

# Is Expected Utility Theory Unfalsifiable? A Response to Hodgson

Risto Uuk

The London School of Economics and Political Science

r.uuk@lse.ac.uk

10 March 2020

## 1 Introduction

It has been stated by economists that expected utility theory is either falsifiable and already falsified or even unfalsifiable. Matthew Rabin and Richard Thaler have said that “there have been repeated demonstrations of the shortcomings of the expected utility model.”<sup>1</sup> Geoffrey Hodgson, on the other hand, has said that “the theory is invulnerable to any empirical attack” and “no amount of evidence can establish non-existence.”<sup>2</sup> I will focus on the latter claim in my essay.

This topic is important, because expected utility theory as a descriptive or predictive theory is commonly used in the field of economics. The aim of it is to explain how people make choices or to predict people’s choices.<sup>3</sup> If it is the case that expected utility theory has been falsified, then it seems that it has a high failure rate and should be replaced by some other theory. If it, however, is unfalsifiable, then it must be play a smaller or no role in economics as it would lack descriptive content.<sup>4</sup>

---

<sup>1</sup>Matthew Rabin and Richard H. Thaler, “Anomalies: Risk Aversion,” *Journal of Economic Perspectives* 15, no. 1 (March 2001): 229.

<sup>2</sup>Geoffrey M. Hodgson, “On the Limits of Rational Choice Theory,” *Economic Thought* 1, no. 1 (2012): 98.

<sup>3</sup>R. A. Briggs, “Normative Theories of Rational Choice: Expected Utility,” in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Fall 2019 (Metaphysics Research Lab, Stanford University, 2019), <https://plato.stanford.edu/archives/fall2019/entries/rationality-normative-utility/>.

<sup>4</sup>Katie Steele, “Choice Models,” in *Philosophy of Social Science: A New Introduction*, ed. Nancy Cartwright and Eleanora Montuschi (Oxford, United Kingdom: Oxford University Press, 2015), 190.

## 2 Hodgson's argument

In this section, I interpret Hodgson's arguments that expected utility theory is unfalsifiable. The first argument Hodgson gives could be standardized in the following way. I will call it *the unobservable utility argument*:<sup>5</sup>

- (1) If utility is unobservable, then expected utility theory is unfalsifiable.
- (2) Utility is unobservable.
- (3)  $\therefore$  Expected utility theory<sup>6</sup> is unfalsifiable.

Hodgson justifies premise (1) by claiming that all kinds of behavior can be explained if utility is unobservable.<sup>7</sup> He says that "no evidence can possibly refute the theory that agents are maximizing some hidden or unknown variable (such as utility)."<sup>8</sup> He also suggests that if behavior is explained by preferences, which are defined by behavior, then there is a kind of circularity that cannot be refuted with counterexamples. It appears from this that agents are maximizing their expected utility by definition.

Hodgson does not justify premise (2) explicitly, but a standard mainstream economic explanation seems plausible. It is commonly thought in economics that pleasure and pain are hard to measure, while preferences can be stated and objectively measured from behaviour.<sup>9</sup> Mental states such as happiness are even thought to be unscientific, because they are considered unobservable according to a dominating view among mainstream economists.<sup>10</sup> Preferences can be directly observed from the choices people make. However, the actual utility of those choices is unknown.

The second argument Hodgson gives could be standardized in the following way. I will call it *the preferences not being compared under identical conditions argument*:<sup>11</sup>

- (1) If preference intransitivity can always be explained away by appealing to preferences not being compared under identical conditions, then expected utility theory is unfalsifiable.
- (2) Preference intransitivity can always be explained away by appealing to preferences not being compared under identical conditions.

---

<sup>5</sup>Hodgson, 97.

<sup>6</sup>In this argument, I interpret expected utility theory to be defined as a descriptive decision theory for rational agents according to which they make choices that give them the most utility in expectation. Utility maximization in this case means maximizing the quantity of utility.

<sup>7</sup>Hodgson, "On the Limits of Rational Choice Theory," 97.

<sup>8</sup>Hodgson.

<sup>9</sup>Erik Angner, "Is It Possible to Measure Happiness?: The Argument from Measurability," *European Journal for Philosophy of Science* 3, no. 2 (May 2013): 222; Daniel Kahneman and Alan B. Krueger, "Developments in the Measurement of Subjective Well-Being," *Journal of Economic Perspectives* 20, no. 1 (March 2006): 18.

<sup>10</sup>Angner, "Is It Possible to Measure Happiness?"

<sup>11</sup>Hodgson, "On the Limits of Rational Choice Theory," 98.

(3)  $\therefore$  Expected utility theory<sup>12</sup> is unfalsifiable.

Hodgson does not explicitly mention premise (1) nor justify it. However, according to the axioms of expected utility theory, a rational agent is someone whose preferences are consistent according to the intransitivity axiom.<sup>13</sup> If humans are utility maximizers in this sense, then showing that someone violates the intransitivity axiom would suffice to show that expected utility theory is false. However, if it is not possible to show that someone violates it, then it cannot be shown that utility theory is false.

Hodgson justifies premise (2) by claiming that “[We can say that preference] comparisons did not take place under identical conditions, or were separated in time or space. Extraneous factors may account for the apparent intransitivity. All we have to do is indicate in some way that the [two preferences being compared] are not quite identical. The [preferences] could be slightly different in timing, substance, or their informational or other contexts.”<sup>14</sup> In other words, every preference is slightly different and inconsistency is impossible.

### 3 Objection 1

In this section, I respond to Hodgson’s argument that expected utility theory is unfalsifiable, because utility is unobservable.

I accept premise (1). Hodgson provides very persuasive examples to show how seeming falsifications do not work when utility is unobservable. He says that, “If experiments show that consumers appear to prefer a monetary reward that is less than the expected outcome, . . . then we can always get round these problems, and make the evidence consistent with utility maximization, by introducing other explanatory variables.”<sup>15</sup>

One of the examples is the following. Suppose that someone is faced with a choice to take \$10 with certainty and \$1,000 with a probability of 2 percent. The latter has an expected monetary payoff of \$20 dollars and appears to be a better choice from an expected utility maximization perspective. Despite that, people sometimes choose the former with lower expected monetary payoff. This seems to be a choice that does not maximize their expected utility. However, since utility is unobservable, it can be claimed that agents get more utility by adding other explanations that are not related to monetary payoffs. If this can always be done, then expected utility theory is unfalsifiable.

I do not accept premise (2), because I do not think that utility is unobservable.

---

<sup>12</sup>In this argument, I interpret expected utility theory to be defined as a descriptive decision theory for rational agents according to which they are consistent in their preferences and do not have intransitive preferences. Utility maximization in this case means to have consistency in preferences.

<sup>13</sup>According to the four axioms of the von Neumann–Morgenstern’s expected utility theory one of which is the axiom of transitivity.

<sup>14</sup>Hodgson, “On the Limits of Rational Choice Theory,” 98.

<sup>15</sup>Hodgson, 97.

Hodgson’s argument works only if it is accepted that third parties are unable to confirm how much utility someone receives from a given choice. I think it can be argued what utility is as well as that utility can be observed. Economists often think that happiness, for example, cannot be observed, but many philosophers think that developing at least approximate measures of happiness is very feasible.<sup>16</sup> There are three main candidates for what utility is, or in other words, what well-being consists of. These are objective values, preferences, and mental states. The objective list account claims that what is best for someone is what fulfills the basic needs and rights of humans.<sup>17</sup> The preference satisfaction account claims that what is best for someone is what satisfies their preferences the most.<sup>18</sup> The mental state account claims that what is best for someone is what provides the most pleasure and the least pain.<sup>19</sup>

All of these accounts of utility are observable. According to the objective list account, a basic human need might, for example, be to consume at least 2,000 calories of nutrient-dense food per day. It is straightforward to observe whether people maximize their utility by consuming enough nutrient-dense food per day. According to the preference satisfaction account, someone might, for example, prefer to graduate from university with specific grades and it could be observed very easily whether they achieve that or not. According to the mental state account, someone might, for example, get a lot of pleasure from watching certain kinds of TV shows and we could ask them how much pleasure they got from watching TV shows last week.

A critic might object to this by claiming that in order to have a proper expected utility theory that guides action in economics, it is necessary to be able to observe cardinal utility and no such thing is possible. Hence, utility is still unobservable. My response to this is that with the objective list account this is a completely irrelevant objection, but it needs to be dealt with by the preference satisfaction and mental state accounts. In the literature of happiness, there are two kinds of methods developed for observing people’s happiness: the experience sampling method and the daily reconstruction method.<sup>20</sup> Both of these methods rely on asking people directly how happy specific activities and circumstances made them.

Furthermore, it is possible to ask people directly how much more pleasure they received from eating a food product A compared to eating a food product B, for example.

---

<sup>16</sup>Dan Haybron, “Happiness,” in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, Winter 2019 (Metaphysics Research Lab, Stanford University, 2019), <https://plato.stanford.edu/archives/win2019/entries/happiness/>.

<sup>17</sup>Paul Dolan and Robert Metcalfe, “Measuring Subjective Wellbeing: Recommendations on Measures for Use by National Governments,” *Journal of Social Policy* 41, no. 2 (April 2012): 411. I referred to these accounts of well-being also in my formative essay for Philosophy and Public Policy third formative essay.

<sup>18</sup>Dolan and Metcalfe.

<sup>19</sup>Dolan and Metcalfe.

<sup>20</sup>Daniel Kahneman and Robert Sugden, “Experienced Utility as a Standard of Policy Evaluation,” *Environmental & Resource Economics* 32, no. 1 (2005): 175.

This could provide an approximation of cardinal utility. In addition, to learn as much about what satisfies people's preferences to be able to observe cardinal utility, a lot of different methods could be implemented, including choices and revealed preferences, willingness to pay, national and regional surveys, randomized evaluations, and many more.

## 4 Objection 2

In this section, I respond to Hodgson's argument that expected utility theory is unfalsifiable, because preference intransitivity can always be explained away by appealing to preferences not being compared under identical conditions.

I accept premise (1). Hodgson provides very persuasive examples to show how seeming falsifications do not work when it is always responded that preferences were not compared under identical conditions. One of them is the following. Suppose that someone visits the same restaurant several times and prefers steak to fish on one day but fish to steak on another. He says that "Is this behaviour inconsistent? Maybe. Maybe not. . . . The two choice occasions were different in terms of circumstances and knowledge. Hence they do not necessarily imply inconsistency."<sup>21</sup> This type of ad-hoc explanations could always be added to explain why someone indeed was consistent and did not have intransitive preferences. I agree that if such strategies can always be used, then expected utility theory is unfalsifiable.

I do not accept premise (2). Intransitivity cannot be explained away if the conditions of the new preferences are plausibly insignificant and should not really influence agent's preferences from the perspective of utility maximization. Suppose a man is going to buy a new car. He prefers car A that costs £20,000 to a car B that costs £30,000 and is planning to buy car A. Right before making the purchase, he speaks with a very beautiful sales woman and decides to buy car B instead. The woman did not give any arguments for why he should buy car B. She was just very polite and nice to the customer. It seems completely implausible that this intransitivity can be explained away by claiming that the preference set turned from  $x > y$  to  $y_2 > x > y_1$ . The fact that the man had a conversation with the salesperson is completely insignificant to a rational agent's preferences. Preference  $y$  is still preference  $y$  as it is unreasonable to claim that  $y_2$  is significantly different from  $y_1$ .

A critic might object to this by claiming that if the agent revealed through their choices that they now prefer car B, then this new circumstance must be relevant enough for them to consider its utility to be higher, because humans are expected utility maximizers. My response to this objection is that it argues that humans are expected utility maximizers by definition. This definition is that whatever humans choose they maximize

---

<sup>21</sup>Hodgson, "On the Limits of Rational Choice Theory," 97.

their expected utility. We have a pragmatic reason to reject this definition, however, because it reduces predictive power. If humans always maximize their expected utility by definition, then there is no way to tell what causal factors influence their decision-making. Determining how to run an economy to allocate scarce resources would become impossible, because whatever circumstances there are, people would make choices and therefore, always maximize their expected utility whether they were living in extreme poverty or be very wealthy.

## 5 Conclusion

Expected utility theory is not unfalsifiable. Hodgson claimed that utility is unobservable and that preference intransitivity can always be explained away by appealing to preferences not being compared under identical conditions. I reject both arguments. Utility is observable as an objective value, preference, or a mental state. Although no preferences are compared under identical conditions, sometimes the new condition should be treated as insignificant and should not influence a rational agent's preferences.

## Bibliography

Angner, Erik. "Is It Possible to Measure Happiness?: The Argument from Measurability." *European Journal for Philosophy of Science* 3, no. 2 (May 2013): 221–40.

Briggs, R. A. "Normative Theories of Rational Choice: Expected Utility." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2019. Metaphysics Research Lab, Stanford University, 2019. <https://plato.stanford.edu/archives/fall2019/entries/rationality-normative-utility/>.

Dolan, Paul, and Robert Metcalfe. "Measuring Subjective Wellbeing: Recommendations on Measures for Use by National Governments." *Journal of Social Policy* 41, no. 2 (April 2012): 409–27. Haybron, Dan. "Happiness." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2019. Metaphysics Research Lab, Stanford University, 2019. <https://plato.stanford.edu/archives/win2019/entries/happiness/>.

Hodgson, Geoffrey M. "On the Limits of Rational Choice Theory." *Economic Thought* 1, no. 1 (2012): 94–108.

Kahneman, Daniel, and Alan B. Krueger. "Developments in the Measurement of Subjective Well-Being." *Journal of Economic Perspectives* 20, no. 1 (March 2006): 3–24.

Kahneman, Daniel, and Robert Sugden. "Experienced Utility as a Standard of Policy Evaluation." *Environmental & Resource Economics* 32, no. 1 (2005): 161–81.

Rabin, Matthew, and Richard H. Thaler. "Anomalies: Risk Aversion." *Journal of Economic Perspectives* 15, no. 1 (March 2001): 219–32.

Steele, Katie. "Choice Models." In *Philosophy of Social Science: A New Introduction*, edited by Nancy Cartwright and Eleanora Montuschi. Oxford, United Kingdom: Oxford University Press, 2015.