BIOGRAPHICAL INFORMATION

Daniel Halliday is a Lecturer in Political Philosophy at Melbourne University. He works mainly on political philosophy, particularly on questions relating to markets and taxation. He is currently completing a book on the problem of inherited wealth.

Risky Curves: On the empirical failure of expected utility theory, Daniel Friedman, R. Mark Isaac, Duncan James and Shyam Sunder. Routledge, 2014, xiii + 137 pages.

Expected utility theory (EUT) is the dominant theory of decision making under uncertainty in economics, despite decades of research that fails to confirm its predictions. In their fascinating new book, *Risky Curves: On the empirical failure of expected utility*, Daniel Friedman, R. Mark Isaac, Duncan James and Shyam Sunder compile and examine systematically the research on EUT, and outline the failure of the theory with respect to both individual decision making and aggregate behaviour. Importantly, they also dig deeper into the evidence to draw conclusions about why the theory fails, and to suggest fruitful directions for future research.

Chapter 1 provides a brief introduction. Here the authors contrast the layman's definition of risk with economists' version. The dictionary definition focuses on the possibility and magnitude of harm, injury or loss, while economists think of risk as the variability of payoffs associated with a particular decision. If you ask someone what risk means to them, the dictionary version is a good approximation of what they are likely to tell you; variability, especially at the high end of the distribution of payoffs, does not immediately come to mind as 'risky'. The distance between the intuitive, vernacular definition of risk and economists' measure is a theme that recurs throughout the book.

This chapter also introduces the EUT model, which is the source of the risk-as-variance definition, and notes its pervasiveness in the economics of decision making under uncertainty. Scientists judge a model by its predictive ability, and the authors promise to lay out the evidence that EUT has not been a big success empirically. The belief in EUT is very strong among economists, and the authors contend that this belief can act as a kind of brainwashing, blinding researchers to its flaws and to the possibility of developing better alternatives. The authors note three main worries about the consequences of maintaining an incorrect model of decision making: It can mislead a young researcher into asking the wrong questions, and therefore can lead to a failed initial research programme

and damaged career; it misleads applied researchers into attributing deviations from the predictions of their models to risk aversion, when other explanations may be more valid, slowing scientific progress; and it impedes the search for better models of decision making.

Chapter 2 traces the history of research on decision making under uncertainty. The essence of the EUT model was developed in Bernoulli's 1738 treatise, which was motivated by the well-known St. Petersburg Paradox.[1] This work introduced the idea that the marginal utility of an additional increment of income diminishes as income increases (diminishing marginal utility), and approximated the phenomenon with a mathematical function mapping income into subjective value. The idea that people have a well-behaved function that maps possible outcomes into a one-dimensional measure is pervasive in economics; the function is often referred to as a Bernoulli function. Diminishing marginal utility implies that an individual will always prefer the expected value of a gamble to the gamble itself – that is the average of the utilities of the payoffs will be below the utility of the expected value (the average of the payoffs). For a gamble with a given expected value, greater variance decreases expected utility. Therefore most individuals, having diminishing marginal utility, will exhibit risk (as variance) aversion, preferring a less-risky gamble. Risk aversion can of course differ across individuals, and some individuals may be risk-neutral or even risk-seeking, though this is thought to be relatively rare.

From the beginning, there was evidence that failed to support the EUT model, and the evidence continued to accumulate over time. Subsequent theorists laboured to expand the model to incorporate observed behaviour, for example by introducing additional inflection points to accommodate simultaneous purchase of insurance and lottery tickets. This second chapter also includes a discussion of a fascinating early study attempting to elicit utility functions using decisions over lotteries. In particular, Grayson (1960) found considerable heterogeneity in the elicited utility functions of his eleven participants, only some of which were consistent with EUT. The authors return to a detailed discussion of this study in Chapter 6.

---

[1] It is curious that the model that comes under attack in this book using empirical evidence is itself empirically motivated. The St. Petersburg Paradox involved a gamble that pays double or nothing for a series of coin tosses. The stakes begin at 2 ducats (though any currency amount will do), and double with every realization of 'heads'. If 'tails' occurs, then the gamble pays off zero. Thus the expected value of the gamble is the sum of a series of terms, consisting of 2^n ducats with probability 2^{-n}: $2*1/2 + 4*1/4 + 8*1/8 \ldots$ etc., or $1+1+1+ \ldots$. Although it has an infinite expected value, in practice no one seems to want to pay more than a small amount for the gamble. Hence the paradox: the valuation of the gamble is much less than its expected value.

In Chapter 3 the authors turn to the elicitation of individual risk preferences, a topic that is close to my own interests. Because risk tolerance is an important parameter in many policy-relevant applied models, measuring risk preferences is critical for calibrating policy. Economists' approach to measuring preferences is embedded in EUT, and essentially amounts to measuring the curvature (concavity) of the utility function. Typically, economists use incentivized tasks – choices between and among lotteries, or valuations of lotteries – to gauge this curvature. Each task invented for this purpose can be thought of as a yardstick, and any yardstick one might use should give the same answer, more or less.

In this chapter the authors review several prominent methods for eliciting risk aversion, and point out the ways in which the results fail to confirm EUT. These include choosing among different gambles or valuing gambles (willingness to pay or accept) in simple or more complex settings. Several disappointing regularities emerge from these studies. First, different measures of risk aversion yield different patterns of risk aversion, with some generally showing that most subjects are risk averse, but others showing a preponderance of risk-seeking behaviour; a related point is that elicited risk aversion parameters show little consistency across measures or over time. Second, there is very little evidence that estimates of risk aversion from any of the measures correlates strongly with behaviour in other lab decisions such as bidding in an auction, or in the major risky decisions that people make in their lives. Third, risk aversion measures are sensitive to aspects of context that, from a theoretical perspective, shouldn't matter; stakes levels, payment procedures, and the composition of sets of decisions all affect elicited risk aversion.

My interest in the measurement of risk aversion dates from a failed experiment from 2004 (unpublished), in which my co-author and I attempted to elicit risk preferences using four different measures, including two incentivized tasks and two survey measures. We had a group of 115 subjects complete all the measures, and then six weeks later they returned to the lab for a retest. Each of the four measures exhibited strong test-retest reliability. But the yardsticks gave different measures. Furthermore, the correlation between the two tasks, and between the tasks and the survey measures, was very low and not always positive (-0.11 to 0.24). That is, if we lined up the 115 subjects according to their measured risk aversion (most risk averse to least risk averse) for one task, and then lined them up again according to another of the tasks, nearly everyone would have to move to a different place in line. Were the yardsticks at fault, or was something else wrong with the study? I discussed these results with a well-known experimental methodologist in another field, and asked him what he made of it. He replied, "I'd say your underlying construct needs a little work." We subsequently searched the literature for similar findings and located quite a few studies showing inconsistency

across measures of risk aversion, including an early paper by Paul Slovic ([1962](#)) where he noted insignificant levels of correlation among a number of risk-aversion measures, incentivized and non-incentivized. The papers we found are published in economics, psychology, and in specialized decision-making journals, and all note low levels of correlations across measures of risk aversion. Yet despite this repeated finding, the result is generally not known by economists, even among those who study decision making under uncertainty. Our underlying construct still is widely used, despite these failures to find empirical support, as academic publications continue with the pervasive assumption of curved Bernoulli utility functions. Perhaps there is a 'best' way to measure risk preferences, and this magic measure will usefully predict behaviour outside the lab, but if so we have not found it yet.

Chapter 4 tackles the question of whether EUT performs well in empirical studies focusing on aggregate behaviour. The authors report that studies using hypothetical or incentivized measures of risk aversion rarely correlate well with self-reported health behaviour or health outcomes; indeed health-related decision making seems to focus more on the probability of a bad outcome, the lay definition of risk, rather than variability of outcomes.[2] Gambling behaviour fails to conform to EUT predictions, and it is clear that something other than the maximization of a standard utility function is going on when gamblers make decisions. When engineers do risk analysis on their projects, they focus on the probability of something going badly wrong, and do not consider risk as variance when making design decisions. Consumers' preferences over insurance contracts seem to defy the assumptions of EUT. Similarly weak or negative findings occur in studies of real estate and financial investments, as well as the aggregate behaviour of financial markets. The equity premium puzzle implies ridiculously high levels of risk aversion.[3] The authors conclude that 'curved Bernoulli functions do not help us gain a better understanding of these diverse and important aspects of the world we live in' (73). Indeed they recommend stepping back from the standard approach and thinking more carefully about risk as the likelihood of harm, injury or loss.

---

[2] Of course it is impossible to cite everything (and the authors of this book do not really attempt a comprehensive literature review), but in addition to the many studies that show no effect, there are some studies that show a positive correlation between incentivized risk aversion measures and health or financial behaviour. Please feel free to contact me for details if you are interested. However, the evidence is not strong or consistent, and the jury is still out on whether incentivized measures are superior to answers to simple survey questions asking subjects whether they see themselves as risk takers.

[3] The equity premium puzzle refers to the fact that the returns on stocks are much higher than lower-risk assets such as government bonds. Stocks are more risky, but the difference in returns is larger than can be justified by estimates of risk aversion from other sources.

In Chapter 5, the authors ask, 'What are risk preferences?' Might the idea of risk preferences as the curvature of a Bernoulli function be 'a figment of the theorists' imagination', the economists' version of phlogiston? In discussing proposed alternatives to EUT, the authors are critical of most such models for two reasons. First, these models (such as prospect theory, among many others) are seemingly able to provide a better fit to existing data mostly by adding more parameters and therefore more flexibility to the estimation; and second, the outputs of these models are no more able to predict the risky behaviour of individuals than the simpler EUT approach. And while, as Neilson (2011: 976), writes, 'For every theory, there exists an experimentalist clever enough to generate evidence violating it' (p. 976), as studies accumulate it should be possible to distinguish a model that is essentially (though perhaps not in every clever instance) correct from others that are less correct.[4] EUT has not fared well in this respect.

Arguing that preferences may not be 'intrinsic', the authors then consider new measures of risk, and therefore of risk aversion, based on the ways in which people perceive risk in practice. These measures revert to the dictionary definition of risk and focus on the likelihood and magnitude of potential losses instead of variance. Their favourite is 'expected loss', which they show to be distinct from variance, and they present some suggestive evidence that this measure of risk may have empirical value. However, it does not give us a measure of riskiness for uncertain prospects that are entirely in the gain domain.

Chapter 6 takes up the topic of context, and how specific aspects of context alter an individual's elicited risk preference. These include aspects of the elicitation procedures themselves, as well as elements of a person's current situation. This is analogous to a shift in focus from 'preferences' to 'constraints' in the standard approach to consumer choice. For example, when eliciting the risk preferences of low-income individuals, whether in developing countries or within the USA, many researchers (myself included) find that these individuals are extremely risk averse and become more so with higher stakes. This could be due to the fact that these folks may face urgent constraints – that is, they may badly need a threshold amount of money to cover a pressingly urgent expenditure – and the fixed payoff option in the choice tasks is enough to cover that expenditure. This chapter includes a careful discussion of a point that should receive more attention than it does. Although there are a number of studies showing that preferences change when situations change, as in the wake of a

---

[4] Neilson's (2011) paper asserts two basic tenets of behavioral economics, of which this quote is the first. The second is 'For every pattern uncovered through experiments, there exists a theorist clever enough to devise a model that can accommodate it' (p. 976). This may explain why there are so many alternative models in the literature.

disaster for example (Imas 2016), there is much work left to systematically explore the effects of specific elements of context.

The properties of the elicitation procedures themselves also may be important. For example, simple elicitation procedures, using only 50/50 gambles and round numbers, might be easier for subjects to understand and respond to accurately as compared with procedures with more complex alternatives. We see some evidence of this in our own experimental research (Dave *et al.* 2010). Indeed, there is growing evidence that intelligence is related to measured risk preferences, and lower math ability affects complex measures more than simple ones. This hints at the possibility that the validity of EUT may be masked by the properties of the tasks used to test it, and that carefully designed tasks taking cognition into account will lead to more consistent results. Simplifying elicitations won't help, of course, if the underlying model on which all of the measures are based is incorrect.

In the final chapter, the authors map out their ideas for moving forward. As they note, researchers have spent a great deal of time and attention coming up with increasingly elaborate ways of eliciting preferences and measuring various possible aspects of utility functions for decisions under uncertainty, with little ultimate success. Perhaps they are carefully measuring things that just aren't there. Instead researchers who really want to understand human decision making should pay close attention to humans as social, psychological and biological creatures. Individuals operate within a broad social and economic context, and context probably plays an important role in risk perceptions and risky behaviour. That context also has shaped evolved human responses to risk over the longer term. The large literature on decision-making heuristics and learning also provides important input, especially considering that heuristics are the result of an evolutionary environment. In addition, the extensive body of research by psychologists on risk perceptions and decision making is worth further evaluation by economists, as it has much in common with the approach the authors recommend.

As economists, we probably should try to remember that the utility function is a theoretical construct, invented not so much as an accurate representation of a human being's valuation process, but rather for convenience in thinking systematically about decisions that people make. Perhaps economists have got into trouble because we are under the sway of the seductive logical beauty of our theories. Economists come under fire all the time for their tendency to give priority to formal models over empirical evidence. Milgrom and Roberts observed that, 'no mere fact ever was a match in economics for a consistent theory' (1978: 185), and that may still be true. Recent critiques by Thomas Piketty (2014) or Paul Romer (2015) echo this sentiment. Perhaps the Bernoulli function is a case in point. In any case EUT is peculiarly resistant to countervailing data.

In the end, this book constitutes a fairly scathing indictment of the standard model of decision making under uncertainty, and by extension the standard model of consumer choice. Perhaps it is not a very strong defence to note that a prominent reason for the model's persistence is the absence of a strong alternative model (a fact noted by the authors). As the authors argue, other models have not proved to be much of an improvement over EUT in terms of their predictive ability. Economists outside the field of decision-making research are probably the most complacent about the validity of the EUT model. One reason for this might be that, twenty years ago, two prominently published papers attempted a kind of horse race between EUT and a subset of the other models available at the time. Harless and Camerer (1994) work with 23 different lab-generated data sets to compare the fit of alternative models and note a trade-off between parsimony (number of parameters) and accuracy, with EUT coming out in the top set of models. And Hey and Orme (1994) conclude that EUT does at least as well as any of the other models they test using lab experiments, if one allows for a bit of decision error. (Neither of these papers attempts to assess predictive accuracy outside the lab.) From these studies it is easy to draw the conclusion that EUT is pretty good after all, and certainly the best we have. As Roth (1996) argued, in the absence of a compelling alternative, EUT is a 'useful approximation' of behaviour. Although there continues to be a great deal of theoretical work on decision making under uncertainty, as far as I know no other similar attempts have been made more recently, and there are no convincing studies showing that other theoretical constructs are better.

The authors are rather quick to suggest that prospect theory is not the path forward, but perhaps it should receive a little more consideration. Prospect theory is motivated by patterns seen in empirical data on decision making, and it encompasses several elements, two of which are loss aversion and probability weighting. I would be quick to agree with the authors that probability weighting is not an appealing theory. One suspects that the pattern it is designed to approximate has some underlying cause that will be discovered and this will provide a better theory. But loss aversion is another story. The data presented in Chapter 6 (from the 1960 experiment by Grayson) show a pervasive 'kink' at a payoff of zero, consistent with loss aversion. Furthermore, the risk measure that the authors propose, expected loss, also seems to indicate that losses loom larger than gains in the mind of the decision maker. Aversion to losses seems to be a real component of human decision making and an unavoidable element of the path forward.[5]

---

[5] Many researchers undoubtedly would disagree with me about probability weighting as well. Google Scholar lists over 30 papers with 'probability weighting' in the title in the last two years, for example.

In sum, the book provides a fascinating argument against EUT, and marshals a wide variety of evidence both from individual decision-making experiments and from aggregate data. If our purpose is to develop a theory that is consistent with empirical evidence, and that can be used to model and predict individual decisions under uncertainty, then we have some work ahead of us. The authors make a strong and quite convincing case for reopening the search for a model of decision making under uncertainty that takes into account the world in which human decision makers live, and the perceptions that individuals have about risks and risky decisions. It is recommended reading for anyone interested in decision making under uncertainty.

Catherine C. Eckel*

REFERENCES

Bernoulli, D. 1954 [1738]. "Specimen Theoriae Novae de Mensura Sortis," Commentarii Academiae Scientiarum Imperialis Petropolitanae, 5, 175–192. English translation in *Econometrica* 22: 23–36.

Dave, C., C. C. Eckel, C. A. Johnson and C. Rojas. 2010. Eliciting risk preferences: when is simple better? *Journal of Risk and Uncertainty* 41(3): 219–243.

Grayson, C.J. 1960. *Decisions under Uncertainty: Drilling Decisions by Oil and Gas Operators*. Cambridge, MA: Harvard University Press.

Harless, D. W. and C. F. Camerer. 1994. The predictive utility of generalized expected utility theories. *Econometrica* 62(6): 1251–1289.

Hey, J. D. and Orme, C., 1994. Investigating generalizations of expected utility theory using experimental data. *Econometrica* 62(6): 1291–1326.

Imas, A. 2016. "The realization effect: risk-taking after realized versus paper losses. *American Economic Review* 106(8): 2086–2109.

Milgrom, P. and J. Roberts. 1987. Informational asymmetries, strategic behavior, and industrial organization. *American Economic Review* 77(2): 184–193.

Neilson, W.S. 2011. Theoretical advances spurred by "stubborn facts:" a defense of the Behavioral Economics Paradigm. *American Behavioral Scientist* 55(8): 976–986.

Piketty, T. 2014. *Capital in the 21st Century*. Cambridge, MA: Harvard University Press.

Romer, P. M. 2015. Mathiness in the theory of economic growth. *American Economic Review* 105(5): 89–93.

Roth, A. E. 1996. Comments on Tversky's 'Rational Theory and Constructive Choice'. In *The Rational Foundations of Economic Behavior*, ed. K. Arrow, E. Colombatto, M. Perlman and C. Schmidt, 198–202. New York, NY: Macmillan.

Slovic, P. 1962. Convergent validation of risk taking measures. *Journal of Abnormal and Social Psychology* 65(1): 68–71.

BIOGRAPHICAL INFORMATION

Catherine C. Eckel is Sarah and John Lindsey Professor in the Liberal Arts and University Distinguished Professor in the Department of Economics at

* Department of Economics, Texas A&M University, TAMU 4228, College Station, TX 77843-4228. Email: ceckel@tamu.edu. URL: https://sites.google.com/site/eckelcatherine/.

Texas A&M University, where she directs the Behavioral Economics and Policy Program. She is an experimental economist, and her research focuses on gender differences, cooperation and charitable giving, and behavioural economics in public policy. In 2013 she was awarded the Carolyn Shaw Bell Award, given annually by the American Economic Association Committee on the Status of Women, for her work developing and participating in mentoring programmes for women faculty.

Reasons without Persons. Rationality, Identity, and Time, Brian Hedden. Oxford: Oxford University Press, 2015. viii + 224 pages.

Reasons without Persons is a brilliant and suggestive book. It will arouse enthusiasm in disciples and critics of Parfit alike, and more generally in all those who study the concept of rationality, including philosophers and economists. Despite the complexity of the topic, the book is accessible, including to those who have never read anything on the subject before. All one needs is logic and reason. The book is divided into 11 chapters, with each chapter short enough to be read in one sitting. Each one is, however, rich, dense and thought provoking.

It is natural to believe that agents are rationally required to coordinate with themselves in a particular way, whilst they are rationally required to coordinate with other people in a different way (2). It is not irrational, for instance, for two distinct agents to have antagonistic preferences whilst it is irrational for a single agent to have such preferences. In *Reasons without Persons*, however, Hedden claims that we are wrong to hold different rational standards for inter-individual and intra-individual preferences, beliefs and actions (2). This claim is not new and has been notably defended by Parfit (1984), but a considerable literature has since developed to further challenge or defend this view. Hedden's book is a good illustration of the evolution of the debate over the last 30 years. References to Parfit are scarce, but his influence is implicit throughout. Parfit's psychological approach to what matters in personal identity, in particular, is not discussed but is tacitly admitted. More importantly, Hedden addresses rationality issues in a form more consistent with the discourse of modern economics. Most of Parfit's idiosyncratic language is consequently dropped.

Hedden embraces an impersonal view of agency, in which agents are reduced to a set of temporally located decision-making units, whose rationality is assessed independently of each other. Hedden calls this Time-Slice Rationality. It rests upon the combination of two basic ideas. First, the locus of rationality should be assigned not to temporally