Recitation 1: Decision Theory Foundations

Xincheng Qiu (qiux@sas.upenn.edu)

First, we review the preference-based approach and the choice-based approach to decision theory using Table 1 as a guide.

Remark 1. Preference-based approach assumes the decision maker chooses the best (according to his rational preference) alternative available (under the constraint). This does not mean that we assume people are profiteering, greed, selfishness, nor ruthlessness; instead, things like appreciation of arts, commitment to religions, and concern for families could be easily incorporated in our completely abstract setting. For example, social preferences have already been incorprated in many research, where people care about their social status, others' welfare, and fairness. But we do assume people pursue some objective. Rationality refers to completeness and transitivity, which are not incredibly crazy. That said, completeness is somewhat technical, but transitivity is indeed violated in two plausible scenarios: aggregation of preferences/internal selves/Condorcet paradox, and imperceptible differences.

Remark 2. Preference-based approach assumes preferences are stable over time. It is not impossible that the environment might affect preferences, but invoking changes in preference can trivially "explain" everything and thus is not compelling. Some may argue addiction provides a counterexample that violates the assumption, as smoking today causes one to like smoking more tomorrow. Usually changing preferences can be resolved by enriching the set of alternatives. For example, the stable preference $(0,0) \succ (0,1) \succ (1,1) \succ (1,0)$ captures the addiction property.

Example 1. The primitives of the preference-based decision theory that we studied were a set X and a complete and transitive binary relation $\succsim$ on X. We could instead have started with X and a binary relation $\succ$ on X satisfying

- Asymmetry: For all x and y, if $x \succ y$ then not $y \succ x$
- Negative Transitivity: For all x, y, and z: not $x \succ y$ and not $y \succ z \Rightarrow$ not $x \succ z$.

The two approaches are equivalent. The interesting direction is the following property: each asymmetric and negatively transitive $\succ$ is the strict preference relation derived from some complete and transitive $\succsim$. We can define $\succsim$ from $\succ$ by: $x \succsim y$ iff not $y \succ x$. Then given an asymmetric and negatively transitive $\succ$, we can prove that $\succsim$ generated by this $\succ$ via the above definition is complete and transitive. Denote the strict part of this $\succsim$ by $\succ^{\#}$, which

Table 1: Two Approaches to Decision Theory

	Preference-Based Approach	Choice-Base Approach
Primitive	Preference: a binary relation $\succsim$ on X	Choice structure: $\langle \mathfrak{B}, C \rangle$
Extract from each other	Derive choice behavior $C^* (B, \succsim) := \{x \in B : x \succsim y, \forall y \in B\}$	Revealed preference relation: $x \succsim^* y$ iff $x = y$ or $\exists B \in \mathfrak{B}$ s.t. $y \in B$ and $x \in C(B)$
"Rational" Assumptions	Completeness $\forall x, y \in X,$ either $x \succsim y,$ or $y \succsim x$	B contains all size-two subsets (Recall by def $\emptyset \neq C(B) \subseteq B$, $\forall B \in \mathfrak{B}$)
	Transitivity $\forall x, y, z \in X, x \succsim y, y \succsim z \Rightarrow x \succsim z$	WARP: $\forall B \in \mathfrak{B}, x, y \in B,$ $y \in C(B)$ and $x \succsim^* y \Rightarrow x \in C(B)$
Connection	The Revealed Preference Theorem	
Standard Assumptions	1. $x(p, m) = C^* (B, \succsim)$ $\Rightarrow$ Homogeneity of degree 0	1. Homogeneity of degree 0 $x(\alpha p, \alpha m) = x(p, m), \forall \alpha > 0$
	2. (SM) $\Rightarrow$ (M) $\Rightarrow$ (LNS) $\Rightarrow$ Walras' law	2. Walras' law $p \cdot x = m, \forall x \in x(p, m)$
	3. Rational: Completeness + Transitivity	3. B contains all size-two subsets + WARP
Advantage	Easier: mathematical optimization tools	Based on observable, measurable objects
Interpretation	A process of introspection	An entirely behavioral foundation

is defined as $x \succ^{\#} y$ iff $x \succsim y$ and not $y \succsim x$. We can show $\succ^{\#} = \succ$ and hence establish the property.

Example 2. Amartya Sen (1969) proposed the following two properties, which give an equivalent characterization of WARP:

- Property α: if $x \in B \subseteq A$ and $x \in C(A)$, then $x \in C(B)$;
- Property β: if $y \in B \subseteq A$ and $y \in C(A)$, then $C(B) \subseteq C(A)$.

Property α can be interpreted as: if the world champion in some game is from Country S, then he must also be the champion in Country S. Property β can be interpreted as: if someone from Country S is a world champion, then all champions of Country S must be champions of the world. Consider a choice structure $\langle \mathfrak{B}, C \rangle$ where $\mathfrak{B}$ contains all subsets of X of size 2. It satisfies WARP if and only if Sen's α and β are satisfied. Proof is left as an exercise.

Then, we look at the revealed preference theorem and examine two related and natural problems.

Theorem 1. *The Revealed Preference Theorem*

1. *If $\langle \mathfrak{B}, C \rangle$ is rationalized by a transitive binary relation $\succsim$, then it satisfies WARP.*
2. *Conversely, if $\langle \mathfrak{B}, C \rangle$ satisfies WARP and B contains all subsets of X of size three, then $\succsim$ is transitive and rationalizes $\langle \mathfrak{B}, C \rangle$. If in addition B contains all subsets of X of size two, then $\succsim$ is also complete and is the only reflexive binary relation that rationalizes $\langle \mathfrak{B}, C \rangle$.*

Remark 3. The revealed preference theorem establishes the "equivalence" (up to some technical assumption, i.e., all subsets of size three) between transitivity and WARP and thus connects the preference-based approach and the choice-based approach. See slides for proof.

Exercise 1. Given a choice correspondence $C : \mathfrak{B} \to 2^X$, we can define the associated revealed preference relation $\succsim^*$, which in turn can generate a choice behavior $C^*(B, \succsim^*)$. A natural question is: does it hold that $C^*(B, \succsim^*) = C(B)$?

The answer is no. But it is always true that $C(B) \subseteq C^*(B, \succsim^*)$. Pick any $x \in C(B)$. By definition of $\succsim^*$, we have $x \succsim^* y, \forall y \in B$. That is, $x \in C^*(B, \succsim^*)$.

The revealed preference theorem tells us that if $\langle \mathfrak{B}, C \rangle$ satisfies WARP, then $\succsim^*$ rationalizes $\langle \mathfrak{B}, C \rangle$, i.e., $C^*(B, \succsim^*) = C(B)$. To show $C^*(B, \succsim^*) \subseteq C(B)$, pick any $x \in C^*(B, \succsim^*)$. Pick any $y \in C(B) \subseteq B$ (this is possible because by our definition $C(B)$ is nonempty), we have $x \succsim^* y$. We want to have $x \in C(B)$, and this is exactly what WARP assumes. The key benefit of going through the proof of a well-established theorem is that it allows you to see clearly where the assumptions matter.

Now construct a counterexample where WARP fails. Let $X = \{a, b, c\}$. Suppose $C(\{a, b\}) = \{a\}, C(\{a, b, c\}) = \{b\}$. We have $a \succsim^* b$, $b \succsim^* a$, and $b \succsim^* c$, as well as $x \succsim^* x$ for all $x \in X$ by definition. Note that WARP is violated, because $a \succsim^* b$, $b \in C(\{a, b, c\})$, but $a \notin C(\{a, b, c\})$. Consider the set $\{a, b\}$, $C^*(\{a, b\}, \succsim^*) = \{a, b\} \neq C(\{a, b\}) = \{a\}$.

Another related question: do we also have $C^*(B, \succsim^*) = C(B), \forall B \in \mathfrak{B}$ implies that $C(B)$ satisfies WARP? Here is a (pathological) counterexample where $C^*(B, \succsim^*) = C(B)$ but WARP is violated: suppose $\mathfrak{B}$ only contains two sets: $B_1 = \{a, b, c\}$ and $B_2 = \{a, b, d\}$, with a choice correspondence $C(\{a, b, c\}) = \{a\}$ and $C(\{a, b, d\}) = \{b\}$, which implies $a \succsim^* b, a \succsim^* c$ as well as $b \succsim^* a, b \succsim^* d$, but no more (except reflexivity). WARP is violated, but you can verify that $C^*(\{a, b, c\}, \succsim^*) = \{a\} = C(\{a, b, c\})$ and $C^*(\{a, b, d\}, \succsim^*) = \{b\} = C(\{a, b, d\})$.

Exercise 2. Start with a given preference relation $\succsim$. It generates a choice correspondence $C^*(B, \succsim)$, which allows us to further define a revealed preference relation, denoted $\succsim^\#$. The "dual" problem to the above: is it true that $\succsim^\# = \succsim$?

$\succsim \subseteq \succsim^\#$: Pick any $(x, y) \in \succsim$ (recall that a binary relation on X is a subset of X^2), or $x \succsim y$. We have $x \in C^*(\{x, y\}, \succsim)$ as long as $\succsim$ is reflexive. That is, $x \succsim^\# y$.

$\succsim^\# \subseteq \succsim$: Pick any (x, y) s.t. $x \succsim^\# y$. By definition of revealed preference, $\exists B \in \mathfrak{B}$ s.t. $x, y \in B$ and $x \in C^*(B, \succsim)$. (It is trivial if $x = y$ as long as $\succsim$ is reflexive.) Therefore we have $x \succsim y$.

But what if $\succsim$ is irreflexive? A weird preference on $\mathbb{R}$: $x \succsim y$ iff $x \geq y + 2$.

Remark 4. Although these two problems are motivated by the revealed preference theorem, the pathological cases may be less interesting than results in the theorem. The real purpose of the above two exercises is to: 1) get familiar with these concepts by playing with the definitions; 2) be comfortable with rigorous proof and constructing counterexamples. The economic side of a theorem is more interesting, but going through the proof makes it clear why the assumptions matter.

Recitation 2: Preferences

Xincheng Qiu (qiux@sas.upenn.edu)

We will make some standard assumptions about consumer preferences, including rationality (completeness and transitivity). Strong monotonicity (if $x \geq y$ and $x \neq y$, then $x \succ y$) implies monotonicity (if $x \gg y$, then $x \succ y$), which in turn implies local nonsatiation (any open neighborhood of any $y \in X$ contains a bundle $x \in X$ such that $x \succ y$), which in turn implies Walras' law. In addition, we often assume that $\succsim$ is convex, i.e., its upper contour sets are convex, which captures the idea of diminishing marginal rate of substitution, or a taste for diversification.

Exercise 1. A consumer in a two-good world demands $x = (1, 2)$ at $(p, m) = (2, 4, 10)$, and he demands $x' = (2, 1)$ at $(p', m') = (6, 3, 15)$. Is he maximizing a locally nonsatiated utility function?

One may want to start with checking if this is consistent with Walra's law, as it is an immediate implication of LNS. Simple calculation shows it does not directly violate Walra's law (eventually we will see that Walra's law is indeed violated). Observe that x is revealed preferred to x' and x' is also revealed preferred to x.

Proof 1: Suppose LNS holds. Because x' is chosen when x is affordable, we have $u(x') \geq u(x)$. Since $p \cdot x' < m$, by continuity of a linear function, there exists a neighborhood N of x' such that $p \cdot y < m$ for all $y \in N$. By LNS, N contains a point y such that $u(y) > u(x') \geq u(x)$. Since y is affordable at (p, m), this contradicts the assumption that x maximizes utility at (p, m).

Proof 2: Because x' and x are each revealed preferred to the other, we must have $u(x') = u(x)$. Hence, x' also maximizes utility at (p, m), but this violates Walras' law.

We have four equivalent (given completeness) definitions for continuity:

1. if $x \succ y$, then $\exists$ neighborhoods N_x and N_y s.t. $x' \succ y'$, $\forall x' \in N_x, y' \in N_y$;
2. the graph of $\succsim$, $\{(x, y) \in X^2 : x \succsim y\}$ is closed (i.e., $\succsim$ is a closed subset of X^2);
3. for any x, the upper and lower contour sets are closed;
4. for any x, the strict upper and lower contour sets are open.

Theorem 1. *Debreu's Representation Theorem*

For any a and $b > a$ in $\mathbb{R}$, a continuous rational preference relation $\succsim$ on a connected set $X \subseteq \mathbb{R}^n$ is representable by a continuous function $u : X \to [a, b]$.

Theorem 2. *Monotone Representation Theorem*

A monotone continuous rational preference relation $\succsim$ on $\mathbb{R}_+^L$ is representable by a continuous function u.

Proof. Key idea: construct the utility function $u(x)$ such that $x \sim (u(x), \ldots, u(x))$. $\square$

Remark 1. We often describe a preference by a utility function, which is easier to deal with mathematically. Utility representation ($x \succsim y \Leftrightarrow u(x) \geq u(y)$) is not unique: any strict increasing transformation of the utility function represents the same preference. Properties that are preserved under such transformations are called ordinal. The above theorems provide conditions under which a utility function representation exists. It turns out that we get an extra bit in the conclusion: not only is $\succsim$ representable, but it is representable by a continuous function. With a continuous utility function (defined over a compact set), we will have a well-defined maximum.

Exercise 2. Preferences are called *homothetic* if they satisfy $x \succsim y \Rightarrow \alpha x \succsim \alpha y, \forall \alpha \geq 0$. Suppose $\succsim$ is a complete, transitive, monotonic, continuous preference relation on $\mathbb{R}_+^L$. Show that $\succsim$ is homothetic if and only if there exists a utility representation u of $\succsim$ such that $u(\alpha x) = \alpha u(x)$ for all $\alpha \geq 0$.

Proof. The ($\Leftarrow$) direction is straightforward. The ($\Rightarrow$) direction is interesting. Suppose $\succsim$ is homothetic, and let u be the function representing $\succsim$ constructed in the proof of the Monotone Representing Theorem. Thus, for any $x \in \mathbb{R}_+^L$, $u(x)$ is the number such that $x \sim (u(x), \ldots, u(x))$. Then you can show that this u is homogeneous of degree 1. $\square$

Next, we discuss constructing counterexamples, as a follow-up of last session in proving results. This also provides a good opportunity to practice with concepts like rationality, continuity, monotonicity.

Exercise 3. Let $\succsim$ be a rational preference relation defined on $\mathbb{R}_+^L$ that is monotone and continuous. We can prove it is weakly monotone (see suggested solutions to PS1), which is defined by "$x \geq y$ implies $x \succsim y$". Provide a counterexample to the above proposition for when $\succsim$ is not continuous. Put it differently, we are going to find a preference that is rational, monotone, but and not weakly monotone and not continuous (this would be redundant given it is rational, monotone, and not weakly monotone) . We collect wisdom from colleagues about how to proceed with constructing counterexamples[1].

[1] Many interesting examples came up. Special thanks to Min, Yoshiki, Tomer, Justin, Rodrigo, Artem for presenting their examples and insights in constructing counterexamples.

- Consider a modified Lexicographic preference $\succsim$ defined on $\mathbb{R}_+^2$: $(x_1, x_2) \succsim (y_1, y_2)$ iff $x_1 > y_1$, or $x_1 = y_1$ and $x_2 \leq y_2$. It is trivial to show completeness and transitivity. To see it is not continuous, consider the upper contour set for a given bundle (x_1, x_2), $\left\{(z_1, z_2) \in \mathbb{R}_+^2 : z_1 > x_1, \text{ or } z_1 = x_1 \text{ and } z_2 \leq x_2\right\}$, which is not closed. To see it is not weakly monotone, consider $x_1 = y_1$ and $x_2 > y_2$, we have $(x_1, x_2) \geq (y_1, y_2)$ but not $(x_1, x_2) \succsim (y_1, y_2)$.

- Let $A = \left\{(x_1, x_2) \in \mathbb{R}_+^2 : (x_1, x_2) \gg 1 \text{ or } (x_1, x_2) = (1, 1)\right\}$. Define $U(x_1, x_2) = x_1 + x_2 + \mathbb{I}\{(x_1, x_2) \in A\}$. Consider the preference $\succsim$ on $\mathbb{R}_+^2$ represented by $U(.)$. Since $\succsim$ is representable, it is rational. We can also verify it is monotonic. But the preference is not continuous at $(1, 1)$. And it is not weakly monotonic, as $U\left(1, \frac{3}{2}\right) = \frac{5}{2} < U(1, 1) = 3$.

- Define $\succsim$ on $\mathbb{R}_+^2$ s.t. if $x_1 x_2 \geq y_1 y_2$ and $(x_1, x_2), (y_1, y_2) \neq (1, 0)$, we have $(x_1, x_2) \succsim (y_1, y_2)$; in addition, $(1, 0) \prec (y_1, y_2)$ for any $(y_1, y_2) \neq (1, 0)$. This is not weakly monotone, as we have $(1, 0) \prec (0, 0)$ by construction.

- Let $\succsim$ defined on $\mathbb{R}_+^2$ be a preference relation represented by

$$U(x_1, x_2) = \begin{cases} \min\{x_1, x_2\} & \text{if } x_1 \neq 0 \\ -x_2 & \text{if } x_1 = 0 \end{cases}$$

It is not continuous since the upper contour set for $(0, 1)$ is not closed. It is not weakly monotone since $(0, 1) \succ (0, 2)$.

- Let $\succsim$ defined on $\mathbb{R}_+^L$ s.t. if $\min\{x_i\} > \min\{y_i\}$ then $x \succ y$; if $\min\{x_i\} = \min\{y_i\}$, then $x \sim y$ except for if $x_1 = x_2 = \cdots = x_L$, then $x \succ y$.

- Consider the preference $\succsim$ defined on $\mathbb{R}_+^2$ that is represented by

$$U(x_1, x_2) = \begin{cases} f(x_2) & \text{if } x_1 < c \\ 10 - f(x_2) & \text{if } x_1 = c \\ 100 + f(x_2) & \text{if } x_1 > c \end{cases}$$

where $f(t) = \frac{t}{t+1} \in [0, 1]$, and c is a given positive constant.

Lastly, we discuss the Kuhn-Tucker NFOC in utility maximization problems. For a UMP

$$\max_{x \in \mathbb{R}_+^L} u(x) \text{ s.t. } p \cdot x \leq m \tag{1}$$

define the Lagrangian function $L(x,\lambda) := u(x) + \lambda(m - p\cdot x)$. Assume u is twice continuously differentiable, the Kuhn-Tucker theorem implies that for any solution $x^* \in \mathbb{R}_+^L$, there exists $\lambda^* \geq 0$ such that (x^*, λ^*) satisfies the following NFOC:

$$u_l(x^*) - \lambda^* p_l \leq 0,\ [u_l(x^*) - \lambda^* p_l]\, x_l^* = 0,\ \forall l \tag{2}$$

$$m - p\cdot x^* \geq 0,\ [m - p\cdot x^*]\,\lambda^* = 0 \tag{3}$$

Remark 2. The intuition[2] behind the Kuhn-Tucker NFOC. From the constrained envelope theorem (ET2), we have $v_m(p,m) = \lambda^*$, so λ^* is often referred to as the marginal utility of income. There is a natural economic interpretation for $u_l(x^*) - \lambda^* p_l$: imagine one increases consumption on good l by a little bit, say, 1 unit, then her utility increases by $u_l(x^*)$. In addition, to buy an additional unit of good l, it is as if her income goes down by p_l, which decreases the indirect utility by $\lambda^* p_l$. Therefore, $u_l(x^*) - \lambda^* p_l$ can be interpreted as the net effect of a small (1 unit) increase in consumption on good l. As long as $u_l(x) - \lambda p_l > 0$, she wants to increase the consumption on good l, which means x cannot be an optimal solution. A similar logic should also apply for when $u_l(x) - \lambda p_l < 0$, she wants to decrease the consumption on good l, except for if x_l is already 0 then there is no way to decrease x_l any more. This is the story behind the complementary slackness condition in equation (2). Equation (3) says that the budget constraint has to be satisfied; the only possibility that one didn't spend all the money is when $\lambda^*(p,m)$, the marginal utility of income, is 0.

Remark 3. The marginal utility of income, however, is not a perfect notion. Suppose at some given $(\bar{p},\bar{m}) \in \mathbb{R}_{++}^{L+1}$, the marginal utility of income derived from UMP of $u(x)$ in (1) is $\lambda(\bar{p},\bar{m}) > 0$, and solution denoted x^*. Consider an increasing transformation of the original utility function: $\hat{u}(x) := [u(x) - u(x^*)]^3$, which represents the same preference as u does. The indirect utility function for $\hat{u}$ at an arbitrary (p,m) is

$$\hat{v}(p,m) = [v(p,m) - u(x^*)]^3$$

Note that $\frac{\partial \hat{v}(p,m)}{\partial m} = 3\,[v(p,m) - u(x^*)]^2\, \frac{\partial v(p,m)}{\partial m}$. In particular, the marginal utility of income

[2] Wikipedia defines intuition to be "the ability to acquire knowledge without proof, evidence, or conscious reasoning." My perspective is, however, in economics, intuition actually refers to reasoning without mathematics (but not talking without reasoning). A model is to formalize the idea using a language free from obscurities, i.e. mathematics, whereas intuition conveys the economic story in English. I never regard intuition as a substitute for a formal proof, but a complement.

at $(\bar{p},\bar{m})$ is now given by

$$\hat{\lambda}(\bar{p},\bar{m})=\frac{\partial\hat{v}(p,m)}{\partial m}\mid_{(p,m)=(\bar{p},\bar{m})}=3\left[v(\bar{p},\bar{m})-u(x^*)\right]^2\frac{\partial v(p,m)}{\partial m}\mid_{(p,m)=(\bar{p},\bar{m})}=0$$

since $v(\bar{p},\bar{m})=u(x^*)$. This unwanted result that $\lambda(\bar{p},\bar{m})>0$ but $\hat{\lambda}(\bar{p},\bar{m})=0$ arises due to that preference is an ordinal concept whereas the marginal utility of income is a cardinal one.[3]

Remark 4. $f(x)$ strictly increasing is not equivalent to $f'(x)>0$. For one thing, $f(x)$ strictly increasing does not imply differentiability; for the other, even if we assume $f(x)$ is differentiable, it might be that $f'(x)=0$ for some points (e.g. $f(x)=x^3$ at $x=0$). In a relevant context: if LNS, we have $v(p,m)$ strictly increasing in m and $v_m(p,m)\geq 0$. To prove $v(p,m)$ strictly increasing in m, two common mistakes might seem attractive at first glance: 1) not to mention the differentiability, the envelop theorem only says $v_m(p,m)\geq 0$, which is not sufficient for $v(p,m)$ strictly increasing; 2) consider $m'>m$, $B(p,m)\subset B(p,m')$ only implies $v(p,m')\geq v(p,m)$, not $>$. Suppose $v(p,m')=v(p,m)$, which means $x'\sim x$, for $x'\in x(p,m'), x\in x(p,m)$. Pick a neighborhood of x s.t. $N\subset B(p,m')$. LNS implies $\exists y\in N$ s.t. $y\succ x$. This means y is affordable at (p,m') but strictly preferred to x', which is a contradiction to x' being optimal.

[3]This point also provides justification for assuming utility functions to be quasiconcave instead of to be concave: quasiconcavity (but not concavity) is an ordinal property, which is preserved by increasing transformations.

Recitation 3: Demand Theory[1]

Xincheng Qiu (qiux@sas.upenn.edu)

In this note, we assume every object is nonempty, i.e., UMP and EMP both have a solution. The existence of solutions is often guaranteed by the Weierstrass extreme value theorem, saying that a continuous function from a nonempty compact space to a subset of the real numbers attains a maximum and a minimum.

Another typical result is that v, e is continuous and x, h is uhc on their domain, for example, when we assume that u is a continuous utility function representing a locally nonsatiated preference. This result follows from Berge's Maximum Theorem[2].

Properties of Marshallian Demand

1. Homogeneous of degree zero: $x(\alpha p, \alpha m) = x(p, m)$, $\forall \alpha > 0$

   Proof. This property does not require any assumption on u. Multiplying all prices and income by the same positive constant does not change the constraint set in the UMP, and even more obviously does not change its objective function.

$$\begin{aligned} x(p,m) &:= \arg\max_{x \in \mathbb{R}^L_+} u(x) \text{ s.t. } p \cdot x \leq m \\ &= \arg\max_{x \in \mathbb{R}^L_+} u(x) \text{ s.t. } (\alpha p) \cdot x \leq (\alpha m) =: x(\alpha p, \alpha m) \end{aligned}$$

   □

2. SM ⇒ M ⇒ LNS ⇒ Walras' Law: $p \cdot x = m$, $\forall x \in x(p, m)$

   Proof. We prove here if LNS then Walras' Law holds. Fix $(p, m) \in \mathbb{R}^{L+1}_{++}$, and let $x \in x(p, m)$. Since x solves the UMP, it is feasible, i.e., $p \cdot x \leq m$. Suppose $p \cdot x < m$.

[1] This note is summarized from the lecture slides by Professor Steven Matthews, but any error is my own. All comments and corrections are welcome.

[2] Berge's Maximum Theorem: Let $f : S \times \Theta \mapsto \mathbb{R}$ be continuous, and $D : \Theta \rightrightarrows S$ is a compact valued and continuous correspondence. Define

$$D^*(\theta) := \arg\max \{f(x,\theta) \,|x \in D(\theta)\}, \quad f^*(\theta) = \max\{f(x,\theta)\,|x \in D(\theta)\}$$

1. $f^*(\theta)$ is continuous.
2. $D^*(\theta)$ is upper hemicontinuous and compact valued.

Then a neighborhood N of x exists such that $p \cdot x' < m$ for all $x' \in N$. LNS implies $x' \in N \cap \mathbb{R}^L_+$ exists such that $u(x') > u(x)$. Thus, x' is affordable at (p, m) but strictly preferred to x, which is a contradiction. □

3. If u is quasiconcave, then $x(p, m)$ is a convex set. If u is strictly quasiconcave, then $x(p, m)$ is a singleton.

   Proof. Suppose u is quasiconcave. Let $x_1, x_2 \in x(p, m)$. By definition, $u(x_1) \geq z$ and $u(x_2) \geq z$ for any $z \in B(p, m) := \left\{z \in \mathbb{R}^L_+ : p \cdot z \leq m\right\}$. Take any $\lambda \in (0, 1)$. By convexity of $B(p, m)$, $(\lambda x_1 + (1 - \lambda) x_2) \in B(p, m)$. By quasiconcavity of u, for any $z \in B(p, m)$,

   $$u(\lambda x_1 + (1 - \lambda) x_2) \geq \min\{u(x_1), u(x_2)\} \geq z$$

   which shows that $(\lambda x_1 + (1 - \lambda) x_2) \in x(p, m)$.

   Suppose u is strictly quasiconcave. Assume $x(p, m)$ contains two points $x_1 \neq x_2$. Let $\lambda \in (0, 1)$ and $\hat{x} = \lambda x_1 + (1 - \lambda) x_2 \in B(p, m)$ by convexity of $B(p, m)$. However, strict quasiconcavity of u implies

   $$u(\lambda x_1 + (1 - \lambda) x_2) > \min\{u(x_1), u(x_2)\}$$

   which is a contradiction to x_1, x_2 both being optimal. □

4. If u is continuously differentiable, an optimal bundle $x^* \in x(p, m)$ can be characterized by the Kuhn-Tucker first-order conditions: $\exists \lambda^* \geq 0$ s.t. $\forall l$

   $$u_l(x^*) - \lambda^* p_l \leq 0, \quad [u_l(x^*) - \lambda^* p_l] x^*_l = 0$$

   $$m - p \cdot x^* \geq 0, \quad [m - p \cdot x^*] \lambda^* = 0$$

From here on, we assume u is twice continuously differentiable, strictly quasiconcave, $\nabla u(x) > 0 \, \forall x \in \mathbb{R}^L_+$ and $x(p, m)$ is continuously differentiable.

5. The Slutsky matrix $S(p, m)$ is negative semidefinite and symmetric, where

   $$S_{ij}(p, m) = \frac{\partial x_i(p, m)}{\partial p_j} + x_j(p, m) \frac{\partial x_i(p, m)}{\partial m}$$

   Proof. Slutsky decomposition implies $S_{ij} = \frac{\partial h_i}{\partial p_j}$. Shepard's lemma implies $\frac{\partial h_i}{\partial p_j} = \frac{\partial^2 e}{\partial p_i p_j}$. Envelop Theorem implies $\left[\frac{\partial^2 e}{\partial p_i p_j}\right]$ is negative semidefinite and Young's theorem implies $\left[\frac{\partial^2 e}{\partial p_i p_j}\right]$ is symmetric. □

6. $p^T S = (0, \ldots, 0)$, $Sp = (0, \ldots, 0)^T$, $p^T S p = 0$.

   Proof. Differentiate $ph(p, U) = e(p, U)$ w.r.t. p_j: $h_j + \sum_i p_i \frac{\partial h_i}{\partial p_j} = \frac{\partial e}{\partial p_j} = h_j$. So $\sum_i p_i \frac{\partial h_i}{\partial p_j} = 0$.

   h is homogeneous of degree zero in p, so $\sum_j \frac{\partial h_i}{\partial p_j} p_j = 0$ by Euler's formula. $\square$

Remark 1. C^1 Marshallian demand functions that arise from monotonic C^2 utility functions on $\mathbb{R}_+^L$ must be: homogeneity of degree 0, Walras' law, symmetric and negative semidefinite substitution matrix (three big testable properties). The Integrability Theorem discusses the reverse question: if a C^1 function satisfies these properties, it must be a demand function arising from a continuous, strictly quasiconcave, and strictly increasing function u. In fact, homogeneity is redundant, since it follows from the other two conditions.

Exercise 1. In a three-good world, suppose a consumer's demands for goods 1 and 2 are given by

$$x_1(p, m) = \frac{p_2}{p_3},\ x_2(p, m) = \frac{p_1}{p_3}$$

Can these demands arise from the maximization of a continuous utility function representing locally nonsatiated strictly convex preferences?

Solution. Suppose they do. Walras' Law then implies

$$x_3(p, m) = \frac{m}{p_3} - \frac{2p_1 p_2}{p_3^2}$$

Thus the $(3, 3)$ element of the Slutsky matrix is

$$s_{33}(p, m) = \frac{\partial x_3(p, m)}{\partial p_3} + x_3(p, m) \frac{\partial x_3(p, m)}{\partial m} = \frac{2p_1 p_2}{p_3^3} > 0$$

So the Slutsky matrix is not negative semidefinite, which is a contradiction.

Properties of Indirect Utility Function

1. $v(p, m)$ is homogenous of degree 0 in (p, m) (since x is so).

2. $v(p, m)$ is nonincreasing in p (since a decrease in p enlarges the budget set).

3. If Walras' law, then $v(p, m)$ strictly increases in m.

Proof. Consider $m' > m$, $B(p, m) \subset B(p, m')$ implies $v(p, m') \geq v(p, m)$. Suppose $v(p, m') = v(p, m)$, which means $x' \sim x$, for $x' \in x(p, m'), x \in x(p, m)$. Walras' law implies $p \cdot x = m < m'$, then x is also optimal under (p, m'), which in turn contradicts with Walras' law. $\square$

4. $v(p, m)$ is quasiconvex in (p, m).

   Proof. Pick any $(p, m), (p', m')$ and $t \in (0, 1)$. Let $(p'', m'') = t(p, m) + (1 - t)(p', m')$. We must prove $v(p'', m'') \leq \max\{v(p, m), v(p', m')\}$. This is because $B(p'', m'') \subset B(p, m) \cup B(p', m')$. To see, pick $x \in B(p'', m'')$, then $t(p \cdot x) + (1 - t)(p' \cdot x) \leq tm + (1 - t)m'$. Hence it must be either $p \cdot x \leq m$ or $p' \cdot x \leq m'$ or both. Therefore

$$\begin{aligned} v(p'', m'') &= \max_{x \in \mathbb{R}_+^L} u(x) \text{ s.t. } x \in B(p'', m'') \\ &\leq \max_{x \in \mathbb{R}_+^L} u(x) \text{ s.t. } x \in B(p, m) \cup B(p', m') \\ &= \max\{v(p, m), v(p', m')\} \end{aligned}$$

   $\square$

Remark 2. Note that this property says $v(p, m)$ is quasiconvex, not quasiconcave. This property does not require $u(\cdot)$ to be quasiconcave. It is saying that consumers prefer one of any two extreme budget sets to any average of them.

Properties of Hicksian Demand

1. $h(p, U)$ is homogeneous of degree 0 in p: $h(\alpha p, U) = h(p, U), \forall \alpha > 0$.

   Proof. The optimal bundle for $\min_{x \in \mathbb{R}_+^L} p \cdot x$ s.t. $u(x) \geq U$ is the same as that for $\min_{x \in \mathbb{R}_+^L} (\alpha p) \cdot x$ s.t. $u(x) \geq U$. $\square$

2. If u is continuous, we have an analog of Walras' law (no excess utility): $u(h) = U, \forall h \in h(p, U)$.

   Proof. Suppose $\exists h \in h(p, U)$ s.t. $u(h) > U$. Since $u(\cdot)$ is continuous, $\exists t \in (0, 1)$ s.t. $u(th) > U$, but $p \cdot (th) < p \cdot h \leq e(p, U)$, contradicting the definition of e. $\square$

3. If u is quasiconcave, then $h(p, U)$ is a convex set. If u is strictly quasiconcave, then $h(p, U)$ is a singleton. (The proof parallels that for x.)

4. Law of Demand: for any p, p', $h \in h(p,U)$ and $h' \in h(p',U)$, we have $(p-p') \cdot (h-h') \leq 0$.

   Proof. Since h' is feasible for $\text{EMP}(p,U)$, $p \cdot h \leq p \cdot h'$. Since h is feasible for $\text{EMP}(p',U)$, $p' \cdot h' \leq p' \cdot h$. Add these two inequalities to obtain $(p-p') \cdot (h-h') \leq 0$. If h is a C^1 function, we can use the Hessian of e being negative semidefinite. $\square$

5. A C^1 Hicksian demand function satisfies the Law of Reciprocity: $\frac{\partial h_l(p,U)}{\partial p_k} = \frac{\partial h_k(p,U)}{\partial p_l}$.

   Proof. Follows from Shepard's lemma and Young's theorem on the symmetry of cross partials of C^2 functions. $\square$

Remark 3. The law of demand and the law of reciprocity of the Hicksian demand corresponds to the Slutsky matrix being negative semidefinite and symmetric.

Properties of Expenditure Function

1. $e(p,U)$ is homogeneous of degree 1 in p (because h is homogeneous of degree 0 in p).
2. $e(p,U)$ is nondecreasing in p.

   Proof. Let $p' > p$. Pick $h' \in h(p',U)$, so $u(h') \geq U$. Then $e(p',U) = p' \cdot h' \geq p \cdot h' \geq e(p,U)$. $\square$

3. If u is continuous, then $e(p,U)$ is strictly increasing U.

   Proof. Suppose to the contrary that for some $U' > U$, $e(p,U') \leq e(p,U)$. Then for some $h' \in h(p,U')$, we have $u(h') \geq U' > U$, but $p \cdot h' \leq e(p,U)$. Since $u(\cdot)$ is continuous, $\exists t \in (0,1)$ s.t. $u(th') > U$, but $p \cdot (th') < p \cdot h' \leq e(p,U)$, contradicting the definition of expenditure function. $\square$

4. $e(p,U)$ is concave in p.

   Proof. Let $\lambda \in (0,1)$ and denote $p = \lambda p' + (1-\lambda) p''$. Pick $h \in h(p,U)$. Since $u(h) \geq U$, we have $p' \cdot h \geq e(p',U)$ and $p'' \cdot h \geq e(p'',U)$. Hence

$$\begin{aligned} e(\lambda p' + (1-\lambda) p'', U) &= (\lambda p' + (1-\lambda) p'') \cdot h \\ &\geq \lambda e(p',U) + (1-\lambda) e(p'',U) \end{aligned}$$

   $\square$

Remark 4. The proof of concavity exploits the following fact: suppose initially the price vector is p and h is an optimal bundle. If prices change to p' while the consumer does not change her consumption, it costs her $p' \cdot h$. But in fact she can adjust her consumption, so her minimized expenditure will be no greater than this amount: $e(p', U) \leq p' \cdot h$. The concavity of e conveys a similar story as the quasiconvexity of v: if the price vector is one extreme one day and another extreme the other day, you will spend less than if it was the average price every day.

Remark 5. Cautions:

1. Without assumptions that guarantee uniqueness, the Marshallian demand and Hicksian demand are typically correspondences, not necessarily functions.

2. Without assumptions on differentiability, arguments based on FOC are incorrect.

3. Making assumptions is an art in research, usually with tradeoffs between being tractable and general, between being simple and rich. If you cannot solve a problem without making an extra assumption, go ahead and make it. And point out, very clearly, which additional assumption you are making.

Exercise 2. (MWG 3.G.16) Consider the expenditure function

$$e(p, U) = \exp\left\{\sum_l (\alpha_l \ln p_l) + \left(\prod_l p_l^{\beta_l}\right) U\right\}$$

What can you say about $\alpha_1, \ldots, \alpha_L, \beta_1, \ldots, \beta_L$?

Solution.

1. Since $e(p, U)$ is homogeneous of degree 1 in p: $\forall (p, U), \forall t > 0, e(tp, U) = te(p, U)$

$$\exp\left\{\ln t \sum_l \alpha_l + \sum_l (\alpha_l \ln p_l) + \left(t^{\sum_l \beta_l}\right)\left(\prod_l p_l^{\beta_l}\right) U\right\} = t \exp\left\{\sum_l (\alpha_l \ln p_l) + \left(\prod_l p_l^{\beta_l}\right) U\right\}$$

In particular, take $p_l = 1, \forall l$ and $U = 1$: $\exp\left\{\ln t \sum_l \alpha_l + t^{\sum_l \beta_l}\right\} = t \exp\{1\}$. This requires $\sum_l \alpha_l = 1$, $\sum_l \beta_l = 0$. Then above equation becomes

$$\exp\left\{\ln t + \sum_l (\alpha_l \ln p_l) + \left(\prod_l p_l^{\beta_l}\right) U\right\} = t \exp\left\{\sum_l (\alpha_l \ln p_l) + \left(\prod_l p_l^{\beta_l}\right) U\right\}$$

which indeed holds at any (p, U) and for any $t > 0$.

2. Since $e(p,U)$ is nondecreasing in p: $\forall (p,U), \forall k$

$$\frac{\partial e(p,U)}{\partial p_k} = \left(\alpha_k + \beta_k \left(\prod_l p_l^{\beta_l}\right) U\right) \frac{e(p,U)}{p_k} \geq 0$$

   In particular, take p sufficiently small, then we will have it must be $\alpha_k \geq 0, \forall k$. Take p sufficiently big, then we have it must be $\beta_k \geq 0, \forall k$. This together with $\sum_l \beta_l = 0$ implies $\beta_k = 0, \forall k$.

3. Thus, $\sum_l \alpha_l = 1$, $\alpha_l \geq 0, \forall l$ and $\beta_l = 0, \forall l$. Now the expenditure function can be simplified as

$$\begin{aligned} e(p,U) &= \exp\left\{\sum_l (\alpha_l \ln p_l) + U\right\} \\ &= \exp\{U\} \prod_l p_l^{\alpha_l} \end{aligned}$$

   which is indeed strictly increasing in U and concave in p.

Recitation 4: Duality

Xincheng Qiu (qiux@sas.upenn.edu)

Assume that a consumer has a continuous and local nonsatiated utility function $u : \mathbb{R}^L_+ \to \mathbb{R}$. Let v, e, x and h be the consumer's indirect utility function, expenditure function, Marshallian demand correspondence and Hicksian demand correspondence, respectively. Prove the four duality identities:

1. $v(p, e(p, U)) \equiv U$

   Proof. Fix any $p \in \mathbb{R}^L_{++}$ and $U \in \big[u(0), \sup_z u(z)\big)$. If $e(p, U) = 0$, then $U = u(0)$ and $x(p, e(p, U)) = 0$. So, $v(p, e(p, U)) = u\big(x(p, e(p, U))\big) = u(0) = U$. Now assume $e(p, U) > 0$. By definition of v, e, and Hicksian demand h, we have

   $$v(p, e(p, U)) = v(p, p \cdot h') \geq u(h') \geq U \quad \text{for some } h' \in h(p, U).$$

   Suppose $v(p, e(p, U)) > U$, i.e., for some $x' \in x(p, e(p, U))$, $u(x') > U$. Since u is local nonsatiated, $p \cdot x' = e(p, U)$ (Walras' law). But, by definition of $e(p, U)$, we know that $u(z) < U$ for all z satisfying $p \cdot z < e(p, U)$. Hence, u is not continuous at x', which is a contradiction. Therefore, $v(p, e(p, U)) = U$. $\square$

2. $x(p, e(p, U)) \equiv h(p, U)$

   Proof. Take any $x' \in x(p, e(p, U))$, by definition of x, we have $p \cdot x' \leq e(p, U)$. For any $h' \in h(p, U)$, by definition of h and e, we have $p \cdot h' = e(p, U)$. Then we have

   $$u(x') \geq u(h') \geq U$$

   The first inequality comes from the fact that $p \cdot h' = e(p, U)$ and $x' \in x(p, e(p, U))$, the second inequality is because h' is in the feasible set of EMP. Since $u(x') \geq U$ and $p \cdot x' \leq e(p, U)$, we have $x' \in h(p, U)$ and thus $x(p, e(p, U)) \subseteq h(p, U)$.

   Take any $h' \in h(p, U)$, by definition of h and result in part (1), we have $u(h') \geq U = v(p, e(p, U))$ and also $p \cdot h' = e(p, U)$ by definition of h and e. Thus h' is a solution of UMP, i.e., $h(p, U) \subseteq x(p, e(p, U))$. Combining results above we have $x(p, e(p, U)) = h(p, U)$. $\square$

3. $e(p, v(p, m)) \equiv m$

Proof. Fix any $p \in \mathbb{R}^L_{++}$ and $m \geq 0$. By definition of v, e, and x, we obtain

$$e(p, v(p,m)) = e\big(p, u(x')\big) \leq p \cdot x' \leq m \quad \text{for some } x' \in x(p,m)\,.$$

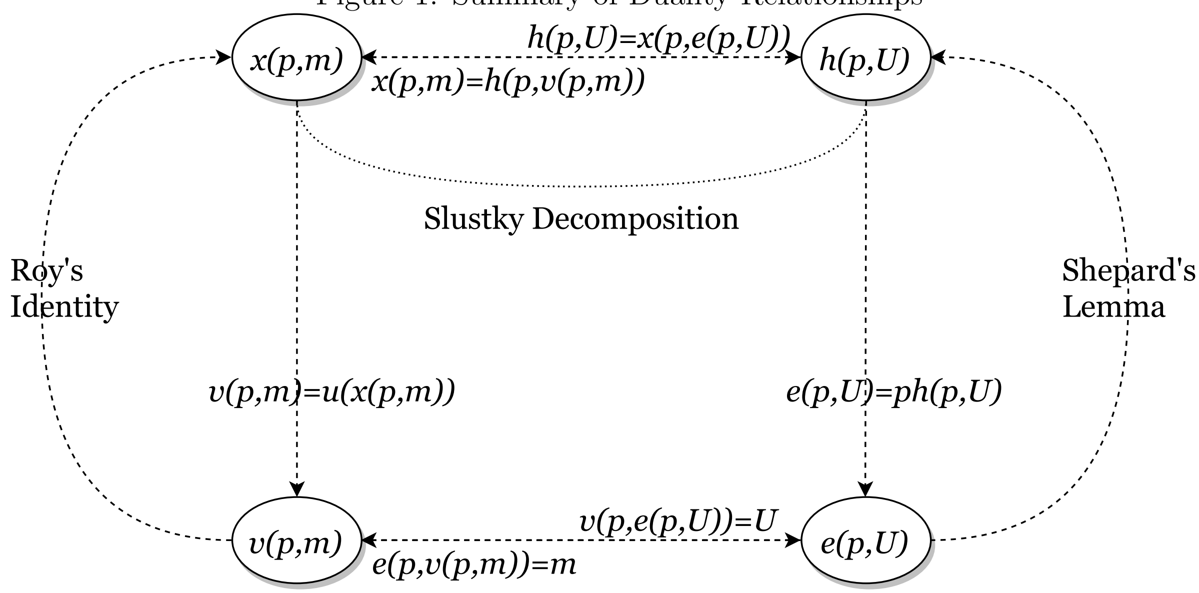
Now suppose $e(p, v(p,m)) < m$. Then, for some $h' \in h(p, v(p,m))$, $p \cdot h' < m$ and $u(h') \geq v(p,m)$. But, these imply that $h' \in x(p,m)$ and, in turn, that Walras' law does not hold. Contradiction. Therefore, $e(p, v(p,m)) = m$. □

4. $h(p, v(p,m)) \equiv x(p,m)$

   Proof. Take any $h' \in h(p, v(p,m))$, by definition of h, e and we have $p \cdot h' = e(p, v(p,m)) \leq m$. Furthermore, by definition of h, $u(h') \geq v(p,m)$. Thus by definition of v and x, $h' \in x(p,m)$, i.e., $h(p, v(p,m)) \subseteq x(p,m)$.

   Take any $x' \in x(p,m)$, by definition of x and v, we have $u(x') = v(p,m)$. Furthermore, by definition of x and result in part (3), we have $p \cdot x' \leq m = e(p, v(p,m))$, thus by definition of $e$and h, $x' \in h(p, v(p,m))$, i.e., $x(p,m) \subseteq h(p, v(p,m))$. Combining results above we have $h(p, v(p,m)) \equiv x(p,m)$. □

Figure 1: Summary of Duality Relationships



Note: This figure is taken from MWG Figure 3.G.3 with minor modifications.

Remark 1. Slutsky decomposition is derived by differentiating the duality identity $h_l\,(p, U) \equiv x_l\,(p, e\,(p, U))$. Make sure you get the arguments right. It has nice economic interpretation: the total effect can be decomposed as the substitution effect and the income effect.

Example 1. (Fall 2014) A strictly increasing utility, quasiconcave function $u : \mathbb{R}^2_+ \to \mathbb{R}_+$ gives rise to the expenditure function

$$e(p,U) = \left(p_1^a + p_2^b + 2p_1^c p_2^c\right) U^2$$

where $a, b, c > 0$. Derive the five functions in demand theory: $x(p,m)$, $v(p,m)$, $h(p,U)$, $e(p,U)$ and $u(x)$.

Solution.

1. $e(p,U)$ is homogeneous of degree 1 in p: $te(p,U) = e(tp,U)$ $\forall t, p_1, p_2 > 0 \Rightarrow a = b = 1$ and $c = \frac{1}{2}$, so
$$e(p,U) = \left(\sqrt{p_1} + \sqrt{p_2}\right)^2 U^2 \tag{1}$$

2. By Shepard's lemma: for $i = 1, 2$
$$h_i(p,U) = \frac{\partial e(p,U)}{\partial p_i} = \frac{\sqrt{p_1} + \sqrt{p_2}}{\sqrt{p_i}} U^2 \tag{2}$$

3. Duality identity $e(p, v(p,m)) = m$ says $v(p,.)$ is the inverse function of $e(p,.)$: $\left(\sqrt{p_1} + \sqrt{p_2}\right)^2 v(p,m)^2 = m$. Hence:
$$v(p,m) = \frac{\sqrt{m}}{\sqrt{p_1} + \sqrt{p_2}} \tag{3}$$

4. By Roy's identity: for $i = 1, 2$
$$x_i(p,m) = -\frac{v_{p_i}(p,m)}{v_m(p,m)} = \frac{m}{\sqrt{p_1} + \sqrt{p_2}} \frac{1}{\sqrt{p_i}} \tag{4}$$

5. For every $x \in \mathbb{R}^2_{++}$, as long as $u(\cdot)$ is quasiconcave and strict increasing, the Supporting Hyperplane Theorem implies that a $\exists p \gg 0$ such that $x = h(p, u(x))$.[1] Letting $r =$

[1] The Supporting Hyperplane Theorem states that: for $D \subseteq \mathbb{R}^n$ convex and $x \in \partial D$, D has a supporting hyperplane at x, that is, $\exists H(p,a), p \neq 0 \in \mathbb{R}^n$ such that $\forall z \in D$, $p \cdot z \geq a$ and $p \cdot x = a$. For any given $x \in \mathbb{R}^2_{++}$, let $D = \{z : u(z) \geq u(x)\}$. D is convex if $u(\cdot)$ is quasiconcave. Since $x \in \partial D$, the Supporting Hyperplane Theorem says $\exists p \in \mathbb{R}^2, p \neq 0$ such that $\forall z \in D$, $p \cdot z \geq p \cdot x$, and hence $x = h(p, u(x))$. We can also prove that $p \gg 0$ if $u(\cdot)$ is strict increasing. Suppose $p_i < 0$ for some i, we could take $\hat{z}$ s.t. $\hat{z}_i > x_i$ but $\hat{z}_j = x_j$ for $j \neq i$. Then $u(\hat{z}) \geq u(x)$ but $p \cdot \hat{z} < p \cdot x$ and we get a contradiction. So $p > 0$. Suppose $p_i = 0$ for some i, we could take $\tilde{z}$ s.t. $\tilde{z}_i$ sufficiently big and $\tilde{z}_j$ smaller than x_j so that $u(\tilde{z}) \geq u(x)$ but $p \cdot \tilde{z} < p \cdot x$ to get a contradiction. So $p \gg 0$.

$\sqrt{\frac{p_2}{p_1}}$ gives us $\begin{cases} x_1 = \frac{\sqrt{p_1}+\sqrt{p_2}}{\sqrt{p_1}}U^2 = (1+r)\,u(x)^2 \\ x_2 = \frac{\sqrt{p_1}+\sqrt{p_2}}{\sqrt{p_2}}U^2 = (1+r^{-1})\,u(x)^2 \end{cases}$. Reducing these to one equation by eliminating r, and then solving for $u(x)$ yields

$$u(x) = \sqrt{\frac{x_1 x_2}{x_1 + x_2}} \tag{5}$$

Example 2. (June 2016) Consider the function $e : \mathbb{R}^2_{++} \times \mathbb{R}_+ \to \mathbb{R}_+$ defined by

$$e(p, U) = U \min\left\{p_1, \frac{p_1 + p_2}{3}, p_2\right\}$$

Say as much as you can about $h(p, U)$, $v(p, m)$, $x(p, m)$, and $u(x)$.

Solution.

1. At any (p, U) where e is differentiable with respect to p, the Hicksian demand correspondence is a singleton, given by the gradient of e with respect to p. Letting $r = \frac{p_1}{p_2}$, $e(p, U)$ can be written as

$$e(p, U) = \begin{cases} p_1 U & r \leq \frac{1}{2} \\ \frac{p_1+p_2}{3}U & \frac{1}{2} \leq r \leq 2 \\ p_2 U & 2 \leq r \end{cases}$$

e is differentiable with respect to p at (p, U) iff $r \notin \left\{\frac{1}{2}, 2\right\}$. Hence,

$$h(p, U) = \begin{cases} \{(U, 0)\} & r < \frac{1}{2} \\ \left\{\left(\frac{U}{3}, \frac{U}{3}\right)\right\} & \frac{1}{2} < r < 2 \\ \{(0, U)\} & 2 < r \end{cases}$$

Because the Hicksian demand correspondence is uhc[2], we also know that $h(p, U)$ con-

[2]Definition and sequential characterization of uhc and lhc:

Definition 1. *Hemicontinuity*

A correspondence $\Phi : \Theta \rightrightarrows S$ is said to be *upper hemicontinuous* at $\theta \in \Theta$ if for all open sets V such that $\Phi(\theta) \subseteq V$, there exists an open set $U \subseteq \Theta$ containing θ such that $\forall \theta' \in U \bigcap \Theta$, $\Phi(\theta') \subseteq V$.

A correspondence $\Phi : \Theta \rightrightarrows S$ is said to be *lower hemicontinuous* at $\theta \in \Theta$ if for all open sets V such that $\Phi(\theta) \bigcap V \neq \varnothing$, there exists an open set $U \subseteq \Theta$ containing θ such that $\forall \theta' \in U \bigcap \Theta$, $\Phi(\theta') \bigcap V \neq \varnothing$.

Theorem 1. *Sequential Characterization of Hemicontinuity*

Let $\Phi : \Theta \rightrightarrows S$ be a compact-valued correspondence. Then Φ is upper hemicontinuous at $\theta \in \Theta$ if and only if $\forall (\theta_n) \subseteq \Theta$ s.t. $\theta_n \to \theta$, and $\forall s_n \in \Phi(\theta_n)$, there exists a subsequence $(s_{n_k}) \subseteq (s_n)$ s.t. $s_{n_k} \to s \in \Phi(\theta)$.

tains $(U,0)$ and $\left(\frac{U}{3},\frac{U}{3}\right)$ if $r=\frac{1}{2}$; $h(p,U)$ contains $\left(\frac{U}{3},\frac{U}{3}\right)$ and $(0,U)$ if $r=2$.

2. $v(p,.)$ is the inverse function of $e(p,.)$:

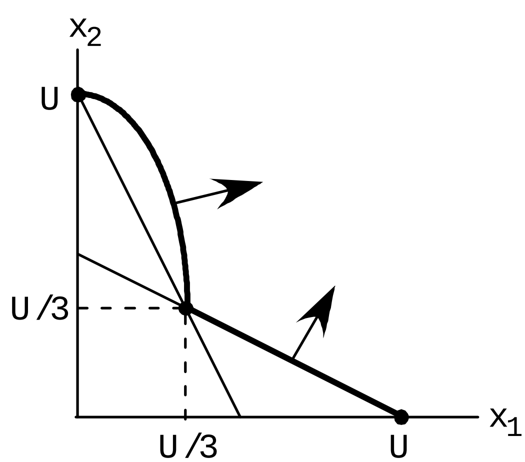
$$v(p,m)=\frac{m}{\min\left\{p_1,\frac{p_1+p_2}{3},p_2\right\}}$$

3. Similar to part (1), Marshallian demand is given by Roy's identity at (p,m) where v is differentiable, hence

$$x(p,m)=\begin{cases}\left\{\left(\frac{m}{p_1},0\right)\right\} & r<\frac{1}{2}\\ \left\{\left(\frac{m}{p_1+p_2},\frac{m}{p_1+p_2}\right)\right\} & \frac{1}{2}<r<2\\ \left\{\left(0,\frac{m}{p_2}\right)\right\} & 2<r\end{cases}$$

Since the Marshallian demand correspondence is uhc, we also know that $x(p,m)$ contains $\left(\frac{m}{p_1},0\right)$ and $\left(\frac{m}{p_1+p_2},\frac{m}{p_1+p_2}\right)$ if $r=\frac{1}{2}$; $x(p,m)$ contains $\left(\frac{m}{p_1+p_2},\frac{m}{p_1+p_2}\right)$ and $\left(0,\frac{m}{p_2}\right)$ if $r=2$.

4. For a given $U>0$, we know from part (1) that the U indifference curve must contain these three points $(U,0),\left(\frac{U}{3},\frac{U}{3}\right),(0,U)$. Since $(0,U)$ and $\left(\frac{U}{3},\frac{U}{3}\right)$ minimize the cost of getting utility U when $r=2$, the U indifference curve must be weakly above the budget line given by $2x_1+x_2=U$. By the same argument, the U indifference curve must also be weakly above the budget line given by $x_1+2x_2=U$.



Several possible examples:

Let $\Phi:\Theta\rightrightarrows S$ be a compact-valued correspondence. Then Φ is lower hemicontinuous at $\theta\in\Theta$ if and only if $\forall(\theta_n)\subseteq\Theta$ s.t. $\theta_n\to\theta$, and $\forall s\in\Phi(\theta)$, there exists $s_n\in\Phi(\theta_n)$ s.t. $s_n\to s$.

- $u(x) = \min\{x_1 + 2x_2, 2x_1 + x_2\}$;
- $\hat{u}(x) = u(x)$ for $x_1 \geq x_2$ and $\hat{u}(x) = u(x) - x_1\left(1 - \frac{x_1}{x_2}\right)$ for $x_1 < x_2$;
- $\tilde{u}(x) = \max\{x_1, x_2\}$ if $x_1 \neq x_2$, and $\tilde{u}(x) = 3x_1$ if $x_1 = x_2$.

Remark 2. If we "know" or "assume" $\underline{u}$, we can derive the other functions by first solving a system of equations to find x. If we "know" or "assume" v, we can easily find the other objects. $e(p, \cdot)$ is just the inverse function of $v(p, \cdot)$. h can be found from e by Shepard's lemma. x can be found from v by Roy's identity. Hence, a modern empirical economist is more likely to start with a functional form for v or e, rather than for u.

Example 3. Translog demand system (Christensen, Jorgenson, and Lau, 1975 AER).

The translog indirect utility function for a consumer with K commodities is

$$-\ln V(p, m) = \theta_0 + \sum_{k=1}^{K} \theta_k \ln\left(\frac{p_k}{m}\right) + \sum_{k=1}^{K}\sum_{j=1}^{K} \gamma_{kj} \ln\left(\frac{p_k}{m}\right) \ln\left(\frac{p_j}{m}\right).$$

Applying Roy's identity to this logarithmic function gives the budget share of the ith good,

$$s_i(p, m) = -\frac{\partial \ln V / \partial \ln p_i}{\partial \ln V / \partial \ln m} = \frac{\theta_i + \sum_{j=1}^{K} \gamma_{kj} \ln\left(\frac{p_j}{m}\right)}{\bar{\theta} + \sum_{j=1}^{K} \bar{\gamma}_j \ln\left(\frac{p_j}{m}\right)},$$

where $\bar{\theta} = \sum_{k=1}^{K} \theta_k$ and $\bar{\gamma}_j = \sum_{k=1}^{K} \gamma_{kj}$.

Example 4. Almost Ideal Demand System (Deaton and Muellbauer, 1980 AER).

$$\ln e(p, U) = (1 - U)\ln a(p) + U \ln b(p)$$

where

$$\ln a(p) = \alpha_0 + \sum_{k=1}^{K} \alpha_k \ln p_k + \frac{1}{2}\sum_{k=1}^{K}\sum_{j=1}^{K} \gamma_{kj}^{*} \ln p_k \ln p_j$$

$$\ln b(p) = \ln a(p) + \beta_0 \prod_{k=1}^{K} p_k^{\beta_k}$$

Using Shepard Lemma we can get shares of expenditure

$$s_i(p, U) = \frac{p_i h_i(p, U)}{e(p, U)} = \frac{\partial \ln e(p, U)}{\partial \ln p_i} = \alpha_i + \sum_{j=1}^{K} \gamma_{ij} \ln p_j + \beta_i \ln\left(\frac{e(p, U)}{a(p)}\right)$$

where $\gamma_{ij} = \frac{1}{2}\left(\gamma_{ij}^{*} + \gamma_{ji}^{*}\right)$ and $a(p)$ could be interpreted as a price index. See also Exercise 2 in Recitation 3.

Recitation 5: Dollar Measures of Welfare Changes[1]

Xincheng Qiu (qiux@sas.upenn.edu)

Remark 1. To evaluate the effects of some changes on consumers' well-being, it does not make too much sense to simply compare utiles. Differences in utiles, or percentage change in utiles, are hart to interpret, because they could change dramatically when doing a strictly increasing transformation on the utility function. We provide some monetary measurements of welfare changes expressed in dollars (euros, yuans).

Compensating Variation

1. The max amount one would be willing to "buy" the change: $v(p^0, m) = v(p^1, m - CV)$

2. Hence $CV = m - e(p^1, u^0) = e(p^0, u^0) - e(p^1, u^0) = e(p^1, u^1) - e(p^1, u^0)$

3. If only the price of good 1 changes, integral representation:

$$\begin{aligned} CV &= e\left(p_1^0, \bar{p}_{-1}, u^0\right) - e\left(p_1^1, \bar{p}_{-1}, u^0\right) \\ &= \int_{p_1^1}^{p_1^0} \frac{\partial e}{\partial p_1}\left(p_1, \bar{p}_{-1}, u^0\right) dp_1 = \int_{p_1^1}^{p_1^0} h_1\left(p_1, \bar{p}_{-1}, u^0\right) dp_1 \end{aligned}$$

4. This argument extends to multiple price changes, using a line integral:

$$CV\left(p^0, p^1, m\right) = \int_0^1 h\left(p(t), u^0\right) \cdot p'(t)\, dt$$

   where p is any differentiable path with $p(0) = p^1$ and $p(1) = p^0$.

Equivalent Variation

1. The min amount one would be willing to "sell" the change: $v(p^0, m + EV) = v(p^1, m)$

2. Hence $EV = e(p^0, u^1) - m = e(p^0, u^1) - e(p^1, u^1) = e(p^0, u^1) - e(p^0, u^0)$

3. If only the price of good 1 changes, integral representation:

[1] This note is summarized from the lecture slides by Professor Steven Matthews, but any error is my own.

$$\begin{aligned} EV &= e\left(p_1^0, \bar{p}_{-1}, u^1\right) - e\left(p_1^1, \bar{p}_{-1}, u^1\right) \\ &= \int_{p_1^1}^{p_1^0} \frac{\partial e}{\partial p_1}\left(p_1, \bar{p}_{-1}, u^1\right) dp_1 = \int_{p_1^1}^{p_1^0} h_1\left(p_1, \bar{p}_{-1}, u^1\right) dp_1 \end{aligned}$$

4. This argument extends to multiple price changes, using a line integral:

$$EV\left(p^0, p^1, m\right) = \int_0^1 h\left(p\left(t\right), u^1\right) \cdot p'\left(t\right) dt$$

Change in Consumer's Surplus

$$\Delta CS = \int_0^1 x\left(p\left(t\right), m\right) \cdot p'\left(t\right) dt = \int_{p_1^1}^{p_1^0} x_1\left(p_1, \bar{p}_{-1}, m\right) dp_1$$

Remark 2. CV uses changing rulers while EV uses a fixed ruler. If we have several alternative projects, CV does not work: it is possible for $u^i < u^k$ even though $CV^i > CV^k$. This is because CV measures utility differences beween a project and the status quo using the price vector of this project, which varies with the project:

$$CV^i = e\left(p^i, u^1\right) - e\left(p^i, u^0\right).$$

$CV^i - CV^k = e\left(p^k, u^0\right) - e\left(p^i, u^0\right)$ does not depend on u^i or u^k. But EV does work: $EV^i > EV^k$ iff $u^i > u^k$. EV measures utility differences beween a project and the status quo using the status quo prices, which does not vary with the project:

$$EV^i = e\left(p^0, u^i\right) - e\left(p^0, u^0\right).$$

$EV^i - EV^k = e\left(p^0, u^i\right) - e\left(p^0, u^k\right)$ is positive iff $u^i > u^k$ as $e\left(p, U\right)$ strictly increases in U.

Remark 3. Since $e\left(p, \cdot\right)$ is strictly increasing, the money metric indirect utility Function $\hat{v}\left(p, m\right) \coloneqq e\left(p^0, v\left(p, m\right)\right)$ represents the same preferences over $\left(p, m\right)$ pairs as does $v\left(p, m\right)$. $EV\left(p^0, p, m\right) = e\left(p^0, v\left(p, m\right)\right) - m = \hat{v}\left(p, m\right) - m$ thus represents the consumer's preferences on prices p.

Example 1. In a two-good world, a consumer has the following Hicksian demand that arises from a strictly increasing utility function:

$$h_i\left(p, U\right) = \left(\frac{p_j}{p_i}\right)^{\frac{1}{2}} U, \ i, j = 1, 2; j \neq i.$$

We can derive the utility function by eliminating the relative price: $u(x_1, x_2) = \sqrt{x_1 x_2}$. A quick implication of Cobb-Douglas utility function is $x_i(p, m) = \frac{m}{2p_i}$, and hence $v(p, m) = \frac{m}{2\sqrt{p_1 p_2}}$. Suppose there is a 10% discount card for good 1. Let B ("buy price") be the maximum price she would pay for this discount card. Let S ("sell price") be the minimum price for which she would be willing to sell the card if she were to already own it. Denote $p^0 = (p_1^0, p_2^0)$, $p^1 = (0.9p_1^0, p_2^0)$ and $u^0 = v(p^0, m) < u^1 = v(p^1, m)$. Since

$$B = \int_{p_1^1}^{p_1^0} h_1\left(p_1, p_2^0, u^0\right) dp_1 = u^0 \int_{0.9p_1^0}^{p_1^0} \sqrt{\frac{p_2^0}{p_1}} dp_1$$

$$S = \int_{p_1^1}^{p_1^0} h_1\left(p_1, p_2^0, u^1\right) dp_1 = u^1 \int_{0.9p_1^0}^{p_1^0} \sqrt{\frac{p_2^0}{p_1}} dp_1$$

We have $B < S$. Actually we can prove a general result. If the price of good 1 drops from p_1^0 to p_1^1, and it is a normal good, then $CV < \Delta CS < EV$. The proof exploits the duality identity $h_1(p, v(p, m)) = x_1(p, m)$. Since good 1 is a normal good, $x_1(p, m)$ increases in m hence so does $h_1(p, v(p, m))$. Since $v(p, m)$ increases in m, we conclude that $h_1(p, U)$ must increase in U. Pick $p_1 \in (p_1^1, p_1^0)$, we have $\underbrace{v\left(p^0, m\right)}_{u^0} < v(p_1, \bar{p}_{-1}, m) < \underbrace{v\left(p^1, m\right)}_{u^1}$, since v decreases in prices. Therefore, $h_1(p_1, \bar{p}_{-1}, u^0) < \underbrace{h_1(p_1, \bar{p}_{-1}, v(p_1, \bar{p}_{-1}, m))}_{x(p_1, \bar{p}_{-1}, m)} < h_1(p_1, \bar{p}_{-1}, u^1)$. The integral expressions now yield $CV < \Delta CS < EV$. The proof suggests that if good 1 is inferior, the reverse is true: $CV > \Delta CS > EV$.

Behavioral economists claim that owning a good makes the consumer attached to it, so that she will not sell it except for a higher price than she would have been willing to pay for it before (the so-called "endowment effect"), and hence $B < S$. An experiment that can distinguish this behavioral hypothesis from the prediction of neoclassical consumer theory must specify an environment in which the neoclassical theory predicts $B > S$. If good 1 is inferior, then

$$\frac{\partial}{\partial U} h_1(p, U) = \frac{\partial}{\partial U} x_1(p, e(p, U)) = \frac{\partial x_1}{\partial m} \frac{\partial e}{\partial U} < 0$$

because $\frac{\partial e}{\partial U} > 0$ and $\frac{\partial x_1}{\partial m} < 0$. Hence, since $u^0 < u^1$, in this case

$$B = \int_{p_1^1}^{p_1^0} h_1\left(p_1, p_2^0, u^0\right) dp_1 > \int_{p_1^1}^{p_1^0} h_1\left(p_1, p_2^0, u^1\right) dp_1 = S.$$

If the consumer states that $B < S$ when the good is inferior, then classical theory can be rejected in favor of the endowment effect hypothesis.

Recitation 6: Production

Xincheng Qiu (qiux@sas.upenn.edu)

1 Properties[1]

1.1 Profit Function $\pi(p) = \sup_{y \in Y} p \cdot y$

1. Homogeneous of degree 1 (simple observation)
2. Continuous (Maximum theorem)
3. Convex (ET1)
4. If Y is closed, convex, and satisfies free disposal, then it can be obtained from the profit function:

$$Y = \left\{ y \in \mathbb{R}^N : p \cdot y \leq \pi(p), \forall p \in \mathbb{R}^N_+ \right\}$$

Proof. Denote $\hat{Y} := \left\{ y \in \mathbb{R}^N : p \cdot y \leq \pi(p), \forall p \in \mathbb{R}^N_+ \right\}$. The definition of π implies $Y \subseteq \hat{Y}$. To show the reverse, pick $\bar{y} \notin Y$, we must prove $\bar{y} \notin \hat{Y}$. A separating hyperplane theorem implies the existence of a nonzero $p \in \mathbb{R}^N$ such that

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

If any $p_i < 0$, this inequality could not hold, since free disposal would imply $\sup_{y \in Y} p \cdot y = \infty$. Hence, $p \in \mathbb{R}^N_+$, and so $\bar{y} \notin \hat{Y}$. □

1.2 Supply Correspondence $y(p) = \arg\sup_{y \in Y} p \cdot y$

1. Homogeneous of degree 0 (simple observation)
2. Upper hemicontinuous (at any $p \in \mathbb{R}^N_+$ for which $y(\cdot) \neq \varnothing$ on a neighborhood of p) (Maximum theorem)
3. Convex-valued at any $p \in \mathbb{R}^N_+$ if Y is convex; Single or empty-valued at any nonzero $p \in \mathbb{R}^N_+$ if Y is strictly convex.
4. Generalized Law of Supply: $(p' - p) \cdot (y' - y) \geq 0$, for any $p, p' \in \mathbb{R}^N_+$, and $y \in y(p)$, $y' \in y(p')$.

[1] This part is summarized from the lecture slides by Professor Steven Matthews, but any error is my own.

1.3 Derivative Properties

Be careful with the assumptions needed.

1. Hotelling's Lemma: $\nabla\pi(p) = y(p)$ if y is single-valued at $p \in \mathbb{R}^N_{++}$ (ET3)

2. $D_p y(p) = D_p^2 \pi(p)$ is symmetric and positive semidefinite, and $D_p y(p)\, p = 0$, If y is single-valued C^1 in a nbd of $p \in \mathbb{R}^N_{++}$.

   Proof. Symmetry by Young's theorem; PSD by the convexity of π; $D_p y(p)\, p = 0$ by homogeneity of degree 0 of y and Euler's theorem. $\square$

2 Two Approaches

Next, assuming a firm produces a single output, one can solve the profit-maximization problem through two approaches.

Algorithm 1. One-step approach.

$$\pi(p, w) = \max_{z \geq 0}\; pf(z) - w \cdot z$$

We can solve for *factor demand* $z(p, w)$ and *profit function* $\pi(p, w) = pf(z(p, w)) - w \cdot z(p, w)$.[2] To recover the *supply curve*, we can plug in the factor demand into the production technology $q(p, w) = f(z(p, w))$.

Algorithm 2. Two-step approach.

1. Step 1: Cost minimization.

$$c(w, q) = \min_{z \geq 0}\; w \cdot z$$
$$\text{s.t.} f(z) \geq q$$

   We can solve for *conditional factor demand* $z(w, q)$ and the *cost function* $c(w, q) = w \cdot z(w, q)$.

2. Step 2: Profit maximization.

$$\pi(p, w) = \max_{q \geq 0}\; pq - c(w, q)$$

[2] If $z(p, w)$ is a correspondence, the appearance of $z(p, w)$ in expressions refers to an element in it.

We can solve for *supply curve* $q(p,w)$ and the *profit function* $\pi(p,w) = pq(p,w) - c(w, q(p,w))$.

Remark 1. The idea of the two-step approach is to split the profit maximization problem into two parts: first, we look at the cost minimization problem at any given amount of output; second, we choose the most profitable output level. Generically, both approaches give the same solution: $\max_{x,y} g(x,y) = \max_x (\max_y g(x,y)) = \max_y (\max_x g(x,y))$. In our case,

$$\begin{aligned}
\max_{q,z\geq 0 \text{ s.t. } q\leq f(z)} pq - w\cdot z &= \max_{z\geq 0}\left(\max_{q \text{ s.t. } 0\leq q\leq f(z)} pq - w\cdot z\right) = \max_{z\geq 0} pf(z) - w\cdot z \\
&= \max_{q\geq 0}\left(\max_{z\geq 0 \text{ s.t. } f(z)\geq q} pq - w\cdot z\right) = \max_{q\geq 0}\left(pq - \min_{z\geq 0 \text{ s.t. } f(z)\geq q} w\cdot z\right)
\end{aligned}$$

Cost minimization is a necessary condition for profit maximization.

Remark 2. Firm's cost minimization problem is isomorphic to Consumer's expenditure minimization problem.

$$\begin{array}{ll}
\min_{z\geq 0} w\cdot z & \qquad \min_{x\geq 0} p\cdot x \\
\text{s.t. } f(z)\geq q & \qquad \text{s.t. } u(x)\geq U
\end{array}$$

with $w \leftrightarrow p, q \leftrightarrow U, f(z) \leftrightarrow u(x)$ and $z(w,q) \leftrightarrow h(p,U)$, $c(w,q) \leftrightarrow e(p,U)$ (except for preference is ordinal but production is cardinal[3]). The same properties for expenditure functions and Hicksian demand also hold for cost functions and conditional factor demand. For example, one can apply the same method in Recitation 4 to recover the conditional input demand functions and hence the production function from the cost function.

Remark 3. Since we can solve the profit-maximization problem by the one-step approach, why do we bother with a two-step approach? The one-step approach is direct, but there are some insights on the cost side when applying the two-step approach.

1. One can derive the cost function, which is relatively measurable.

2. The CMP depends only on input prices but not on output price, so it is the same no matter whether the output market is competitive or not (e.g., a monopolistic firm).

3. The choice of q facilitates the understanding of trade-offs: MR and MC.

4. For some production technologies (e.g., CRS and IRS), the profit maximization problem does not have a well-defined solution, but the cost minimization problem still has.

[3]The cardinal properties of f generate new results:

1. If f is homogeneous of degree 1 (CRS), then $c(w,q)$ and $z(w,q)$ are linear in q (constant MC).
2. If f is concave, then $c(w,q)$ is convex in q (increasing MC).

Recitation 7: Uncertainty

Xincheng Qiu (qiux@sas.upenn.edu)

1 Expected Utility

Remark 1. Primitives: in the basic consumer theory, a preference relation is defined on X; in the expected utility theory, a preference relation is defined on $\hat{\mathcal{L}}$ (not C). Another way to phrase this idea is: now the set of alternatives X becomes $\hat{\mathcal{L}}$, the set of all compound lotteries[1].

Definition 1. $\precsim$ defined on $\hat{\mathcal{L}}$ satisfies Expected Utility Hypothesis (EUH) iff a function $u : C \to \mathbb{R}$ s.t. $\precsim$ is represented by

$$U\left(\hat{L}\right) := \sum_{n=1}^{N} p_n u\left(c_n\right)$$

where $(p_1, \ldots, p_N) = R\left(\hat{L}\right)$. Here u is called the Bernoulli utility function and U is called the von-Neumann-Morgenstern (vNM) expected utility function.

Remark 2. It seems natural to think of the vNM utility as the expectation of u_n (hence called expected utility). However, note that the object of interest is $R\left(\hat{L}\right) = (p_1, \ldots, p_n)$, so it is deeper to think of $U\left(\hat{L}\right)$ as a linear function in p_n's, weighted by u_n's, not the other way around. In fact, the expect utility form is equivalent to $U\left(\sum_{i=1}^{K} \alpha_i L_i\right) = \sum_{i=1}^{K} \alpha_i U\left(L_i\right)$ for $\alpha_i \geq 0$ and $\sum_{i=1}^{K} \alpha_i = 1$.

Theorem 1. *Bernoulli utility is unique up to a positive affine transformation.*

Remark 3. Cardinal properties of u matters. Differences of Bernoulli utility have meaning:: $u_1 - u_2 > u_3 - u_4 \Leftrightarrow \frac{1}{2}u_1 + \frac{1}{2}u_4 > \frac{1}{2}u_2 + \frac{1}{2}u_3$. This ranking of differences is preserved by positive affine transformations (but not by an arbitrary monotonic transformations).

Axiom 1. *(R) Reduction (Consequentialism). For any $L, L' \in \hat{\mathcal{L}}$, $L \precsim L'$ iff $R\left(L\right) \precsim R\left(L'\right)$.*

[1] Compound lottery is a recursive definition $\hat{L} = \left(\hat{L}_i, \alpha_i\right)_{i=1,\ldots,K}$ where $\alpha \in \Delta^{K-1}$. Simple Lotteries are $\mathcal{L} = \Delta^{N-1}$ on consequences (outcomes) $C = \{c_1, \ldots, c_N\}$. A compound lottery $\hat{L}$ could be reduced to a simple lottery $R\left(\hat{L}\right) \in \mathcal{L}$ that generates the sample ultimate distribution over outcomes.

Axiom 2. *(C) Continuity (Archimedian Axiom). Given any three simple lotteries* L, L', L'' *for which* $L \precsim L' \precsim L''$, *there exists* $a \in [0,1]$ *such that* $aL + (1-a)L'' \sim L$.

Axiom 3. *(I) Independence. Given any simple lotteries* L, L', L'' *and any number* $a \in [0,1]$, $L \precsim L'$ *iff* $aL + (1-a)L'' \precsim aL' + (1-a)L''$.

Remark 4. Reduction axiom says only the reduced lottery over outcomes is relevant to the decision maker. Continuity implies small changes in probabilities will not change the ordering between two lotteries. Independence is a big assumption (Allais paradox violates the independence axiom), meaning that if we mix each lottery with a third one, the preference ordering over the two mixtures does not change. It is unlike anything encountered in the basic choice theory (we do not assume consumer's preference over various bundles of good 1 and 2 to be independent of good 3) and of central importance for the decision theory under uncertainty (it is closely linked to the expected utility representation).

Theorem 2. *Expected Utility Theorem*

A rational $\precsim$ *on* $\mathcal{L}$ *satisfies Axioms C and I iff* $u_1, \ldots, u_N$ *exist such that* $\precsim$ *is represented by the* $U : \mathcal{L} \to \mathbb{R}$ *defined by* $U(p_1, \ldots, p_N) \coloneqq \sum_{n=1}^{N} p_n u(c_n)$. *A rational* $\precsim$ *on* $\hat{\mathcal{L}}$ *satisfies Axioms R, C and I iff it satisfies EUH.*

Remark 5. This theorem gives the conditions under which we can represent preferences by the expected utility form, which is extremely convenient. Continuity axiom guarantees the existence of utility representation. Independence axiom implies the indifference curves are convex (hence straight lines if not thick) and parallel (draw graphs).

Example 1. Suppose $\precsim$ is represented by the median function, where for $L = (p_1, \ldots, p_N)$,

$$m(L) \coloneqq \min \left\{ c \in C : \sum_{k \le c} p_k \ge 0.5 \right\}$$

The preference relation m represents is convex, but independence is violated. Consider $L = (1,0,0)$, $L_1 = (.4, 0, .6)$, $L_2 = (0, .6, .4)$. Note that $m(L_1) = 3 > m(L_2) = 2$. Take $L_1^\lambda = .6L_1 + .4L = (.64, 0, .36)$, $L_2^\lambda = .6L_2 + .4L = (.4, .36, .24)$. But $m\left(L_1^\lambda\right) = 1 < m(L_2) = 2$.

2 Risk Aversion

2.1 Definition and Measures

We have four ways to describe risk aversion:[2]

[2]Although Section 1 assumes finite outcomes in order to avoid technical complications, the theory can be extended to an infinite domain. From here on, we can work with distribution functions to describe lotteries

1. The definition of risk aversion, *per se*, does not presume EUH: $\precsim$ is risk averse if $F \precsim \delta_{\mathbb{E}_F x}$ for all F.

2. If EUH is satisfied, $F \precsim \delta_{\mathbb{E}_F x}$ can be written as $\mathbb{E}_F u(x) \leq u(\mathbb{E}_F x)$, which is the defining property (Jensen's Inequality) of a concave function. So $\precsim$ is risk averse iff the Bernoulli utility is concave (given EUH).

3. Assume now EUH is satisfied with a strictly increasing Bernoulli utility, then $\precsim$ is risk averse iff $c(F, u) \leq \mathbb{E}_F x$ for all F. This is because $u(c(F, u)) \coloneqq \mathbb{E}_F u(x) \leq u(\mathbb{E}_F x)$.

4. For a twice differentiable u, the Arrow-Pratt coefficient of absolute risk aversion is the function $A(x) = -\frac{u''(x)}{u'(x)}$. If u is concave and strictly increasing, then $A(x) \geq 0$.

Remark 6. A provides a measure of the extent of risk aversion. Note that it is invariant to positive affine transformations of u.

2.2 Comparisons of Risk Aversion

2.2.1 Comparison Across Individuals

Theorem 3. *Pratt's Theorem.*

Let $\precsim_a$ and $\precsim_b$ have twice differentiable and strictly increasing Bernoulli utility functions u_a and u_b. The following are equivalent:

1. *For all F and x, $F \succsim_a \delta_x \Rightarrow F \succsim_b \delta_x$ ($\precsim_a$ is more risk averse than $\precsim_b$)*

2. *$u_a = h \circ u_b$ for some concave and strictly increasing function h (u_a more concave)*

3. *$c(F, u_a) \leq c(F, u_b), \forall F$ (a has a smaller certainty equivalent)*

4. *$A_a \geq A_b$ (a has a higher Arrow-Pratt coefficient of absolute risk aversion)*

2.2.2 Comparison at Different Levels of Wealth

Definition 2. (DARA, CARA, IARA) u exhibits decreasing, constant, or increasing absolute risk aversion if $A(x)$ is decreasing, constant, or increasing, respectively.

Example 2. Assume $u' > 0, u'' < 0, \mathbb{E}\tilde{r} > 0, w > 0$ and DARA.

$$x^*(w) = \arg\max_{x \geq 0} \mathbb{E}u(w + \tilde{r}x)$$

over monetary outcomes, which can be either a discrete or continuous random variable. We therefore take $\mathcal{L}$ as the set of all distribution functions over some interval.

- Claim: $x^*(w) \neq 0$. This is because $\mathbb{E}u'(w+\tilde{r}x)\tilde{r}|_{x=0} = \mathbb{E}u'(w)\tilde{r} = u'(w)\mathbb{E}\tilde{r} > 0$.
- The FOC for interior solutions is $\mathbb{E}u'(w+\tilde{r}x^*(w))\tilde{r} = 0$. Differentiate this w.r.t. w

$$x^{*\prime}(w) = \frac{-\mathbb{E}u''(w+\tilde{r}x^*(w))\tilde{r}}{\mathbb{E}\left[u''(w+\tilde{r}x^*(w))\tilde{r}^2\right]} := \frac{N}{D}$$

- $D < 0$ because $u'' < 0$.
- $N = \mathbb{E}\tilde{r}A(w+\tilde{r}x^*(w))u'(w+\tilde{r}x^*(w))$. DARA together with $x^*(w) > 0$ implies

$$r > 0 \Rightarrow A(w+rx^*(w)) < A(w) \Rightarrow rA(w+rx^*(w)) < rA(w)$$

$$r < 0 \Rightarrow A(w+rx^*(w)) > A(w) \Rightarrow rA(w+rx^*(w)) < rA(w)$$

  Hence $N < \mathbb{E}\tilde{r}A(w)u'(w+\tilde{r}x^*(w)) = A(w)\underbrace{\mathbb{E}\tilde{r}u'(w+\tilde{r}x^*(w))}_{\text{F.O.C.}} = 0$. So $x^{*\prime}(w) > 0$.

Remark 7. To obtain comparative statics, write down the FOC, then take derivative of both sides with respect to the parameter of interest. This example shows DARA can capture that wealthier people tend to invest more in risky assets. A related concept (which is not covered in the slides) is relative risk aversion $R(y) := -\frac{yu''(y)}{u'(y)}$. Kenneth Arrow claimed, as an empirical matter, that as an investor becomes wealthier, she invests a smaller *proportion* of her wealth in risky assets. IRRA (i.e., $R(\cdot)$ increasing) can capture this idea.

Remark 8. To formalize the comparison of risk aversion at different levels of wealth, we could define $u_1(x) := u(x+w_1)$ and $u_2(x) := u(x+w_2)$. Then comparing one individual's risk attitudes at different levels of her wealth is as if we were comparing the Bernoulli utility functions $u_1(x)$ and $u_2(x)$, which has been studied in Section 2.2.1.

Example 3. The sale price $s(a)$ for a gamble $a+\tilde{\varepsilon}$ is the minimum amount one would sell the gamble for:

$$u(s(a)) = \mathbb{E}u(a+\tilde{\varepsilon})$$

If the C^2 Bernoulli utility function with $u' > 0, u'' \leq 0$ exhibits DARA. Say as much as you can about the derivative $s'(a)$.

Solution. You may want to try differentiating the identity defining $s(a)$

$$s'(a) = \frac{\mathbb{E}u'(a+\tilde{\varepsilon})}{u'(s(a))}$$

It is obvious that $s'(a) > 0$ because $u' > 0$, but it turns out that we can actually obtain a stronger conclusion. It seems hard to proceed thought it is still possible. Note that the above equation can be rewritten as

$$s'(a) = \frac{\mathbb{E}u'\left(u^{-1}\left(u\left(a+\tilde{\varepsilon}\right)\right)\right)}{u'\left(u^{-1}\left(u\left(s\left(a\right)\right)\right)\right)} = \frac{\mathbb{E}u'\left(u^{-1}\left(u\left(a+\tilde{\varepsilon}\right)\right)\right)}{u'\left(u^{-1}\left(\mathbb{E}u\left(a+\tilde{\varepsilon}\right)\right)\right)} \triangleq \frac{\mathbb{E}g\left(\tilde{\eta}\right)}{g\left(\mathbb{E}\tilde{\eta}\right)}$$

by defining $\tilde{\eta} \coloneqq u\left(a+\tilde{\varepsilon}\right)$ and $g = u' \circ u^{-1}$. The RHS is a familiar expression, which is same as the form of Jensen's inequality. Indeed, we could prove g is a convex function.

$$g'(\cdot) = \frac{u''\left(u^{-1}(\cdot)\right)}{u'\left(u^{-1}(\cdot)\right)} = -A\left(u^{-1}(\cdot)\right)$$

Since u is strictly increasing, so is u^{-1}. DARA implies A is decreasing, so g' is increasing, which means g is convex. $u' > 0$ guarantees $g > 0$ and hence $s'(a) = \frac{\mathbb{E}g(\tilde{\eta})}{g(\mathbb{E}\tilde{\eta})} \geq 1$.

Now let's try the useful trick in Remark 8, which makes things much easier. For any a, define $u_a(z) \coloneqq u(a+z)$. Then

$$u_a\left(s\left(a\right)-a\right) = u\left(s\left(a\right)\right) = \mathbb{E}u\left(a+\tilde{\varepsilon}\right) = \mathbb{E}u_a\left(\tilde{\varepsilon}\right)$$

Hence $s(a) - a$ is the the certainty equivalent of the risk $\tilde{\varepsilon}$ for the utility function u_a. DARA implies u_a becomes less risk averse as a increases, so this certainty equivalent increases in a: $s'(a) - 1 \geq 0$.

Economics 701A **Steven A. Matthews**
Fall 2018 **University of Pennsylvania**

Suggested Solutions to the Quiz

30 points, 40 minutes. Closed books, notes, calculators.
Indicate your reasoning, using clearly written words as well as math.

1. (20 pts) Preferences are called *homothetic* if they satisfy the following property:

$$x \succsim y \Rightarrow \alpha x \succsim \alpha y \quad \forall \alpha \geq 0$$

Suppose $\succsim$ is a complete, transitive, monotonic, continuous preference relation on $\mathbb{R}^L_+$. Show that $\succsim$ is homothetic if and only if there exists a utility representation u of $\succsim$ such that $u(\alpha x) = \alpha u(x)$ for all $\alpha \geq 0$.

Soln: ($\Leftarrow$) Suppose $u(\cdot)$ represents $\succsim$ and is homogeneous of degree 1. Let $x, y \in \mathbb{R}^L_+$ be such that $x \succsim y$, and let $\alpha \geq 0$. Then

$$\begin{aligned} u(x) &\geq u(y) \text{ (Since } u \text{ represents } \succsim) \\ \Rightarrow \alpha u(x) &\geq \alpha u(y) \text{ (Since } \alpha \geq 0) \\ \Rightarrow u(\alpha x) &\geq u(\alpha y) \text{ (Since } u \text{ is homogeneous of degree 1)} \\ \Rightarrow \alpha x &\succsim \alpha y \text{ (Since } u \text{ represents } \succsim). \end{aligned}$$

This shows that any preference relation represented by a utility function that is homogeneous of degree 1 is homothetic.

($\Rightarrow$) Suppose $\succsim$ is homothetic, and let u be the function representing $\succsim$ constructed in the proof of the Monotone Representing Theorem we sketched in class. Thus, for any $x \in \mathbb{R}^L_+$, $u(x)$ is the number such that

$$x \sim (u(x), \cdots, u(x)). \tag{1}$$

We show that this u is homogeneous of degree 1. Fixing x, applying the definition of homothetic preference to (1) (twice, once in each direction) yields

$$\alpha x \sim (\alpha u(x), \cdots, \alpha u(x)).$$

We also know, by the definition of u, that

$$\alpha x \sim (u(\alpha x), \cdots, u(\alpha x)).$$

Transitivity now implies

$$(\alpha u(x), \cdots, \alpha u(x)) \sim (u(\alpha x), \cdots, u(\alpha x)).$$

This and monotonicity imply $\alpha u(x) = u(\alpha x)$, as desired. ∎

2. (10 pts) The UMP for some $u : \mathbb{R}^2_+ \to \mathbb{R}$ yields the indirect utility function

$$v(p, m) = \left(\frac{1}{p_1} + \frac{1}{p_2} \right) m.$$

Find the consumer's demand function $x(p, m)$.

Soln: Using the Envelope Theorem (ET2) twice, we obtain

$$x_i(p, m) = -\frac{v_{p_i}(p, m)}{v_m(p, m)} \text{ for } i = 1, 2.$$

(This is Roy's Identity.) For our v we have

$$v_{p_i} = -p_i^{-2} m, \quad v_m = p_1^{-1} + p_2^{-1}.$$

Hence,

$$x_i(p, m) = -\frac{-p_i^{-2} m}{p_1^{-1} + p_2^{-1}} = \frac{p_j m}{p_i^2 + p_1 p_2}$$

for $i, j \in \{1, 2\}$, $i \neq j$. ∎

Economics 701A **Steven A. Matthews**
Fall 2018 **University of Pennsylvania**

Suggested Solutions to the Exam

100 points, 75 minutes. Closed books, notes, calculators.
Indicate your reasoning.

1. (30 pts) Let X be a nonempty set. Consider a choice structure $\langle \mathfrak{B}, C \rangle$, where $\mathfrak{B} \subseteq 2^X \setminus \varnothing$ and $C : \mathfrak{B} \rightrightarrows X$ satisfies $\varnothing \neq C(B) \subseteq B$ for all $B \in \mathfrak{B}$.

   (a) (10 pts) Say what it means for $\langle \mathfrak{B}, C \rangle$ to satisfy the Weak Axiom of Revealed Preference (WARP).
   Soln: First define the revealed preference relation $\succsim^*$: for all $x, y \in X$,

   $$x \succsim^* y \text{ iff } x = y \text{ or } B \in \mathfrak{B} \text{ exists such that } y \in B \text{ and } x \in C(B).$$

   Then we say $\langle \mathfrak{B}, C \rangle$ satisfies WARP iff for all $B \in \mathfrak{B}$ and $x, y \in B$,

   $$y \in C(B) \text{ and } x \succsim^* y \;\Rightarrow\; x \in C(B).$$

   ■

   (b) (20 pts) Show that if $\langle \mathfrak{B}, C \rangle$ satisfies WARP, then a binary relation exists that rationalizes it.
   Soln: The binary relation that rationalizes the choice structure is the revealed preference relation $\succsim^*$. To show this, let $B \in \mathfrak{B}$. We must show

   $$C(B) = C^*(B, \succsim^*) := \{x \in B : x \succsim^* y \; \forall y \in B\}.$$

   If $x \in C(B)$, then by definition $x \succsim^* y$ for all $y \in B$. Hence,

   $$C(B) \subseteq C^*(B, \succsim^*). \tag{1}$$

   To show the reverse inclusion, let $x \in C^*(B, \succsim^*)$. Since $C(B) \neq \varnothing$, there exists $y \in C(B)$. As $x \in C^*(B, \succsim^*)$, we have $x \succsim^* y$. Hence, $x \in C(B)$ by WARP. We thus have

   $$C(B) \supseteq C^*(B, \succsim^*). \tag{2}$$

   From (1) and (2) we obtain $C(B) = C^*(B, \succsim^*)$. ■

2. (30 pts) Jane lives in a two-good world. Her income is $m = 18$. She has a continuous increasing utility function on $\mathbb{R}^2_+$ that gives rise to the expenditure function

   $$e(p, u) := (p_1 + 2p_2)u.$$

   (a) (10 pts) Find Jane's compensating variation for the price change from $p^0 = (4, 1)$ to $p^1 = (1, 1)$.
   Soln: $CV = 9$.
   Proof. Since $e(p, \cdot)$ and $v(p, \cdot)$ are inverse functions, we can find v from e :

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

Recall that CV is the maximum amount Jane would be willing to pay for the price change:

$$v(p^1, m - CV) = v(p^0, m),$$

which here is

$$\frac{m - CV}{p_1^1 + 2p_2^1} = \frac{m}{p_1^0 + 2p_2^0}.$$

Using $m = 18$, $p^0 = (4, 1)$, and $p^1 = (1, 1)$, this becomes

$$\frac{18 - CV}{1 + 2} = \frac{18}{4 + 2}.$$

Thus, $CV = 9$. ■

(b) (20 pts) Find Jane's change in consumer surplus, $\triangle CS$, for this same price change and the same income $m = 18$.

Soln: $\triangle CS = 18 \ln 2 \approx 12.477$.

Proof. From Shepard's lemma we find the Hicksian demand function for good 1:

$$h_1(p, u) = \frac{\partial e}{\partial p_1} = u.$$

This, the duality identity $x(p, m) = h(p, v(p, m))$, and the $v(p, m)$ found in (a) give us the Marshallian demand function for good 1:

$$\begin{aligned} x_1(p, m) &= h_1(p, v(p, m)) \\ &= v(p, m) \\ &= \frac{m}{p_1 + 2p_2}. \end{aligned}$$

The change in consumer surplus is the area under this demand curve between the two prices for good 1:

$$\begin{aligned} \triangle CS &= \int_1^4 \left(\frac{m}{p_1 + 2p_2} \right) dp_1 \\ &= \int_1^4 \left(\frac{18}{p_1 + 2} \right) dp_1 \\ &= 18 \left[\ln(p_1 + 2) \right]_{p_1 = 1}^4 \\ &= 18 \ln 2. \end{aligned}$$

■

3. (10 pts) Suppose a production function $f : \mathbb{R}_+^{L-1} \to \mathbb{R}_+$ gives rise to a cost function $c(w, q)$. Prove that if f is concave, then $c(w, q)$ is convex in q.

Soln: Let $q^1, q^2 \in \mathbb{R}_+$, $\lambda \in [0, 1]$, $z^1 \in z(w, q^1)$, and $z^2 \in z(w, q^2)$. Then, since $f(\cdot)$ is concave, we have

$$f\left(\lambda z^1 + (1 - \lambda) z^2\right) \geq \lambda f\left(z^1\right) + (1 - \lambda) f\left(z^2\right) \geq \lambda q^1 + (1 - \lambda) q^2.$$

Thus, $z = \lambda z^1 + (1 - \lambda) z^2$ is feasible for the problem of minimizing the cost of producing at least $\lambda q^1 + (1 - \lambda) q^2$. Thus,

$$\begin{aligned} c\left(w, \lambda q^1 + (1 - \lambda) q^2\right) &\leq w \cdot \left(\lambda z^1 + (1 - \lambda) z^2\right) \\ &= \lambda c\left(w, q^1\right) + (1 - \lambda) c\left(w, q^2\right). \end{aligned}$$

This proves $c(\cdot)$ is convex in q. ■

4. (30 pts) A consumer has a Bernoulli utility function $u : \mathbb{R} \to \mathbb{R}$ that is C^2, with $u' > 0$. Furthermore, u exhibits DARA, i.e., $A = -u''/u'$ is a strictly decreasing function. The proof of Pratt's theorem can be modified to show the following for this DARA u :

Useful Fact. For $a \in \mathbb{R}$, define $u_a(x) = u(x + a)$. Then, if F is a nondegenerate distribution, the certainty equivalent $c(F, u_a)$ is strictly increasing in a.

Suppose this consumer may buy an asset that will generate a random income $\tilde{x}$ that has a nondegenerate distribution F. Her income if she does not purchase the asset is m. Her *buy price*, $b(w)$, is the maximum amount she would be willing to pay for the asset, defined by

$$\mathbb{E}u(\tilde{x} + w - b(w)) = u(w).$$

Show that $b'(w) \in [0, 1]$.

Soln: The following propositions show that in fact, $b'(w) \in (0, 1)$, and that $b'(w) < 1$ even if u does not exhibit DARA.

Proposition 1. If $u : \mathbb{R} \to \mathbb{R}$ is differentiable and satisfies $u' > 0$, then $b' < 1$.

Proof. Differentiate both sides of the identity defining $b(w)$ to obtain

$$1 - b'(w) = \frac{u'(w)}{\mathbb{E}u'(\tilde{x} + w - b(w))} > 0. \tag{3}$$

This gives us $b'(w) < 1$. ■

Proposition 2. If $u : \mathbb{R} \to \mathbb{R}$ is twice differentiable with $u' > 0$, and it exhibits DARA, then $b' > 0$.

Proof. From (3) we have

$$\begin{aligned} 1 - b'(w) &= \frac{u'(w)}{\mathbb{E}u'(\tilde{x} + w - b(w))} \\ &= \frac{u'(u^{-1}(u(w)))}{\mathbb{E}u'(u^{-1}(u(\tilde{x} + w - b(w))))} \\ &= \frac{u'(u^{-1}(\mathbb{E}u(\tilde{x} + w - b(w))))}{\mathbb{E}u'(u^{-1}(u(\tilde{x} + w - b(w))))}. \end{aligned}$$

Hence,

$$1 - b'(w) = \frac{g(\mathbb{E}\tilde{\eta})}{\mathbb{E}g(\tilde{\eta})}, \tag{4}$$

where $\tilde{\eta} := u(\tilde{x} + w - b(w))$ and $g := u' \circ u^{-1}$. Note that

$$g'(\cdot) = \frac{u''(u^{-1}(\cdot))}{u'(u^{-1}(\cdot))} = -A(u^{-1}(\cdot)).$$

As u is strictly increasing, so is u^{-1}. By DARA, A is strictly decreasing. Hence, g' is strictly increasing, and so g is strictly convex. Noting that $\tilde{\eta}$ is a nondegenerate random variable because u is strictly increasing and $\tilde{x}$ is nondegenerate, the strict version of Jensen's inequality gives us $g(\mathbb{E}\tilde{\eta}) < \mathbb{E}g(\tilde{\eta})$. This and (4) imply that $1 - b'(w) < 1$, and so $b'(w) > 0$. ■

Remark. The following argument using the Useful Fact proves the somewhat weaker result that $b'(w) \geq 0$: The definition of $b(w)$ gives us

$$\mathbb{E}u_{w-b(w)}(\tilde{x}) = \mathbb{E}u(\tilde{x} + w - b(w)) = u(w) = u_{w-b(w)}(b(w)).$$

This shows that $b(w) = c(F, u_{w-b(w)})$.

Assume $b'(w) < 0$. Then $(w - b(w))' = 1 - b'(w) > 0$. This and the Useful Fact imply $c(F, u_{w-b(w)})$ is increasing in w. But this and $b(w) = c(F, u_{w-b(w)})$ together imply $b'(w) \geq 0$, contrary to the initial assumption that $b'(w) < 0$. This contradiction proves that $b'(w) \geq 0$. ■