

# Decision Theory

E. Glen Weyl

University of Chicago

First Guest Lecture  
Price Theory  
Fall 2011

# Defining rationality

Rationality perhaps central foundation of economics

- But what does it mean?
  - ① Common: acting to achieve ends as well as possible
    - I don't like this: "well" by whose standards?
  - ② Acting in a manner to which you can respect
    - The person has reflected carefully on their action
    - Has goals and beliefs that seem worth listening to
    - Sort of person with whom you could converse
    - Need not agree, but not clearly confused or mistaken
- Persuade you on definition, some elements of rationality
  - ① Classical rationality and why it is meaningless on its own
  - ② Requirements beyond consistency for rationality
  - ③ Rational attitudes to uncertainty and monetary risk
  - ④ Empirical data on the right form of risk aversion

# Roles of rationality

Rationality serves two basic purposes in economics

- 1 Normative: how *should we* act and think?
  - Normative role of decision theory useful to us
  - Help us resist confused, poor reasoning
  - Help us analyze arguments, think through problems
    - Form the foundations of econometrics
    - Help formulate normative planning in theoretical models
- 2 Positive: how *do people* act and think?
  - Fundamental postulate of economics: people act rationally
    - Gary would argue this is what *defines* economics
  - Most of this course interprets/predicts behavior as rational
  - This is a powerful approach, much like evolution
    - In evolution, traits interpreted as adaptive to environment
    - Economics: behavior optimal s.t. constraints and information

# Revealed preference and classical axioms

Classical notions of economic rationality based on consistency

- 1 Of preferences ( $x \succ y$  means  $x$  is preferred to  $y$ )?
  - Asymmetry: If  $x \succ y$  then  $y \not\succeq x$
  - Transitivity: If  $x \succ y$  and  $y \succ z$  then  $x \succ z$ 
    - If you reflect, you have to come to such a preference
    - Inconsistency is impossible to sustain on reflection
- 2 Or of choices ( $x \in c(A)$  means  $x$  is a choice from  $A$ )?
  - WARP:  $x, y \in A, B, x \in c(A)$  and  $y \in c(B) \implies x \in c(B)$ 
    - $x$  chosen when  $y$  available
  - $\implies$  Whenever  $y$  is chosen,  $x$  must be as well

Equivalent ways of expressing same idea

- *Revealed preference* is link between these
- $\implies$  Interpret choice as preference and vice-versa

# Preferences, choices and utility

To link, we say  $x \succ y$  if  $x \in S \implies y \notin c(S)$

- Then WARP  $\implies$  transitivity; suppose  $x \succ y$  and  $y \succ z$ 
  - ①  $x \succ z$  is  $z \notin c(\{x, z\})$  or  $z \notin c(\{x, y, z\})$
  - ② Suppose that either of these is contradicted
  - ③ Either directly or by WARP  $z \in c(\{x, y, z\})$ , violating  $y \succ z$
- Utility function *may* be seen as a representation?
  - If so, defined only up to positive monotone transformations
  - Simply means  $U(x) > U(y) \iff x \succ y$
- For some purposes, this is sufficient
  - But obviously makes impossible inter-personal comparisons
  - This is often unsatisfying from a normative perspective
- Also this "minimalist" notion of rationality unsatisfying
  - In fact, unclear whether it has any content at all

# Sen: does consistency ask too little?

Let's consider Kahneman and Tversky's paradigmatic example

- Set-up #1: choose between
  - 1 Let 200 die for sure
  - 2 Have 600 die with  $\frac{1}{3}$  or none with  $\frac{2}{3}$
- Almost everyone chooses option 2
- Set-up #2: 600 people will die if nothing done:
  - Save 400 for sure
  - Save 600 with  $\frac{2}{3}$ , no one with  $\frac{1}{3}$
- Almost everyone chooses first
- Obvious on reflection that these contradict...
  - Or is it? Couldn't someone argue framing part of decision?
  - Who's to say this isn't relevant to the person?
  - Yet if we start going down this road, rationality loses content

⇒ It is never "just" internal consistency we care about

# Some methodological take-aways

## Positive

- 1 Rationality is a language more than an assumption
  - So are all scientific theories, as Putnam points out
    - Universal gravitation, evolution by natural selection, etc.
- 2 What matter is how fruitfully we use it
  - Question not "is this behavior rational or irrational?"
  - Instead: "find plausible and useful rationalist account"
- 3 Most uses involve strong restrictions on what people want
  - Standard mechanism design assumption  $u = v - p$
  - Powerful because requires more than "internal consistency"

## Normative

- 1 Criticism always based on more than consistency
  - Tax subsidies v. subsidies in congressional debates
- 2 Consistency still disciplines *combined with substance*

# Some requirements beyond consistency

Some common substantive requirements: examples?

- 1 Plausible preferences
  - Psychology, biology, sociology, history, literature all useful
  - Often long-term material gain important component
  - Conditions, not preferences, differ primarily
- 2 Reasonable beliefs
  - Not necessarily what you'd think (different information)
  - But beliefs you wouldn't be embarrassed about holding
- 3 Consistency across time
  - Life goals executed consistently (exponential discounting)
- 4 Induction and learning
  - Einstein's definition of insanity...nothing internal about this
- 5 Behavioral: break people into multiple (conflicting) motives

Not only paradigm in economics, but primary one

# Probability

What are reasonable/desirable attitudes to uncertainty?

- You should have probabilities and obey laws of probability
  - 1 Assign number between 0 and 1 to any event
  - 2 The probability of *something happening* is 1
  - 3 If  $E_1, E_2 \dots$  are *disjoint*  $\implies P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$
- These  $\implies$  Bayes's rule, etc. (see stats class)
- Thus I am asking you be a *Bayesian*
  - Probabilities represent *how plausible an event is*
  - "Subjective" in that they represent individual beliefs
  - Bayesianism hard: how to assign probability to everything?
    - Beliefs may be hard to pin down
    - What is belief about US GDP in 2020?
    - What is distribution of GDP growth each year?
  - $\implies$  You won't always be fully coherent
- Nonetheless, compelling arguments for constantly *trying*

# Dutch books

If you don't, an opponent can arbitrage you: example

- Consider Nobel winner next year; two possibilities
  - 1 Lars Hansen wins alone
  - 2 Jean Tirole wins alone
- Someone asks you to set fair odds on each
  - Your prices  $p_H, p_T$  for dollar if Tirole/Hansen win
  - Also  $p_{HT}$  if *either wins*
  - Price where you'd be indifferent to buying or selling
- Disjoint so that probability laws require  $p_{HT} = p_H + p_T$
- If you violate this, someone can mint money off you
- Suppose  $p_{HT} < p_H + p_T$ 
  - ⇒ Someone can buy joint, sell you both and make money
- Presumably you don't want to be taken advantage of
- Every other violation of Bayesian allows similar exploitation

# A Sure-Thing story

A reasonable person uses probability in specific way: story

- A researcher goes into meeting to interview politician
  - He starts asking her about loyalty to husband
  - She's unsure of his intentions:
    - 1 He may be hitting on her
      - In this case she wants to rebuff him
    - 2 He may be trying to test her morals
      - In this case she wants to pass the test
  - In either case, best to respond loyally
- ⇒ She should, regardless of ignorance, act loyal
- This reasoning is known as

## Savage's Sure Thing Principle

*If  $\alpha \succ \beta$  if you knew  $E$  occurred, and same if you knew  $\neg E$  occurred, then you should do  $\alpha \succ \beta$  regardless.*

# The essence of expected utility

These principles (probability and sure thing) imply EU

$$U(\alpha) = \sum_{s \in \mathcal{S}} u(\alpha, s) = \sum_{s \in \mathcal{S}} u(\alpha, s)p(s)$$

- 1 Sure-thing makes utility *independent across states*
- 2 Probability makes *utility weights* constant across acts
  - In fact, probability *can be derived* from sure-thing
  - Need a few other technical conditions, but this is heart
  - Savage's proofs are quite involved, but this is basic idea

Again, we have quite a bit of flexibility in utility

- Von Neumann-Morgenstern utility index  $u$  semi-cardinalized
  - Defined up to *level and scale*  $\zeta u + \nu$
- As before, we can choose to cardinalize it fully to compare

Upshot is that you probably (should) want to maximize EU

# Ellsberg's paradox

Nonetheless, important objections have been made

- Most fundamental is *Ellsberg's Paradox*
  - An urn has 100 balls, some red, some black
  - Don't know how many of each color
  - One ball will be picked at random; three choices:
    - 1 Get \$101 if red ball pulled
    - 2 Get \$101 if black ball pulled
    - 3 Get \$100 if I get heads on a fair coin
  - Most people prefer the third option
  - But note that this is inconsistent with probability:
    - If  $p(R) \geq p(B)$  then choose first option
    - If  $p(B) \geq p(R)$  then choose second option
    - NEVER choose third option

# Allais's paradox

Another famous challenge is from Allais; two experiments:

- Experiment 1, two choices:
    - 1 Get \$1 million for sure
    - 2 Get \$5 with 10% chance, \$1 with 89%, and nothing with 1%
  - Most people choose the first option
  - Experiment 2, two choices
    - 1 \$1 with 11%, nothing with 89%
    - 2 \$5 with 10%, nothing with 90%
  - Most people choose the second option
  - Note this is inconsistent with Sure-Thing
    - $E \equiv$  89% chance that in first experiment second gets \$1
    - Second exp. simply makes this into nothing in that event
    - Just as indifferent between \$0 and \$0 as \$1 and \$1
- ⇒ Your preferences should be consistent

# Is this rational behavior?

Three common responses to these results?

- 1 People don't actually behave this way
  - Some of these may be artificial hypothetical settings
  - John List has does great work on this
  - Something to this, but many believe real-world examples
- 2 These are mistakes
  - Basically my view; think about Ellsberg
  - Generate coin flip by flipping for choice of Red or Black
  - But this means you are *worse knowing outcome of flip*
    - Once you do, you are back to plain Red or Black
  - This seems absurd, Dutch Bookable etc.
  - People likely confused, think experimenter is hostile
  - I don't think you'd respect someone who behaves this way
- 3 It is reasonable not to try to maximize EU
  - Long line of recent literature on this (but I am skeptical)

# Utility over money

Many risky situations affect assets fungible with money

- In these cases, we can focus on *utility over money*
  - Let us denote this  $u(w)$ ; we standardly assume concave
    - First exhaust best things to do with money, then move on
    - Intuition that poor value marginal dollar more than rich
    - However, easy to find counterexamples?
      - To start a business, need minimum capital
      - In this case utility is convex over range
    - Often easy to find fair lotteries (betting etc.)
      - If so, concavifies utility over everything else
- ⇒ We will almost universally assume weakly concave
- Also argument based on rationality

# The St. Petersburg Paradox: the case for risk-aversion

You might be asking, why not just be risk-neutral?

- Seems a natural given EU; possible bet shows wrong
  - Flip coin. If tails, you get \$2; if heads, flip again
  - If tails, you get \$4; if heads, flip again
  - \$8, \$16, \$32, ...
- How much would you be willing to pay for this?
- How much would a risk-neutral person be willing to pay?
  - $\frac{1}{2} \cdot 2 + \frac{1}{4} \cdot 4 + \dots = \infty!$

⇒ Apparently you're not risk neutral
- What is the maximum a reasonable person would pay?
  - If finite, you believe rationality requires risk aversion
  - Or at least *eventual* risk-aversion, as I do

# Common measures of risk-aversion

One common measure is *absolute risk-aversion*(ARA):

$$A(w) = -\frac{u''(w)}{u'(w)}$$

- If constant at  $\alpha$ ,  $u(w) = 1 - e^{-\alpha w}$ , *constant ARA* (CARA)
- ⇒ Willingness-to-pay for %50 chance of \$100,000 constant
- Much more likely rich would pay more for this than poor
  - Thus CARA not very plausible...
    - But tractable with normal distribution, so often used

Another common measure is *relative risk-aversion* (RRA)

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

- If constant at  $\rho$ ,  $u(w) = \frac{w^{1-\rho}}{1-\rho}$  (CRRA)
- ⇒ Willing to pay constant  $\beta w$  for 50% at  $\gamma w$

# What's the best guess we have?

CRRA seems much more plausible than CARA

- But what do we know within this class?
- Most convenient is when  $\rho \rightarrow 1$ 
  - By L'Hôpital's rule,  $u(w) = \log(w)$
- Super-intuitive (value of x% increase in wealth constant)
- Mathematically ultra-convenient in many cases
- Most important, has more empirical support than any other
  - Most striking example is from Sacks et al. (2011) data
  - Study "subjective well-being" data (questionable, obviously)
    - Ladder of life, 0 to 10 (worst to best life imaginable)
  - Letting  $h$  be well-being and  $w$  be wealth they find

$$h = .35 \log(w) + \epsilon$$

fits extremely well within and across countries

⇒ Log-utility good baseline specification



# Rabin's Paradox: the case for practical risk-neutrality

So, both positively and normatively, risk-aversion is useful

- “Risk aversion” is from declining marginal utility of wealth
  - ⇒ Only if lifetime wealth substantially affected does risk matter
    - When you think about investing, very relevant, but...
    - In everyday life, you should be risk-neutral
- Matthew Rabin gives sharp expression, how?
  - If  $w = \$290k$  and  $\forall w < \$300k$  reject 50-50 lose \$100 v.
  - Gain top row ⇒ reject 50-50 lose L, gain entries
- Reasonable: somewhere between 1st and 2nd column
  - ⇒ Over stakes of \$100 or less, risk premium no more than 5%
- Thus irrational to be risk-averse over modest stakes

## Rabin's results

TABLE II

TABLE I REPLICATED, FOR INITIAL WEALTH LEVEL \$290,000,  
WHEN  $l/g$  BEHAVIOR IS ONLY KNOWN TO HOLD FOR  $w \leq \$300,000$ .

$L$	$g$			
	\$101	\$105	\$110	\$125
\$400	400	420	550	1,250
\$600	600	730	990	36,000,000,000
\$800	800	1,050	2,090	90,000,000,000
\$1,000	1,010	1,570	718,190	160,000,000,000
\$2,000	2,320	69,930	12,210,880	850,000,000,000
\$4,000	5,750	635,670	60,528,930	9,400,000,000,000
\$6,000	11,510	1,557,360	180,000,000	89,000,000,000,000
\$8,000	19,290	3,058,540	510,000,000	830,000,000,000,000
\$10,000	27,780	5,503,790	1,300,000,000	7,700,000,000,000,000
\$20,000	85,750	71,799,110	160,000,000,000	540,000,000,000,000,000,000

## Common “risk-aversion” and prospect theory

Nonetheless, it seems in practice people often are

- Common explanation: “Prospect Theory”
- Many formulations, but share common features
  - ① Changes relative to reference point, not final wealth
    - Current wealth, current expected wealth, anchor, etc.
  - ② Losses hurt more than gains help
    - “Kink” generates risk aversion over very small stakes
  - ③ Risk-seeking over losses, risk-averse over gains
    - Helps rationalize pattern with the disease earlier
  - ④ Non-linear probability weighting
- Empirical relevance
  - Most successful theory in laboratory (very influential)
  - Much more mixed in field; experience may cure
- Viewed by most as irrational behavior
  - Normatively, yes, but still predictable

## What not to insure against

Risk aversion also does not explain/justify “false” insurance

- Often hear people discuss “insuring” against anything bad
  - You don’t get into expensive school, do badly on test etc.
- But bad news only weakly related to insurance under EU
- Key is *marginal utility of wealth* (MUW)
  - You may want to insure *against* good events
    - You may get an opportunity to make great investment
    - If you do, this is clearly good news
    - But you also want to “insure” against this, as MUW higher
  - Many types of bad news you don’t want to insure against
    - Your girlfriend leaves you
    - Your team loses at a sports match
    - You do poorly on job market, take high-paying consulting job
- Only MUW should be reason for insurance
  - Everything else a form of betting/gambling