Rational Choice and its Limits

Antony Eagle

Choices, Models and Morals » Lecture 3

Contents

Expected Utility

- Last time we introduced the expected utility framework for modelling rational decisions under uncertainty.
- When confronted with a choice between what you subjectively regard as lotteries, choose the act that has the **greatest subjective expected utility**.
- Subjective in two ways: the values assigned to outcomes are yours, not objectively given; and the probabilities assigned to outcomes are yours, not necessarily objective chances.
  - Many people are comfortable with subjectivity of value: 'Reason is, and ought only to be the slave of the passions' (Reiss 2013, p. 50, Hume 2023, sec. 2.3.3.4).
  - But some are uncomfortable using subjective judgments of how likely outcomes are. (Is there any legitimacy to this asymmetry?)
- This theory nevertheless faces challenges:
  1. Even as a normative theory of what makes choices **rational**, there are concerns about whether standard expected utility theory gives the wrong results.
  2. And as a theory used to **explain behaviour**, the theory must be **descriptively adequate**. But is it?

Puzzling Choices: Sure Things and Managing the News

Dominance

- Consider the rule of **dominance**: if act A leads to an equal or better outcome in every state of nature than any other action, and a strictly better outcome in at least one state of nature, then choose act A (Sainsbury 2009, p. 74).
  - This rule is simple – it doesn't even mention probability! – but it only applies to a **restricted** set of decision problems. In many decision scenarios, there is no dominant act.
- Note that the standard expected utility theory **explains why** dominance is a good rule.
  - If the utilities assigned to A in every state of nature are at least as high than those assigned to B, then any weighted average of those utilities that assigns the same probabilities to states of nature will assign A an expected utility at least as high as that assigned to B.

Dominance Illustrated

- Consider this case, where Israel is deciding whether to withdraw from its occupied territories. (The example is very old but unfortunately still timely.) In some choice of utility units, the payoffs for this decision may look like this:

Table 1: Israeli War (Bar-Hillel and Margalit 1972, pp. 296–7)

Acts\States	Egypt declares war D	No war $\neg D$
Israel withdraws W	0	2
Israel remains R	1	3

- In this case, act R **dominates**: it is higher utility in every state of nature.
  - Suppose the Israeli generals think the probability that Egypt declares war is p. Then $EU(W) = 2(1-p) = 2-2p$, and $EU(R) = (1p + 3(1-p) = 3-2p$.
  - No matter what p is, $EU(R) - EU(W) = 1$: this entails that $R \succ W$.
- So dominance reasoning, and standard expected utility theory, suggests that Israel ought to remain in the occupied territories.

Puzzling Choices

- But in a case like this, dominance can lead to puzzling choices.

> Clearly, remaining in the occupied territories is the dominant strategy.... Suppose, however, that you believe that with a high probability withdrawal will be conducive to peace while remaining in the territories will eventually lead to war. Then you might prefer to withdraw and end up in the 2 cell than to remain and end up in the 1 cell. (Bar-Hillel and Margalit 1972, pp. 296–7)

- Here, the states of nature are not **independent** of the possible acts, and then dominace – and standard EU theory! – gives intuitively the wrong results: it suggests Israel would be **irrational** if they preferred to withdraw, and that doesn't seem irrational.
  - This intuition also leads to a failure of Savage's **sure-thing principle** (Hausman *et al.* 2017, p. 61), which is very close to dominance: these prospects differ only in that the prizes are sweetened uniformly in R versus W.

Act-Outcome Dependence

- When the probabilities of outcomes depend on the act chosen, dominance and the sure-thing principle give incorrect advice.
- The standard alternative is to tweak the definition of expected utility to use **conditional probability**, reflecting how the probability of an outcome varies conditional on acts. $\Pr(p \mid q)$ is read 'the probability of p given the supposition that q'.

(EU redefined) The **expected utility** of action a in some decision scenario for a subject is (re)defined as follows, where $s_1, \ldots, s_n$ are the possible states of nature, Pr is a degree of belief function, and U is a subjective utility function:

$$EU(a) = \Pr(s_1 \mid a)U(s_1 \wedge a) + \ldots + \Pr(s_n \mid a)U(s_n \wedge a)$$

$$= \sum_{i=1}^{n} \Pr(s_i \mid a)U(s_i \wedge a).$$

Refined Expected Utility in the Egypt-Israel Case

- Suppose we have more realistic probabilities in the Egypt-Israel case (these are credences of the Israeli decision makers):

$$\Pr(D \mid R) = 0.8, \Pr(\neg D \mid W) = 0.2, \Pr(D \mid W) = 0.1, \Pr(\neg D \mid W) = 0.9.$$

- Then we calculate the **refined expected utility** of R and $\neg R$ as follows:

$$\begin{aligned} EU(R) &= \Pr(D \mid R) \times U(D \wedge R) + \Pr(\neg D \mid R) \times U(\neg D \wedge R) \\ &= 0.8 \times 1 + 0.2 \times 3 = 1.4 \\ EU(W) &= \Pr(D \mid W) \times U(W \wedge W) + \Pr(\neg D \mid W) \times U(\neg D \wedge W) \\ &= 0.1 \times 0 + 0.9 \times 2 = 1.8. \end{aligned}$$

- And we see that the expected value of remaining is lower than that of withdrawing, according to Israel's values and probabilities, **despite** the fact that remaining dominates withdrawing.
- This approach takes utilities as **basic**; we can approach things the other way, and take preference as basic – but we'd need some other assumptions on the structure of rational preference (Jeffrey 1983, pp. 8–10).

Newcomb's Problem

- Another puzzling case is Newcomb's problem (Sainsbury 2009, p 69, Hausman *et al.* 2017, p. 62). I present a 'medical' variant on the standard case

> Susan is debating whether or not to smoke. She believes that smoking is strongly correlated with lung cancer, but only because there is a common cause—a condition that tends to cause both smoking and cancer. Once we fix the presence or absence of this condition, there is no additional correlation between smoking and cancer. Susan prefers smoking without cancer to not smoking without cancer, and she prefers smoking with cancer to not smoking with cancer. Should Susan smoke? It seems clear that she should. (Set aside your theoretical commitments and put yourself in Susan's situation. Would you smoke? Would you take yourself to be irrational for doing so?) (Egan 2007, p. 94)

Payoffs and Probabilities in Medical Newcomb

Table 2: A Medical Newcomb Puzzle; S dominates $\neg S$

Acts\States	Cancer gene C	No cancer gene $\neg C$
Smoke S	2	20
Refrain $\neg S$	1	10

- Probabilities reflecting Susan's judgment on the basis of the statistics: they don't necessarily reflect the real **causal structure**:

$$\Pr(C \mid S) = 0.8, \Pr(\neg C \mid S) = 0.2, \Pr(C \mid \neg S) = 0.2, \Pr(\neg C \mid \neg S) = 0.8.$$

$$EU(S) = 0.8 \times 2 + 0.2 \times 20 = 5.6;$$
$$EU(\neg S) = 0.2 \times 1 + 0.8 \times 10 = 8.2.$$

- But Susan doesn't want to simply **act like** a non-cancer-gene-haver; she wants to **be one already**. And regardless of whether she is, smoking looks more fun and – *she believes* – it won't influence her health. So why refrain, even though it maximises expected utility?

Explaining the Intuitions in the Medical Newcomb case

- In the Medical Newcomb case, the recommendation of (refined) expected utility theory seems to be this: **act like the person you want to be**.
  - Looking at the statistics, it is clear to Susan that most smokers end up with cancer and obtain 2 utiles, and most non-smokers end up with 10 utiles. But not smoking for this reason, in Susan's view, is like interfering with the weather forecast and hoping thereby to prevent a thunderstorm.
- This diverges from this attractive recommendation:

  > do what's most likely to bring about the best results. (Egan 2007, p. 113)

- So we aim for a further revision to decision theory, aiming to replace non-causal conditional probabilities by something which distinguishes mere correlation from causation.
  - Lewis (1981) proposes that, instead of weighting utilities by conditional probabilities of outcomes given acts, we ought to weight them by the probabilities of hypotheses concerning how outcomes **causally depend** on acts.
- The debate over how to respond to Newcomb cases is vast and we cannot survey it here (Sainsbury 2009, pp. 69–82).

Borderline Intuitions: Allais and Ellsberg

Intuition and Decision Theory

- A successful theory of rational decision should **systematise** our pre-theoretical judgments ('intuitions') about what we ought to do.
- The discussion of the failures of dominance reasoning in the Israel-Egypt case, was intended to elicit a very clear judgment that it is at least **rationally permissible** not to follow Dominance in cases where act-outcome independence is violated.
  - Very common in strategic interactions. But note in a 'regular' smoking case, you might prefer to smoke whether or not you have lung cancer, but whether you have lung cancer is causally determined by whether you smoke. We wouldn't want to recommend the dominant act; it promotes states you don't want to be in.
- But perhaps further intuitions, in Newcomb's problem cases, prompt further refinements as we attempt to systematize intuition.
- In these cases the intuitions are clearly correct. But that's not always obvious. Some widespread intuitive judgments may conflict with standard decision theory, but we might respond that it is the intuitions which have to go, rather than the theory which needs refinement.

Allais' Paradox

- Some famous paradoxes of decision theory were designed as challenges to the normative status of standard expected utility theory.
- Consider Allais' paradox (Reiss 2013, pp. 46–7), in which ball is drawn from an urn of known composition: 1 red ball, 89 white, 10 blue. The 100 possible drawn balls are the states of nature.
- The decision problem is given by this payoff matrix:

Table 3: Allais' Paradox payoffs

Acts \ States	White (0.89)	Blue (0.1)	Red (0.01)	Expectation
A1	€1,000	€1,000	€1,000	€1,000
A2	€1,000	€5,000	€0	€1,390
A3	€0	€1,000	€1,000	€110
A4	€0	€5,000	€0	€500

- You face just one of two scenarios, with different outcomes: either a choice between A1 and A2, or a choice between A3 and A4.

Inconsistent with Expected Utility Theory

- Many people share the intuition that A1 is preferable to A2 (€1,000 no matter what!), and that A4 is preferable to A3 (if we're playing a lottery anyway, go for the bigger prize – after all, in A4 you have a 90% chance of nothing, but a 10% chance of €5000; while in A3 you have an slightly lower 89% chance of nothing, and a slightly higher 11% chance of a much worse prize of only €1000).
- Expected utility theory cannot **rationalize** this intuition.
  - The standard theory says that the utilities assigned to these monetary values are up to the agent – they can have diminishing marginal utility for money, or incorporate various kinds of risk aversion (Reiss 2013, p. 45) – but that **ratios of differences** between utilities assigned to acts are meaningful.
  - In this case the utility difference between A2 and A1, and A4 and A3, is each €390, so the ratio of their differences is 1 : 1. On any utility function at all, whether risk averse or otherwise, A4 will be preferred to A3, and A2 to A1.

Intuition and Sure Thing

- Intuitions here again violate the sure thing principle: someone who prefers A4 to A3 should continue to prefer it when we add to each lottery the condition that you get €1000 for sure if a white ball is drawn.
- But is intuition at fault here?

> there is certainly a relevant difference between the pairs A1/A2 and A3/A4. If I were to end up [with a red ball] after choosing A2, I will regret my choice a great deal. I could have had a good amount of money for sure. I chose to gamble and lost. That was silly. In the choice between A3 and A4, the odds that I end up with nothing are overwhelming anyway. I'd consider myself lucky if I did win but not winning wasn't silly. Quite to the contrary, it would have been unreasonable to forfeit a good chance of a considerably higher gain for a minimally smaller chance of losing. I would not regret my choice. (Reiss 2013, pp. 47–48)

A RCT-preserving Response: Manipulate the Outcomes

- One response: Maybe we have **individuated** the outcomes poorly. If people do prefer certainty and want to minimise regret, we ought to include regret **in the outcomes**.
- We should actually have this payoff matrix:

Table 4: Allais' Paradox refined

Acts \ States	White (0.89)	Blue (0.1)	Red (0.01)	Expectation
A1	€1,000	€1,000	€1,000	€1,000
A2	€1,000	€5,000	–€x	€(1,390 - x/100)
A3	€0	€1,000	€1,000	€110
A4	€0	€5,000	€0	€500

- You can verify: if the agent values regret at not getting €1,000 in A2 low enough, we break the symmetry between A1/A2 and A3/A4.
- Yet is it plausible to fiddle with the outcomes in this way? Especially to give them **holistic** features that depend on an **unchosen** option?

Ellsberg's Paradox

- Another similar challenge to standard expected utility theory is Ellsberg's Paradox (Reiss 2013, pp. 48–49).
- There is again an urn, but this time of **unknown** composition. It contains 90 balls, 30 red, and the remaining 60 are some unknown mix of yellow and black.
- You face two decisions:

  Choice 1 Option A – You receive €100 if you draw a red ball.
  Option B – You receive €100 if you draw a black ball.
  Choice 2 Option C – You receive €100 if you draw a red or yellow ball.
  Option D – You receive €100 if you draw a black or yellow ball.

- In these cases, expected utility says you should prefer A to B, and C to D, if your subjective probability for *black* is higher than that of *red*, i.e., 1/3.
- But again, **intuition** suggests that D is preferable to C – D has an expectation of no less than €66.67, while C's unknown expectation could be as little as €33.33. Intuition recommends we 'avoid gambles with unknown probabilities' (Reiss 2013, p. 49).

Alternative Theories of Rational Choice

Two main options in response to these contestable intuitions (Kahneman 2011, p. 314).

1. They can simply **ignore** these intuitions that conflict with standard rational choice theory. Not every intuition must be respected; maybe there is a psychological story to be told about why people have these irrational intuitions, but we don't need to accomodate them in our theory.

   › Compare: some people appear to find gay sex repugnant; but this intuition, or gut reaction, is not one we should change our moral theories to accommodate. The right response to such intuition is therapy, not theory.

2. We could try to develop **alternative** axioms on rational preference that allow the patterns of preference exhibited in Allais and Ellsberg cases to be rationally permissible.

   › Current versions of this sort of approach often involve non-standard ways to understand risk aversion (Buchak 2013, Peterson 2017, sec. 8.5).

Neither option is easy: either we need to **explain away** the intuitions, or we need to formulate a theory respecting them, which isn't trivial (the more patterns of judgment we need to accomodate the harder it is to systematize them).

Behavioural Economics and the Empirical Basis of Economic Explanation

Economic Prediction and Explanation

- The prevalence of Allais- and Ellsberg-like intuitions poses a challenge for EU theory as a descriptively accurate model of behaviour. Remember that we came across EU theory as a way of representing people's reasons for action in answer to *why* questions about why those actions were performed: e.g., *why did Bob take the train? Because taking the train had, according to him, the highest expected value among his options*.
- If enough people act on Allais-like preferences, then this kind of explanation falls over. *Why did Alice opt for A1 over A2 and A4 over A3?* cannot be answered by citing any facts about maximal expected value.
- Perhaps this is a fringe issue:

> the assumption that people are rational leads us to some clear and testable theories about the way the world works. … Of course, there isn't much use in producing clear and testable theories if the theories are always wrong. But they aren't – economists' faith in people's rationality is usually about right. … people are sufficiently rational often enough to make the assumption of rational choice a very useful one. (Harford 2008, p. 7)

- Is there any evidence other than mere 'faith' that rational choice theory holds 'often enough'?

The ‘Unusual Disease’

› Suppose we are confronted with a novel infectious disease, and these acts (particular epidemiological interventions) and their prospective outcomes:

Table 5: Tversky and Kahneman’s ‘Unusual Asian Disease’ (Tversky and Kahneman 1981, p. 453, Kahneman 2011, pp. 368–9)

Act	Outcome
No drug	600 people dead
Plan A	200 people saved
Plan B	600 people saved with probability 1/3 No one saved with probability 2/3
Plan C	400 people dead
Plan D	No one dead with probability 1/3 600 people dead with probability 2/3

Framing Effects

- Experimentally, **most** people (72%) prefer Plan A to Plan B, and most (78%) prefer plan D to plan C (Tversky and Kahneman 1981, p. 453).
  - The experimental protocol meant that different subjects were exposed to the A/B choice as to the C/D choice; no subjects were confronted with both
- Yet it is clear that these are just **different ways of describing** only two plans: plan A/C and plan B/D. To prefer A to B is to prefer A to D; to prefer D to C is to prefer D to A; hence, the common pattern of preference agent forces the irrational (money-pumpable) preference for A strictly over A.
- People seem to simultaneously evaluate the same outcomes as having different costs, depending on how they were **framed**: as lives saved, or deaths resulting, from an act (Slovic and Lichtenstein 1983, p. 601).
  - The psychological explanation offered is that people prefer to 'lock in' sure gains, but will gamble to avoid sure losses. The different framing triggers these different risk attitudes:

> choices involving gains are often risk averse and choices involving losses are often risk taking. … The change [in description] is accompanied by a pronounced shift from risk aversion to risk taking. (Tversky and Kahneman 1981, p. 453)

Preference Between Framed Choices

- In the case of the 'Unusual Disease', no model which involves probability and utility could rationalize these preferences between outcomes.
- Since we cannot argue that *200 people saved from a population of 600* is a different outcome than *400 people dying from a population of 600*, the only way to salvage people's attitudes as utility maximising is if they are choosing not between outcomes, but between outcomes **under a particular description**.
- Could it be rational to vary one's choices based merely on how things are described?
- It's reasonable to prefer *calling on Batman for help* to *calling on Bruce Wayne for help*, because it need not be *a priori* knowable that Bruce Wayne is Batman. (Recall Davidson (1963), pp. 686–7: acting intentionally is under a description of the intended outcome.)
- But surely 200 saved from 600 is *a priori* equivalent to 400 dead from 600. So the exhibited preferences in this case still seem **irrational**.

> The message about the nature of framing is stark: framing should not be viewed as an intervention that masks or distorts an underlying preference. … in the problem of the Asian disease … there is no underlying preference that is masked or distorted by the frame. Our preferences are about framed problems, and our moral intuitions are about descriptions, not about substance. …(Kahneman 2011, p. 370)

Irrational Preference: Preference Reversals

- Tversky and Kahneman followed this chain of thought to its natural end:

> We retained utility theory as a logic of rational choice but abandoned the idea that people are perfectly rational choosers. We took on the task of describing the choices people make, regardless of whether they are rational. (Kahneman 2011, p. 314)

- Since their pioneering work, many further studies have seemed to confirm this sceptical opinion about human rationality. Consider **preference reversals**:

> Preference reversals occur when individuals are presented with two gambles, one featuring a high probability of winning a modest sum of money (the *P* bet), the other featuring a low probability of winning a large amount of money (the $ bet). The typical [experimental] finding is that people often chose the *P* bet but assign a larger monetary value to the $ bet. (Slovic and Lichtenstein 1983, p. 596)

- This strategy looks irrational, given that the bets pay off in money.

Example

> choices among pairs of gambles appeared to be influenced primarily by probabilities of winning and losing, whereas buying and selling prices were primarily determined by the dollar amounts that could be won or lost. … the bet with the most predicted reversals was: *P* bet, 9/12 to win $1.20, 3/12 to lose $.10; $ bet, 3/12 to win $9.20 and 9/12 to lose $2.00. (Slovic and Lichtenstein 1983, p. 597)

Table 6: Payoff matrix for preference reversal scenario

Acts \ States	Tickets 1–9	Tickets 10–12	Expectation
P	$1.2	-$0.1	87.5¢
$	-$2	$9.2	80¢

- People generally prefer the *P* bet (which does in fact have a higher expectation).
- But people also state the $ bet to be worth more, when asked to price them.

Money pumps again

- An agent with this pricing structure can be exploited, because these preferences violate **transitivity** (Reiss 2013, pp. 41–2).
  - They are asked to set a fair price: x for the $ bet (e.g., 90¢), and invited to pay $x - \delta$ for it (e.g., 89¢), which they accept as advantageous.
  - They are given the option to switch, cost-free, to a P bet; as they prefer P to $, they accept.
  - They are then asked to set a fair price the P bet, which they now own. They do, at $y < x$ – e.g., 87.5¢.
  - They are then offered $y + \delta$ to sell the bet (e.g., 88.5¢), which they accept.
- The result? They **end up where they started**, having paid $x - \delta$, and received $y + \delta$, for a net gain of $(y - x) + 2\delta$.
- If the difference between x and y is big enough, $(x - y) > 2\delta$, so their net gain is **negative**, i.e., they suffer a net loss in cycling through a sequence of states that leaves them where they started – in our case, a net loss of 0.5¢.
- An agent with these preferences is evaluating the **very same outcome** as having different utilities, depending on how it is presented; again, not possibly rationalizable.

The Endowment Effect

Members of different groups were initially endowed with either a coffee mug, a chocolate bar, or nothing. They were then given opportunities to exchange their initial endowment for the other option, or choose if unendowed. Results:

Group	Mug Over Candy (%)	Candy Over Mug (%)	*N*
Give up mug to obtain candy	89	11	76
Give up candy to obtain mug	10	90	87
No initial endowment	56	44	55

> In this simple experiment, participants' preferences were not independent of the direction of the exchanges, as is commonly assumed. They expressed a dramatic asymmetry in valuations by weighing the loss of giving up their initial or reference endowment far more heavily than the foregone gains of not obtaining the alternative endowment. (Knetsch 1989, pp. 1278–9, see also Kahneman 2011, pp. 292–3)

- Mere possession (granted arbitrarily and with no significance to the subject) suffices make the subject prefer an option; this undermines the existence of any **stable**

Options in light of Behavioural Economics

Confronted with these experimental results, we have several options:

1. We can follow Tversky and Kaheman, **yield** the field of behavioural explanation to psychologists, and retreat to orthodox rational choice theory as **purely normative**.
   - This is the typically response of philosophers interested in rational decision making, who have long considered *akrasia* or weakness of the will: to act against one's judgment of the best (Davidson 1980).
2. We might attempt to **query** the experimental results – most of the experimental economics discussed in the introduction to Slovic and Lichtenstein (1983, pp. 596–597) were attempts to show the psychologists were doing bad science.
   - The replication crisis in psychology has affected many results in psychology – but these results have largely been successfully replicated (Klein *et al.* 2014).
3. Proponents of rational choice theory can strategically **retreat**: e.g., to the claim that people are 'often' rational. (In Harford (2008)'s words, 'people are sufficiently rational often enough'.)
   - But are they? Almost all of the choices we really face are framed in particular ways (perhaps by marketers who want to exploit our irrational tendencies), or involve choices involve exchanging goods for money which are systematically distorted as the endowment effect and preference reversals show.

Alternative Models of Reason

- If we join Kahneman as 'sceptics about rationality' (Kahneman 2011, p. 374), then do we abandon the hope of predicting or explaining human behaviour?
- Not necessarily. The sceptic about human rationality need not be a **sceptic about reasons**.
- They can appeal to other **psychophysical mechanisms** in explaining behaviour. These can still be theories of intentional action, because they can give alternative accounts of **motivating reasons** (Davidson 1963).
  - So we might cite the reasons that an agent was themselves causally moved by, without thinking that causal explanation to involve reasons that are **rational** according to any external norm. We are already committed to such explanation in the case of weakness of the will, where an agent is **knowingly** moved by reasons they regard as not rationally optimal. How much easier to be caused to act by irrational reasons of which one is **unaware**?

Prospect Theory

One example of a worked out version of this account is **prospect theory**:

> In prospect theory, outcomes are expressed as positive or negative deviations (gains or losses) from a neutral reference outcome, which is assigned a value of zero. Although subjective values differ among individuals and attributes, we propose that the value function is commonly S-shaped, concave above the reference point and convex below it.... the response to losses is more extreme than the response to gains. The displeasure associated with losing a sum of money is generally greater than the pleasure associated with winning the same amount....
>
> The second major departure of prospect theory from the expected utility model involves the treatment of probabilities. ... In prospect theory the value of an uncertain outcome is multiplied by a decision weight $\pi(p)$, which is a monotonic function of p but is not a probability. ... for low probabilities $\pi(p) > p$, but $\pi(p) + \pi(1 - p) \leq 1$. Thus low probabilities are overweighted, moderate and high probabilities are underweighted, and the latter effect is more pronounced than the former. (Tversky and Kahneman 1981, p. 454, see also Kahneman and Tversky 1979, Kahneman 2011, pp. 278–288)

Are Most Acts Rational, Even If Most Actors Aren't?

- The experimental studies often involve **undergraduate psych students** as subjects. Are these naive subjects **representative** of economic activity?
- There is some evidence that experience improves rationality. Experimental work on the **endowment effect** suggests that

> individual behavior converges to the neoclassical prediction as market experience intensifies, [yet] it remains an open question as to whether the endowment effect is absent for practiced consumers because of experience (treatment effect), or because a prior disposition toward having no such gap leads them to trade more often (selection effect). … I find that market experience significantly attenuates the endowment effect. (List 2003, p. 45)

- And who participates in the market more? **High volume traders**. So even if most **market actors** are irrational, most **market actions** might still be rational.
  - This wouldn't help with the explanation of most individual choices, but would help explain aggregate economic behaviour which is driven by high volume agents.

References

References

Bar-Hillel, M. and Margalit, A., 1972. Newcomb's paradox revisited. *The British Journal for the Philosophy of Science*, 23 (4), 295–304.

Buchak, L., 2013. *Risk and rationality*. Oxford University Press.

Davidson, D., 1963. Actions, reasons, and causes. *The Journal of Philosophy*, 60 (23), 685–700.

Davidson, D., 1980. How is weakness of the will possible? *In*: *Essays on actions and events*. Oxford University Press, 21–42.

Egan, A., 2007. Some counterexamples to causal decision theory. *The Philosophical Review*, 116 (1), 93–114.

Harford, T., 2008. *The logic of life*. Abacus.

Hausman, D., McPherson, M., and Satz, D., 2017. *Economic analysis, moral philosophy, and public policy*. 3rd ed. Cambridge University Press.

Hume, D., 2023. *A treatise of human nature*.

References (cont.)

Jeffrey, R.C., 1983. *The logic of decision*. University of Chicago Press.

Kahneman, D., 2011. *Thinking fast and slow*. Penguin.

Kahneman, D. and Tversky, A., 1979. Prospect theory: An analysis of decision under risk. *Econometrica*, 47 (2), 263.

Klein, R.A., Ratliff, K.A., Vianello, M., Adams, R.B., Bahník, Š., Bernstein, M.J., Bocian, K., Brandt, M.J., Brooks, B., Brumbaugh, C.C., Cemalcilar, Z., Chandler, J., Cheong, W., Davis, W.E., Devos, T., Eisner, M., Frankowska, N., Furrow, D., Galliani, E.M., Hasselman, F., Hicks, J.A., Hovermale, J.F., Hunt, S.J., Huntsinger, J.R., IJzerman, H., John, M.-S., Joy-Gaba, J.A., Kappes, H.B., Krueger, L.E., Kurtz, J., Levitan, C.A., Mallett, R.K., Morris, W.L., Nelson, A.J., Nier, J.A., Packard, G., Pilati, R., Rutchick, A.M., Schmidt, K., Skorinko, J.L., Smith, R., Steiner, T.G., Storbeck, J., Swol, L.M.V., Thompson, D., Veer, A.E. van 't., Vaughn, L.A., Vranka, M., Wichman, A.L., Woodzicka, J.A., and Nosek, B.A., 2014. Investigating variation in replicability. *Social Psychology*, 45 (3), 142–152.

References (cont.)

Knetsch, J.L., 1989. The endowment effect and evidence of nonreversible indifference curves. *American Economic Review*, 79, 1277–1284.

Lewis, D., 1981. Causal decision theory. *Australasian Journal of Philosophy*, 59 (1), 5–30.

List, J.A., 2003. Does market experience eliminate market anomalies? *The Quarterly Journal of Economics*, 118 (1), 41–71.

Peterson, M., 2017. *An introduction to decision theory*. 2nd ed. Cambridge University Press.

Reiss, J., 2013. *Philosophy of economics*. Routledge.

Sainsbury, R.M., 2009. *Paradoxes*. 3rd ed. Cambridge University Press.

Slovic, P. and Lichtenstein, S., 1983. Preference reversals: A broader perspective. *The American Economic Review*, 73, 596–605.

Tversky, A. and Kahneman, D., 1981. The framing of decisions and the psychology of choice. *Science*, 211 (4481), 453–458.