Microeconomics 1

Lecture Notes from Microeconomics 1 Course Institute of Economics, Federal University of Rio de Janeiro (IE/UFRJ)

Pedro Hemsley PPGE IE/UFRJ

2025-01-20

Table of contents

Syllabus 2024.1

> **Note**
>
> Based on Alexander Wolitzky lecture slides, available at MIT Opencourseware.
> CC BY-NC-SA 4.0 Legal Code | Attribution-NonCommercial-ShareAlike 4.0 International | Creative Commons. See here for more information.

> Join the 2024.1 class WhatsApp group for the first part of the course.

LEC #	TOPICS	READINGS
1, 2	Choice, Preference, and Utility	[MWG] Chapter 1. [Kreps] Chapters 1 and 2.
3–5	Consumer Theory and its Applications	**Consumer Theory** [MWG] Chapters 2 and 3. [Kreps] Chapters 3, 10, and 11. **Applications of Consumer Theory** [MWG] Chapter 4.
6, 7	Producer Theory and Monotone Methods	[MWG] Chapter 5.
8–11	Choice Under Uncertainty	[MWG] Chapter 6. [Kreps] Chapters 5 and 6.
12	Dynamic Choice	[Kreps] Chapter 7.

Part I

Lectures 1 – 2

1 Choice, Preference and Utility

> **Microeconomics is decision theory plus a theory of interaction among decision-makers, in different environments**

This is our working definition, and is not restricted to any specific field of science. This setup is very general and is used in other fields: political science, sociology, psychology, biology…

We will see specific models of decision theory, and specific models of social interaction, that are often used in economics. But it's important to notice the difference between general theory and particular models.

1.1 Decision Theory

We begin with decision theory. To build a theory of individual choice, we need some assumptions:

1. Choices are possible

- Seems obvious, but sets us apart from many models from 19th century sociology
- Big discussions here. See *Determined: a science of life without free will* (Sapolsky, 2023)
- Important link to theory of agency in biology
- We usually model this through restrictions (see point 4 below)

2. There is a defined subject that makes decisions, in some circumstance

- (Usually) It's and individual, not a set of individuals such as family nor parts of an individual such as Id and Ego – but there has been much development in this direction.
- We just need to separate individual from the whole society (or there would be no need to study interactions)
- This is not the selfish agent assumption, which we can also make in particular settings

3. Choices are made according to some criteria

- They are not purely random (at least not necessarily)
- This is the same as saying as there is some objective. Could be anything: profit, utility, leisure, stability, time with loved ones, watching as many world cup matches as possible...

4. Choices are subject to constraints

- One cannot simply choose 'everything'. At the very least there is the opportunity cost of time.
- This is not necessarily a budget restriction: could be time, attention, memory, information, ability to process information, or external factors such as physical environment or laws, or simply other agents.
- In this course, we will consider a budget constraint, but this is just a particular case, and the absent restrictions are as important as the one(s) we will consider!

5. Consistency requirement: equilibrium

- Everyone is choosing according to the four previous assumptions at any point in time.
- In other words: given the criteria and the restrictions at some point in time, no one would like to change their decision at that moment.
- For a given individual, part of the environment is 'other individuals making choices, and maybe they're even taking my actions into account'
- This is **NOT** about absence of movement or change. There could be mistakes and regrets.

We typically write the first four points as

> *"Choose x to solve* $\max u(x)$ *subject to restrictions on x"*

Then we move on to interaction: no individual incentive to deviate, given what others are doing.

In short, our structure is optimization + equilibrium.

This is very general. Hard to work without it.

1.2 About assumptions

We've barely begun and already have some big assumptions.

We'll make many other assumptions along the way. Pay attention to them: assumptions must be clearly understood. We usually write $A \Rightarrow B$: we must be able to understand what happens (B) under some circunstances (A). This is **NOT** them same as stating B (think of the first welfare theorem).

Assumptions have a tradeoff. On one hand, they take away generality: if we assume economic agents have perfect memory, we must be cautious when applying our model to agents without perfect memory. On the other hand, they allow us to better understand a (more restricted) setting.

This is the tradeoff of the lab rat. It's easier to study, and we learn a lot from it, but must be aware of its limits. In our case, 'homo economicus', for example:

- Simplest economic model: maximize utility subject only to a budget constraint.

We study the homo economicus hoping to learn something about its distant cousin: homo sapiens. The more realist economic agent from behavioral economics is our 'lab monkey': the same tradeoff applies.

Is there a way around this tradeoff? No: it comes from our own limitations, as we cannot pay attention and process all available information. Think for a moment far from economics: zebras and lions. Where are their eyes placed? They "choose" to focus on some specific type of information, and give up on many things. It works according to some specific criterion: **survival**.

In economics (and any other field of science), we will make assumptions, whether it's clear or not. The problem is the criterion we use to evaluate them. No "given" criterion such as survival. Major discussion about criteria, for every research question.

We may put it another way: there is a tradeoff to realism. If it's too much: can't really understand what's going on, can't make predictions. Borges' map. If it's too little: you'll understand clearly something not relevant to your research question. Drunkard looking for his keys under the lamp post, not where he lost them.

In the end of the day, we use approximations, and we want predictions that can be tested to see whether these approximations are good enough. Keep in mind that external validity is always an issue, even when we have good empirical results, and this depends on how restrictive our assumptions are (must be extra careful with observational data!).

Lastly, a point about representation: a model of reality is different from reality. Implication: solving a model in decision theory is different from actually making that decision (same reasoning holds for anything else in science). Think of catching an object thrown to you (no actual functional analysis problem) or buying stuff at the grocery store (no actual lagrangian).

This goes back to theory of agency in biology: consciousness, representation and abstraction. Important: if the modeler knows more than the agents in the model, this must be modeled too! It's some form of restriction.

As for criticism: when we criticize a model, we either say "A is not the relevant setup to consider" or "A does not really imply B". Considering a different setup (A' instead of A) is not really a criticism – it's just 'doing something else'. Think of political and relief maps, or flat maps and globes.

> We always have to ask: *What are we building our models and theories for? What do we want to understand?*
> **Models are made to be used, not to be believed**
> If they help us with something relevant, they're doing their job: increase food production, cure a disease, decrease unemployment, etc.

Mathematics

Our assumptions often generate a setup that may be analyzed mathematically. Our optimization problem will be written as something similar to:

$$\max_{\{x\}} f(x)$$

subject to some restriction

$$g(x) < 0$$

Yet, it is important to notice the difference between language and tools. We use mathematical language as in the example above. This does not mean we're using any mathematical tools yet: this could be written in plain English. Often we will use actual mathematical tools: if f and g are differentiable functions, then we may use first-order conditions.

An additional point

- As science advances, it becomes more continuous, and less discrete.
- In economics, we fight too often over petty details (Freud: narcissism of small differences.)

We will start with the most basic model: the traditional *homo economicus*, our lab rat, whose only restriction is budgetary.

An example

Lionel Page fascinating explanation of Prospect Theory:

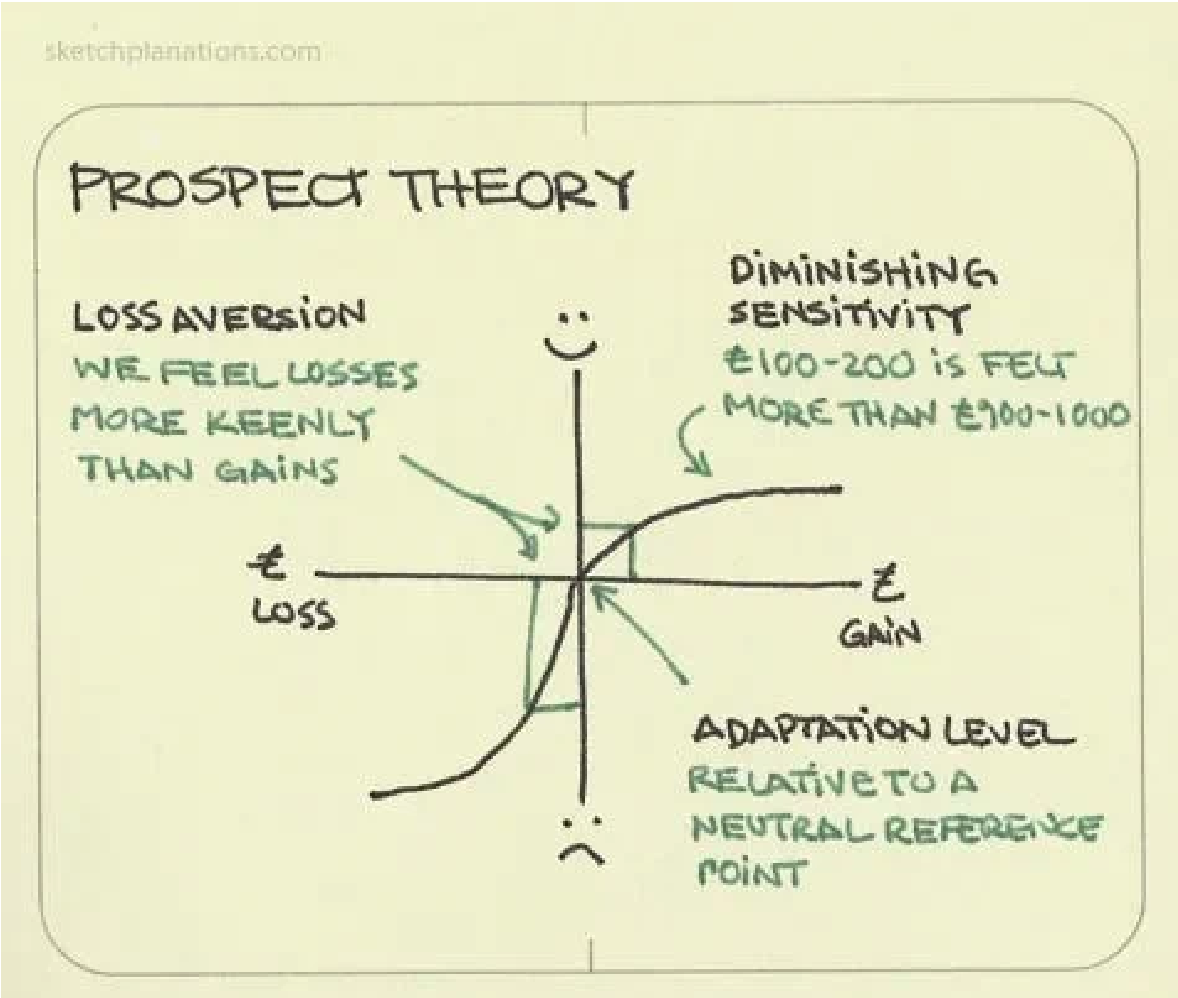
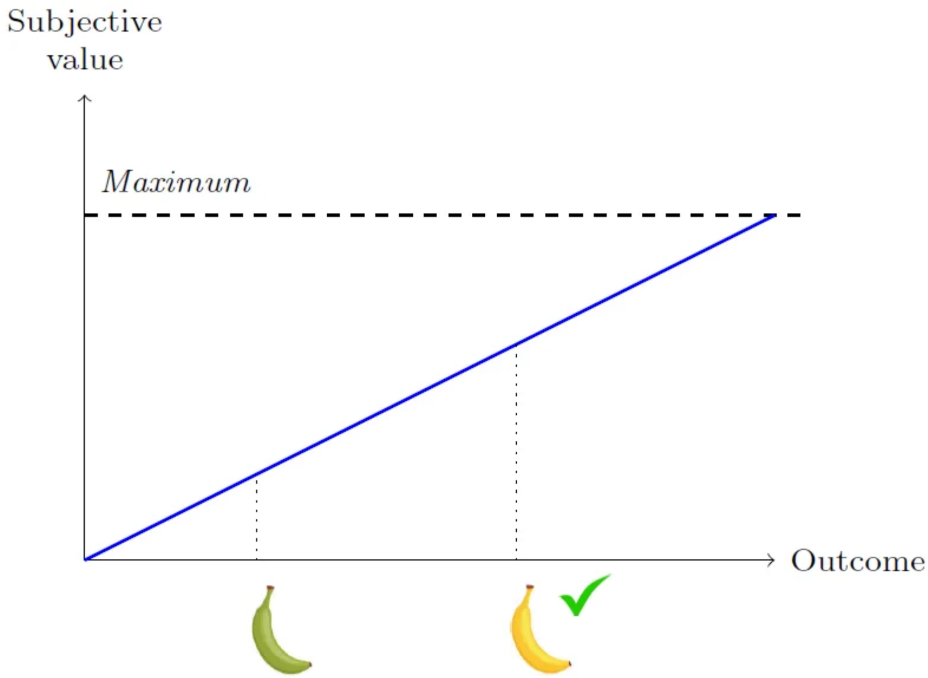


Figure 1.1: Prospect Theory

This is the most cited paper in economics: people value what they have by considering it as a gain or a loss relative to a subjective "reference point". This point may be status quo, expectations, aspirations: not defined / agreed upon.

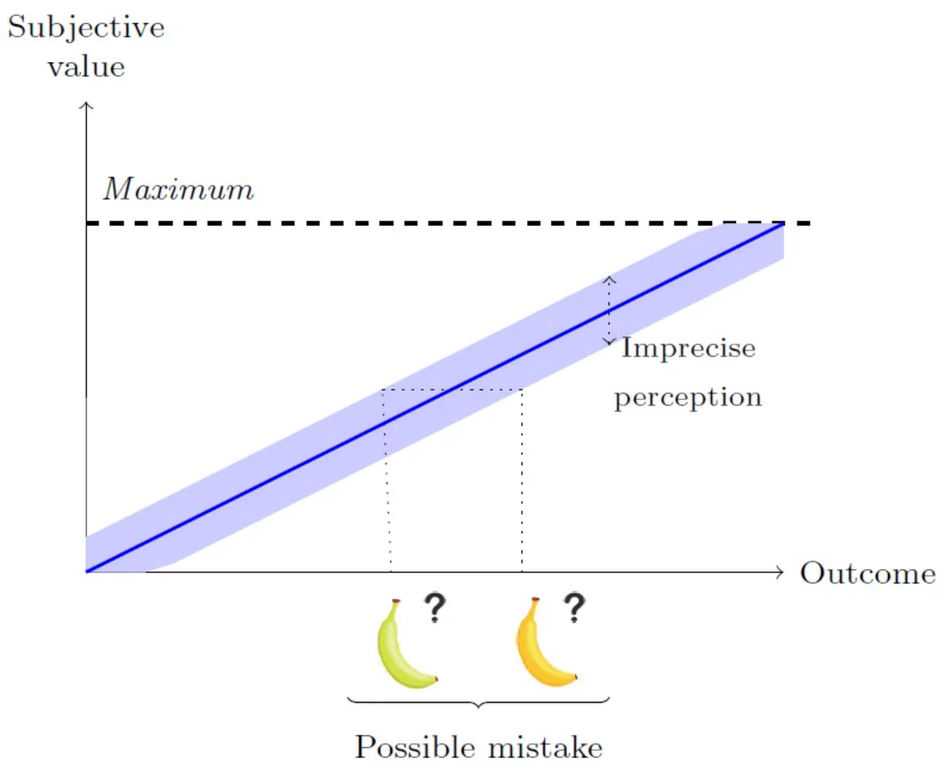
Reference dependent preferences are a cognitive flaw? The biological basis of economic behavior (Robson, 2001): *subjective satisfaction can be seen as an informative signal that helps us identify the best option.* Eating, sleeping and having sex 'feel good' because they help us survive: subjective satisfaction is informative signal.



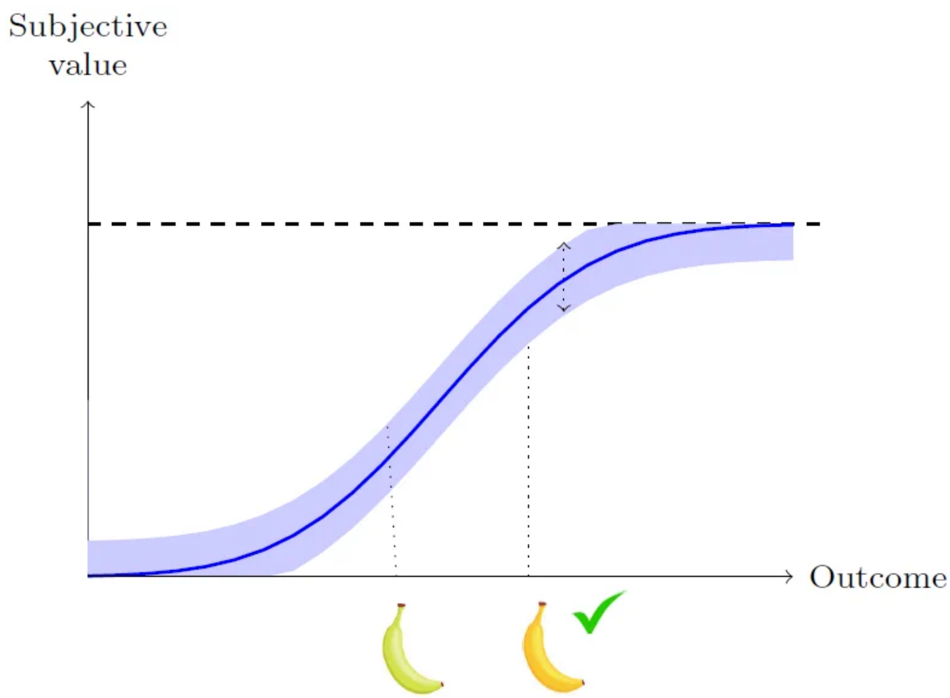
But brain (neurological processes, more generally) have constraints when generating signals of satisfaction:

- Signals must be bounded because there is a limited number of neurons to process them
- Signals are not perfectly precise (drugs as hijacking)

Hence: we are more likely to make mistakes when options are close.

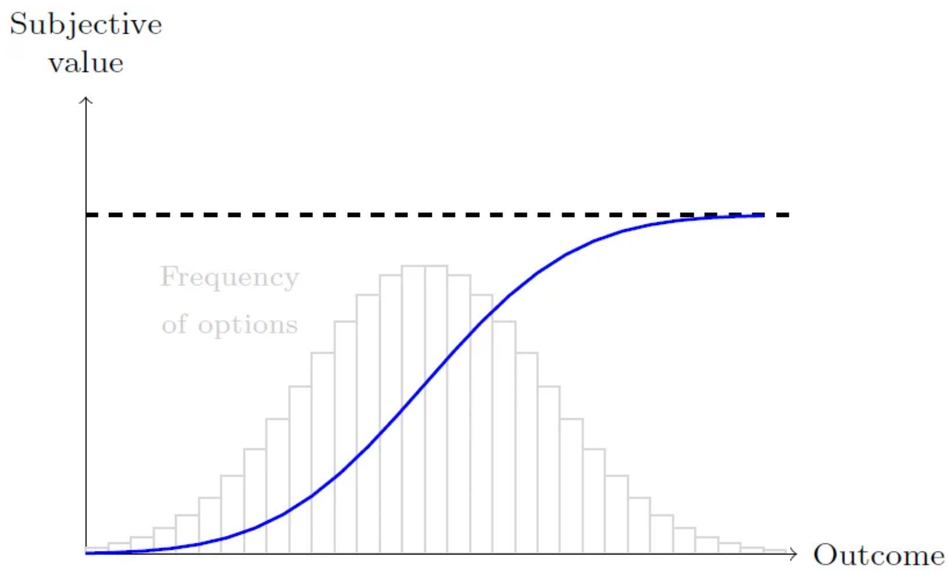


Can your system of perception be improved to reduce mistakes?
Yes. When the slope of subjective satisfaction increases, it reduces mistakes between options that are close.



But satisfaction is bounded: physical limit to pleasure (that is, to informativeness of signals). Then satisfaction cannot have a super high slope everywhere.

The question becomes: Where should the slope of subjective satisfaction be steeper to limit mistakes? The optimal solution is that it should be steeper where you are more likely to face options to choose from!



See also Rayo and Becker (2007) and Netzer (2009).

- Support from neuroscience: sensory systems respond to stimuli by following their distribution.
- Reference-dependent preferences are not a cognitive flaw. "They are an optimal solution, under irreducible biological constraints faced by our perceptual systems." $\rightarrow$ Efficient coding.
- Reference point is an expectation.
- Recent literature in economics and neuroscience.

In short: again, we have the problem *"Choose x to solve* $\max u(x)$ *subject to restrictions on x"*. Restrictions that don't show up are as relevant as those that show up.

Far from the only example... in fact, hard to think of economics without this structure (and hard in social sciences in general). What decisions are automatic, and what are well-thought? (Do you think about where to brush your teeth every day?)

Some difficult words

This setup is an approach based on **optimization** and **equilibrium**. Sometimes we talk about **rational** agents. All these words have multiple meanings and lead to confusion.

- Optimization simply means that individuals make choices according to some criteria, given the relevant restrictions. But sometimes used as 'perfect optimization' or 'hyper rationality', which needs not be the case.

- Rationality has different uses within microeconomics. First: similar to optimization under restrictions. Second: a particular set of basic assumptions on the decision-maker. We will use the latter. Sometimes (very often!) used in the sense of 'super computational power'. This is usually the case in basic microeconomics: we only have a budget constraint, meaning there are no cognitive constraints. This is often referred to as 'homo economicus' of 'homo rationalis' and is simply our lab rat: we study it not because it's realistic, but because it's much easier to study, and many things we learn carry over to 'actual' humans.
- Equilibrium means there is no unilateral incentive to change in a given context. Sometimes used, even within economics, in the sense of physics: lack of movement, or some very stable movement. This is NOT the meaning of the word in microeconomics.

1.3 Utility Maximization

Basic model of individual choice:

- A decision-maker (DM) must choose one alternative x from a set X.
- Chooses to maximize a utility function u.
- u specifies how much utility DM gets from each alternative: $u : X \to \mathbb{R}$

Example: DM chooses whether to eat an apple or a banana.

$$X = \{apple, banana\}$$

Utility function might say $u(apple) = 7, u(banana) = 12$. Observe that we already started to use mathematics – but only as language.

What do Utility Levels Mean? Hedonic Interpretation

Utility is an objective measure of individual's well-being.

> *Nature has placed mankind under the governance of two sovereign masters, pain and pleasure. It is for them alone to point out what we ought to do... By the principle of utility is meant that principle which approves or disapproves of every action whatsoever according to the tendency it appears to have to augment or diminish the happiness of the party whose interest is in question: or, what is the same thing in other words to promote or to oppose that happiness. I say of every action whatsoever, and therefore not only of every action of a private individual, but of every measure of government.*
>
> Jeremy Bentham

"$u(apple) = 7, u(banana) = 12$" $\rightarrow$ apple gives 7 units of pleasure, banana gives 12 units of pleasure. **This is not the standard way economists think about utility.**

What do Utility Levels Mean? Revealed-Preference Interpretation

Utility represents an individual's choices.

- Individual choices are primitive data that economists can observe.
- Choices are taken to reveal individual's preferences.
- Utility is a convenient mathematical construction for modeling choices and preferences.

> "$u(apple) = 7, u(banana) = 12$" $\rightarrow$ individual prefers bananas to apples.
> "$u(apple) = 2, u(banana) = 15$" $\rightarrow$ individual prefers bananas to apples.

1.4 Choice

How can an individual's choices reveal her preferences? A choice structure (or choice dataset) $(\mathcal{B}, C)$ consists of:

1. A set $\mathcal{B}$ of choice sets $B \subseteq X$.
2. A choice rule C that maps each $B \in \mathcal{B}$ to non-empty set of chosen alternatives $C(B) \subseteq B$. C is a correspondence.

Interpretation: $C(B)$ is the set of alternatives the DM might choose from B.

1.5 Preference

Goal: relate observable choice data to preferences over X.

A preference relation $\succcurlyeq$ is a binary relation on X.

> Meaning
>
> "$x \succcurlyeq y$" means "*x is weakly preferred to y*"

Given preference relation $\succcurlyeq$, define:

- Strict preference $(\succ) : x \succ y \iff x \succcurlyeq y$ but not $y \succcurlyeq x$.
- Indifference $(\sim) : x \sim y \iff x \succcurlyeq y$ and $y \succcurlyeq x$.

Think a little bit about logic and set theory here.

Rational Preferences

To make any progress, need to impose some restrictions on preferences.
Most important: rationality.

Definition 1.1. A preference relation $\succsim$ is rational if it satisfies:

1. Completeness: for all $x, y, x \succsim y$ or $y \succsim x$.
2. Transitivity: for all x, y, z, if $x \succsim y$ and $y \succsim z$, then $x \succsim z$.

If $\succsim$ is rational, then $\succ$ and $\sim$ are also transitive. (Prove this!)

Hard to say much about behavior of irrational DM.

Maximizing a Preference Relation

Optimal choices according to $\succsim$:

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

$\succsim$ rationalizes choice data $(\mathcal{B}, C)$ if $C(B) = C^*(\mathcal{B}, \succsim)$ for all $B \in \mathcal{B}$.

Fundamental Question of Revealed Preference Theory

When does choice data reveal that individual is choosing according to rational preferences?

Definition 1.2. *Given choice data $(\mathcal{B}, C)$, the revealed preference relation $\succsim^*$ is defined by $x \succsim^* y \iff$ there is some $B \in \mathcal{B}$ with $x, y \in B$ and $x \in C(B)$.*

- *x is weakly revealed preferred to y if x is ever chosen when y is available. Notice that this allows for $y \in C(B)$ as one may have $x \sim y$.*
- *x is strictly revealed preferred to y if there is some $B \in \mathcal{B}$ with $x, y \in B, x \in C(B)$ and $y \notin C(B)$.*

WARP

Key condition on choice data for $\succsim^*$ to be rational and generate observed data: **weak axiom of revealed preference (WARP)**.

Definition 1.3. *Choice data (B, C) satisfies WARP if whenever there exists $B \in \mathcal{B}$ with $x, y \in B$ and $x \in C(B)$, then for all $B' \in \mathcal{B}$ with $x, y \in B'$, it is not the case that both $y \in C(B')$ and $x \notin C(B')$.*

> Meaning
>
> *"If x is weakly revealed preferred to y, then y cannot be strictly revealed preferred to x"*

> Example
>
> $$X = \{x, y, z\}$$
>
> $$\mathcal{B} = \{\{x, y\}, \{x, y, z\}\}$$
>
> - Choice rule $C_1 : C_1(\{x, y\}) = \{x\}, C_1(\{x, y, z\}) = \{x\}$
>   Satisfies WARP: x is weakly revealed preferred to y and z, nothing is strictly revealed preferred to x
> - Choice rule $C_2 : C_2(\{x, y\}) = \{x\}, C_2(\{x, y, z\}) = \{x, y\}$.
>   Violates WARP: y is weakly revealed preferred to x, x is strictly revealed preferred to y. This is Exercise 1C1 (MWG).

Fundamental Theorem of Revealed Preference Theory

Theorem 1.1. If choice data $(\mathcal{B}, C)$ satisfies WARP and includes all subsets of X of up to 3 elements, then $\succsim^*$ is rational and rationalizes the data: that is, $C^*(B, \succsim^*) = C(B)$. Furthermore, this is the only preference relation that rationalizes the data (MWG Proposition 1.D.2). Conversely, if the choice data violates WARP, then it cannot be rationalized by any rational preference relation. (MWG Proposition 1D1).

For the first part: Remember $\mathcal{B}$ is a set of sets: this condition states that it must include all sets of up to three elements. Check MWG example 1D1 to see that we cannot drop this assumption.

! Proof

- Let's prove the first part.

We need to show that:

i. $\succsim^*$ is rational;
ii. $C^*(B, \succsim^*) = C(B)$;
iii. $\succsim^*$ is the only preference relation that satisfies ii.

For item (i), we must show that $\succsim *$ is complete and transitive.
Complete. Take some $\{x, y\} \in \mathcal{B}$. This holds because $\{x, y\}$ has only two elements. Then either $x \in C(x, y)$ or $y \in C(\{x, y\})$ (or both). In the first case, $x \succsim^* y$. In the second case, $y \succsim^* x$. Hence $\succsim^*$ is complete.
Transitive. Take $x \succsim^* y$ and $y \succsim^* z$. Consider $\{x, y, z \ \in \mathcal{B}\}$ (again, it has no more than three elements, so it belongs to $\mathcal{B}$). We have to show that $x \in C(\{x, y, z\})$ because this implies $x \succsim^* y$: transitivity.
We know that $C(\{x, y, z\}) \neq \emptyset$.

1. If $y \in C(\{x, y, z\})$: since $x \succsim^* y$, the weak axiom yields $x \in C(\{x, y, z\})$.
2. If $z \in C(\{x, y, z\})$: since $y \succsim^* z$, the weak axiom yields $y \in C(\{x, y, z\})$, and from the previous line we have $x \in C(\{x, y, z\})$. In any case, $x \in C(\{x, y, z\})$, as we wanted to show.

This concludes the proof of item (i).
For item (ii), we proceed in two steps.
First: suppose $x \in C(B)$. Then $x \succsim^* y$ for all $y \in B$. Hence $x \in C^*(B, \succsim^*)$. In short: every element x that belongs to $C(B)$ also belongs to $C^*(B, \succsim^*)$. In other words, $C(B) \subset C^*(B, \succsim^*)$.
Second: suppose now $x \in C^*(B, \succsim^*)$.
Then $x \succsim^* y$ for all $y \in B$, as above.
Hence for each $y \in B$, there exists some $B_y \in \mathcal{B}$ such that $x, y \in B_y$ and $x \in C(B_y)$: at the very least, one may choose $B_y = \{x, y\}$, which has only two elements and therefore belongs to $\mathcal{B}$.
Since $C(B) \neq \emptyset$, the weak axiom implies $x \in C(B)$, by the same reasoning as in part i: whatever y may belong to $C(B)$, it cannot be revealed as preferred to x because $x \in C(B_y)$ and hence $x \in C(B)$.
In short: $x \in C^*(B, \succsim^*) \implies x \in C(B)$, or $C^*(B, \succsim^*) \subset C(B)$.
Taking these two steps together, we conclude that $C(B) = C^*(B, \succsim^*)$, finishing the proof of item ii.
For item (iii), remember that $\mathcal{B}$ includes all two-element subsets of X. Hence the choice structure $C(\quad)$ determines the pairwise preference over X of any rationalizing preference.
QED.

- The second part may be written as: if preferences are rational, then choice data $(\mathcal{B}, C)$ satisfies WARP. Let's prove this.

Consider $B \in \mathcal{B}$ such that $x, y \in B$, and $x \in C^*(B, \succsim)$. Then $x \succsim y$. ("x is weakly revealed preferred to y")
Consider $B' \in \mathcal{B}$ such that $x, y \in B'$, and $y \in C^*(B', \succsim)$. Then $y \succsim z$ for all $z \in B'$.
Transitivity then implies $x \succsim z$ for all $z \in B'$.
But this is the same as saying $x \in C^*(B', \succsim)$.
That is, we cannot find any B' such that $y \succ x$. ("y) cannot be strictly revealed preferred to x"). Hence, WARP is satisfied.
QED.

Theorem tells us how individual's choices reveal her preferences: as long as choices satisfy WARP, can interpret choices as resulting from maximizing a rational preference relation.

We may conclude that if $\mathcal{B}$ includes all subsets of X, then choice and preferences work together just fine. But this is too restrictive: think of budget sets.

We use then the Strong Axiom of Revealed Preference, a "recursive closure" of the weak axiom. If x is directly or indirectly revealed preferred to y, then y cannot be directly revealed preferred to x. The Strong Axiom is more restrictive in general than the Weak Axiom (but they are equivalent for two goods). The Strong Axiom is a necessary and sufficient condition for choices to be generated by rational preferences.

Question: choices usually follow WARP?

Yes – even more than that. Bedi and Burghart (2018):

> *"Choices made under the influence of THC, MDMA, and placebo were all GARP compliant. Thus, even when participants were acutely intoxicated with THC or MDMA, their choices remained consistent with the tenets of neoclassical choice theory."*

GARP is the Strong Axiom that allows for non-unique optimal choices.

Preference and Utility

Now that know how to infer preferences from choice, next step is representing preferences with a utility function.

Definition 1.4. *A utility function $u : X \to \mathbb{R}$ represents preference relation $\succsim$ if, for all x, y,* $x \succsim y \iff u(x) \geq u(y)$

banana $\succcurlyeq$ *apple* is represented by both:

$$u(apple) = 7, \quad u(banana) = 12$$
$$u(apple) = 2, \quad u(banana) = 15$$

If u represents $\succcurlyeq$, so does any strictly increasing transformation of u.

Representing a given preference relation is an ordinal property. The numerical values of utility are cardinal properties.

What Preferences have a Utility Representation?

Theorem 1.2. Only rational preferences relations can be represented by a utility function (MWG Proposition 1B2). Conversely, if X is finite, any rational preference relation can be represented by a utility function (MWG exercise 1B5 - 'X finite' is only one possibility).

! Proof

- Let's prove the first part.

We may write it as: $\succcurlyeq$ is represented by utility function implies $\succcurlyeq$ is rational.
To show it's rational, we have to show $\succcurlyeq$ are complete and transitive.
Let's show first $\succcurlyeq$ are complete. Consider $x, y \in X$.
$u(.) \in \mathbb{R}$ implies that necessarily either $u(x) \geq u(y)$ or $u(y) \geq u(x)$. In the first case, by definition of u, we have $x \succcurlyeq y$. Analogously, in the second case we have $y \succcurlyeq x$. Hence $\succcurlyeq$ is complete.
Let's show now $\succcurlyeq$ are transitive. Take $x, y, z \in X$ such that $x \succcurlyeq y$ and $y \succcurlyeq z$. We have to show that $x \succcurlyeq z$.
By definition of $u(.)$, it follows that $u(x) \geq u(y)$ and $u(y) \geq u(z)$. It then follows from the structure of the real numbers that $u(x) \geq u(z)$. Again, use the definition of $u(.)$ to conclude that $x \succcurlyeq z$. This is what we wanted to show, concluding the proof.
QED.

What Goes Wrong with Infinitely Many Alternatives?

Lexicographic preferences: dictionary system – e.g., "I'm not going by plane".

$$X = [0, 1] \times [0, 1]$$

$(x_1, x_2) \succcurlyeq (y_1, y_2)$ if either

- $x_1 > y_1$ or
- $x_1 = y_1$ and $x_1 \geq y_2$

Maximize first component. In case of tie, maximize second component.

Theorem 1.3. Lexicographic preferences cannot be represented by a utility function.

This is on MWG page 46.

> **!** Proof
>
> Let's prove this by contradiction. Assume there is a utility function representing the lexicographic preferences $\succcurlyeq$.
> Fix some $x_1 \in \mathbb{R}$. Then:
>
> $$u(x_1, 2) > u(x_1, 1)$$
>
> Both $u(x_1, 2)$ and $u(x_1, 1)$ are real number. We will use the following mathematical result: we can find a rational number between any two real numbers.
> Let's call this rational number $r(x_1)$:
>
> $$u(x_1, 2) > r(x_1) > u(x_1, 1)$$
>
> Notice now that if $x_1 > x_1^{'}$, then $r(x_1) > u(x_1, 1) > u(x_1^{'}, 2) > r(x_1^{'})$.
> In short, $x_1 > x_1^{'} \implies r(x_1) > r(x_1^{'})$.
> This means that $r(.)$ is a strictly increasing function, and hence it is a bijection from $\mathbb{R}$ to $\mathbb{Q}$.
> But this is not possible.
> *QED.*

Continuous Preferences

What if rule out discontinuous preferences?

Definition 1.5. *For $X \subseteq \mathbb{R}^n$, preference relation $\succcurlyeq$ is continuous if whenever $x^m \to x, y^m \to y$, and $x^m \succcurlyeq y^m$ for all m, we have $x \succcurlyeq y$.*

> **!** Proof
>
> Lexicographic preferences are not continuous: see example 3C1 cont.
> Let's show this. Consider two sequences of bundles:

$$x_n = \left(\frac{1}{n}, 0\right)$$

$$y_n = (0, 1)$$

For any n we choose, we have $1/n > 0$, and hence $x_n \succ y_n$. But $lim_{n\to\infty} x_n = (0,0) \prec (0,1) = lim_{n\to\infty} y_n$
That is, preference reverts in the limit: continuity does not hold, and hence $\succcurlyeq$ are not continuous.
QED.

Theorem 1.4. For $X \subseteq \mathbb{R}^n$, any continuous, rational preference relation can be represented by a (continuous) utility function.

This is MWG Proposition 3.C.1 – a bit advanced.

! Proof

The general proof is difficult. Let's show a sketch, assuming additionally that preferences are monotone.
Consider only two goods (this is without loss of generality). For any $\alpha \geq 0$, define the bundle (α, α).
Pick some $x = (x_1, x_2) \in \mathbb{R}^2_+$. Notice that monotonicity implies $x \succcurlyeq (0,0)$.
Notice also that if $(\overline{\alpha}, \overline{\alpha}) \gg (x_1, x_2)$, then monotonicity implies $(\overline{\alpha}, \overline{\alpha}) > (x_1, x_2)$.
One may show that there is only one value of $\alpha \in [0, \overline{\alpha}]$ such that indifference holds.
We will call it $\alpha(x)$:

$$(\alpha(x), \alpha(x)) \sim (x_1, x_2)$$

Take $u(x) = \alpha(x)$. This is our utility function.
QED.

Notice that this is very general. Assumptions are not very restrictive (not even $X \subseteq \mathbb{R}^n$).

Review of Revealed Preference Theory

- If choice data satisfies WARP, can interpret as resulting from maximizing a rational preference relation.
- If set of alternatives is finite or preferences are continuous, can represent these preferences with a utility function.
- Utility function is just a convenient mathematical representation of individual's ordinal preferences.
- Utility may or may not be correlated with pleasure/avoidance of pain.

2 Properties of Preferences and Utility Functions

Doing useful analysis entails making assumptions. Try to do this carefully: make clearest, simplest, least restrictive assumptions. Understand what assumptions about utility correspond to in terms of preferences, since utility is just a way of representing preferences. We now cover some of the most important assumptions on preferences (and, implicitly, on choices).

Setting/Notation

For rest of lecture, assume $X \subseteq \mathbb{R}^n$.

Example

Consumer Problem: given fixed budget, choose how much of n goods to consume.
Notation: For vectors $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$,

- $x \geq y$ means $x_k \geq y_k$ for all $k = 1, ..., n$
- $x > y$ means $x_k \geq y_k$ for all k and $x_k > y_k$ for some k
- $x \gg y$ means $x_k > y_k$ for all k

For example, $(2, 3, 4) > (3, 3, 4)$ and $(4, 4, 5) \gg (3, 3, 4)$

- For $\alpha \in [0, 1]$,

$$\alpha x + (1 - \alpha)y = (\alpha x_1 + (1 - \alpha)y_1, \dots, \alpha x_n + (1 - \alpha)y_n)$$

This is a convex combination for each coordinate.

2.1 Monotonicity

Meaning

"All goods are desirable"

Preferences

Definition 2.1 (MWG 3B2)**.** $\succsim$ *is monotone if* $x \geq y$ *implies* $x \succsim y$

$\succsim$ *is strictly monotone if* $x > y$ *implies* $x \succ y$

For example, strict monotonicity implies $(2, 3, 4) \succ (1, 3, 4)$.

Utility

If preferences are monotone, what does that mean for the utility function?

Theorem 2.1 (MWG Exercise 3B1)**.** Suppose utility function u represents preferences $\succsim$. Then:

u non-decreasing $\iff \succsim$ monotone
u strictly increasing $\iff \succsim$ strictly monotone

Proof

Let's prove this.

- For the first part:

u non-decreasing $\iff$
$[x \geq y \iff u(x) \geq u(y)] \iff$
$[x \geq y \iff x \succsim y] \iff$
$\succsim$ monotone
The third line uses the definition of utility function: $u(x) \geq u(y)$ if and only if $x \succsim y$.

- Analogously for the second part:

u strictly increasing $\iff$
$[x > y \iff u(x) > u(y)] \iff$
$[x > y \iff x \succ y] \iff$
$\succsim$ striclty monotone
QED.

2.2 Local-Nonsatiation

Meaning

"No bliss points" (not even local ones)

Let $B_\varepsilon(x) = \{y : |x - y| < \varepsilon\}$

Definition 2.2. *$\succcurlyeq$ is locally non-satiated if for any x and $\varepsilon > 0$ there exists $y \in B_\varepsilon(x)$ with $y \succ x$*

If u represents $\succcurlyeq$, then t is locally non-satiated if and only if u has no local maximum. (Prove this!)

2.3 Convexity

Meaning

"Diversity is good"

Definition 2.3. *$\succcurlyeq$ is convex if $x \succcurlyeq y$, $x' \succcurlyeq y$ and $\alpha \in [0, 1]$ imply:*

$$\alpha x + (1 - \alpha)x' \succcurlyeq y$$

$\succcurlyeq$ is strictly convex if $x \succcurlyeq y$, $x' \succcurlyeq y$, $\alpha \in (0, 1)$ and $x \neq x'$ imply:

$$\alpha x + (1 - \alpha)x' \succ y$$

Does this make sense? Is $\left(\frac{1}{2}\right)$ *beer* + $\left(\frac{1}{2}\right)$ *wine* a good thing?
We now discuss several properties of convex preferences.

Contour Sets

For $x \in X$, the upper contour set of x is

$$S(x) = \{y \in X : y \succcurlyeq x\}$$

Theorem 2.2. $\succcurlyeq$ is convex $\iff$ $U(s)$ is a convex set for every $x \in X$.

! Proof

- Proof of sufficiency ($\Rightarrow$):

  Assume $\succcurlyeq$ convex. We have to show that $S(x)$ is convex. To do so, take any two elements $y, y' \in S(x)$.

  By definition of $S(x)$, $y, y' \in S(x)$ means that $y \succcurlyeq x$ and $y' \succcurlyeq x$

  Convexity then implies that for all $\alpha \in [0,1]$, $\alpha y + (1-\alpha)y' \succcurlyeq x$.

  But this implies that $\alpha y + (1-\alpha)y' \in U(x)$.

  In short: we showed that $y, y' \in S(x) \Rightarrow \alpha y + (1-\alpha)y' \in U(x)$ for all $\alpha \in [0,1]$.

  This means that $S(x)$ is convex.

- Proof of necessity ($\Leftarrow$):

  We will prove by contradiction. Assume $S(x)$ is convex, but $\succcurlyeq$ is not convex. We have to find an absurd conclusion.

  Assume $\succcurlyeq$ is not convex. Then there are $\alpha \in (0,1), y, y'$ and x such that $y \succcurlyeq x, y' \succcurlyeq x$ but $x \succ \alpha y + (1-\alpha)y'$. This implies that $S(x)$ is not convex.

  In short: we showed that if $\succcurlyeq$ is not convex, then $S(x)$ is not convex for some x. This is equivalent to showing that $S(x) \Rightarrow \succcurlyeq$ convex.

QED.

That's why convex preferences are called convex: for every x, the set of all alternatives preferred to x is convex. When the context is clear, we will write simply S instead of $S(x)$, and interpret it as the set of alternatives preferred to some unspecified x.

Set of Maximizers

Theorem 2.3. If $\succcurlyeq$ is convex, then for any convex choice set B, the set $C^*(B, \succcurlyeq)$ is convex. If $\succcurlyeq$ is convex, then for any convex choice set B, the set $C^*(B, \succcurlyeq)$ is single-valued (for empty).

! Proof

- First part:

  Take any $x \in B$

  If $y, y' \in C^*(B, \succcurlyeq)$, then $y \succcurlyeq x$ and $y' \succcurlyeq x$.

  Convexity then implies that for all $\alpha \in [0,1]$, $\alpha y + (1-\alpha)y' \succcurlyeq x$.

  Since this holds for any $x \in B$, it follows that $\alpha \in [0,1]$, $\alpha y + (1-\alpha)y' \in C^*(B, \succcurlyeq)$.

That is, $C^*(B, \succcurlyeq)$ is convex.

- Second part:

  Assume there are $x, x' \in C^*(B, \succcurlyeq)$ with $x \neq x'$.

  Then $x \succcurlyeq x$ and $x' \succcurlyeq x$: definition of optimal choices.

  Strict convexity then implies that for all $\alpha \in [0,1]$, one has $\alpha x + (1-\alpha)x' \succ x$.

  But the first part of the theorem implies $\alpha x + (1-\alpha)x' \in C^*(B, \succcurlyeq)$.

In words: if there are two different optimal choices with strictly convex preferences, then it is possible to find an alternative that is strictly better than at least one of them. This is absurd as the DM should have chosen this alternative.
QED.

Convexity: Utility Functions

The characteristic of utility functions that represent convex preferences is quasi-concavity.

Definition 2.4. *A function $u : X \to \mathbb{R}$ is quasi-concave if, for every x, y with $u(x) \geq u(y)$ and every $\alpha \in (0,1)$,*

$$u(\alpha x + (1-\alpha)y) \geq u(y)$$

A function $u : X \to \mathbb{R}$ is strictly quasi-concave if, for every x, y with $u(x) \geq u(y), x \neq y$ and every $\alpha \in (0,1)$,

$$u(\alpha x + (1-\alpha)y) > u(y)$$

Theorem 2.4. u is quasi-concave $\iff$ for every $r \in \mathbb{R}$ the upper contour set $S = \{x \in X : u(x) \geq r\}$ is convex.

Notation: this is the same set of 'preferred alternatives' we used before. We may write $S(x)$ if we want to highlight that these alternatives are preferred to some specific x. Analogously, we may write $S(r)$ to highlight that these alternatives achieve a level of utility no lower than r. If the alternative x or the choice r are generic, we may write simply S.

! Proof

- Proof of sufficiency ($\Rightarrow$):

Take any $r \in \mathbb{R}$.

Take two elements of $S = \{x \in X : u(x) \geq r\}$. That is, take x, x' such that $u(x), u(x') \geq r$.
Assume without loss of generality that $u(x) \geq u(x')$.
Then:

$$u(\alpha x + (1-\alpha)x') \underbrace{\geq}_{quasi-concavity} u(x') \geq r$$

This implies $\alpha x + (1-\alpha)x' \in \{x \in X : u(x) \geq r\}$
This holds for any two elements in the upper contour set $S = \{x \in X : u(x) \geq r\}$, and any $\alpha \in [0,1]$. It follows that this set is convex.

- Proof of necessity ($\Leftarrow$):

Assume there is some $r \in \mathbb{R}$, some $\alpha \in (0,1)$ and $x, x' \in S = \{x \in X : u(x) \geq r\}$ such that:

$$\alpha x + (1-\alpha)x' \notin S = \{x \in X : u(x) \geq r\}$$

That is, S is not convex. This meand that $u(x) \geq r, u(x') \geq r$ but $u(\alpha x + (1-\alpha)x') < r$, for some $r \in \mathbb{R}$ and some $\alpha \in (0,1)$.
This means that u is not quasi-concave.
In short, we showed that if S is not convex for all $r \in \mathbb{R}$, then u is not quasi-concave. This is equivalent to showing that if u is quasi-concave, then S is convex for all u.
QED.

Theorem 2.5. Suppose utility function u represents preferences $\succcurlyeq$. Then:

u quasi-concave $\iff \succcurlyeq$ convex

u strictly quasi-concave $\iff \succcurlyeq$ strictly convex

Exercise

Prove this theorem. Proof follows directly from Theorem 2.4.

Warning: convex preferences are represented by quasi-concave utility functions. Convex preferences get that name because they make upper contour sets convex. Quasi-concave utility functions get that name because quasi-concavity is a weaker property than concavity.

2.4 Separability

Often very useful to restrict ways in which a consumer's preferences over one kind of good can depend on consumption of other goods. If allowed arbitrary interdependencies, would need to observe consumer's entire consumption bundle to infer anything. Properties of preferences that separation among different kinds of goods are called *separability properties*.

Weak Separability: Preferences

Meaning

"Preferences over one kind of goods don't depend on what other goods are consumed"

Definition 2.5. *$\succcurlyeq$ is weakly separable in J_1, J_2 if for $k = 1, 2$ and for every x_{J_k} and x'_{J_k} and for every $x_{J_k^c}$ and $x'_{J_k^c}$ one has:*

$$(x_{J_k}, x_{J_k^c}) \succcurlyeq (x'_{J_k}, x_{J_k^c}) \iff (x_{J_k}, x'_{J_k^c}) \succcurlyeq (x'_{J_k}, x'_{J_k^c})$$

May extend for than two subsets of X.

Weak Separability: Utility

Theorem 2.6. Preferences are weakly separable in J_1, J_2 if and only if utility function (if it exists) is:

$$u(x) = v(u_1(x_{J_1}), u_2(x_{J_2}), u(x_{(J_1 \cup J_2)^C}))$$

Other Kinds of Separability

- *Strong separability: utility is additively separable*

$$u(x) = u_1(x_{J_1}) + u_2(x_{J_2})$$

- *Quasi-linear utility*

$$u(x) = x_1 + v(x_2, \dots, x_n)$$

Part II

Lectures 3 – 4

3 Consumer Theory

Consumer theory studies how rational consumer chooses what bundle of goods to consume.

Rational has a new meaning now: unbounded rationality!

Special case of general theory of choice.

Key new assumption: choice sets defined **ONLY** by prices of each of n goods, and income (or wealth).

3.1 Consumer Problem (CP)

$$\max_{x \in \mathbb{R}^n_+} \quad u(x)\, s.t. \quad p \cdot x \leq w$$

Restrictions that don't show up are the most important ones!

Notation: $p \in \mathbb{R}^n$, $x \in \mathbb{R}^n$, $p \cdot x = p_1 x_1 + ... + p_n x_n$ (inner product)

- Consumer chooses consumption vector $x = (x_1, \ldots, x_n)$
- x_k is consumption of good k
- Each unit of good k costs p_k
- $p \cdot x$ is total expenditure
- Total available income is w

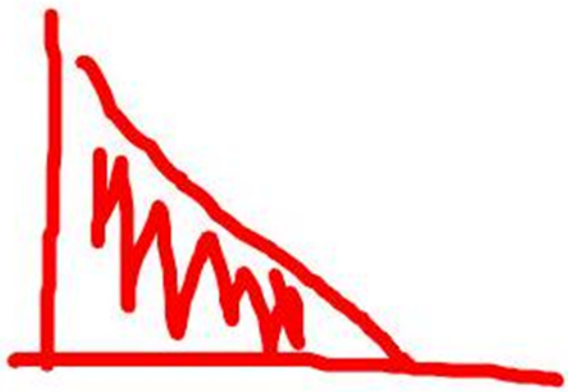
Now discuss some implicit assumptions underlying (CP).

3.1.1 First: prices are Linear

Each unit of good k costs the same.

No quantity discounts or supply constraints.

Consumer's choice set (or budget set) is:



$$B(p, w) = \{x \in \mathbb{R}^n_+ : p \cdot x \leq w\}$$

Set is defined by single line (or hyperplane): the budget line

$$p \cdot x = w$$

Assume $p \geq 0$.

(If we're talking about bad things, so that price cannot be positive, think of garbage collection with positive price)

3.1.2 Second: Goods are Divisible

$x \in \mathbb{R}^n_+$ and consumer can consume any bundle in budget set

Can model indivisibilities by assuming utility only depends on integer part of x.

3.1.3 Third: Set of Goods is Finite ($n < \infty$)

This is not obvious: think of a dynamic economy with no known final date.

Debreu (1959): *A commodity is characterized by its physical properties, the date at which it will be available, and the location at which it will be available.*

In practice, set of goods suggests itself naturally based on context.

3.2 Marshallian Demand

The solution to the (CP) is called the Marshallian demand (or Walrasian demand).

May be multiple solutions, so formal definition is:

Definition 3.1. The Marshallian demand correspondence $x : \mathbb{R}_+^n \times \mathbb{R} \rightrightarrows \mathbb{R}_+^n$ is defined by

$$\begin{aligned} x(p,w) &= argmax_{x \in B(p,w)}\ u(x) \\ &= \left\{ z \in B(p,w) : u(z) = \max_{x \in B(p,w)} \ u(x) \right\} \end{aligned}$$

Heavy notation for simple idea!

Domain is $\mathbb{R}_+^n \times \mathbb{R}$: n prices, one level of income.

Start by deriving basic properties of budget sets and Marshallian demand.

Example: Cobb-Douglas Marshallian demand: $u\left(x_1, x_2\right) = x_1^\alpha x_2^{1-\alpha}$

$$x_1\left(p_1, p_2, w\right) = \alpha \frac{w}{p_1} x_2\left(p_1, p_2, w\right) = (1-\alpha)\frac{w}{p_2}$$

3.3 Budget Sets

Theorem 3.1. *Budget sets are homogeneous of degree 0 : that is, for all* $\lambda > 0$, $B(\lambda\, p, \lambda\, w) = B(p,w)$.

! Proof

$$\begin{aligned} B(\lambda p, \lambda w) &= \{x \in \mathbb{R}_+^n \mid \lambda p \cdot x \leq \lambda w\} \\ &= \{x \in \mathbb{R}_+^n \mid p \cdot x \leq w\} = B(p,w). \end{aligned}$$

Nothing changes if scale prices and income by same factor.
QED.

Theorem 3.2. *If* $p \gg 0$, *then* $B(p,w)$ *is compact.*

! "Proof" (Write it formally as an exercise)

A subset of $\mathbb{R}^n$ is compact if and only if it is closed and limited.
For any $p, B(p, w)$ is closed. (Notice the weak inequality in the definition of B.)
If $p \gg 0$, then $B(p, w)$ is also bounded.
QED.

3.4 Marshallian Demand: Existence

Theorem 3.3. *If u is continuous and $p \gg 0$, then (CP) has a solution.*

(That is, $x(p, w)$ is non-empty.)

! Proof

A continuous function on a compact set attains its maximum (Weirstrass theorem).
QED.

3.5 Marshallian Demand: Uniqueness?

The Marshallian Demand needs not be unique

Example: perfect substitutes – write it out as an exercise.

Generally, the Marshallian demand is a correspondence, or a set-valued function: for each (p, w), it associates a set of optimal choices $x\,(p, w)$.

We've seen before the following results (we'll just rewrite them in our context):

Since the budget set is convex, the Marshallian demand is a convex correspondence if preferences are convex.

The Marshallian demand is unique (that is, a function) if preferences are strictly convex.

3.6 Marshallian Demand: Homogeneity of Degree 0

Theorem 3.4. *For all $\lambda > 0$, $x(\lambda\ p, \lambda\ w) = x(p, w)$.*

! Important

$B(\lambda\, p, \lambda\, w) = B(p, w)$, so (CP) with prices $\lambda\, p$ and income $\lambda\, w$ is same problem as (CP) with prices p and income w, since utility function is not affected by λ.
QED.

3.7 Marshallian Demand: "Continuity"

Theorem 3.5. *Assume u is continuous and strictly quasi-concave (that is, preferences are strictly convex). Then the Marshallian demand $x(p, w)$ is a continuous function.*

More generally: if u is continuous (but not necessarily quasi-concave), then the Marshallian demand is upper hemicontinuous.

! Proof

Follows directly from the theorem of the maximum.

3.8 Marshallian Demand: Walras' Law

Reminder: Preferences are locally non-satiated if for all $x \in X$ and all $\varepsilon > 0$, there exists $y \succ x$ such that $y \in B_{\varepsilon}(x) = \{y \in X : d(x, y) < \varepsilon\}$.

Theorem 3.6. *If preferences are locally non-satiated, then for every (p,w) and every $x \in x(p, w)$, we have $p \cdot x = w$.*

! Proof

If $p \cdot x < w$, then there exists $\varepsilon > 0$ such that $B_{\varepsilon}(x) \subseteq B(p, w)$. By local non-satiation, for every $\varepsilon > 0$ there exists $y \in B_{\varepsilon}(x)$ such that $y \succ x$.
Hence, there exists $y \in B(p, w)$ such that $y \succ x$.
But then $x \notin x(p, w)$, that is, x is not an optimal choice: contradiction.
QED.

Walras' Law lets us rewrite (CP) as

$$\max_{x \in \mathbb{R}^n_+} \; u(x) \text{ s.t. } p \cdot x = w$$

3.9 Marshallian Demand: Differentiable Demand

Implications if demand is single-valued and differentiable:

- A proportional change in all prices and income does not affect demand:

$$\sum_{j=1}^{n} p_j \frac{\partial}{\partial p_j} x_i(p,w) + w \frac{\partial}{\partial w} x_i(p,w) = 0$$

- A change in the price of one good does not affect total expenditure:

$$\sum_{j=1}^{n} p_j \frac{\partial}{\partial p_i} x_j(p,w) + x_i(p,w) = 0.$$

- A change in income leads to an identical change in total expenditure:

$$\sum_{i=1}^{n} p_i \frac{\partial}{\partial w} x_i(p,w) = 1$$

3.10 The Indirect Utility Function

Can learn more about set of solutions to (CP) (Marshallian demand) by relating to the **value** of (CP).

Value of (CP) = welfare of consumer facing prices p with income w.

The value function of (CP) is called the indirect utility function.

Definition 3.2 (The Indirect Utility Function)**.** $v : \mathbb{R}_+^n \times \mathbb{R} \to \mathbb{R}$ is defined by $v(p,w) = \max_{x \in B(p,w)} u(x)$.

So we have $(p,w) \longrightarrow\ x(p,w) \longrightarrow\ u(x(p,w)) =\ v(p,w)$.

Notice that demand x is the image of (p,w) and also the argument of u.

Example: Cobb-Douglas indirect utility function:

$$u(x_1,x_2) = x_1^{\alpha} x_2^{1-\alpha}$$

$$x_1(p_1,p_2,w) = \alpha \frac{w}{p_1}$$

$$x_2(p_1,p_2,w) = (1-\alpha)\frac{w}{p_2}$$

$$v(p_1,p_2,w) = \left(\alpha\frac{w}{p_1}\right)^{\alpha}\cdot\left((1-\alpha)\frac{w}{p_2}\right)^{1-\alpha} = w\left(\frac{\alpha}{p_1}\right)^{\alpha}\cdot\left(\frac{(1-\alpha)}{p_2}\right)^{1-\alpha}$$

In our context, the Theorem of the Maximum implies:

Marshallian demand is upper hemicontinuous (if it's a function, it's continuous)

Indirect utility is continuous

3.11 Indirect Utility Function: Properties

Theorem 3.7. *The indirect utility function has the following properties:*

1. *Homogeneity of degree 0: for all* $\lambda > 0$, $v(\lambda\, p, \lambda\, w) = v(p, w)$.
2. *Continuity: if* u *is continuous, then* v *is continuous on* $\{(p, w) : p \gg 0, w \geq 0\}$.
3. *Monotonicty:* $v(p, w)$ *is non-increasing in* p *and non-decreasing in* w. *If* $p \gg 0$ *and preferences are locally non-satiated, then* $v(p, w)$ *is strictly increasing in* w.
4. *Quasi-convexity: for all* $\bar{v} \in R$, *the set* $\{(p, w) : v(p, w)\} \leq \bar{v}\}$ *is convex. (Consumer is worse off at average prices/income.)*

! Proof

1. Follows from homogeneity of degree zero of Marshallian demand.
2. Follows directly from the Theorem of the Maximum.
3. Left as an exercise.
4. Pick two elements in the domain of v: (p,w) *and* (p',w')

   Assume:

   $v(p,w) \leq \bar{v}$

   $v(p',w') \leq \bar{v}$

   Define $(p'',w'') = \alpha\cdot(p,w) + (1-\alpha)\cdot(p',w')$

   We need to show that $v(p'',w'') \leq \bar{v}$

   We will show something stronger: for all x such that $p''\cdot\ x'' \leq\ w''$, one has $u(x) \leq \bar{v}$.

Use the definition of (p'', w''):

$$p'' \cdot \ x'' \leq \ w'' \Leftrightarrow (\alpha p + (1-\alpha)\, p') \cdot \ x \leq \alpha \ w + (1-\alpha)\, w'$$

This holds if and only either $p \cdot \ x \leq \ w$ of $p' \cdot \ x \leq \ w'$.

Then we have:

$$p \cdot \ x \leq \ w \Rightarrow \ u\,(x) \leq \ v\,(p, w) \leq \bar{v} p' \cdot \ x \leq \ w' \Rightarrow \ u\,(x) \leq \ v\,(p', w') \leq \bar{v}$$

In any case, $u\,(x) \leq \bar{v}$, as we wanted to show.

QED.

Interpretation of quasi-convexity: consumer prefers extreme prices/income than average ones.

Extreme prices allow consumers to explore substitution: something desirable must be cheap enough.

(Income is one dimensional and hence it follows immediately that the average of two income levels is lower than the highest of them.)

GRAPHIC

3.12 Indirect Utility Function: Derivatives

When indirect utility function is differentiable, its derivatives are very interesting.

Q: When is indirect utility function differentiable? A: When u is (continuously) differentiable and Marshallian demand is unique.

Theorem 3.8. *Suppose (1) u is locally non-satiated and continuously differentiable, and (2) Marshallian demand is unique in an open neighborhood of (p, w) with $p \gg 0$ and $w\,0$. Then v is differentiable at (p, w).*

We'll skip the proof. For details if curious, see Milgrom and Segal (2002), "Envelope Theorems for Arbitrary Choice Sets."

Or check chapters 3 and 4 in Stokey and Lucas (1989).

Furthermore, letting $x = x(p, w)$, the derivatives of v are given by:

$$\frac{\partial}{\partial w} v(p, w) = \frac{1}{p_j} \frac{\partial}{\partial x_j} u(x)$$

and

$$\frac{\partial}{\partial p_i} v(p,w) = -\frac{x_i}{p_j}\frac{\partial}{\partial x_j} u(x),$$

where j is any index such that $x_j > 0$.

- Suppose consumer's income increases by \$1.
- Should spend this dollar on any good that gives biggest "bang for the buck."
- Bang for spending on good j equals $\frac{1}{p_j}\frac{\partial u}{\partial x_j}$: can buy $\frac{1}{p_j}$ units, each gives utility $\frac{\partial u}{\partial x_j}$.
- Finally, $x_j > 0$ for precisely those goods that maximize bang for buck. $\implies$ marginal utility of income equals $\frac{1}{p_j}\frac{\partial u}{\partial x_j}$, for any j with $x_j > 0$.

3.13 Indirect Utility Function: Derivatives

$$\frac{\partial}{\partial w} v(p,w) = \frac{1}{p_j}\frac{\partial}{\partial x_j} u(x) \frac{\partial}{\partial p_i} v(p,w) = -\frac{x_i}{p_j}\frac{\partial}{\partial x_j} u(x)$$

Suppose price of good i increases by \$1.

This effectively makes consumer x_i poorer.

Just saw that marginal effect of making \$1 poorer is $-\frac{1}{p_j}\frac{\partial u}{\partial x_j}$, for any j with $x_j > 0$.

$\implies$ marginal disutility of increase in p_i equals $-\frac{x_i}{p_j}\frac{\partial u}{\partial x_j}$, for any j with $x_j > 0$.

3.14 Kuhn-Tucker Theorem

Theorem 3.9 (Kuhn-Tucker)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable functions (for some $i \in 1, \dots, I$), and consider the constrained optimization problem*

$$\max_{x\in\mathbb{R}^n} \quad f(x) s.t. \quad g_i(x) \geq 0 \text{ for all } i$$

If x^* is a solution to this problem (even a local solution) and a condition called constraint qualification is satisfied at x^*, then there exists a vector of Lagrange multipliers $\lambda = (\lambda_1, \dots, \lambda_l)$ such that

$$\nabla f(x^*) + \sum_{i=1}^{l} \lambda_i \nabla g_i(x^*) = 0$$

and

$$\lambda_i \geq 0 \;\; \text{and} \;\; \lambda_i g_i(x^*) = 0 \text{ for all } i$$

3.15 Kuhn-Tucker Theorem: Comments

1. Any local solution to constrained optimization problem must satisfy first-order conditions of the Lagrangian

$$\mathcal{L}(x) = f(x) + \sum_{i=1}^{l} \lambda_i g_i(x)$$

2. Condition that $\lambda_i g_i(x^*) = 0$ for all i is called complementary slackness.

- Says that multipliers on slack constraints must equal 0.
- Consistent with interpreting λ_i as marginal value of relaxing constraint i.

3. There are different versions of constraint qualification. Simplest version: vectors $\nabla g_i\left(x^*\right)$ are linearly independent for binding constraints.

Exercise: check that constraint qualification is always satisfied in the (CP) when $p \gg 0, w > 0$, and preferences are locally non-satiated.

3.16 Lagrangian for (CP)

For two goods:

$$\mathcal{L}\left(x_1, x_2\right) = u\left(x_1, x_2\right) + \lambda\left[w - p_1 \cdot x_1 - p_2 \cdot x_2\right] + \mu_1 x_1 + +\mu_2 x_2$$

Generally:

$$\mathcal{L}(x) = u(x) + \lambda[w - p \cdot x] + \sum_{k=1}^{n} \mu_k x_k$$

$\lambda \geq 0$ is multiplier on budget constraint. $\mu_k \geq 0$ is multiplier on the constraint $x_k \geq 0$.

FOC with respect to x_i:

$$\frac{\partial u}{\partial x_i} + \mu_i = \lambda p_i$$

Complementary slackness: $\mu_i = 0 if x_i\, 0$. So:

$$\frac{\partial u}{\partial x_i} = \lambda p_i \text{ if } x_i > 0 \frac{\partial u}{\partial x_i} \leq \lambda p_i \text{ if } \ x_i = 0$$

What's the intuition of $\frac{\partial u}{\partial x_i} < \lambda p_i$?

Implication: marginal rate of substitution $\frac{\partial u}{\partial x_i} / \frac{\partial u}{\partial x_j}$ between any two goods consumed in positive quantity must equal the ratio of their prices p_i/p_j.

In other words: slope of indifference curve between goods i and j must equal slope of budget line.

Intuition: equal "bang for the buck" $\frac{1}{p_i}\frac{\partial u}{\partial x_i}$ among goods consumed in positive quantity.

3.17 Back to Derivatives of v

When v is differentiable, we have the following result:

Theorem 3.10.

$$\frac{\partial v}{\partial w} = \lambda \quad \text{(marginal utility of income)}$$

$$\frac{\partial v}{\partial p_i} = -\lambda x_i \quad \text{(marginal disutility of price)}$$

! Proof

- For the first part:

Without loss of generality, take only two goods: $x = (x_1, x_2)$, $p = (p_1, p_2)$.
Then:

$$v\left(p_1, p_2, w\right) = u\left(x_1\left(p_1, p_2, w\right), x_2\left(p_1, p_2, w\right)\right)$$

Take the derivative with respect to income w:

$$\frac{\partial v}{\partial w} = \frac{\partial u}{\partial x_1} \cdot \frac{\partial x_1}{\partial w} + \frac{\partial u}{\partial x_2} \cdot \frac{\partial x_2}{\partial w}$$

But $\frac{\partial u}{\partial x_i} = \lambda p_i$.

Rewrite the previous equation:

$$\frac{\partial v}{\partial w} = \lambda p_1 \cdot \frac{\partial x_1}{\partial w} + \lambda p_2 \cdot \frac{\partial x_2}{\partial w} = \lambda \cdot \left[p_1 \cdot \frac{\partial x_1}{\partial w} + p_1 \cdot \frac{\partial x_1}{\partial w} \right]$$

We also know that $p_1 x_1 + p_2 x_2 = w$ for all w.
This allows us to differentiate both sides with respect to w, and get:

$$p_1 \cdot \frac{\partial x_1}{\partial w} + p_2 \cdot \frac{\partial x_2}{\partial w} = 1$$

This is exactly the term in square brackets in the previous equation, which becomes:

$$\frac{\partial v}{\partial w} = \lambda p_1 \cdot \frac{\partial x_1}{\partial w} + \lambda p_1 \cdot \frac{\partial x_1}{\partial w} = \lambda \cdot \underbrace{\left[p_1 \cdot \frac{\partial x_1}{\partial w} + p_1 \cdot \frac{\partial x_1}{\partial w} \right]}_1 = \lambda$$

In short: $\frac{\partial v}{\partial w} = \lambda$

- For the second part:

$$v\left(p_1, p_2, w\right) = u\left(x_1\left(p_1, p_2, w\right), x_2\left(p_1, p_2, w\right)\right) + \lambda^* \cdot \left[w - p_1 x_1 - p_2 x_2\right]$$

This holds because we always have $\lambda^* \cdot [w - p_1 x_1 - p_2 x_2] = 0$ (Kuhn-Tucker).
Differentiate with respect to p_1:

$$\frac{\partial v}{\partial p_1} = \frac{\partial u}{\partial x_1} \cdot \frac{\partial x_1}{\partial p_1} + \frac{\partial u}{\partial x_2} \cdot \frac{\partial x_2}{\partial p_1} + \frac{\partial \lambda}{\partial p_1} \cdot \underbrace{\left[w - p_1 x_1 - p_2 x_2\right]}_0 - \lambda \cdot \left[x_1 + p_1 \cdot \frac{\partial x_1}{\partial p_1} + p_2 \cdot \frac{\partial x_2}{\partial p_1} \right]$$

Collect terms to get:

$$\frac{\partial v}{\partial p_1} = \frac{\partial x_1}{\partial p_1} \cdot \underbrace{\left[\frac{\partial u}{\partial x_1} - \lambda p_1 \right]}_0 + \frac{\partial x_1}{\partial p_1} \cdot \underbrace{\left[\frac{\partial u}{\partial x_1} - \lambda p_1 \right]}_0 - \lambda x_1 = -\lambda x_1$$

QED.

These are applications of the envelope theorem: ignore indirect effect of changes in parameters (that is, impact through changes in optimal decisions).

3.18 Envelope Theorem

Theorem 3.11 (Envelope Theorem). *For $\Theta \subseteq \mathbb{R}$, let $f : X \times \Theta \to \mathbb{R}$ be a differentiable function, let $V(\theta) = \max\_x \in X f(x, \theta)$, and let $X^*(\theta) = x \in X : f(x, \theta) = V(\theta)$. If V is* *differentiable at θ then, for any $x^* \in X^*(\theta), V'(\theta) = \frac{\partial}{\partial\theta} f(x^*, \theta)$.*

3.19 Back again to Derivatives of v

$$\frac{\partial v}{\partial w} = \lambda$$
$$\frac{\partial v}{\partial p_i} = -\lambda x_i$$

Combining with $\frac{\partial u}{\partial x_j} = \lambda p_j$ if $x_j > 0$, obtain

$$\begin{aligned} \frac{\partial v}{\partial w} &= \frac{1}{p_j}\frac{\partial u}{\partial x_j} \\ \frac{\partial v}{\partial p_i} &= -\frac{x_i}{p_j}\frac{\partial u}{\partial x_j} \end{aligned}$$

for any j with $x_j > 0$.

This proves above theorem on derivatives of v.

We've already seen the intuition.

3.20 Roy's Identity

Under conditions of last theorem, if $x_i(p, w) > 0$ then

$$x_i(p, w) = -\frac{\frac{\partial}{\partial p_i} v(p, w)}{\frac{\partial}{\partial w} v(p, w)}$$

3.21 Key Facts about (CP), Assuming Differentiability

- Consumer's marginal utility of income equals multiplier on budget constraint: $\frac{\partial v}{\partial w} = \lambda$.
- Marginal disutility of increase in price of good i equals $-\lambda x_i$.
- Marginal utility of consumption of any good consumed in positive quantity equals λp_i.

3.22 The Expenditure Minimization Problem

In (CP), consumer chooses consumption vector to maximize utility subject to maximum budget constraint.

Also useful to study "dual" problem of choosing consumption vector to minimize expenditure subject to minimum utility constraint.

This expenditure minimization problem (EMP) is formally defined as:

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

3.23 Hicksian Demand

Hicksian demand is the set of solutions $x = h(p, u)$ to the EMP.

The expenditure function is the value function for the EMP:

$$e(p, u) = \min_{x \in \mathbb{R}+^n (x) \geq u} p \cdot x$$

$e(p, u)$ is income required to attain utility u when facing prices p. Each element of $h(p, u)$ is a consumption vector that attains utility u while minimizing expenditure given prices p.

Hicksian demand and expenditure function relate to EMP just as Marshallian demand and indirect utility function relate to CP.

Exercise: find the Hicksian demand and the expenditure function for the Cobb-Douglas utility function.

3.24 Why Should we Care about the EMP?

For this course, 2 reasons:

(1) Hicksian demand useful for studying effects of price changes on "real" (Marshallian) demand.

In particular, Hicksian demand is key concept needed to decompose effect of a price change into income and substitution effects.

(2) Expenditure function important for welfare economics.

In particular, use expenditure function to analyze effects of price changes on consumer welfare.

3.25 Hicksian Demand: Properties

Theorem 3.12 (MWG 3E3)**.** *Assume $X = \mathbb{R}+^n$, preferences are locally non-satiated, and $p \gg 0$. Then the Hicksian demand satisfies:*

1. *Homogeneity of degree 0 in* $\mathbf{p}$*: for all $\lambda > 0$, $h(\lambda p, u) = h(p, u)$.*
2. *No excess utility: if $u(\cdot)$ is continuous and $p \gg 0$, then $u(x) = u$ for all $x \in h(p, u)$.*
3. *Convexity/uniqueness: if preferences are convex, then $h(p, u)$ is a convex set. If preferences are strictly convex and "no excess utility" holds, then $h(p, u)$ contains at most one element.*

! Proof

1. Minimizing $p \cdot x$ or $\alpha p \cdot x$ yields the same result for any $\alpha > 0$. QED.
2. The proof is by contradiction.

Assume by contradiction that at the solution, $u(x) > u$. Take $x' = \alpha x$, for $\alpha \in (0, 1)$. The continuity of u implies that for α close enough to one, $u(x') > u$, and $px' < px = w$. That is, it is possible to find some x that respects the constraint and decreases expenditure. Contradiction. Hence one cannot have $u(x) > u$ at the solution. QED.

3. Left as an exercise.

3.26 Expenditure Function: Properties

Theorem 3.13 (MWG 3E2)**.** *The expenditure function satisfies:*

1. *Homogeneity of degree 1 in* $\mathbf{p}$*: for all $\lambda > 0$, $e(\lambda p, u) = \lambda e(p, u)$.*
2. *Continuity: if $u(\cdot)$ is continuous, then e is continuous in p and u.*
3. *Monotonicity: $e(p, u)$ is non-decreasing in p and non-decreasing in u. If "no excess utility" holds, then $e(p, u)$ is strictly increasing in u.*
4. *Concavity in p: e is concave in p.*

Proof

1. For all $\alpha > 0$, we know from the previous proposition that $h(\lambda p, u) = h(p, u)$.

Then one may write:

$$e\left(\alpha p, u\right) = \left(\alpha p\right) h = \alpha \underbrace{\left(p \cdot h\right)}_{e(p,u)} = \alpha \cdot e\left(p, u\right)$$

2. Follows from the Maximum Theorem.

3. Show first that $e\left(p, u\right)$ is strictly increasing in u

The proof is by contradiction. Assume by contradiction that $e\left(p, u\right)$ is not strictly increasing in u. That is:
Let x' and x'' be optimal to achieve utility levels u' and u'', respectively.
Assume $u'' > u'$ and $px' \geq px'' > 0$. This is the contradiction.
Build a new bundle: $\tilde{x} = \alpha x''$ for $\alpha \in (0, 1)$.
$u\left(\right)$ is continuous implies that there is some α close enough to one such that:
$u\left(\tilde{x}\right) > u'$ $px' > p\tilde{x}$
Then x' is not optimal to achieve u'.
Let's show now that $e\left(p, u\right)$ is non-decreasing in p_l.
Take two price vectors p'', p' such that $p_l'' \geq p_l'$, and, for all $k \neq l$, $p_k'' \geq p_k'$.
Let x'' be optimal for prices p''. Then:
$e\left(p'', u\right) = p''x'' \geq p'x'' \geq e\left(p', u\right)$
The first inequality follows from the previous line: $p_l'' \geq p_l'$.
The second inequality follows from the definition of $e\left(p, u\right)$.
It follows that $e\left(., .\right)$ is non-decreasing in p_l. QED.

4. To show concavity, fix some level $\bar{u}$.

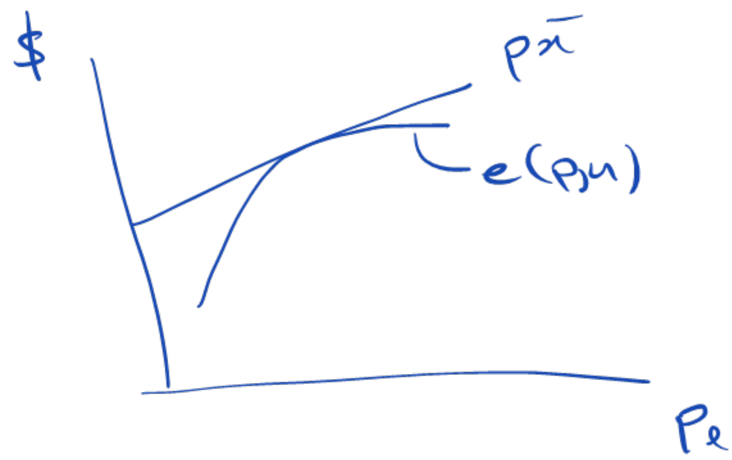
Define $p'' = \alpha p + \left(1 - \alpha\right) p'$ for some $\alpha \in [0, 1)$.
Let x'' be optimal for p''. Then:
$e\left(p'', \bar{u}\right) = p''x'' = \alpha p x'' + \left(1 - \alpha\right) p'x'' \geq \alpha e\left(p, \bar{u}\right) + \left(1 - \alpha\right) e\left(p', \bar{u}\right)$
The inequality comes from the definition of $e\left(p, u\right)$ and from the fact that $u\left(x''\right) \geq \bar{u}$.
QED.

3.27 Intuition for concavity

Start with $\bar{p}$ and an optimal bundle $\bar{x}$.

If prices change to p but $\bar{x}$ is fixed, new expenditure is $p\bar{x}$: linear in $\bar{x}$.

If consumer may adjust $\bar{x}$ to minimize px, new expenditure cannot be larger.



3.28 Expenditure Function: Derivatives

Shephard's Lemma: if Hicksian demand is single-valued, it coincides with the derivative of the expenditure function.

Theorem 3.14. *If $u(\cdot)$ is continuous and $h(p,u)$ is single-valued, then the expenditure function is differentiable in p at (p,u), with derivatives given by*

$$\frac{\partial}{\partial p_i} e(p,u) = h_i(p,u).$$

Intuition: If price of good i increases by \$1, unique optimal consumption bundle now costs $h_i(p,u)$ more.

! Important

Recall that

$$e(p,u) = \min_{x(x)\geq u} ;p \cdot x$$

Given that e is differentiable in p, envelope theorem implies that

$$\frac{\partial}{\partial p_i} e(p,u) = \frac{\partial}{\partial p_i} p \cdot h_i(p,u) = h_i(p,u) \text{ for any } x^* \in h(p,u).$$

3.29 Comparative Statics

Comparative statics are statements about how the solution to a problem change with the parameters.

(CP): parameters are (p, w), want to know how $x(p, w)$ and $v(p, w)$ vary with p and w.

(EMP): parameters are (p, u), want to know how $h(p, u)$ and $e(p, u)$ vary with p and u.

Turns out that comparative statics of (EMP) are very simple, and help us understand comparative statics of (CP).

3.30 The Law of Demand

"Hicksian demand is always decreasing in prices."

Theorem 3.15 (Law of Demand)**.** *For every $p, p' \geq 0, x \in h(p, u)$, and $x' \in h(p', u)$, we have $(p' - p)(x' - x) \leq 0$*

Example: if p' and p only differ in price of good i, then

$$(p'_i - p_i)(h_i(p', u) - h_i(p, u)) \leq 0$$

Hicksian demand for a good is always decreasing in its own price.

Graphically, budget line gets steeper $\implies$ shift along indifference curve to consume less of good 1.

! Proof

By definition:

$$p'' h(p'', u) \leq p'' h(p', u) p' h(p', u) \leq p' h(p'', u)$$

Add these inequations:

$$p'' h(p'', u) + p' h(p', u) \leq p'' h(p', u) + p' h(p'', u)$$

Factor out $h(p'', u) + p'$ and $h(p', u) \leq p''$:

$$(p'' - p') \cdot h(p'', u) + (p' - p'') \cdot h(p', u) \leq 0$$

Change $(p' - p'')$ for $-(p'' - p')$:

$$(p'' - p') \cdot h(p'', u) - (p'' - p') \cdot h(p', u) \leq 0$$

Now factor out $(p'' - p')$:

$$(p'' - p') \cdot (h(p'', u) - h(p', u)) \leq 0$$

This is the law of demand.
QED

3.31 The Slutsky Matrix

If Hicksian demand is differentiable, can derive an interesting result about the matrix of price-derivatives

$$D_p h(p, u) = \begin{pmatrix} \frac{\partial h_1(p,u)}{\partial p_1} & \dots & \frac{\partial h_n(p,u)}{\partial p_1} \\ \vdots & & \vdots \\ \frac{\partial h_1(p,u)}{\partial p_n} & \dots & \frac{\partial h_n(p,u)}{\partial p_n} \end{pmatrix}$$

This is the Slutsky matrix.

A $n \times n$ symmetric matrix M is negative semi-definite if, for all $z \in \mathbb{R}^n, z \cdot Mz \leq 0$.

Theorem 3.16. *If $h(p, u)$ is single-valued and continuously differentiable in p at (p, u), with $p \gg 0$, then the matrix $D_p h(p, u)$ is symmetric and negative semi-definite.*

❗ Proof

Follows from Shephard's Lemma $\left(\frac{\partial}{\partial p_i} e(p, u) = h_i(p, u)\right)$ and Young's Theorem.

3.32 The Slutsky Matrix

What's economic content of symmetry and negative semi-definiteness of Slutsky matrix?

Negative semi-definiteness: differential version of law of demand.

Ex. if $z = (0, \dots, 0, 1, 0, \dots, 0)$ with 1 in the j^{th} component, then $z \cdot D_p h(p, u) z = \frac{\partial h_i(p,u)}{\partial p_i}$, so negative semi-definiteness implies that $\frac{\partial h_i(p,u)}{\partial p_i} \leq 0$.

Symmetry: derivative of Hicksian demand for good i with respect to price of good j equals derivative of Hicksian demand for good j with respect to price of good i.

Not true for Marshallian demand, due to income effects.

3.33 Relation between Hicksian and Marshallian Demand

Approach to comparative statics of Marshallian demand is to relate to Hicksian demand, decompose into income and substitution effects via Slutsky equation.

First, relate Hicksian and Marshallian demand.

Let's ask a simple question: what's the Marshallian demand $x(p, w)$ evaluated at some exogenous utility level u?

At first, this seems nonsense: Marshallian demand is a function of prices p and income w, not of utility u.

In fact, utility is not even exogenous: it is endogenous when one solves for the Marshallian demand, which is computed exactly maximizing utility.

Yet, we can evaluate $x(p, w)$ at any level of income w… that is, at any monetary value w.

So we can choose in particular $w = e(p, u)$, since $e(p, u)$ is a monetary value: it is the minimum expenditure to reach the exogenous utility level u.

Hence $x(p, w)$ and $w = e(p, u)$, so we can write $x(p, e(p, u))$: Marshallian demand as a function of some exogenous utility level u.

Next question: what's the Marshallian demand evaluated at u?

Intuition is simple: it is simply the Hicksian demand: $x(p, e(p, u)) = h(p, u)$.

Analogously, we may write $h(p, u) = h(p, v(p, w))$: Hicksian demand as a function of income.

Then $h(p, v(p, w)) = x(p, w)$.

So these equalities hold for the solutions to UMP e EMP.

A similar reasoning applies to the value functions of these problems:

$$e(p, v(p, w)) = wv(p, e(p, u)) = u$$

If $v(p, w)$ is the most utility consumer can attain with income w, then consumer needs income w to attain utility $v(p, w)$.

If need income $e(p, u)$ to attain utility u, then u is most utility consumer can attain with income $e(p, u)$.

The next result formalizes these relationships.

Theorem 3.17. *Suppose $u(\cdot)$ is continuous and locally non-satiated. Then:*

For all $p \gg 0$ and $w \geq 0, x(p,w) = h(p, v(p,w))$ and $e(p, v(p,w)) = w$.

For all $p \gg 0$ and $u \geq u(0), h(p,u) = x(p, e(p,u))$ and $v(p, e(p,u)) = u$.

❗ Proof

Left as an exercise.

3.34 The Slutsky Equation

Theorem 3.18. *Suppose $u(\cdot)$ is continuous and locally non-satiated. Let $p \gg 0$ and $w = e(p,u)$. If $x(p,w)$ and $h(p,u)$ are single-valued and differentiable, then, for all i, j,*

$$\underbrace{\frac{\partial x_i(p,w)}{\partial p_j}}_{\text{total effect}} = \underbrace{\frac{\partial h_i(p,u)}{\partial p_j}}_{\text{substitution effect}} - \underbrace{\frac{\partial x_i(p,w)}{\partial w} x_j(p,w)}_{\text{income effect}}$$

❗ Proof

For all i, $h_i(p,u) = x_i(p, \underbrace{e(p,u)}_{w})$

Since this holds for all goods i and for all prices, one may differentiate both sides with respect to some price p_j:

$$\frac{\partial h_i}{\partial p_j} = \frac{\partial x_i}{\partial p_j} + \frac{\partial x_i}{\partial w} \frac{\partial e}{\partial p_j}$$

But we know that $\frac{\partial e}{\partial p_j} = h_j(p, v(p,w)) = x_j(p,w)$. Substitute into the previous equation to get:

$$\frac{\partial h_i}{\partial p_j} = \frac{\partial x_i}{\partial p_j} + \frac{\partial x_i}{\partial w} x_j(p,w)$$

QED

Intuition: If p_j increases, two effects on demand for good i:

- Substitution effect: $\frac{\partial h_i(p,u)}{\partial p_j}$
- Movement along original indifference curve.

- Response to change in prices, holding utility fixed.
- Income effect: $-\frac{\partial x_i(p,w)}{\partial w} x_j(p,w)$
- Movement from one indifference curve to another.
- Response to change in income, holding prices fixed.

3.35 Terminology for Consumer Theory Comparative Statics

Definition 3.3. Good i is a normal good if $x_i(p,w)$ is increasing in w. It is an inferior good if $x_i(p,w)$ is decreasing in w.

Definition 3.4. Good i is a regular good if $x_i(p,w)$ is decreasing in p_i. It is a Giffen good if $x_i(p,w)$ is increasing in p_i.

Definition 3.5. Good i is a substitute for good j if $h_i(p,u)$ is increasing in p_j. It is a complement if $h_i(p,u)$ is decreasing in p_j.

Definition 3.6. Good i is a gross substitute for good j if $x_i(p,u)$ is increasing in p_j.

It is a gross complement if $x_i(p,u)$ is decreasing in p_j.

3.36 Comparative Statics: Remarks

- Both the substitution effect and the income effect can have either sign.
- Substitution effect is positive for substitutes and negative for complements.
- Income effect is negative for normal goods and positive for inferior goods.
- By symmetry of Slutsky matrix, i is a substitute for $j \Leftrightarrow j$ is a substitute for i.
- Not true that i is a gross substitute for $j \Leftrightarrow j$ is a gross substitute for i.
- Income effects are not symmetric.

Part III

Lectures 5 – 6

4 Expected Utility Theory

Course so far introduced basic theory of choice and utility, extended to DM and producer theory.

Last topic extends in another direction: choice under uncertainty.

4.1 Choice under Uncertainty

All choices made under some kind of uncertainty.

Sometimes useful to ignore uncertainty, focus on ultimate choices.

Other times, must model uncertainty explicitly.

Examples:

- Insurance markets.
- Financial markets.
- Game theory.

4.1.1 Overview

Impose extra assumptions on basic choice model of Lectures 1-2.

> **Lottery**
>
> Rather than choosing outcome directly, decision-maker chooses uncertain prospect (or lottery).
> A lottery is a probability distribution over outcomes.

Leads to von Neumann-Morgenstern expected utility model.

4.2 Consequences and Lotteries

Two basic elements of expected utility theory: consequences (or outcomes) and lotteries.

4.2.1 Consequences

Finite set C of consequences.

Consequences are what the decision-maker ultimately cares about.

Example: "I have a car accident, my insurance company covers most of the costs, but I have to pay a $500 deductible."

4.2.2 Decision-maker (DM) does not choose consequences directly.

DM chooses a lottery, p.

Lotteries are probability distributions over consequences:

$p : C \to [0, 1]$ with $\sum_{c \in C} p(c) = 1$.

Set of all lotteries is denoted by P.

Example: "A gold-level health insurance plan, which covers all kinds of diseases, but has a $500 deductible."

Makes sense because DM assumed to rank health insurance plans only insofar as lead to different probability distributions over consequences.

4.3 Choice

Decision-maker makes choices from set of alternatives X.

What's set of alternatives here, C or P ?

Answer: P

4.3.1 DM does not choose consequences directly, but instead chooses lotteries

Assume decision-maker has a rational preference relation $\gtrsim$ on P.

Natural to assume this?

4.3.2 Convex Combinations of Lotteries

Given two lotteries p and p', the convex combination $\alpha p + (1-\alpha)p'$ is the lottery defined by

$$(\alpha p + (1-\alpha)p')\,(c) = \alpha p(c) + (1-\alpha)p'(c) \text{ for all } c \in C$$

One way to generate it:

- First, randomize between p and p' with weights α and $1-\alpha$.
- Second, choose a consequence according to whichever lottery came up.

Such a probability distribution over lotteries is called a compound lottery.

In expected utility theory, no distinction between simple and compound lotteries: simple lottery $\alpha p + (1-\alpha)p'$ and above compound lottery give same distribution over consequences, so identified with same element of P.

So, no problem if DM doesn't know exactly the distribution for something. We'll come back to this.

4.3.3 The Set P

As $\alpha p + (1-\alpha)p'$ is also a lottery, P is convex.

P is also closed and bounded (why?).

$\Rightarrow P$ is a compact subset of $\mathbb{R}^n$, where $n = |C|$.

Whenever $\succsim$ is rational and continuous, can be represented by continuous utility function $U : P \to \mathbb{R}$:

$$p \succsim q \Leftrightarrow U(p) \geq U(q)$$

We're just applying it to lotteries because that's what the DM chooses now.

Intuitively, want more than this.

Want not only that DM has utility function over lotteries, but also that somehow related to "utility" over consequences.

Only care about lotteries insofar as affect distribution over consequences, so preferences over lotteries should have something to do with "preferences" over consequences.

4.4 Expected Utility

Best we could hope for is representation by utility function of following form:

Definition: a utility function $U : P \to \mathbb{R}$ has an expected utility form if there exists a function $u : C \to \mathbb{R}$ such that

$$U(p) = \sum_{c \in C} p(c)u(c) \text{ for all } p \in P$$

In this case, the function U is called an expected utility function, and the function u is call a von Neumann-Morgenstern utility function.

If preferences over lotteries happen to have an expected utility representation, it's as if DM has a "utility function" over consequences (and chooses among lotteries so as to maximize expected "utility over consequences").

Remarks

$$U(p) = \sum_{c \in C} p(c)u(c)$$

Expected utility function $U : P \to \mathbb{R}$ represents preferences $\gtrsim$ on P just as we had before

$U : P \to \mathbb{R}$ is an example of a standard utility function.

von Neumann-Morgenstern utility function $u : C \to \mathbb{R}$ is not a standard utility function.

4.4.1 Can't have a "real" utility function on consequences, as DM never chooses among consequences.

If preferences over lotteries happen to have an expected utility representation, it's as if DM has a "utility function" over consequences.

This "utility function" over consequences is the von Neumann-Morgenstern utility function.

4.5 Example

Suppose hipster restaurant doesn't let you order steak or chicken, but only probability distributions over steak and chicken.

How should you assess menu item $(p(steak), p(chicken))$?

One way: ask yourself how much you'd like to eat steak, $u(steak)$, and chicken, $u(chicken)$, and evaluate according to

$$p(\text{ steak }) \cdot u(\text{ steak }) + p(\text{ chicken }) \cdot u(\text{ chicken })$$

If this is what you'd do, then your preferences have an expected utility representation.

Suppose instead you choose whichever menu item has p (steak) closest to $\frac{1}{2}$.

Your preferences are rational, so they have a utility representation.

But they do not have an expected utility representation - we'll see this.

4.6 Property of EU: Linearity in Probabilities

Theorem 4.1. *If $U : P \to \mathbb{R}$ is an expected utility function, then*

$$U\left(\alpha p+(1-\alpha) p^{\prime}\right)=\alpha U(p)+(1-\alpha) U\left(p^{\prime}\right)$$

In fact, a utility function $U : P \to \mathbb{R}$ has an expected utility form iff this equation holds for all p, p', and $\alpha \in [0, 1]$.

! Proof

Appendix.

4.7 Property of EU: Invariant to Affine Transformations

Suppose $U : P \to \mathbb{R}$ is an expected utility function representing preferences $\succsim$.

Any increasing transformation of U also represents $\succsim$.

Not all increasing transformations of U have expected utility form.

Theorem 4.2. *Suppose $U : P \to \mathbb{R}$ is an expected utility function representing preferences $\succsim$. Then $V : P \to \mathbb{R}$ is also an expected utility function representing $\gtrsim$ iff there exist $a, b > 0$ such that*

$$V(p) = a + bU(p) \text{ for all } p \in P$$

If this is so, we also have $V(p) = \sum_{c \in C} p(c)v(c)$ for all $p \in P$, where

$$v(c) = a + bu(c) \text{ for all } c \in C$$

! Proof

Appendix.

4.8 What Preferences have an Expected Utility Representation?

Preferences must be rational to have any kind of utility representation.

Preferences on a compact and convex set must be continuous to have a continuous utility representation.

Besides rationality and continuity, what's needed to ensure that preferences have an expected utility representation?

4.9 The Independence Axiom

Definition 4.1. A preference relation $\gtrsim$ satisfies independence if, for every $p, q, r \in P$ and $\alpha \in (0, 1)$,

$$p \succsim q \Leftrightarrow \alpha p + (1 - \alpha)r \succsim \alpha q + (1 - \alpha)r$$

Can interpret as form of "dynamic consistency."

Doesn't need to hold for consequences.

4.10 Back to Example

Suppose choose lottery with p (steak) closest to $\frac{1}{2}$.

Let $p = \left(\frac{1}{2}, \frac{1}{2}\right), q = (0, 1), r = (1, 0)$, and $\alpha = \frac{1}{2}$.

Then

$$p = \left(\frac{1}{2}, \frac{1}{2}\right) > (0, 1) = q$$

but

$$\alpha q + (1-\alpha)r = \left(\frac{1}{2}, \frac{1}{2}\right) > \left(\frac{3}{4}, \frac{1}{4}\right) = \alpha p + (1-\alpha)r$$

Does not satisfy independence.

4.11 Expected Utility: Characterization

A preference relation $\gtrsim$ has an expected utility representation iff it satisfies rationality, continuity, and independence.

Intuition: both having expected utility form and satisfying independence boil down to having straight, parallel indifference curves.

4.12 Subjective Expected Utility Theory

So far, probabilities are objective.

In reality, uncertainty is usually subjective.

Subjective expected utility theory (Savage, 1954): under assumptions roughly similar to ones form this lecture, preferences have an expected utility representation where both the utilities over consequences and the subjective probabilities themselves are revealed by decision-maker's choices.

Thus, expected utility theory applies even when the probabilities are not objectively given.

(To learn more, a good starting point is Kreps (1988), "Notes on the Theory of Choice.")

Again, no problem if DM doesn't know the exact distribution.

The same holds in general equilibrium: allows for different individual priors.

One may go beyond and assume DM has some rule to deal with set of priors - e.g., DM may assume that nature will choose the worst possible prior, conditional on his optimal choice, leading to a mini-max structure that deals with fear of misspecification and relates to sub-rational behavior.

See nice discussion in Hansen and Sargent (2000) and a critique by Sims (AER 2001).

4.13 Attitudes toward Risk

4.13.1 Money Lotteries

Turn now to special case of choice under uncertainty where outcomes are measured in dollars.

Set of consequences C is subset of $\mathbb{R}$.

A lottery is a cumulative distribution function F on $\mathbb{R}$.

(Now we use F instead of p)

Assume preferences have expected utility representation:

$$U(F) = E_F[u(x)] = \int u(x)f(x)dx$$

More generally, we could write $\int u(x)dF(x)$.

This is useful if we do not know whether a density f exists.

We'll assume it does and make $dF(x)/dx = f(x)$, so that $dF(x) = f(x)dx$, leading to our representation above.

(But everything holds for a general $F(x)$.)

Assume u increasing, differentiable.

Question: how do properties of von Neumann-Morgenstern utility function u relate to decision-maker's attitude toward risk?

4.14 Expected Value vs. Expected Utility

Expected value of lottery F is

$$E_F[x] = \int x f(x) dx$$

Expected utility of lottery F is

$$E_F[u(x)] = \int u(x) f(x) dx$$

Can learn about DM's risk attitude by comparing $E_F[u(x)]$ and $u\left(E_F[x]\right)$.

4.15 Risk Attitude: Definitions

Definition 4.2. A decision-maker is risk-averse if she always prefers the sure wealth level $E_F[x]$ to the lottery F : that is,

$$\int u(x) f(x) dx \leq u\left(\int x f(x) dx\right) \text{ for all } F$$

A decision-maker is strictly risk-averse if the inequality is strict for all nondegenerate lotteries F.

A decision-maker is risk-neutral if she is always indifferent:

$$\int u(x) f(x) dx = u\left(\int x f(x) dx\right) \text{ for all } F$$

A decision-maker is risk-loving if she always prefers the lottery:

$$\int u(x) f(x) dx \geq u\left(\int x f(x) dx\right) \text{ for all } F$$

4.16 Risk Aversion and Concavity

Statement that $\int u(x)dF(x) \leq u\left(\int xdF(x)\right)$ for all F is called Jensen's inequality.

Fact: Jensen's inequality holds iff u is concave.

This implies:

Theorem 4.3. *A decision-maker is (strictly) risk-averse if and only if u is (strictly) concave.*

A decision-maker is risk-neutral if and only if u is linear.

A decision-maker is (strictly) risk-loving if and only if u is (strictly) convex.

4.17 Certainty Equivalents

Can also define risk-aversion using certainty equivalents.

Definition 4.3. The certainty equivalent of a lottery F is the sure wealth level that yields the same expected utility as F: that is,

$$u[CE(F,u)] = \int u(x)f(x)dx$$

That is,

$$CE(F,u) = u^{-1}\left(\int u(x)dF(x)\right)$$

Theorem 4.4. *A decision-maker is risk-averse iff $CE(F,u) \leq E_F(x)$ for all F.*

A decision-maker is risk-neutral iff $CE(F,u) = E_F(x)$ for all F.

A decision-maker is risk-loving iff $CE(F,u) \geq E_F(x)$ for all F.

4.18 Quantifying Risk Attitude

We know what it means for a DM to be risk-averse.

What does it mean for one DM to be more risk-averse than another?

Two possibilities:

- u is more risk-averse than v if, for every $F, CE(F, u) \leq CE(F, v)$.
- u is more risk-averse than v if u is "more concave" than v, in that $u = g \circ v$ for some increasing, concave g.

One more, based on local curvature of utility function: u is more-risk averse than v if, for every x,

$$-\frac{u''(x)}{u'(x)} \geq -\frac{v''(x)}{v'(x)}$$

$A(x, u) = -\frac{u''(x)}{u'(x)}$ is called the Arrow-Pratt coefficient of absolute risk-aversion.

4.19 An Equivalence

The following are equivalent:

- For every F, $CE(F, u) \leq CE(F, v)$.
- There exists an increasing, concave function g such that $u = g \circ v$.
- For every $x, A(x, u) \geq A(x, v)$.

4.20 Risk Attitude and Wealth Levels

How does risk attitude vary with wealth?

Natural to assume that a richer individual is more willing to bear risk: whenever a poorer individual is willing to accept a risky gamble, so is a richer individual.

Captured by decreasing absolute risk-aversion:

A von Neumann-Morenstern utility function u exhibits decreasing (constant, increasing) absolute risk-aversion if $A(x, u)$ is decreasing (constant, increasing) in x.

4.21 Risk Attitude and Wealth Levels

Theorem 4.5. *Suppose u exhibits decreasing absolute risk-aversion.*

If the decision-maker accepts some gamble at a lower wealth level, she also accepts it at any higher wealth level:

that is, for any lottery $F(x)$, if

$$E_F[u(w+x)] \geq u(w)$$

then, for any $w' > w$,

$$E_F\left[u\left(w'+x\right)\right] \geq u\left(w'\right)$$

4.22 Multiplicative Gambles

What about gambles that multiply wealth, like choosing how risky a stock portfolio to hold? Are richer individuals also more willing to bear multiplicative risk? Depends on increasing/decreasing relative risk-aversion:

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

Suppose u exhibits decreasing relative risk-aversion.

If the decision-maker accepts some multiplicative gamble at a lower wealth level, she also accepts it at any higher wealth level: that is, for any lottery $F(t)$, if

$$E_F[u(tw)] \geq u(w)$$

then, for any $w' > w$,

Relative Risk-Aversion vs. Absolute Risk-Aversion

$$R(x) = xA(x)$$

decreasing relative risk-aversion $\Rightarrow$ decreasing absolute risk-aversion

increasing absolute risk-aversion $\Rightarrow$ increasing relative risk-aversion

Ex. decreasing relative risk-aversion $\Rightarrow$ more willing to gamble 1% of wealth as get richer.

So certainly more willing to gamble a fixed amount of money.

4.23 Application: Insurance

Risk-averse agent with wealth w, faces probability p of incurring monetary $lossL$.

Can insure against the loss by buying a policy that pays out a if the loss occurs.

Policy that pays out a costs qa.

How much insurance should she buy?

4.24 Agent's Problem

$$\max_a pu(w - qa - L + a) + (1 - p)u(w - qa)$$

u concave $\Rightarrow$ concave problem, so FOC is necessary and sufficient.

FOC:

$$p(1 - q)u'(w - qa - L + a) = (1 - p)qu'(w - qa)$$

Equate marginal benefit of extra dollar in each state.

4.25 Actuarily Fair Prices

Insurance is actuarily fair if expected payout qa equals cost of insurance $pa:$ that is, $p = q$.

With acturarily fair insurance, FOC becomes

$$u'(w - qa - L + a) = u'(w - qa)$$

Solution: $a = L$

A risk-averse DM facing actuarily fair prices will always fully insure.

4.26 Actuarily Unfair Prices

What if insurance company makes a profit, so $q > p$?

Rearrange FOC as

$$\frac{u'(w-qa-L+a)}{u'(w-qa)} = \frac{(1-p)q}{p(1-q)} > 1$$

Solution: $a < L$

A risk-averse DM facing actuarily unfair prices will never fully insure.

Intuition: u approximately linear for small risks, so not worth giving up expected value to insure away last little bit of variance.

4.27 Comparative Statics

$$\max_a pu(w-qa-L+a) + (1-p)u(w-qa)$$

Bigger loss $\Rightarrow$ buy more insurance (a^* increasing in L) Follows from Topkis' theorem.

If agent has decreasing absolute risk-aversion, then she buys less insurance as she gets richer.

Prove it as an exercise!

4.28 Application: Portfolio Choice

Risk-averse agent with wealth w has to invest in a safe asset and a risky asset.

Safe asset pays certain return r.

Risky asset pays random return z, with cdf F.

Agent's problem

$$\max_{a\in[0,w]} \int u(az+(w-a)r)dF(z)$$

First-order condition

$$\int (z-r)u'(az+(w-a)r)dF(z) = 0$$

4.29 Risk-Neutral Benchmark

Suppose $u'(x) = \alpha x$ for some $\alpha > 0$.

Then

$$U(a) = \int \alpha(az + (w - a)r)dF(z)$$

so

$$U'(a) = \alpha(E[z] - r)$$

Solution: set $a = w$ if $E[z] > r$, set $a = 0$ if $E[z] < r$.

Risk-neutral investor puts all wealth in the asset with the highest rate of return.

$r > E[z]$ Benchmark

$$U'(0) = \int (z - r)u'(w)dF = (E[z] - r)u'(w)$$

If safe asset has higher rate of return, then even risk-averse investor puts all wealth in the safe asset.

4.30 More Interesting Case

What if agent is risk-averse, but risky asset has higher expected return?

$$U'(0) = (E[z] - r)u'(w) > 0$$

If risky asset has higher rate of return, then risk-averse investor always puts some wealth in the risky asset.

4.31 Comparative Statics

Does a less risk-averse agent always invest more in the risky asset?

Sufficient condition for agent v to invest more than agent u :

$$\int (z-r)u'(az+(w-a)r)dF = 0$$
$$\Rightarrow \int (z-r)v'(az+(w-a)r)dF \geq 0$$

u more risk-averse $\Rightarrow v = h \circ u$ for some increasing, convex h. Inequality equals

$$\int (z-r)h'(u(az+(w-a)r))u'(az+(w-a)r)dF \geq 0$$

$h'(\cdot)$ positive and increasing in z

$\Rightarrow$ multiplying by $h'(\cdot)$ puts more weight on positive $(z > r)$ terms, less weight on negative terms.

A less risk-averse agent always invests more in the risky-asset.

5 Comparing Risky Prospects

5.1 Risky Prospects

We've studied decision-maker's subjective attitude toward risk.

Now: study objective properties of risky prospects (lotteries, gambles) themselves, relate to individual decision-making.

Topics:

- First-Order Stochastic Dominance
- Second-Order Stochastic Dominance
- (Optional) Some recent research extending these concepts

5.2 First-Order Stochastic Dominance

When is one lottery unambiguously better than another?

Natural definition: F dominates G if, for every amount of money x, F is more likely to yield at least x dollars than G is.

For any lotteries F and G over $\mathbb{R}, F$ first-order stochastically dominates (FOSD) G if

$$F(x) \leq G(x) \text{ for all } x$$

5.3 FOSD and Choice

Main theorem relating FOSD to decision-making:

Theorem 5.1. *F FOSD G $\iff$ every decision-maker with a non-decreasing utility function prefers F to G.*

That is, the following are equivalent:

- $F(x) \leq G(x)$ for all x.
- $\int u(x)dF \geq \int u(x)dG$ for every non-decreasing function $u : \mathbb{R} \to \mathbb{R}$.

❗ Proof

Preferred by Everyone $\Longrightarrow$ FOSD
If F does not FOSD G, then there's some amount of money x^* such that G is more likely to give at least x^* than F is.
Consider a DM who only cares about getting at least x^* dollars.
She will prefer G.

FOSD $\Rightarrow$ Preferred by Everyone

Main idea: F FOSD $G \Rightarrow F$ gives more money "realization-by-realization."
Suppose draw x according to G, but then instead give decision-maker

$$y(x) = F^{-1}(G(x))$$

Then:

- $y(x) \geq x$ for all x, and
- y is distributed according to F.

$\Rightarrow$ paying decision-maker according to F just like first paying according to G, then sometimes giving more money.
Any decision-maker who likes money likes this.
QED.

Second-Order Stochastic Dominance

Q: When is one lottery better than another for any decision-maker?
A: First-Order Stochastic Dominance.

Q: When is one lottery better than another for any risk-averse decision-maker?
A: Second-Order Stochastic Dominance.

Theorem 5.2. *F second-order stochastically dominates (SOSD) G iff every decision-maker with a non-decreasing and concave utility function prefers F to G: that is,*

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

for every non-decreasing and concave function $u : \mathbb{R} \to \mathbb{R}$.

SOSD is a weaker property than FOSD.

5.4 SOSD for Distributions with Same Mean

If F and G have same mean, when will any risk-averse decision-maker prefer F ?

When is F "unambiguously less risky" than G ?

5.5 Mean-Preserving Spreads

Definition 5.1. G is a mean-preserving spread of F if G can be obtained by first drawing a realization from F and then adding noise.

G is a mean-preserving spread of F iff there exist random variables x, y, and ε such that

$$y = x + \varepsilon$$

x is distributed according to F, y is distributed according to G, and $E[\varepsilon \mid x] = 0$ for all x.

Formulation in terms of cdfs:

$$\int_{-\infty}^{x} G(y)dy \geq \int_{-\infty}^{x} F(y)dy \text{ for all } x$$

5.6 Characterization of SOSD for CDFs with Same Mean

Assume that $\int xdF = \int xdG$. Then the following are equivalent:

- F SOSD G.
- G is a mean-preserving spread of F.
- $\int_{-\infty}^{x} G(y)dy \geq \int_{-\infty}^{x} F(y)dy$ for all x.

5.7 General Characterization of SOSD

Theorem 5.3. *The following are equivalent:*

- *F SOSD G.*
- $\int_{-\infty}^{x} G(y)dy \geq \int_{-\infty}^{x} F(y)dy$ *for all x.*
- *There exist random variables x, y, z, and ε such that*

$$y = x + z + \varepsilon$$

x is distributed according to F, y is distributed according to G, z is always nonpositive, and $E[\varepsilon \mid x] = 0$ for all x.

- There exists a cdf H such that F FOSD H and G is a mean-preserving spread of H.

Complete Dominance Orderings [Optional]

FOSD and SOSD are partial orders on lotteries:
"most distributions" are not ranked by FOSD or SOSD.
To some extent, nothing to be done:
If F doesn't FOSD G, some decision-maker prefers G.
If F doesn't SOSD G, some risk-averse decision-maker prefers G.
However, recent series of papers points out that if view F and G as lotteries over monetary gains and losses rather than final wealth levels, and only require that no decision-maker prefers G to F for all wealth levels, do get a complete order on lotteries (and index of lottery's "riskiness").

5.8 Acceptance Dominance

Consider decision-maker with wealth w, has to accept or reject a gamble F over gains / losses x.
Accept iff

$$E_F[u(w+x)] \geq u(w)$$

F acceptance dominates G if, whenever F is rejected by decision-maker with concave utility function u and wealth w, so is G.
That is, for all u concave and $w > 0$,

$$\begin{aligned} E_F[u(w+x)] &\leq u(w) \\ E_G[u(w+x)] &\leq u(w) \end{aligned}$$

5.9 Acceptance Dominance and FOSD/SOSD

F SOSD G

$\Rightarrow E_F[u(w+x)] \geq E_G[u(w+x)]$ for all concave u and wealth w

$\Rightarrow F$ acceptance dominates G.

If $E_F[x] > 0$ but x can take on both positive and negative values, can show that F acceptance dominates lottery that doubles all gains and losses.

Acceptance dominance refines SOSD.

But still very incomplete.

Turns out can get complete order from something like: acceptance dominance at all wealth levels, or for all concave utility functions.

5.10 Wealth Uniform Dominance

F wealth-uniformly dominates G if, whenever F is rejected by decision-maker with concave utility function u at every wealth level w, so is G.

That is, for all $u \in U^*$,

$$E_F[u(w+x)] \leq u(w) \text{ for all } w > 0$$
$$E_G[u(w+x)] \leq u(w) \text{ for all } w > 0$$

5.11 Utility Uniform Dominance

F utility-uniformly dominates G if, whenever F is rejected at wealth level w by a decision-maker with any utility function $u \in U^*$, so is G.

That is, for all $w > 0$,

$$E_F[u(w+x)] \leq u(w) \text{ for all } u \in U^*$$
$$E_G[u(w+x)] \leq u(w) \text{ for all } u \in U^*$$

5.12 Uniform Dominance: Results

Hart (2011):

- Wealth-uniform dominance and utility-uniform dominance are complete orders.
- Comparison of two lotteries in these orders boils down to comparison of simple measures of the "riskiness" of the lotteries.
- Measure for wealth-uniform dominance: critical level of risk-aversion above which decision maker with constant absolute risk-aversion rejects the lottery.
- Measure for utility-uniform dominance: critical level of wealth below which decision-maker with log utility rejects the lottery.

5.13 Appendix: some proofs

5.13.1 U has expected utility form $\Leftrightarrow U$ linear in probabilities

Theorem 5.4. *$U : P \rightarrow \mathbb{R}$ has an expected utility form if and only if*

$$U\left(\alpha p+(1-\alpha) p^{\prime}\right)=\alpha U(p)+(1-\alpha) U\left(p^{\prime}\right)$$

holds for all p, p', and $\alpha \in [0,1]$.

Notice: this is MWG proposition 6B1. It uses the following notation: $U\left(\sum \alpha_k p_k\right)=\sum \alpha_k U\left(p_k\right)$, just substituting p for L (which stands for 'lottery').

! Proof

Without loss of generality, we will assume only two consequences, c^1 and c^2.
Hence any lottery p may be written as $p=\left(p^1, p^2\right)$, in which $p^1=\operatorname{Prob}\left(c^1\right)$ and $p^2=\operatorname{Prob}\left(c^2\right)$.
All arguments below hold unchanged for $p=\left(p^1, \ldots, p^n\right)$, that is, for n consequences $c^1, \ldots, c^n$. This extension is shown in red below; you may simply ignore it in your first reading.
The arguments below also hold for $c \in\left[c^1, c^n\right] \in \mathbb{R}$, but the math is not exactly the same.

- Necessity: U linear in probabilities $\Rightarrow U$ has expected utility form

Write lottery $p=\left(p^1, p^2\right)$ as a convex combination of degenerate lotteries $\left(C^1, C^2\right)$:

$$p=p^1 C^1+p^2 C^2+\cdots+p^n C^n$$

That is, $C^1=(1,0)$, meaning that consequence $1\left(c^1\right)$ happens with probability 1 , and $C^2=(0,1)$, meaning that consequence $2\left(c^2\right)$ happens with probability 1 . The\ equation above is simply $p=\left(p^1, p^2\right)=\left(p^1, 0\right)+\left(0, p^2\right)=p^1 \cdot(1,0)+p^2 \cdot(0,1)=p^1 C^1+p^2 C^2$.
Then:

$$U(p)=U\left(p^1 C^1+p^2 C^2+\cdots+p^n C^n\right)=p^1 U\left(C^1\right)+p^2 U\left(C^2\right)+\cdots+p^n U\left(C^n\right)$$

The second equality follows from our assumption: U is linear in probabilities.
But $U\left(C^1\right)$ is the utility from a degenerate lottery, that is, it's simply the vNM utility of consequence $c_1: U\left(C^1\right)=u\left(c^1\right)$.
Remember our notation: big U is for DM's actual utility; small u is form DM's vNM utility. Big C denotes a lottery; small c denotes a consequence.
The last equation may be rewritten as:

$$
\begin{aligned}
U(p) = U\left(p^1 C^1 + p^2 C^2 + \dots + p^n C^n\right) &= p^1 U\left(C^1\right) + p^2 U\left(C^2\right) + \dots + p^n U\left(C^n\right) \\
&= p^1 u\left(c^1\right) + p^2 u\left(c^2\right) + \dots + p^n u\left(c^n\right)
\end{aligned}
$$

In short:

$$
U(p) = p^1 u\left(c^1\right) + p^2 u\left(c^2\right) + \dots + p^n u\left(c^n\right)
$$

This is exactly the expected utility property, concluding the proof.

- Sufficiency: U has expected utility form $\Rightarrow U$ linear in probabilities

Consider a compound lottery:

$$
(p_1, p_2, \dots, p_k; \alpha_1, \alpha_2, \dots, \alpha_k)
$$

Notice that now we have p_1 instead of p^1; and p_2 instead of p^2. Superscripts refer to consequences; subscripts refer to specific lotteries.
That is, p_1 and p_2 are different lotteries, and each one is a vector assigning probabilities for each of the two possible consequences c^1 and c^2 :

$$
p_i = \left(p_i^1, p_i^2, \dots, p_i^n\right)
$$

For $i = 1, \dots, k$. That is, we have k lotteries, and each one is chosen with probability α_i in our compound lottery.
We will allow k to be generic. If you want, just take $k = 2$ in the following computations - again, it's without loss of generality, but be careful not to confuse the number of consequences with the number of lotteries.
Consider now the following (reduced) lottery:

$$
\alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_k p_k
$$

Consider the utility of this lottery:

$$
U\left(\alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_k p_k\right)
$$

We may now use our assumption: U has the expected utility form. That is, one may rewrite this utility as:

$$
U\left(\alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_k p_k\right) = u^1 \cdot \operatorname{Prob}\left(u^1\right) + u^2 \cdot \operatorname{Prob}\left(u^2\right) + \dots + u^n \cdot \operatorname{Prob}\left(u^n\right)
$$

What are these u^i s? We just need to know that there are some u^i s that make this equation hold - our assumption guarantees this is the case. But we do have an interpretation for them: u^i is just the vNM utility of consequence c^i. Analogously, $\operatorname{Prob}\left(u^i\right)$ is simply the probability of this consequence, computed from the compound lottery $\alpha_1 p_1 + \alpha_2 p_2 + \dots + \alpha_k p_k$.

This explains why we have subscripts on the LHS, but superscripts on the RHS. In the LHS, we have lotteries (that generate a compound lottery). On the RHS, we have consequences with vNM utilities $u^1, u^2, \ldots, u^n$. If you want, you may think of the lottery on the LHS as any given lottery p.
Let's develop this equation:

$$
\begin{aligned}
&U(\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k) = \\
&u^1 \cdot \text{Prob}(u^1) + u^2 \cdot \text{Prob}(u^2) = \\
&u^1 \cdot \underbrace{(\alpha_1 p_1^1 + \alpha_2 p_2^1 + \cdots + \alpha_k p_k^1)}_{\text{Prob}(u^1)} + u^2 \cdot \underbrace{(\alpha_1 p_1^2 + \alpha_2 p_2^2 + \cdots + \alpha_k p_k^2)}_{\text{Prob}(u^2)} + \cdots + u^n \\
&\underbrace{(\alpha_1 p_1^n + \alpha_2 p_2^n + \cdots + \alpha_k p_k^n)}_{\text{Prob}(u^2)} = \\
&\alpha_1 \cdot \underbrace{(u^1 \cdot p_1^1 + u^2 \cdot p_1^2 + \cdots + u^n \cdot p_1^n)}_{U(p_1)} + \alpha_2 \cdot \underbrace{(u^1 \cdot p_2^1 + u^2 \cdot p_2^2 + \cdots + u^n \cdot p_2^n)}_{U(p_2)} + \cdots \\
&+ \alpha_k \cdot \underbrace{(u^1 \cdot p_k^1 + u^2 \cdot p_k^2 + \cdots + u^n \cdot p_k^n)}_{U(p_k)} = \\
&\alpha_1 \cdot U(p_1) + \alpha_2 \cdot U(p_2) + \cdots + \alpha_k \cdot U(p_k)
\end{aligned}
$$

From the second to the third line, we use the definition of $\text{Prob}(u^1)$: it is simply the first coordinate of the vector $\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k$, ie, the compound lottery. It is analogous for $\text{Prob}(u^2)$ to $\text{Prob}(u^n)$.
From the third to the fourth line: we use again our assumption: U has the expected utility form. Hence we may write $U(p_1) = u^1 \cdot p_1^1 + u^2 \cdot p_1^2 + \cdots + u^n \cdot p_1^n$. It is analogous for $U(p_2)$ to $U(p_k)$.
In short:

$$U(\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k) = \alpha_1 \cdot U(p_1) + \alpha_2 \cdot U(p_2) + \cdots + \alpha_k \cdot U(p_k)$$

That is, U is linear in probabilities, concluding the proof.
QED.

5.13.2 Expected utility form is preserved under positive affine transformations

Theorem 5.5. *$U, \widetilde{U}$ have expected utility form (and represent the same preferences) $\Leftrightarrow$ there are $\beta > 0, \gamma$ such that for all $p, \widetilde{U}(p) = \beta U(p) + \gamma$.*

This is MWG proposition 6B2.

! Proof

Choose $\bar{p}, \underline{p}$ such that for all lottery $p, \bar{p} \succcurlyeq p \succcurlyeq \underline{p}$.
If $\bar{p} \sim \underline{p}$, then all utility functions are constant, and the result follows immediately. Assume now $\bar{p} > \underline{p}$.

- Sufficiency: If U has expected utility form, then $\widetilde{U}(p) = \beta U(p) + \gamma$ also has expected utility form.

Consider a compound lottery $\alpha_1 p_1 + \alpha_2 p_2$. That is, we have two lotteries (p_1 and p_2), and each is chosen with probability α_1 and α_2, respectively.
Without loss of generality, we consider only two lotteries, but the argument is unchanged for k lotteries.
Compute the utility of this compound lottery under $\widetilde{U}$:

$$\begin{gathered}
\widetilde{U}\left(\alpha_1 p_1+\alpha_2 p_2\right)= \\
\beta \cdot U\left(\alpha_1 p_1+\alpha_2 p_2\right)+\gamma= \\
\beta \cdot\left[\alpha_1 U\left(p_1\right)+\alpha_2 U\left(p_2\right)\right]+\gamma= \\
\alpha_1 \beta \cdot U\left(p_1\right)+\alpha_2 \beta \cdot U\left(p_2\right)+\underbrace{\left[\alpha_1 \gamma+\alpha_2 \gamma\right]}_{=\gamma}= \\
\alpha_1 \cdot \underbrace{\left[\beta \cdot U\left(p_1\right)+\gamma\right]}_{\widetilde{U}\left(p_1\right)}+\alpha_2 \cdot \underbrace{\left[\beta \cdot U\left(p_2\right)+\gamma\right]}_{\widetilde{U}\left(p_2\right)}= \\
\alpha_1 \cdot \widetilde{U}\left(p_1\right)+\alpha_2 \cdot \widetilde{U}\left(p_2\right)
\end{gathered}$$

From the first to the second line: we use the definition $\widetilde{U}(p) = \beta U(p) + \gamma$.
From the second to the third line: we use the assumption that U has the expected utility form: hence, $U\left(\alpha_1 p_1+\alpha_2 p_2\right)=\alpha_1 U\left(p_1\right)+\alpha_2 U\left(p_2\right)$.
From the third to the fourth line: we simply write $\gamma = \alpha_1 \gamma + \alpha_2 \gamma$, which is true because $\alpha_1 + \alpha_2 = 1$ (it's a probability distribution, so it must sum up to one).
From the fourth to the fifth line: we factor out α_1 and α_2.
In short:

$$\widetilde{U}\left(\alpha_1 p_1+\alpha_2 p_2\right)=\alpha_1 \cdot \widetilde{U}\left(p_1\right)+\alpha_2 \cdot \widetilde{U}\left(p_2\right)$$

That is, $\widetilde{U}$ has the expected utility form, as we wanted to show.

- Necessity: U and $\widetilde{U}$ have the expected utility form (and represent the same preferences) implies that for some $\beta > 0, \gamma$, one has $\widetilde{U}(p) = \beta U(p) + \gamma$

Fix a lottery p.
Choose $\lambda_p \in [0,1]$ such that:

$$U(p) = \lambda_p \cdot U(\bar{p}) + \left(1 - \lambda_p\right) \cdot U(\underline{p})$$

This equation has two implications.
First implication: $p \sim \lambda_p \bar{p} + \left(1 - \lambda_p\right) \underline{p}$
This holds because we're assuming U has the expected utility form. The previous theorem states that if this is the case, then U is linear in probabilities. Hence the RHS of this last equation may be rewritten as:

$$\lambda_p \cdot U(\bar{p}) + \left(1 - \lambda_p\right) \cdot U(\underline{p}) = U\left(\lambda_p \cdot \bar{p} + \left(1 - \lambda_p\right) \cdot \underline{p}\right)$$

Hence $U(p) = U\left(\lambda_p \cdot \bar{p} + \left(1 - \lambda_p\right) \cdot \underline{p}\right)$. By definition of a utility function, the arguments on each side must be indifferent for the DM.
Second implication:

$$\lambda_p = \frac{U(p) - U(\underline{p})}{U(\bar{p}) - U(\underline{p})}$$

This is just a rearrangement of the equation above.
We know that $\widetilde{U}$ is linear in probabilities (previous theorems) and represents the same preferences. Hence:

$$\begin{gathered}
\widetilde{U}(p) = \lambda_p \cdot \widetilde{U}(\bar{p}) + \left(1 - \lambda_p\right) \cdot \widetilde{U}(\underline{p}) = \\
\lambda_p \cdot [\widetilde{U}(\bar{p}) - \widetilde{U}(\underline{p})] + \widetilde{U}(\underline{p}) = \\
\underbrace{\left. \frac{U(p) - U(\underline{p})}{U(\bar{p}) - U(\underline{p})} \right)}_{\lambda_p} \cdot [\widetilde{U}(\bar{p}) - \widetilde{U}(\underline{p})] + \widetilde{U}(\underline{p})
\end{gathered}$$

In short:

$$\widetilde{U}(p) = \underbrace{\left(\frac{U(p) - U(\underline{p})}{U(\bar{p}) - U(\underline{p})} \right)}_{\lambda_p} \cdot [\widetilde{U}(\bar{p}) - \widetilde{U}(\underline{p})] + \widetilde{U}(\underline{p})$$

In this last expression, only $U(p)$ depends on p. All other terms are parameters. Rearrange this expression to get the following:

$$\widetilde{U}(p) = \left[\frac{\widetilde{U}(\bar{p}) - \widetilde{U}(\underline{p})}{U(\bar{p}) - U(\underline{p})} \right] \cdot U(p) + \widetilde{U}(\underline{p}) - U(\underline{p}) \cdot \left[\frac{\widetilde{U}(\bar{p}) - \widetilde{U}(\underline{p})}{U(\bar{p}) - U(\underline{p})} \right]$$

Again: except $U(p)$, everything in this expression is a parameter, built from functions (U or $\widetilde{U}$) evaluated at specific arguments ($\bar{p}$ or $\underline{p}$). We can label them as we want. Let's choose:

$$\beta = \left[\frac{\widetilde{U}(\bar{p})-\widetilde{U}(\underline{p})}{U(\bar{p})-U(\underline{p})}\right]$$
$$\gamma = \widetilde{U}(\underline{p}) - U(\underline{p}) \cdot \underbrace{\left[\frac{\widetilde{U}(\bar{p})-\widetilde{U}(\underline{p})}{U(\bar{p})-U(\underline{p})}\right]}_{\beta}$$

Then one has:

$$\widetilde{U}(p) = \beta U(p) + \gamma$$

This is what we wanted to show, concluding the proof.
QED.

5.13.3 Expected utility theorem

Theorem 5.6. *(Rational and continuous) Preferences may be represented by an utility function with the expected utility form if and only if it respects the axiom of independence.*

! Proof

- Proof of Necessity: if $\succcurlyeq$ respect the axiom of independence, then it may be represented by a utility function with the expected utility form.

Assume there are lotteries $\bar{p}$ and $\underline{p}$ such that for all p, one has $\bar{p} \succcurlyeq p \succcurlyeq \underline{p}$.
If $\bar{p} \sim \underline{p}$, the result follows immediately: use a constant utility function.
Assume from now on $\bar{p} > \underline{p}$.
Step 1
Take α and β such that $1 > \beta > \alpha > 0$.
Write:

$$\begin{gathered}
\bar{p} = \\
\beta\bar{p} + (1-\beta)\bar{p} > \\
\beta\bar{p} + (1-\beta)\underline{p} = \\
(\beta-\alpha)\bar{p} + \alpha\bar{p} + (1-\beta)\underline{p} > \\
(\beta-\alpha)\underline{p} + \alpha\bar{p} + (1-\beta)\underline{p} = \\
\alpha\bar{p} + (1-\alpha)\underline{p} > \\
\alpha\underline{p} + (1-\alpha)\underline{p} = \\
\underline{p}
\end{gathered}$$

From the first to the second line: $\bar{p}$ is the average of $\bar{p}$ and $\bar{p}$!
From the second to the third line: we apply the axiom of independence. Observe that

we keep $\bar{p}$ in the first term of the sum, but substitute $\bar{p}$ for $\underline{p}$ in the second term. Since $\bar{p} > \underline{p}$, the axiom of independence implies the strict preference.
From the third to the fourth line: add and subtract $\alpha\bar{p}$.
From the fourth to the fifth line: again, we just substitute $\bar{p}$ for $\underline{p}$ in one term of the sum, and leave the rest unchanged. The axiom of independence applies again.
From the fifth to the sixth line: we cancel out $\beta\underline{p}$.
From the sixth to the seventh line: we repeat the argument of the 2^{nd} to 3^{rd} line, in reverse order.
From the seventh to the eighth line: we repeat the argument of the 1^{st} to 2^{nd} line, in reverse order.
Step 2: for all p, there is only one λ_p such that $\lambda_p\bar{p} + \left(1-\lambda_p\right)\underline{p} \sim p$
Existence follows from continuity. For any lottery p, define the sets:

$$\{\lambda \in [0,1] : \lambda\bar{p} + (1-\lambda)\underline{p} \succcurlyeq p\}$$
$$\{\lambda \in [0,1] : \lambda\bar{p} + (1-\lambda)\underline{p} \preccurlyeq p\}$$

Continuity and completeness of $\succcurlyeq$ imply that both sets are closed (why?). Moreover, any λ belongs to at least one of these sets. Since both sets are non-empty and $[0,1]$ is connected, there must by some λ belonging to both (again: why?). Define it as λ_p.
Uniqueness follows from the previous step. If we were to slightly increase the value of λ_p (from α to β in the notation of the previous step), we would get a new lottery strictly preferred to the DM, breaking indifference.
Step 3: $U(p) = \lambda_p$ is a utility function that represents $\succcurlyeq$
Consider two lotteries p and q.
From steps 1 and 2, we may write:

$$p \succcurlyeq q \Leftrightarrow \lambda_{\mathrm{p}}\bar{p} + \left(1-\lambda_{\mathrm{p}}\right)\underline{p} \succcurlyeq \lambda_{\mathrm{q}}\bar{p} + \left(1-\lambda_{\mathrm{q}}\right)\underline{p} \Leftrightarrow \lambda_{\mathrm{p}} \succcurlyeq \lambda_{\mathrm{q}}$$

The first $\Leftrightarrow$ comes from step 2: use $p \sim \lambda_{\mathrm{p}}\bar{p} + \left(1-\lambda_{\mathrm{p}}\right)\underline{p}$, and analogously $q \sim \lambda_{\mathrm{q}}\bar{p} + \left(1-\lambda_{\mathrm{q}}\right)\underline{p}$.
The second $\Leftrightarrow$ comes from step 1 , taking $\lambda_{\mathrm{p}} = \beta$ and $\lambda_{\mathrm{q}} = \alpha$.
In short, $p \succcurlyeq q \Leftrightarrow \lambda_{\mathrm{p}} \succcurlyeq \lambda_{\mathrm{q}}$. This is the definition of an utility function representing $\succcurlyeq$.
Step 4: $U(p) = \lambda_p$ has the expected utility form.
We have to show that for all $\alpha \in [0,1]$, and for any lotteries p, p', one has:

$$U\left[\alpha p + (1-\alpha)p'\right] = \alpha U(p) + (1-\alpha)U\left(p'\right)$$

From step 2, we have:

$$p \sim \lambda_{\mathrm{p}}\bar{p} + \left(1-\lambda_{\mathrm{p}}\right)\underline{p}$$
$$p' \sim \lambda_{\mathrm{p}'}\bar{p} + \left(1-\lambda_{\mathrm{p}'}\right)\underline{p}$$

From step 3, we have : $U(p) = \lambda_p$. These two relations become:

$$
\begin{aligned}
p &\sim U(p)\bar{p}+(1-U(p))\underline{p} \\
p' &\sim U(p')\,\bar{p}+(1-U(p'))\,\underline{p}
\end{aligned}
$$

Take a convex combination $\alpha p+(1-\alpha)p'$. Given the two relations above, we have:

$$
\begin{aligned}
&\alpha p+(1-\alpha)p' \sim \alpha[U(p)\bar{p}+(1-U(p))\underline{p}]+(1-\alpha)\left[U(p')\,\bar{p}+(1-U(p'))\,\underline{p}\right] \text{ Factor out } \bar{p} \text{ and } \underline{p} \\
&\sim [\alpha U(p)+(1-\alpha)U(p')]\cdot\bar{p}+[\alpha(1-U(p))+(1-\alpha)(1-U(p'))]\cdot\underline{p} \\
&\quad\sim [\alpha U(p)+(1-\alpha)U(p')]\cdot\bar{p}+[1-(\alpha U(p)+(1-\alpha)U(p'))]\cdot\underline{p}
\end{aligned}
$$

Notice now that the terms in red are the same. We may denote it by any letter - for example, λ (without the subscript to distinguish it from both λ_p and $\lambda_{p'}$): $\lambda=\alpha U(p)+(1-\alpha)U(p')$.
The last line becomes:

$$
\lambda\cdot\bar{p}+[1-\lambda]\cdot\underline{p}
$$

In short, and without all the colors:

$$
\alpha p+(1-\alpha)p' \sim \lambda\cdot\bar{p}+[1-\lambda]\cdot\underline{p}
$$

But this is the very definition of λ_p as defined in step 2 , only applied to $\alpha p+$ $(1-\alpha)p'$.
(If you want, you may write $\lambda_{\alpha p+(1-\alpha)p'}$ instead of λ.)
Or, using step 4, we have the more intuitive notation $\lambda=U[\alpha p+(1-\alpha)p']$.
But we just defined $\lambda=\alpha U(p)+(1-\alpha)U(p')$.
Since these are the same λ, one has:

$$
U[\alpha p+(1-\alpha)p']=\alpha U(p)+(1-\alpha)U(p')
$$

That is, U has the expected utility property, concluding the proof.
QED.