The Role of the Independence and Betweenness Axioms in EU and Non-EU Models

Alessio Appruzzese

September 2025

UNIVERSITÀ
DI TORINO

Contents

1 Introduction **3**

2 Theoretical Framework **3**
2.1 Expected Utility Theory (EUT) 3
2.2 Non-EU Models 5

3 Experimental Evidence **6**
3.1 Bending the Curves in the Triangle (Bernasconi, 1994) 7
3.2 Disappointment and Mixture Aversion (Camerer and Ho, 1994) . 8
3.3 Not Every Preference Is Clear-Cut(Butler and Loomes, 2011) . . 8
3.4 Common Consequences and Probability Weighting (Starmer, 1992) 8
3.5 Timing Matters (Cubitt, Starmer, and Sugden, 1998) 9

4 Discussion **9**

5 Conclusion **11**

1 Introduction

Expected Utility Theory (EUT) has been one of the main models in economics to explain how people make decisions when facing risk and uncertainty. It's a model that looks very clean and logical: you define your preferences over lotteries, you respect a few axioms, and then you behave as if you're maximizing the expected value of some utility function. We will focus on the most important axioms: Independence Axiom and Betweenness Axiom, which is basically a softened version of independence.

At the beginning, I used to take these axioms as given. They seemed obvious. But then I started looking at actual experiments (both in the literature and some small classroom ones), and I saw that things are more complicated. People don't always follow these rules. Sometimes they prefer certainty too much, sometimes they dislike ambiguity even when probabilities are balanced, and sometimes they just don't mix options in a "rational" way. And these behaviors are not just random mistakes, most of the time, they follow a clear psychological pattern.

This is what I will try to do in this essay: I look at what theory says, and then check how that compares to what people actually do when making choices. The idea is to understand why people tend violate these axioms: is it because they are irrational? Or maybe they are following some other logic that standard EUT doesn't capture?

So, the essay will follow this structure. First, I introduce the main theoretical models, starting from EUT. Then, I go through the experiments and explain what they tested and what they found. Finally, I try to make sense of all of it: what these results tell us, what kind of models we need to describe real behavior, and whether EUT can still be saved, or at least adapted.

2 Theoretical Framework

The foundations of decision theory are built around a small number of rationality principles, and two of the most important ones are the Independence and Betweenness axioms. In Expected Utility Theory (EUT), these axioms ensure that preferences are stable, transitive, and linear in probabilities. But over time, many researchers have shown that people do not always behave in a way that satisfies these assumptions, and the experiments that test these ideas have raised a lot of doubts about how descriptive these axioms really are.

2.1 Expected Utility Theory (EUT)

Expected Utility Theory (EUT), from von Neumann and Morgenstern, has been the standard model for decisions under risk. The idea behind it is very intuitive: when we face lotteries, we don't just look at the possible outcomes, but also how likely they are. So, we multiply each outcome by its probability, but not using the money directly — we use a utility function $u(x)$ instead, which captures how much we like each amount.

Formally, if a lottery L gives outcomes $x_1, x_2, ..., x_n$ with probabilities $p_1, p_2, ..., p_n$, the utility of that lottery is:

$$U(L) = \sum_{i=1}^{n} p_i \cdot u(x_i)$$

We don't assume that $u(x)$ is linear in money. If someone is risk-averse, then the utility function is concave. But the probabilities always enter linearly, and that's a key feature.

Now, this model is not just a formula. It's based on a few assumptions about how rational choices should behave. The two that we focus on here are the Independence Axiom and the Betweenness Axiom . These are what we will try to test later with experiments.

The Independence Axiom

This axiom says the following: if you prefer A over B, then mixing both with a third lottery C shouldn't change your mind. So, if you like A more than B, then you should also prefer $pA+(1-p)C$ over $pB+(1-p)C$, for any probability p between 0 and 1.

$$A \succeq B \;\Rightarrow\; pA + (1-p)C \succeq pB + (1-p)C \quad \text{for } 0 < p < 1$$

The intuition is that if C is just some background risk, it shouldn't affect the comparison between A and B. This also leads to a couple of nice properties that make the model very clean:

- If two lotteries share the same outcome in some branch, we can ignore that part and just compare what's different. This is usually called separability.
- If a lottery is made in two steps you can reduce it to a one-step lottery.

These properties allow EUT to treat complicated situations in a very simple way. But actually, this is also where problems start. In many experiments, people violate this Independence rule. The most famous case is the Allais paradox, where choices change even when the shared outcome is the same. Also, when lotteries are presented in two stages, people often don't reduce them mentally, and that breaks the assumption.

The Betweenness Axiom

This second axiom is a bit more relaxed. It says that if you are indifferent between A and B, then any mix between the two should also be in the middle. So if $A \sim B$, then the mixture $pA + (1-p)B$ should be also equally good, or at least not outside the range.

$$A \sim B \;\Rightarrow\; A \succeq pA + (1-p)B \succeq B$$

This is more intuitive maybe, because if two options are equally attractive, then a mix of them should not be a bad deal. You're just hedging between them. This idea appears in some non-EU models too. For example, Chew's Weighted Utility or Fishburn's bilinear utility both satisfy Betweenness, but not full Independence.

So Betweenness can be seen like a relaxed version of Independence. It doesn't require full linearity in preferences, but still assumes that preferences are smooth enough to make mixtures reasonable.

But again, even this axiom doesn't always survive when tested in real decisions. Many experiments show violations, especially when we look at how people behave inside the Marschak-Machina triangle (Figure 2). In theory, if Betweenness holds, indifference curves in that triangle should be straight. But what we find, both in the literature and in our classroom data, is that people's indifference curves are often bent, either fanning in or fanning out, depending on the region. This tells us that something more complex is going on. And that's what the next section of this essay will explore.

So, do people really behave according to Expected Utility? Or more precisely, do they respect axioms like Independence and Betweenness when making choices under risk? Many experiments tried to answer this, and the results are quite clear: very often, these axioms don't hold.

2.2 Non-EU Models

When experiments began to reveal systematic violations of the Independence and Betweenness axioms, many researchers tried to develop new models that could capture this behavior. These are often called Non-EU models. Most of them try to keep the idea that people evaluate outcomes and probabilities, but they allow for more flexibility in how probabilities are treated or how preferences are formed.

One of the most influential models is **Prospect Theory**, developed by Kahneman and Tversky. This model keeps the idea of a utility function, but it changes how probabilities are perceived. People tend to overweight small probabilities and underweight large ones. For example, they may prefer a 1% chance to win a big prize more than they should. Prospect Theory also includes *loss aversion* and a *reference point*: people care about gains and losses relative to that point, not absolute wealth. This helps explain why they react so strongly to certainty and to small changes around 0.

A more refined version is **Cumulative Prospect Theory**, which applies probability weighting to the cumulative distribution of outcomes instead of each outcome separately. This model can explain both the common consequence effect and violations of Betweenness, especially when the outcomes are ranked and the probabilities are transformed nonlinearly. Starmer and others found that the fanning-in and fanning-out of indifference curves in the Marschak–Machina triangle are consistent with this model.

Another important class is the **Rank-Dependent Utility** (RDU) models. These keep the idea of expected utility, but apply a transformation to the prob-

abilities based on their rank. So outcomes are ordered, and then the cumulative probability of each is transformed. This also leads to non-parallel and curved indifference curves in the triangle, allowing to capture the same violations seen in experiments. RDU models usually satisfy Betweenness, but not Independence.

Other models focus on how people feel about uncertainty itself. For example, **Ambiguity Aversion** models try to capture the discomfort that people feel when probabilities are unknown or not well defined. **MaxMin Expected Utility** (Gilboa and Schmeidler) assumes people evaluate each act by the worst expected utility over a set of priors. The **Choquet Expected Utility** model uses non-additive probabilities, allowing for ambiguity attitudes. Another one, the **Anticipated Utility** model (Segal), assumes people anticipate the feeling of disappointment or satisfaction, and take that into account when making choices.

In the same direction, Gul's **Disappointment Aversion** model (1991) keeps Expected Utility but introduces disappointment directly into the utility function: people dislike outcomes that fall below a reference point, especially if they were expecting more. This model can fit many of the observed violations of Betweenness and even some of the Allais-type choices, with just one extra parameter.

Finally, we have a very different idea: instead of assuming that people always have clear and sharp preferences, Butler and Loomes (2011) suggest that **preferences are sometimes imprecise**. When two lotteries are very similar, a subject might not really know which one they prefer. So, if they choose one over the other today, they might reverse it tomorrow — not because they are irrational, but because the difference is not meaningful to them. This model of *imprecise preferences* predicts that indifference curves may fan out or bend in strange ways, exactly like what we observe in the triangle. It also helps explain why people might violate Betweenness in a systematic way, especially when the lotteries are close together.

To sum up, Non-EU models offer a wide range of tools to explain why people violate axioms like Independence and Betweenness. Some of them add psychological realism (like Prospect Theory), others change how probabilities are treated (like RDU and Choquet), and some even question the idea that preferences are always well-defined (like Butler and Loomes). What they all have in common is that they fit the experimental data better than classical EUT, and they open the door to a more flexible and realistic way to model human behavior under risk and uncertainty.

3 Experimental Evidence

Theoretical models can be very clean on paper, but to be useful, they need to describe how people actually make decisions. That's why experimental economics is so important. In this section, I go through a series of experiments that tested whether the Independence and Betweenness axioms really hold when people make choices. I focus on some well-known studies from the literature and com-

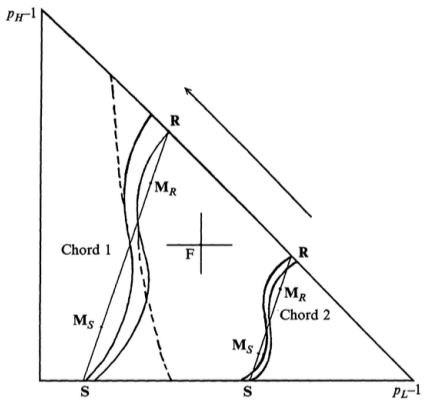

Figure 1: Indifference curves in the Marschak–Machina triangle. Linear (EUT) vs. curved (actual data). Source: Bernasconi (1994)

pare them with our own classroom experiments. Even though our sample was small and not incentivized with real money, we still found patterns that match what the literature shows. This tells us that the deviations from the standard model are not just statistical noise or artifacts. They reflect something deeper in the way we think about risk and uncertainty.

3.1 Bending the Curves in the Triangle (Bernasconi, 1994)

Bernasconi ran an experiment based on the Marschak-Machina triangle, which is a way of representing three-outcome lotteries. His goal was to test the Betweenness axiom, which predicts that if a person is indifferent between two lotteries, then they should also be indifferent to any mix of those lotteries. If that's true, indifference curves should be straight lines.

Instead, Bernasconi found something else. The curves were not straight. They were convex in some areas and concave in others. He called this the squiggle pattern. It means that people don't treat mixtures as neutral. Sometimes they even see the mix as worse than both original lotteries, which breaks the basic logic of Betweenness.

He also tried presenting the lotteries in two ways: reduced form and compound form. The reduced form is a simple, one-stage lottery, while the compound version involves two stages. He noticed that people violated Betweenness more often in the reduced form. This suggests that the structure or presentation of the gamble matters. It also casts doubt on the Reduction of Compound Lotteries axiom. If people behave differently just because a lottery is described in a different way, then EUT's predictions are clearly at risk.

3.2 Disappointment and Mixture Aversion (Camerer and Ho, 1994)

Camerer and Ho also worked on testing Betweenness, but from another angle. They gave participants choices between lotteries and their mixtures, again using both reduced and compound forms. Just like in Bernasconi's experiment, they saw more violations when the lotteries were presented in reduced form.

What's interesting is that the violations weren't random. They followed a pattern. In particular, when one of the original lotteries had a sure gain, people were reluctant to give it up in the mixture. Even if the mix had the same expected value, it didn't feel as good.

To explain this, they used models like Disappointment Aversion and non-linear probability weighting. In those models, people compare outcomes to expectations. A mixture might disappoint because it removes the certainty that was in one of the original lotteries. This helps explain why people often prefer A and B over a mix of A and B. The mix feels less safe or emotionally worse, even if it makes sense mathematically.

They also ruled out simple mistakes or inconsistencies. The violations they found were systematic and stable. That makes the case stronger for rejecting Betweenness as a general rule.

3.3 Not Every Preference Is Clear-Cut(Butler and Loomes, 2011)

Butler and Loomes took a different approach. Instead of assuming that people have clear and consistent preferences, they asked whether some violations might come from fuzziness in those preferences. Their idea is that sometimes people just don't know which option they prefer, especially when the options are very similar.

In their experiment, participants had to choose between lotteries that were almost the same. The differences were small, sometimes just a few percentage points in probability. What they found is that people sometimes chose one lottery, and other times chose the other, even when nothing had really changed.

They argued that this isn't irrational. It's a sign that preferences can be imprecise. People don't always have a strict ranking in their heads. Sometimes they're undecided, and that can lead to what looks like inconsistency. But actually, it's just indifference. Their model predicts that indifference curves won't be straight, and may fan out in areas where people are unsure. This helps explain some of the patterns seen in earlier experiments.

3.4 Common Consequences and Probability Weighting (Starmer, 1992)

Starmer used the common consequence test to evaluate both Independence and Betweenness. This test involves adding the same outcome to both options in

a choice. If people are consistent, their preferences should stay the same. But that's not what happened.

Many participants changed their preferences when the same outcome was added. That's a violation of both axioms. But again, it wasn't just noise. The way preferences shifted followed a clear pattern. The shape of the indifference curves matched what we would expect if people were overweighting small probabilities and underweighting large ones.

This behavior doesn't fit Expected Utility Theory. But it fits models like Rank-Dependent Utility or Prospect Theory, where people transform probabilities before evaluating outcomes. Starmer's work shows that these violations aren't random. They're predictable and connected to the way we process risk.

3.5 Timing Matters (Cubitt, Starmer, and Sugden, 1998)

Cubitt, Starmer, and Sugden went a step further and looked at how people behave in dynamic situations. They designed a set of five problems that, in theory, should all lead to the same choice if four principles hold: separability, timing independence, frame independence, and reduction.

In practice, the principle that failed was timing independence. People chose differently depending on when they had to make the decision. If they had to decide before a random event, they chose one thing. If they could wait until after, they chose something else.

This is a big deal because timing independence is one of the pillars of dynamic Expected Utility Theory. If it fails, the whole dynamic structure of the model becomes shaky. Even if people still try to maximize utility, they clearly don't do it in the way EUT expects when timing and information are involved.

Across all these experiments, the same conclusion comes up again and again. The Independence and Betweenness axioms don't describe how people actually choose. The violations are not random. They follow patterns that make sense when we assume people care about certainty, fear disappointment, dislike ambiguity, and sometimes don't have fixed preferences.

These are all things that Expected Utility Theory can't explain, but other models can. Prospect Theory, Rank-Dependent Utility, models with ambiguity aversion, and the imprecise preferences approach all do a better job. If we want to understand decision-making under uncertainty, we need to go beyond the clean logic of expected utility and include the messier psychology behind real choices.

4 Discussion

After reviewing both the theoretical models and the experimental evidence, it becomes clear that the structure of Expected Utility Theory (EUT), although elegant and theoretically useful, is often too rigid to describe how people actually make choices. We saw in the previous section that violations of the Independence and Betweenness axioms are frequent and consistent. What matters now is

understanding what these violations tell us. Are people behaving irrationally, or are they simply using different internal models that EUT does not capture? And how much of this depends on the way the decision is framed? These are the questions I will try to answer here.

Let's start from what we saw in Bernasconi's experiment. The squiggle pattern in the Marschak-Machina triangle shows that people do not treat mixtures as neutral, and their preferences cannot be described by linear indifference curves. This is not just an isolated finding, it's part of a larger pattern. Camerer and Ho confirm it by showing that Betweenness is violated more often when lotteries are presented in reduced form. Our own Report II supports this with a clear ranking of D over C, C over H, H over A, even though all four lotteries have the same expected value. These examples all suggest that mixtures are not treated as "between" their components. This is a direct challenge to the Betweenness axiom.

Now, if these violations were random, we could call them noise. But they are not. They follow clear structures. The way probabilities are distributed, the shape of the lottery tree, the presence or absence of ambiguity, and even whether the subject gets to choose the bag or not, all of these details influence choices. In Report I, participants often violated Independence when a certain outcome was present, then switched to the risky option when that certainty was removed. This matches exactly what Prospect Theory predicts: the Certainty Effect, the overweighting of small probabilities, and the general sensitivity to framing.

So, are people irrational? I don't think that's the right way to put it. What we saw in Butler and Loomes' model gives a different perspective. If preferences are imprecise, meaning that people do not always have a clear ranking between two similar options, then some of these "violations" might just be cases where the subject is truly undecided. Their model predicts fanning-out patterns, which are exactly what we saw in the data. This suggests that what looks like inconsistency may just be the result of fuzziness, not error. In this sense, people are not irrational, they just have limits in how sharply they can distinguish between close lotteries.

The role of framing is also much stronger than it might seem at first. This was obvious in our classroom experiments, especially in Report II. When lotteries with the same expected value were presented differently, as compound or reduced, as self-selected or randomly chosen, people changed their evaluations. Starmer shows something similar with common consequence designs. Just by adding a background outcome to both options, preferences shift. This breaks Independence, but it also tells us that people react not only to the formal properties of the lottery, but also to how the situation is presented. This adds another layer of complexity that EUT cannot account for, but other models can.

Prospect Theory handles many of these observations quite well. It allows for probability weighting and loss aversion. The shape of the utility function and the weighting of probabilities explain why people behave differently in gain and loss domains, and why small probabilities are sometimes overweighted. Rank-Dependent Utility gives another approach, especially useful for explaining the

fanning-in and fanning-out patterns in the triangle, as seen in Starmer's work and in the triangle-based evidence from Bernasconi.

So, what can we learn from all this? First of all, it becomes obvious that classical EUT is too restrictive as a descriptive model. The assumptions it makes, that people have sharp preferences, reduce compound lotteries perfectly, treat mixtures neutrally, and evaluate only expected value, do not hold in real data. But more importantly, we also learn that these deviations are not random. They are predictable, structured, and often psychologically motivated.

That's why we need behavioural models. Not because people are irrational, but because they process information differently from what EUT assumes. They care about the way options are presented, they feel discomfort with ambiguity, they overweight rare events, and they do not always reduce complex lotteries into simpler ones.

In short, we have to recognise that economic behaviour under uncertainty is shaped by more than just outcomes and probabilities. It is shaped by how people see the problem, how confident they feel about the information, and how close or different the options are. That's what the evidence tells us. And if our models do not capture that, then they are not really models of behaviour, just mathematical ideals.

The next and final section will summarise these points and suggest where we can go from here.

5 Conclusion

The aim of this essay was to understand whether the Independence and Betweenness axioms really describe how people behave when making decisions under risk and uncertainty. These axioms are central to Expected Utility Theory, which has been the dominant model in economics for decades. But what we have seen, both in theory and in the experimental evidence, is that these assumptions are often not respected in practice.

From the theoretical side, many alternative models have been developed to explain why these violations happen. Prospect Theory introduced psychological elements like loss aversion and probability weighting, which help explain patterns like the certainty effect and risk-seeking in low-probability gains. Rank-Dependent Utility went in a more technical direction, modifying how probabilities are treated while keeping the utility structure. Butler and Loomes offered a more behavioural explanation, saying that preferences can be imprecise and fuzzy rather than sharply defined. These models all have different starting points, but they share one idea in common: the behaviour we see in experiments cannot be explained by standard EU theory alone.

On the experimental side, the violations of Independence and Betweenness are not only frequent, but also systematic. We saw this in the studies by Bernasconi, Starmer, Camerer and Ho, as well as in our own Reports I and II. Participants consistently reacted to certainty, to ambiguity, and to the framing of information. In Report I, people changed their choices when certainty was

removed, even though expected values stayed the same. In Report II, participants gave very different evaluations to lotteries with identical expected value, depending on whether the uncertainty was ambiguous or purely risky. These are not just small deviations. They are robust patterns that show up in different settings, with different designs, and even in simple classroom experiments.

So, to answer the main question: do Independence and Betweenness describe how people behave? The answer is clearly no, at least not in general. These axioms are useful for building models, and they may still work in some situations, but they are not good descriptions of actual decision-making. People are influenced by certainty, by ambiguity, by framing, and by how close the options are. Their preferences are not always linear, or stable, or perfectly transitive. And that's not a sign of irrationality. It just means that human decision-making is more complex than classical theory assumes.

What we can learn from all this is that economic theory needs to be more flexible. Behavioural models are not a rejection of rationality, they are an attempt to model it in a way that is closer to reality. The experiments give us a clear signal: if we want to understand how people make choices under uncertainty, we need to move beyond the clean structure of EUT and include the elements that actually shape behaviour. That is what theory should do — not just simplify, but explain.

References

Bernasconi, M. (1994). Nonlinear preferences and two-stage lotteries: Theories and evidence. *The Economic Journal*, *104*(422), 54–70.

Butler, D., & Loomes, G. (2011). Imprecision as an account of the preference reversal phenomenon. *The American Economic Review*, *101*(2), 724–743.

Camerer, C. F., & Ho, T.-H. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty*, *8*(2), 167–196.

Cubitt, R. P., Starmer, C., & Sugden, R. (1998). Dynamic choice and the common ratio effect: An experimental investigation. *The Economic Journal*, *108*(450), 1362–1380.

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, *18*(2), 141–153.

Segal, U. (1990). Two-stage lotteries without the reduction axiom. *Econometrica*, *58*(2), 349–377.

Starmer, C. (1992). Testing new theories of choice under uncertainty. *The Review of Economic Studies*, *59*(4), 719–746.