On the Foundations of Decision Theory

Ken Binmore[1]

Abstract Bayesian decision theory was invented by Leonard Savage, who is on record as saying that it would be "preposterous" and "utterly ridiculous" to apply his theory except in a small world. But modern Bayesians proceed as though Savage's theory is always the rational way to make choices in all circumstances. What is a small world? Why did Savage restrict his theory to small worlds? Where did Savage think priors come from? What are the implications for behavioral applications? How could Savage's theory be generalized to apply to at least some large worlds? This paper offers some partial answers to such questions.

1 Axiomatic Method

A standard methodology in decision theory is to propose axioms from which theorems can be deduced about how rational decisions should be made. Such a definition-axiom-proof-theorem format closes the mind to factors that are not written into the formalism. When nothing that matters is omitted, the success of the approach speaks for itself across the whole scientific spectrum. But what if axioms

✉ Ken Binmore
k.binmore@ucl.ac.uk

1 Economics Department, Bristol University, 8 Woodland Road, Bristol BS8 1TB, UK

that have been found to work well in one kind of environment or with one kind of agent are held to apply in all environments and for all agent types without any further evidence?

To set the scene, it may be helpful to review how such a question plays out with the first of all axiom systems—Euclidean geometry. The historical focus has always been on the Parallel Postulate in Euclid's system, just as the focus is on the Sure-Thing Principle in Savage's axiom system.[1] The postulate implies among much else that the angles of a triangle add up to 180°. Experiments that Greek geometers could have carried out locally would have confirmed this prediction very closely.

If the earth were flat, the same would be true no matter how large a triangle we drew upon its surface. But the angles of a triangle drawn upon a round earth can sum to a lot more than 180°. Euclidean geometry therefore fails when applied outside the "small world" for which it was originally constructed. To insist otherwise is to proceed as though the earth were flat.[2]

Geometry on the surface of the earth is better approximated by a (Riemannian) axiom system in which Euclid's assumption that there is exactly one line through a given point parallel to a given line is replaced by the assumption that there are no such parallel lines. Flat-earthers are more successful when they follow Isaac Newton in taking for granted that Euclidean geometry holds in outer space, but even here they are in trouble close to a gravitating body, where the angles of a triangle add up to *less* than 180° because the geometry of space is then better approximated by a (Lobachevskian) axiom system in which there are many lines parallel to a given line through a given point.

The lesson is not only that axioms need to be tailored to the environment in which they are to be applied, but that implicit assumptions taken for granted when an axiom system is formulated need to be reviewed when considering modifications intended to widen the range of application of the system. This paper argues that Savage's (1954) chose to restrict his theory of rational choice to what he called a small world because he was anxious to avoid this kind of mistake, but that modern Bayesianism has fallen headlong into the errors he sought to evade by proceeding as though his descriptive theory of subjective probability can be reinterpreted as a prescriptive theory of logical probability without any need for further foundational discussion—the same error made by Immanuel Kant when he used Euclidean geometry as his leading example of a synthetic a priori.[3]

The paper therefore holds that the widely held belief among economists that rational learning always consists of no more than Bayesian updating is a castle in the air created by ignoring everything in Savage's theory not given formal expression. But no books need be burned. Bayesian decision theory remains available as a modelling tool that is arguably as good or better than any other

---

[1] Or the more elegant independence or replacement axioms more usually quoted nowadays.

[2] Not that Greek scientists were unaware that the world is round. Eratosthenes came up with a pretty good estimate of its diameter by measuring the length of shadows at noon at different places in Egypt. Aristarchus even knew that the planets orbit the sun.

[3] A synthetic a priori is something that might be wrong but we somehow know is right without needing any evidence.

currently on offer.[4] All we need to give up is the claim that Bayesian decision theory is the only rational way to make decisions in all situations. Otherwise posterity will classify us with the flat-earthers who could not see that there might be more than one geometry.

2 Probability

This section reviews the theory of probability from the perspective of the preceding section. As Gillies (2000) explains, there are multiple ways in which probabilities can be interpreted, depending on the context in which the concept is to be applied. Other possibilities are canvassed in the literature, but this paper considers only objective, subjective and logical probabilities.

Objective probability The small world for which Kolmogorov (1950) tailored his famous probability axioms is the set of long-run frequencies.[5] Kolmogorov's world is small, but still large enough to accommodate the expected utility theory of Von Neumann and Morgenstern (VN&M), whose own small world consists of rational agents who choose between gambles of the form

$$\mathbf{G} \; = \; \begin{array}{|c|c|c|c|c|} \hline \mathcal{P}_1 & \mathcal{P}_2 & \mathcal{P}_3 & \cdots & \mathcal{P}_m \\ \hline Q_1 & Q_2 & Q_3 & \cdots & Q_m \\ \hline \end{array} \tag{1}$$

in which Q_i is an event to which a probability can be *objectively* assigned (as a long-run frequency), and $\mathcal{P}_i$ is the outcome for a decision-maker that results when Q_i occurs. Betting in a casino is an obvious context. If an agent's preferences among such gambles satisfy VN&M's (Von Neumann and Morgenstern 1944) consistency axioms, then the agent will necessarily choose as though maximizing the expected value of a utility function defined over the set of all relevant outcomes.

Subjective probability Ramsey (1931) invented what might be called a theory of revealed or attributed belief, in which Alice's beliefs are deduced from her betting behavior—just as her preferences are deduced from her buying behavior in Samuelson's (1947) later theory of revealed preference.

---

[4] In particular, no criticism is offered of Bayesian statistics when the prior used is defended on the empirical grounds that similar priors have worked out well in similar contexts in the past. No criticism at all is directed at the use of Bayesian decision theory in game theory, as a game counts as a small world in Savage's sense almost by definition.

[5] von Mises (1928) offers a formal account of such long-run frequencies. de Finetti (1974, p. 124) gleefully points out that von Mises' long-run frequencies do not satisfy Kolmogorov's countable additivity axiom, but Binmore (2009, Chapter 6) shows how strengthening von Mises convergence requirement can eliminate this imperfection. Although we follow Kolmogorov in this paper in taking probabilities to be countably additive, nothing significant would change if we followed de Finetti (and the later Savage) in assuming only finite additivity.

If Alice's betting behavior satisfies certain consistency requirements,[6] the theory of subjective probability argues that she will act as though she believes that each relevant state of the world has a probability. These probabilities are said to be subjective, because Alice's beliefs may be based on little or no objective data, and so there is no particular reason why another person's subjective probabilities should resemble hers.

The axiom systems usually quoted nowadays in support of the subjective interpretation of probability are derived from that of Savage (1954), which extends the VN&M theory of expected utility to cases in which the events in a gamble (1) may not have objective probabilities. Horse-racing is usually quoted as a relevant context where objective probabilities are unavailable.

A person who honors the consistency requirements of Savage's axiom system necessarily behaves as though maximizing expected utility relative to a utility function defined on a set of relevant outcomes and a (subjective) probability measure defined on a set of relevant states of the world. The theory is usually called Bayesian decision theory because probabilities are updated using Bayes' rule as described in Sect. 4.

Logical probability A logical probability for a proposition is a measure of the rational degree of belief it derives from the available evidence. A full-blown theory of logical probability would solve the age-old problem of scientific induction, and so allow an objective assessment of questions like: What happens inside a black hole? Is there life elsewhere in the universe? What will Trump do next? Will economists soon be replaced by robots?[7] How much will cabbages sell for next week?

In considering the difference between subjective and logical probabilities, it is important to recognise that logical probabilities are like truth values in logic—they are objective to the extent that they are the same for everybody. Treating Savage's subjective probabilities as though they were logical probabilities therefore faces the problem that something that might be different for different people is treated as though it were necessarily the same for everybody. This error is built into the "Principal Principle" of the philosopher Lewis (1980), which says that credences must coincide with chances when the latter are to be found. Here the chance of an event is its objective probability and its credence is its subjective probability.[8]

If there were a viable theory of logical probabilities, they would doubtless have to coincide with objective probabilities in the small world where the latter exist, but there is no reason why Alice's subjective probabilities should coincide with the objective probabilities of events, even when she knows what they are. Gambling in

---

[6] The requirements include the very strong assumption that Alice always be ready to take one side or the other of any bet. One can then summarize the consistency requirements by saying that such an Alice must be invulnerable to Dutch books (a collection of bets that guarantees that she will lose no matter what happens). One cannot but wonder whether there has ever been a financier or industrialist who felt subject to such a constraint!

[7] The probability is exactly 42.9% according to today's London *Times*.

[8] Elsewhere I have mistakenly taken for granted that Lewis intended a credence to be a logical probability, since the Principal Principle is plainly false if credences are no more than subjective probabilities for the reasons given in the next paragraph.

casinos is an obvious example. Alice may well place her bets at Roulette in an entirely consistent manner and hence reveal subjective probabilities for where the little ball will end up that differ from their objective values. If asked why, she would likely say—in common with most gamblers—that she was feeling lucky that day. One may think that she is really playing for the thrill that some people get from gambling, but if Alice's betting behavior satisfies Savage's axioms, her subjective probabilities are no less respectable than those of people who consistently play to minimize their objective expected loss.

In any case, it needs to be said that *nothing* follows from the claim that *if* there were an adequate theory of logical probability it *would* coincide with theories of objective or subjective probability. An argument that derived a logical theory this way would be circular. Kolmogorov's axioms similarly lack any leverage here. They work for objective and subjective probabilities, but the need to provide new arguments if they are also to describe rational degrees of belief is often overlooked altogether—the error of the flat-earthers of Sect. 1. Personally, I am skeptical that a viable theory of rational degrees of belief is possible at all when restricted to such a limited tool as a probability measure.

3 Savage's Small World

The expression *small world* used in this paper is borrowed from Savage (1954, p. 16). Why did Savage (1954) restrict the application of his theory to small worlds? Modern authors commonly have no such compunction. Bayesianism is nowadays taken to be the doctrine that Savage's axioms always apply—that rationality requires employing Bayesian decision theory everywhere.

Savage insisted his theory be applied only in a small world because his axioms capture only a *thin* notion of rationality, which demands no more than that decisions be made consistently with each other. But why be consistent? After all, physicists know that quantum theory and relativity are inconsistent where they overlap, but they live with this inconsistency rather than abandon the accurate predictions that each makes within its own domain. Physicists strive to create a "theory of everything" from which such inconsistencies have been removed, but everybody recognizes that this is a problem of immense difficulty. Physicists and other scientists would therefore seem committed to a *thick* theory of rationality that somehow weighs the virtue of consistency against the vice of inaccuracy.

Savage's formal theory says nothing whatever about accuracy. It simply describes what systems of bets ensure that a Dutch book cannot be made against Alice on the assumption that she is somehow bound to take one side or the other of any bet. How accurate Alice's subjective probabilities may be is left entirely hanging in the air. In brief, Savage's theory merely provides a *descriptive* theory for a certain kind of idealized decision-maker. But what Alice would prefer—and a theory of logical probability would provide—is a *prescriptive* theory that tells her how best to choose among those bets that are available.

Savage nevertheless attached much importance to having opinions that are consistent with each other. A major reason is that by exploring what she would

believe in the future if various hypotheses were to be realized, Alice is able to test whatever unformalized set of considerations—usually characterized as gut feelings—she relies on when deciding what to believe in the present. If these future contingent beliefs turn out to be inconsistent (in that they together violate Savage's axioms) then she has grounds for distrusting her gut feelings and refining them until the inconsistency is removed.

Savage's famous encounter with Maurice Allais in Paris illustrates the point. When it was pointed out that his answers to what is now called the Allais paradox were inconsistent, he revised them until they became consistent. Luce and Raiffa (1957, p. 302) summarize Savage's views as follows:

> Once confronted with inconsistencies, one should, so the argument goes, modify one's initial decisions so as to be consistent. Let us assume that this jockeying—making snap judgments, checking up on their consistency, modifying them, again checking on consistency etc—leads ultimately to a bona fide, prior distribution.

It is to this unformalized massaging process that Savage (1954, p. 16) is referring when he says that it would be both "preposterous" and "utterly ridiculous" to use his theory except in situations in which the decision-maker can look before she leaps. For only in a sufficiently small world can Alice examine every bridge that she might conceivably have to cross in advance of knowing what route she will find herself following in the future. I have borrowed this geographical metaphor from Savage, but it fails to capture the enormity of the cognitive task Alice would usually face in seeking to examine the implications of all possible future information she might receive when seeking to massage her contingent snap beliefs into consistency. In the large world of scientific endeavor, where we cannot even guess what future state spaces may be thought reasonable, it is evidently quite impossible.

4 Bayes' Rule

Savage's theory reduces statistical inference to applying Bayes' rule for updating conditional probabilities. It is ironic that his theory was first called Bayesian for this reason by classical statisticians in a feeble attempt at ridicule.

Kolomogorov's definition of a conditional probability makes

$$\mathsf{prob}\,(E\,|\,F) = \frac{\mathsf{prob}(E\,\mathsf{and}\,F)}{\mathsf{prob}\,(F)}. \tag{2}$$

Bayes' rule then follows because

$$\mathsf{prob}(E\,|\,F)\,\mathsf{prob}\,(F) = \mathsf{prob}(F\,|\,E)\,\mathsf{prob}\,(E). \tag{3}$$

Why this approach to conditioning on new information and not another approach, perhaps involving more than simply updating a single measure? The answer depends on the small world in which Kolmogorov's definition is to be applied.

Objective probability Granted the experimental validity of the law of large numbers, no doubts about Kolmogorov's definition arise in the small world of long-run frequencies, as the following example illustrates.

A fair dice is to be rolled. Alice wins in the event E that the dice shows less than 3 or less. What is her probability of winning, conditional on the event F that the result is odd? Roll the dice $6n$ times. If n is large enough, it is very likely that each number on the dice will appear in the record about n times. If Alice now deletes all the even numbers, she will be left with a record containing about $3n$ odd numbers. She will lose when one of these numbers is 5, and win when it is 1 or 3. The number of times that the latter event occurs is about $2n$. The frequency with which Alice wins when the dice shows an odd number is therefore about $2n / 3n$. It is for this reason that Kolmogorov joined the rest of the world in endorsing the traditional definition that makes $\mathsf{prob}(E \mid F) = \frac{2}{3}$.

Subjective probability The reasons why Bayes' rule holds in the small world of objective probability are easily understood. But the same reasons will not suffice for other contexts in which probability is useful.

Savage ([1954](#), p. 43) therefore had to find another argument to justify Kolmogorov's definition for the small world of subjective probability. The argument is tucked away in Section 3.5 of his book, where he implicitly employs the Sure-Thing Principle to defend the proposition that Alice's preferences over pairs of gambles *given that F has occurred* should be taken to be the same as her preferences over pairs of gambles that each yield the same outcome if F does not occur and otherwise are determined by equation ([1](#)) with Q_i replaced by $Q_i \cap F$. Kolmogorov's definition then follows because his axioms imply that the preference relation over gambles that his theory takes as its primitive determines a subjective probability measure uniquely (up to a set of measure zero).

Logical probability Why will Savage's argument not suffice for logical probabilities? A clue is provided by its dependence on the Sure-Thing Principle—whose various incarnations as independence or replacement axioms are usually fingered as being too strong by those seeking to generalize the theory. However, the crux is whether it can be rational for Alice ever to change her mind when the state space with which she begins is replaced by a new state space B as a result of her receiving a new piece of information (Howson [2000](#)).[9]

Suppose, for example, that E and F are subsets of B, which is itself a subspace of a state space C. Are there situations in which it might make sense for Alice to hold that E is more likely than F when her state space is C, but that F is more likely than E when her state space becomes B?

Black swan events If Alice has exhaustively examined in advance the implications of all possible new pieces of information she might receive—as is taken for granted in Savage's small world—then such a change of mind would clearly be irrational. But suppose she has not carried through such an exhaustive analysis because her world is too large for such an analysis to be possible for all future events. As Simon

---

[9] Renyi ([1955](#)) and Popper ([1959](#)) and others consider the possibility of replacing Kolmogorov's axioms with a "relativized" system in which all probabilities are conditional, thereby leaving room for deviations from ([2](#)). Unfortunately, discussion seems to center around Popper's notion of probability as "causal propensity".

([1976](#)) argues, she will then need to decide (by some process for which we currently have no formal model) how best to use whatever cognitive or computing capacity she has at her disposal.[10] She will then presumably explore the implications of events that she thinks likely more deeply than those she thinks lie just on the edge of possibility. But she would be foolish not to bring her computing capacity to bear again when an event later occurs whose implications she had not previously explored adequately because of her computing constraint. In doing so, she may be lucky enough to enjoy an Archimedean eureka moment, after which anything might happen.

We are talking now about what Taleb ([2007](#)) calls *black swan* events. The credit crunch of 2007–2008 provides the most spectacular recent examples, but Shackle ([1949](#)) was drawing attention to the inevitability of such surprising events in macroeconomics and finance long before. One cannot evade their implications for rational decision theory by saying that experts should bring more computing power to bear. Doubtless the experts would have done better if they had tried harder, but it is in the nature of large worlds that they exhaust our capacity to fully plumb their depths. Nobody, for example, would argue that Isaac Newton should have explored the possibility that velocities near the speed of light do not add linearly, or that a photon might be in two places at the same time.

So where does this leave Bayes' rule for logical probabilities? The economics literature—where Bayesianism is most entrenched—actually offers *no defense at all*. Nobody seems to notice that this elephant is missing from the room when reinterpreting Savage's subjective probabilities as logical probabilities. Adopting a definition-axiom-proof-theorem format has closed our minds to the unmodeled massaging process that Savage assumed Alice to have completed before his axioms could reasonably be applied.

Many scholars would respond by saying that no defense is needed—that rational degrees of belief *must* be represented as probabilities and Kolmogorov's definition of a conditional probability comes as part of the package. But to insist that a theory of rational degrees of belief one must always require that $\mathsf{prob}\,(E \cap F) = \mathsf{prob}\,(E\,|\,F)\,\mathsf{prob}\,(F)$ is to proceed as though the event F and the "conditional event" $E\,|\,F$ are independent—which is exactly what must not be taken for granted if F is a black-swan event.

5 How are Priors Chosen?

Bayesian updating transforms a prior probability distribution into a posterior distribution. Such an updating procedure guarantees that consistency is preserved. But what of accuracy? Bayesianism addresses this question by quoting theorems which guarantee that the posterior distribution converges on the "true" probability

---

[10] In game theory at least, one cannot wave away such problems by assuming that Alice can costlessly carry out all computations instantaneously. It then sometimes matters that some propositions are mathematically undecidable (Binmore [2009](#), p. 147).

distribution as the sample size increases.[11] This response makes good sense when the probabilities in question are objective, and so it is reasonable to take for granted the existence of a true distribution from which new information is sampled. But even in the small world of objective probabilities, the rate of convergence of posterior probabilities on their true values will depend on the choice of prior. In the large world of logical probabilities, there is no guarantee that it is even meaningful to speak of a "true" probability distribution. Either way, the choice of prior is a problem for Bayesian decision theory.

Within Savage's theory of subjective probability, the choice of prior emerges from the massaging process that Alice uses to organize her initial disorganized impressions usually referred to as gut feelings. My reading of Savage's approach goes like this (Binmore 2009, p. 130).

Savage on constructing priors Everybody would presumably agree that it would be better for Alice to consult her gut feelings when she has more evidence rather than less. For each possible future course of events, she should therefore ask herself what subjective probabilities her gut would come up with *after* experiencing these events. In the likely event that these posterior probabilities turn out to be inconsistent with each other, she should then take account of the confidence she has in her current snap judgements to massage her posterior probabilities until they become consistent. After the massaging is over, Alice would have eliminated the possibility of being surprised by a black-swan event. She would *already* have taken account of the impact that *all* future information might have on the unformalized internal model that she uses in determining her beliefs. She would not only have adjusted the subjective probabilities she attaches to events in the small world with which she begins, but potentially expanded her state space to a larger small world in light of the new possibilities that black-swan events commonly suggest (Gillies 2001; Williamson 2003).

The end-product of Alice's massaging process is therefore a bunch of consistent posteriors defined on a state space that she will never need to revise. With the consistency axioms of Savage's theory of subjective probability, her massaged posteriors can all be deduced by Bayes' rule from a single prior. In this story, Bayes' rule is therefore reduced to a mere book-keeping tool that saves Alice from having to remember all her massaged posterior probabilities. The prior that Savage attributes to Alice therefore squeezes all the juice that can be squeezed from the disorderly set of impressions with which she comes to the problem.

Bayesianism and priors It is ironic that the avowed prophet of Bayesianism should advocate a procedure that reverses what Bayesianism recommends. Instead of deducing her prior from a massaged set of posteriors, the latter takes for granted that Alice begins with a prior and simply deduces her posteriors by updating using Bayes' rule.

There is no consensus on where such a logical prior comes from, but scholars who see Bayesianism as embodying the essence of rational learning usually argue that we should place the choice of prior at a time in the past when Alice is

---

[11] Provided the state space is finite and the prior attaches positive probability to all states.

completely ignorant.[12] An appeal is then sometimes made to the "Harsanyi doctrine", which says that different rational agents independently placed in a situation of complete ignorance will necessarily formulate the same common prior (Harsanyi 1977). But even if one accepts this doubtful principle, one is left wondering what this "rational prior" would look like.

In practice, some version of Laplace's principle of insufficient reason is usually employed.[13] Others take this position further by arguing that the complete ignorance assumption implies, for example, that the rational prior will maximize entropy (Jaynes and Bretthorst 2003). We are then as far from Savage's view on constructing priors as it is possible to be. Instead of using all potentially available information in the small world to be studied in formulating a prior, one treats all such information as irrelevant.

The literature offers many examples that draw attention to the difficulties implicit in appeals to versions of Laplace's principle. Similar objections can be made to any proposal for choosing a logical prior that depends only on the state space. Even in small worlds we seldom know the "correct" state space. In a large world, we have to update our state space as we go along—just as physicists are trying to do right now as they seek to create a theory of everything.[14]

In brief, Bayesianism has no solid answer to the question of where logical priors come from. Attempts to apply versions of the principle of insufficient reason fail because they take for granted that the state space is somehow known a priori beyond any need for revision. The same goes for any proposal that makes the prior a function of the state space alone. But how can Alice be so certain about the state space when she is assumed to know nothing whatever about how likely different states may be?

6 Rational Versus Behavioral

Economists pursue theories of rationality partly because they hope that they will assist in predicting how real people make economic decisions. It seems to be generally agreed that people frequently do behave as rational theories predict in market situations, but the advent of experimental economics has dealt a body blow to the idea that economic theories of rationality are good at predicting individual decision-making as was once taken for granted (Smith 1991; Kahneman and Tversky 1981). Expected utility theory is particularly unsuccessful (Friedman et al. 2014).

This section argues that there is nevertheless a positive role for theories of rationality in modeling individual behavior provided that the theories are not overly refined (Binmore 1975). However, it is necessary to give up the ambition of finding

---

[12] Bayesian statisticians choose whatever prior has worked out well in similar problems in the past. They might therefore be seen as using history as an (imperfect) substitute for Savage's massaging process.

[13] Equal probabilities should be assigned to two events if no reason exists to suppose that one is more likely than the other (de Laplace 1902).

[14] Binmore (2016) proposes the beginnings of a theory in which the state space may be updated.

one single behavioral theory that applies to all agents (Stahl 2014). That is to say, it is not only the world in which a problem is set that needs to be small, but the class of agents to which a theory is to be applied.

Ellsberg paradox The Ellsberg paradox is widely regarded as a final nail in the coffin of expected utility theory. One version involves an urn that contains 10 red balls and 20 balls that are black or white in an unknown proportion. Ellsberg (1961a, b) suggested that decision-makers asked to express preferences between various reward schedules that depend on which ball is drawn at random from the urn would not behave as Bayesianism recommends. Many experiments have confirmed this conjecture in a variety of settings.[15] As a result, it is now regarded as a stylized fact that most people have an inbuilt aversion to ambiguity over probabilities (Wakker 2017). But the Ellsberg world seems about as small as a world can get. So what behavioral role can there be for Savage's theory of expected utility if it does not even work here?

A possible answer relies on Kahneman and Tversky's (1981) demonstration of the importance of framing in economic experiments. Because of the applications that experimenters have in mind, the framing employed in many Ellsberg experiments poses the decision problem in precisely the kind of large-world setting to which Savage thought it ridiculous that his theory be applied. He would therefore not have been surprised to find that his theory did not work in such experiments. But what of small-world settings? Binmore et al. (2012) then find very little ambiguity aversion.[16] Charness et al. (2013) and Stahl (2014) similarly report little ambiguity aversion.

But such failures to observe much ambiguity aversion do not arise because of a triumph of expected utility theory. For expected utility theory to perform reasonably well, it also seems necessary to screen the subject pool using some other rationality criterion. For example, Halevy (2007) screens out subjects who do not understand how to simplify compound lotteries. Voorhoeve et al. (2016) screen out subjects who do not consistently give the same answer to questions that differ only in their framing.

In brief, it would seem premature to deny Bayesian decision theory any behavioral role at all. However, to make it work, we need not only to restrict the domain of application to a world that Savage would regard as small, but also to restrict the class of agents to those who pass some kind of rationality test. But most worlds of interest are large, and most people behave irrationally much of the time.

---

[15] In particular, if offered a choice between winning if a red ball is drawn and winning if a black ball is drawn, subjects usually prefer the former.

[16] The crucial factor seems to be minimizing the possibility that the experimenter may be manipulating the supposedly random draw behind the scenes. Pulford (2009) reports that merely asking subjects whether they trust the experimenter increases ambiguity aversion.

7 Ambiguity

Although Bayesianism is the prevailing orthodoxy in economics, its failures in the laboratory have fueled an ongoing literature in which alternatives to the Bayesian axioms are proposed and evaluated. The leading proposals tackle the problem of determining where the prior comes from by relaxing the requirements from which the existence of a unique prior follows. Theories with multiple priors then become possible, with the result that all probabilities are liable to be ambiguous. Such ambiguities are usually dealt with by computing the expected utility of a gamble for each prior and then aggregating the resulting collection of expected utilities somehow. Motivated by the ambiguity aversion reported in experiments on the Ellsberg paradox, the most popular aggregating method is to take the minimum value of the collection of expected utilities (Gilboa 2004).

Savage's restriction of his theory to small worlds is seldom mentioned in such attempts to generalize his theory to situations in which probabilities may not be uniquely determined. The position taken is that a fully informed Alice is wholly characterized by her VN&M utility function and her prior. Assuming her utility function is fixed, her ignorance can then only be about which prior should be employed.

Such a formulation in terms of multiple priors takes for granted that probabilities are logical, but Savage's probabilities are subjective. What happens if we reverse Bayesianism's assumption that one starts with a prior, and return to Savage's idea that a massaging process is required to convert Alice's original disorderly impressions into a consistent set of posteriors from which a single prior can be deduced?

If we add the all-too-likely proviso that Alice may not be able to bring the massaging process to a determinate conclusion, her problem ceases to be that of choosing between various probability distributions. We have no reason to suppose that her attempt to massage away inconsistencies in her snap judgements will end up with anything so simple as a set of possible priors between which she is unable to distinguish. She needs some new way of organizing her thoughts when probability is inadequate as a measure of her degree of belief in some events. Knight's (1921) notion of uncertainty is wide enough to capture this possibility, but it is necessary to live with the fact that ambiguity is nowadays often used as a synonym for Knightian uncertainty.

There is a large literature on how to proceed in situations when probability is thought inadequate, in which Schmeidler (1989) has taken a leading part. Much of this literature works with "imprecise probabilities" represented by the interval between a lower and an upper probability (Binmore 2009, p. 90). One may then hope to construct a decision theory in which expected utility is replaced by some way of assessing a gamble (1) in terms of the upper and lower probabilities of the events Q_i.

Various axiom systems have been proposed to this end, in which Savage's axioms are tweaked to see what emerges. But seldom is there any attempt to establish a context within which it is appropriate to apply one set of axioms rather

than another. The remainder of this section is therefore devoted to proposing a scenario in which Savage's concept of a small world can be used to explain where Alice's upper and lower probabilities for an event might come from, and why no other properties of the event should be regarded as relevant.

After a black-swan event Suppose that Alice has followed Savage's advice and so massaged her way to a subjective probability measure defined in some small world. She now updates her probabilities using Bayes' rule until surprised by a black-swan event—an event that she had not previously considered or thought relevant to her decision problem. It is now necessary for her to refine her state space lest she be surprised by similar shocks in the future. But how does she quantify her beliefs about the new events that now need to be assessed?

In the immediate aftermath of the black-swan event, she will have nothing to guide her at all beyond the beliefs that are summarized by her current subjective probabilities defined on the events of her original small world. Binmore (2016) proposes that these events be regarded as an algebra of measurable sets within her new state space. The new events of which she had not previously taken account are then regarded as unmeasurable. Their inner and outer measures with respect to her old subjective probability can then be identified with her new lower and upper probabilities.[17]

It remains to ask how gambles in which some of the determining events may be unmeasurable should be evaluated. In an attempt to minimize deviations from the Bayesian orthodoxy, Binmore (2016, 2017) proposes axioms that result in the gamble being assessed in terms of an average of the upper and lower probabilities of the relevant events.[18] The result is something quite far from expected utility theory, which is perhaps not so surprising when one takes account of the fact that unmeasurable sets can have very paradoxical properties.[19]

In brief, taking Knight seriously about the difference between risk and uncertainty would seem to call for a great deal more than some tinkering with the Bayesian axioms. At the very least, we need some kind of classification of large worlds so that new decision theories can be tailored to the context in which they are to be applied.

---

[17] In a finite space, the inner measure of a set is simply the measure of its largest measurable subset. Its outer measure is the measure of its smallest measurable superset. For example, in the Ellsberg paradox, the algebra of measurable sets would be $\{\emptyset, R, \sim R, B\}$, where R is the event that a red ball is drawn. The largest measurable subset of the unmeasurable event W that a white ball is drawn is $\emptyset$, and so the inner measure of W is 0. The smallest measurable superset of W is $\sim R$, and so the outer measure of W is 2/3.

[18] Either a weighted arithmetic or geometric mean. The latter is more attractive as ambiguity aversion comes built in and Bayes' rule survives for updating upper and lower probabilities.

[19] For example, in the Banach–Tarski paradox, a solid sphere is split into five unmeasurable subsets that can be reassembled into a sphere of twice the volume of the original sphere.

8 Conclusion

This paper argues that playing around with axioms without asking when it makes sense to apply one set of axioms rather than another is an unpromising approach to the general problem of scientific induction—to which very naive defenders of Bayesianism think their creed is a solution.

Even restricting the type of problem to be considered to its barest bones poses problems comparable to those faced by Fermat and Pascal when they groped their way toward the concept of probability. It took several hundred years before their hard-won insights were finally set in stone by Kolmogorov's axiom system. In refusing to open our minds in the same kind of way that the founding fathers of probability theory opened theirs when they turned their minds to solving various gambling problems, we risk making the same mistakes as flat-earthers like Immanuel Kant who thought Euclid's axioms were beyond criticism. The time for closing down debate with axioms is *after* the conceptual issues have been addressed and the context in which the theory is to be applied is properly understood.

References

Binmore, K. (1975). Rationality. In *Handbook of game theory* (vol. 4, pp. 1–25). Elsevier, Oxford.

Binmore, K. (2009). *Rational decisions*. Princeton: Princeton University Press.

Binmore, K., Stewart, L., & Voorhoeve, A. (2012). How much ambiguity aversion? Finding indifferences between Ellsberg's risky and ambiguous bets. *Journal of Risk and Uncertainty*, *45*, 215–238.

Binmore, K. (2016). A minimal extension of Bayesian decision theory. *Theory and Decision*, *80*, 341–362.

Binmore, K. (2017). Ambiguity in games? http://www.homepages.ucl.ac.uk/ uctpa97

Charness, D., Karni, E., & Levin, D. (2013). Ambiguity attitudes and social interactions: An experimental investigation. *Journal of Risk and Uncertainty*, *46*, 1–25.

de Finetti, B. (1974). *Theory of probability* (Vol. I). London: Wiley.

de Laplace, P.-S. (1902). A philosophical essay on probabilities. New York: Wiley. (Translated by R. Brambrough).

Ellsberg, D. (1961a). The crude analysis of strategy choices. *American Economic Review*, *51*, 472–478.

Ellsberg, D. (1961b). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, *75*, 643–669.

Friedman, D., Isaac, R., James, D., & Sunder, S. (2014). *Risky curves: On the empirical failure of expected utility*. New York: Routledge.

Gilboa, I. (2004). *Uncertainty in economic theory: Essays in honor of David Schmeidler's 65th birthday*. London: Routledge.

Gillies, D. (2001). Bayesianism and the fixidity of the theoretical framework. In *Foundations of Bayesianism* (pp. 363–379). Dordrecht: Kluwer.

Gillies, D. (2000). *Philosophical theories of probability*. London: Routledge.

Halevy, Y. (2007). Ellsberg revisited: An experimental study. *Econometrica*, *75*, 503–506.

Harsanyi, J. (1977). *Rational behavior and bargaining equilibrium in games and social situations*. Cambridge: Cambridge University Press.

Howson, C. (2000). *Hume's problem*. Oxford: Oxford University Press.

Jaynes, E., & Bretthorst, L. (2003). *Probability theory: The logic of science*. Cambridge: Cambridge University Press.

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

Knight, F. (1921). *Risk, uncertainty, and profit*. Boston: Houghton-Miffin.

Kolmogorov, A. (1950). *Foundations of the theory of probability*. New York: Chelsea.

Lewis, D. (1980). A subjectivist's guide to objective chance. In R. Jeffries (Ed.), *Studies in inductive logic and probability*. Berkeley: University of California Press.

Luce, R., & Raiffa, H. (1957). *Games and decisions*. New York: Wiley.

Popper, K. (1959). *The logic of scientific discovery*. London: Hutchison.

Pulford, B. (2009). Is luck on my side? Optimism, pessimism, and ambiguity aversion. *Quarterly Journal of Experimental Psychology*, *62*, 1079–1087.

Ramsey, F. (1931). Truth and probability. In F. Ramsey (Ed.), *Foundations of mathematics and other logical essays*. New York: Harcourt.

Renyi, A. (1955). On a new axiomatic theory of probability. *Acta Mathematica Academiae Scientiarum Hungaricae*, *6*, 286–385.

Samuelson, P. (1947). *Foundations of economic analysis*. Cambridge: Harvard University Press.

Savage, L. (1954). *The foundations of statistics*. New York: Wiley.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, *57*, 571–585.

Shackle, G. (1949). *Expectation in economics*. Cambridge: Cambridge University Press.

Simon, H. (1976). From substantive to procedural rationality. In *Method and appraisal in economics*. Cambridge: Cambridge University Press.

Smith, V. (1991). *Papers in experimental economics*. Cambridge: Cambridge University Press.

Stahl, D. (2014). Heterogeneity of ambiguity preferences. *Review of Economics and Statistics*, *96*, 609–617.

Taleb, N. (2007). *The black swan: Impact of the highly improbable*. New York: Random House.

von Mises, R. (1928). *Probability, statistics, and truth*. London: Allen and Unwin.

Von Neumann, J., & Morgenstern, O. (1944). *The theory of games and economic behavior*. Princeton: Princeton University Press.

Voorhoeve, A., Binmore, K., Stefansson, A., & Stewart, L. (2016). Ambiguity attitudes, framing, and consistency. *Theory and Decision*, *81*, 313–337.

Wakker, P. (2017). Annotated bibliography. http://personal.eur.nl/wakker/refs/webrfrncs.docx.

Williamson, J. (2003). Bayesianism and language change. *Journal of Logic, Language and Information*, *12*, 53–97.