOSD G if and only if $F(x) \leq G(x)$ for every x.*

Proof: ($\Leftarrow$) Define $H(x) := F(x) - G(x)$ so that $H(a) = H(b) = 0$ and $H(x) \leq 0$ for every x. $\int u(x) dH(x) = u(x)H(x)|_a^b - \int u'(x)H(x)dx = -\int u'(x)H(x)dx \geq 0$.
($\Rightarrow$) Suppose there is $\bar{x}$ such that $F(\bar{x}) > G(\bar{x})$. Let $u(x)$ be an indicator function of $x \geq \bar{x}$. Then, $\int u(x) dF(x) = 1 - F(\bar{x}) < 1 - G(\bar{x}) = \int u(x) dG(\bar{x})$, which is contradictory. □

That is, F FOSD G means that the probability of getting at least x is higher under F than under G. Figure 7 illustrates FOSD.

We can also compare two lotteries with the same mean in terms of their riskiness.

Definition 11. *For two lotteries F and G with the same mean, F second-order stochastically dominates (SOSD) G if for every nondecreasing concave function $u : \mathbb{R} \to \mathbb{R}$,*

$$\int u(x) dF(x) \geq \int u(x) dG(x).$$

Remarks:

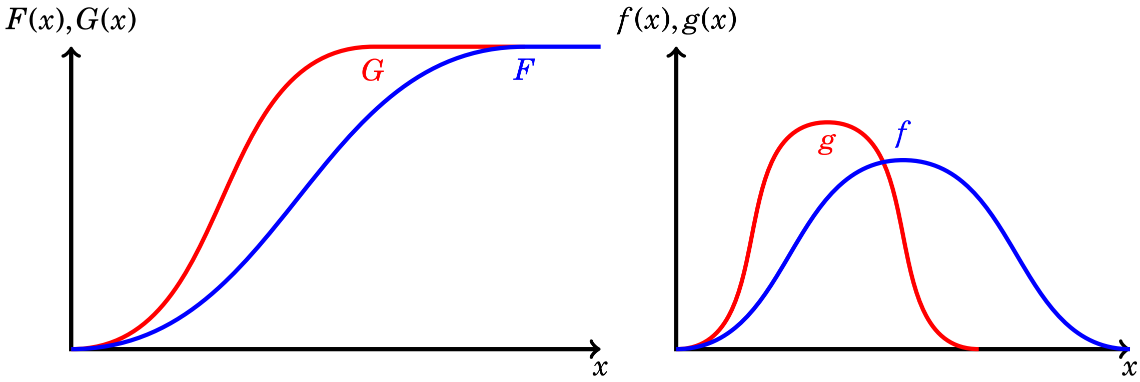

Figure 7: First-Order Stochastic Dominance

- For arbitrary lotteries, neither may dominate the other in either FOSD or SOSD senses. That is, both of these dominance concepts are partial orderings of lotteries, not complete orderings.

- F FOSD G does not mean that the payoff drawn from F is always higher than the one from G. We can only say it in a probabilistic manner.

- FOSD is stronger than requiring the mean of F to be higher than G. (**Exercise**: Provide an example of two uni-modal symmetric distributions F and G that the mean of F is higher than that of G, but F does not FOSD G.)

- FOSD implies SOSD.

- Given a lottery $F(\cdot)$, its *mean-preserving spread* (MPS) is another lottery $G(\cdot)$ whose payoff is $y = x + z$, where x is drawn from $F(\cdot)$ and z is from a distribution $H(\cdot)$ with $\int z dH(z) = 0$. (That is, G can be obtained by perturbing F without changing its mean.) See Figure 8. Since $F(a) = G(a) = 0$ and $F(b) = G(b) = 1$, $\int_a^t F(x)dx$ has to be smaller than $\int_a^t G(x)dx$. So $\inf x dF(x) = \int x dG(x)$ and $\int_a^t F(x)dx \leq \int_b^t G(x)dx$ are two conditions that define the MPS.

Proposition 12. *For F and G such that $\int x dF(x) = \int x dG(x)$, the following are equivalent.*

1. *F SOSD G.*

2. *G is a mean-preserving spread of F.*

3. *$\int_a^t G(x)dx \geq \int_a^t F(x)dx$ for all $t \in [a, b]$.*

Proof: (2) implies (1). $\int u(y)dG(y) = \int\int u(x+z)dH(z)dF(x) \leq \int u\left(\int (x+z)dH(z)\right)dF(x) = \int u(x)dF(x)$, where the inequality is due to the Jensen's inequality.

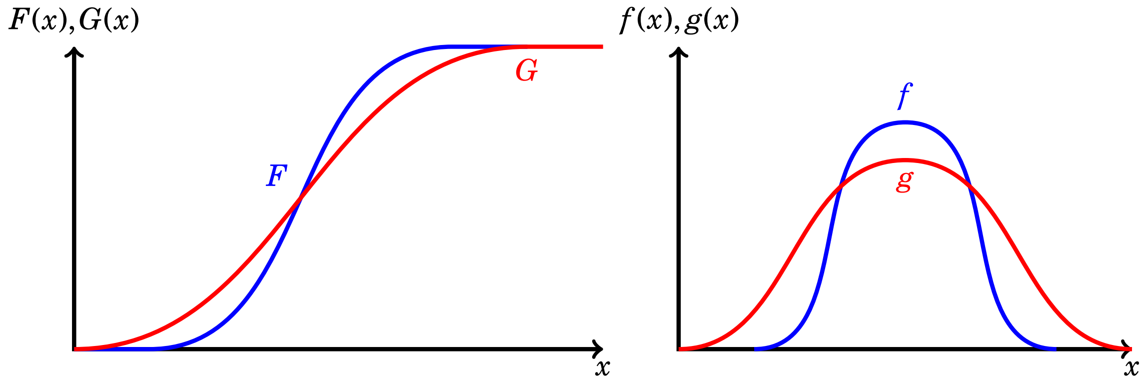

Figure 8: Second-Order Stochastic Dominance

(1) implies (3). Let $u(x) = \min\{x-t, 0\}$. Note that it is increasing and concave in x. Since $\int u(x)dF(x) \geq \int u(x)dG(x)$, $\int_a^t (x-t)dF(x) \geq \int_a^t (x-t)dG(x)$. Note that $\int_a^t (x-t)dF(x) = \int_a^t xdF(x) - \int_a^t tdF(x) = xF(x)|_a^t - \int_a^t F(x)dx - tF(t) = -\int_a^t F(x)dx$, and similarly, $\int_a^t (x-t)dG(x) = -\int_a^t G(x)dx$, implying (3). (3) with $\int xdF(x) = \int xdG(x)$ implies (1) by definition. $\square$

4 Producer Theory

What is a firm? Simply put, a firm is an entity created by individuals for "some" purpose. This entity typically purchases inputs and combines them to produce and sell output. What's the purpose of a firm? Profit maximization is the most common answer an economist would give. Profit—the difference between revenue the firm earns from selling its output and expenditure it makes buying its inputs—is income to owners of the firm. Since these owners are also consumers, and consumers accrue their utility from the goods and services purchased with their income, the more profit the firm can make, the happier its owners/consumers.

Profit maximization may not be the only motive behind firm behavior, and economists have considered many others; sales, market share, social impact, or even prestige maximization. Each of these alternatives to profit maximization have at least some superficial appeal. Yet the majority of economists continue to embrace the hypothesis of profit maximization most of the time in most of their work. There are good reasons for it. Empirically, firms seem to maximize their profits. Theoretically, analyzing profit-maximizing firms is simpler than other those with alternative motives, and it corresponds well to the consumers' utility maximization theory. Also, the market competition coerces the firm towards profit maximization.

4.1 Production Set

Consider an economy with L commodities.

- $y = (y_1, \cdots, y_L) \in \mathbb{R}^L$ is a *production plan*. If $y_l > 0$, then y_l units of lth commodity are produced as an output, and if $y_l < 0$, then $|y_l|$ units of lth commodity are used as an input. For example, when $L = 2$, $y = (5, -7)$ means that the firm produces 5 units of commodity 1 by using 7 unit of commodity 2.

- A given technology is described by the *production set* denoted $Y \subset \mathbb{R}^L$, which is the set of all feasible production plans.

We assume that

- Y is nonempty and closed.

- Free disposal: if $y \in Y$ and $y' \leq y$ (so that y' produces at most the same amount of outputs using at least the same amount of inputs), then $y' \in Y$.

With only one output and m inputs, the technology can also be described by the *input requirement set*:

$$V(y) := \{x \in \mathbb{R}^m_+ | (y, -x) \in Y\},$$

which is the set of all input bundles x that produce at least y units of output.

- $f(x) := \max\{y \in \mathbb{R}_+ | x \in V(y)\}$: Production function (the largest units of output it can produce.)
- $Q(y) := \{x \in \mathbb{R}^m_+ | f(x) = y\}$: Isoquant for output level y.
- $f_l(x) := \frac{\partial f(x)}{\partial x_l}$: Marginal product of input l.
- $MRTS_{lk}(x) := \frac{f_l(x)}{f_k(x)}$: Marginal rate of technical substitution (MRTS) between inputs l and k when the current income vector is x.

[RESUME IT HERE.]

4.2 Profit Maximization

(To be added)

4.3 Cost Minimization

(To be added)

(introduce Topkis theorem here. You may want to check Milgrom and Shannon (1994) and Milgrom and Roberts (1996).)

5 Markets

5.1 2-person 2-good Exchange Economy

(To be added)

5.2 Competitive Equilibrium

(To be added)

5.3 Imperfect Competition: Monopoly and Oligopoly

(To be added)

6 Market Failures

6.1 Public Goods Provision

(To be added)

6.2 Externalities

(To be added)

6.3 Asymmetric Information

(To be added)

7 Further Topics in Consumer Choice

7.1 Critiques of Expected Utility Theory

(refer to MIT OCW Micro Theory III)

7.2 Prospect Theory

(refer to MIT OCW Micro Theory III)

7.3 Rabin's Critiques of Exponential Discounting

7.4 Dynamic choice and hyperbolic discounting

7.5 Costly Contemplation Model—Rational Inattention

7.6 Preference for Flexibility, Temptation, and Self Control

Convex preference.

References

Andreoni, James and John Miller, "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism," *Econometrica*, 2002, *70* (2), 737–753.

Choi, Syngjoo, Shachar Kariv, Wieland Müller, and Dan Silverman, "Who Is (More) Rational?," *American Economic Review*, June 2014, *104* (6), 1518–1550.

Milgrom, Paul and Chris Shannon, "Monotone Comparative Statics," *Econometrica*, 1994, *62* (1), 157–180.

— **and John Roberts**, "The LeChatelier Principle," *The American Economic Review*, 1996, *86* (1), 173–179.