



BERLIN SCHOOL OF ECONOMICS
DISCUSSION PAPERS

Discussion Paper #37

April 2024

**Bounded Rationality, Beliefs,
and Behavior**

Sebastian Schweighofer-Kodritsch

Bounded Rationality, Beliefs, and Behavior

Sebastian Schweighofer-Kodritsch, Humboldt-Universität zu Berlin

ORCID: 0000-0003-4370-1942

Abstract

This chapter presents a microeconomic, behavioral perspective on bounded rationality and beliefs. It begins with an account of how research on belief biases, in particular via probabilistic belief elicitation, has become mainstream in economics only relatively recently and late, even in behavioral economics (aka “psychology and economics”). The chapter then offers a review of the decision-theoretic foundations of modeling and eliciting (subjective) beliefs as probabilities, as well as selected—both classic and recent—evidence on humans’ bounded rationality from related research in psychology and economics. In doing so, it connects the historical debates within decision theory, on the one hand, and within psychology, on the other, concerning the normative status of expected utility and Bayesianism, as well as its methodological implications. A conclusion draws lessons for the practice of belief elicitation and future research.

Bounded rationality is a term proposed by Herbert Simon, who could be described as the first psychologist to be awarded the Nobel Memorial Prize in Economic Sciences (Camerer, Loewenstein, and Rabin 2004, xii). Simon refers to the discrepancies between human decision-making and that of “economic man” (*homo economicus*) as modeled in traditional economic theory (Simon 1955, 99). Since its inception, the term has received multiple interpretations, corresponding to different approaches to the implications of Simon’s proposition. Most prominently, bounded rationality is associated with the use of heuristics in judgment in the face of uncertainty, though there is disagreement on whether evidence of bias should be interpreted as irrationality (e.g., compare Gigerenzer and Selten (2002) with Kahneman (2003)).¹

The most important economic decisions we face (savings, education, fertility, etc.) are so because of their long-term and accordingly uncertain consequences. It seems obvious that our beliefs are a key component of our economic decision-making calculus. It is therefore rather surprising that research in economics, which seeks to understand observed outcomes as the

result of optimal individual behavior, has until rather recently shunned the direct elicitation of individuals' beliefs.²

Quite tellingly, the research in behavioral economics, which seeks to better explain and predict individual behavior by incorporating more realistic psychological assumptions, initially focused on and in the early 1990s gained momentum with firmly establishing and rigorously exploring deviations from standard assumptions on risk, social, and time preferences, as opposed to belief formation. Pioneers of this approach recount the hostility with which it was met. In view of this response to initial work on non-standard preferences, for instance “a model of nonbelief in the law of large numbers” (Benjamin, Rabin, and Raymond 2016) would have seemed unpublishable, if not unthinkable, in those early days.

Incidentally, macroeconomic modeling for a long time assumed that (representative) agents' consumption and saving behavior follows adaptive expectations, a form of bounded rationality. The rational expectations revolution overturned this and instilled an insistence on microeconomic foundations in macroeconomic modeling (see this Handbook's Chapter 2 on “Rational Expectations”). Against this backdrop, one can imagine that deviating from rational expectations in *micro*-economic modeling would have felt like a step backward. In any case, even the psychology-oriented behavioral economics approach initially maintained not only that beliefs are probabilistic but also that they are unbiased relative to “a” rational benchmark. Depending on the type of analysis, this would correspond to probabilities that are either objectively given or theoretically determined by an equilibrium notion or empirically estimated from actual data.

Over the past decade or so, however, this has changed drastically. Non-Bayesian decision models as well as models of agents whose beliefs and tastes are non-separable are increasingly developed, and belief elicitation techniques nowadays belong to the standard toolbox of experimental economists. Illustrating this change, the first economics experiment on belief updating (Grether 1980) had fewer citations during the whole of its first ten years after publication than in the year 2022 alone (53 vs. 63).³ My own research has followed the same development, from studying non-standard preferences to studying non-standard belief formation, in “fast forward” so to speak.⁴

However, what “really” are beliefs, as a mental representation of uncertainty, and how do responses to probabilistic belief elicitation relate to them? In particular, how could one interpret elicited “belief bias” relative to some objective probability as a sign of irrationality? These questions should be fundamental for any researcher studying beliefs and anyone relying on such research. This chapter does not offer definitive answers, but instead some key insights as to how the answers depend on what one is willing to assume. Thus, it may also shed light on the above documented historical development of “beliefs research,” in particular its otherwise surprisingly late adoption in behavioral economics.

The perspective presented in this chapter is a microeconomic, behavioral one, which is worthwhile clarifying. Economics generally, but especially *micro*-economics, is foremost concerned with explaining and predicting *behavior*, such as economic agents’ responses to different policies. Its approach of developing formal and broadly applicable models of decision-making for this purpose aspires to be fully transparent with regards to the assumptions underlying any welfare conclusions and policy recommendations. Strictly speaking, any psychological construct that is not directly observable, including an

individual's "beliefs," therefore becomes meaningful to economics only insofar as it can be understood in terms of this individual's observable behavior; indeed, a strict formulation of this methodological stance is the following: "Economic models connect to reality through economic variables, prices, quantities etc. and not through their modeling of the individual's decision-making process. Evidence of the sort cited in neuroeconomics may inspire economists to write different models but it cannot reject economic models" (Gul and Pesendorfer 2008, 22). While one may simply ask people to state a probability that an event will occur, like "The Dow Jones will increase by more than 5% by next week," the difference in economic meaning between responses of 43% and 29% depends on what that difference says about the behavior of interest, e.g., demand for various financial assets.

I first review the decision-theoretic foundations of expected utility and subjective probability ("Bayesianism") in terms of observable behavior, including the debate among decision theorists about their normative status and related evidence from economics, both classic and recent. Then, I revisit the historical debate in psychology regarding Kahneman and Tversky's "Heuristics and Biases" approach, specifically that between these two researchers and their fierce critic Gerd Gigerenzer, where I add some recent related evidence from experimental economics, including my own work. I conclude with what I consider the main lessons from this chapter.

Decision-Theoretic Foundations: Beliefs as Probabilities, and Rationality

In the modeling of agents that make economic decisions under uncertainty, the standard assumption is that they maximize expected utility (EU). This means their subjective beliefs and tastes are well-defined separable objects, with beliefs being probability distributions.

Indeed, in economics, there is never any need to explicitly mention, let alone defend, the assumption that beliefs are probabilistic.

There are good practical reasons for this assumption, of course. Being able to rely on probability theory, in particular Bayes' Theorem, is immensely powerful, similar to how modeling tastes via utility functions is useful in that it enables us to employ mathematical optimization techniques. In particular, using probabilities for beliefs provides a clean, mathematically well-understood structure for thinking about and illustrating agents' uncertainty, again similar to how utility functions are useful representations of tastes for this purpose. Economists therefore impose such basic assumptions, on which they then build equilibrium/solution concepts (e.g., some version of competitive equilibrium, or of Nash equilibrium) in order to derive implications for economic behavior and outcomes, which is their ultimate interest.

Decision theorists are specifically concerned with the foundations of such basic assumptions, both from a positive and from a normative point of view (Gilboa 2009). The dominant approach behaviorally characterizes models of decision-making in terms of preference axioms that directly correspond to choices, hence (at least in principle) observable behavior. On the one hand, these axioms therefore offer direct ways to testing the descriptive validity of a given decision model and potentially refining or generalizing it. On the other hand, the axioms allow us to better assess the model's prescriptive appeal as a decision criterion directly on this behavioral level.

Students of economics encounter this approach early on. Indeed, I would expect the majority of undergraduate students to know of an alternative to utility functions in modeling consumer

choice (under certainty), namely preference relations over consumption bundles. Some might even be familiar with the axiomatic foundation of EU by von Neumann and Morgenstern (1947) for choice among so-called lotteries, where uncertainty is given as probabilistic, as part of the data. By contrast, I would expect that even many actual economists would be puzzled if asked about an alternative to modeling a human decision-maker's beliefs as a probability distribution. (Preference relations, of course.) Yet, most economic applications involve *subjective* uncertainty, i.e., uncertainty that is not merely due to pure chance or the use of randomization devices. Modeling beliefs as probability distributions in this setting is an assumption indeed, and a—if not the—major achievement of decision theory has been to clarify what it means in terms of behavior. Moreover, it also allows us to clarify how given probabilistic beliefs should and do relate to behavior.

Expected Utility as Model

A classic distinction in decision theory is between types of uncertainty: objective uncertainty aka “risk” and subjective uncertainty aka “ambiguity” (Gilboa 2009). Work in the former domain assumes probabilities are a given, as observable data. The seminal behavioral characterization result in decision theory with risk is von Neumann and Morgenstern's (1947) theorem. It offers the first axiomatic foundation of maximizing EU:

$$U(P) = \sum_{x \in X} p(x)u(x),$$

in choosing between lotteries P that assign probabilities $p(x)$ to various outcomes $x \in X$, which are usually taken to correspond to monetary amounts. Since probabilities are a given here, the behavioral characterization concerns how these probabilities enter the decision-making calculus.

In Leonard J. Savage’s *tour de force* to revolutionize statistics (Giocoli 2013), he obtained a behavioral characterization of EU for uncertainty about which nothing is assumed or given (Savage 1954). Instead, the uncertainty and its probabilistic assessment are subjective and part of the behavioral characterization, considering choices between acts $f: S \rightarrow X$ whose outcomes depend on which state of the world $s \in S$ realizes. For his fully subjective setting, Savage requires an infinite state space, however. Anscombe and Aumann (1963) show how assuming that there also exist randomization devices—so that “acts,” in every state, may result in lotteries with given probabilities of various outcomes (Savage acts are then special cases of degenerate lotteries). This greatly simplifies the characterization of *subjective* EU (SEU) and allows to obtain it also for a finite state space. Denoting by $f(x|s)$ the objective probability of outcome x under act f if state s realizes, EU takes the form

$$U(f) = \sum_{s \in S} p(s) \sum_{x \in X} f(x|s)u(x).$$

This latter characterization result provides the foundation for the standard assumption of EU in economic modeling with uncertainty, and its generalizations in decision theory.

Objective Uncertainty and Rationality

What is meant by *objective* probability of a particular outcome, and why even probability?

The most common approach is to define the probability of an event in a frequentist manner, as the relative frequency with which it would occur in a large number of trials “under the same conditions,” so by turning the law of large numbers that concerns convergence to a pre-supposed true probability on its head (Gilboa 2009, 21). As such, it is merely the definition of

a precise quantitative measure of uncertainty, and what it results in is objective to the extent that we all agree on this number (Anscombe and Aumann 1963, 204).

This notion shifts the question to when we all shall definitely agree. For randomization devices like coins that are tossed, dice that are rolled, or roulette wheels that are spun—at least when this is done vigorously (Diaconis, Holmes, and Montgomery 2007)—but also for computer-generated (quasi-) random numbers, we could actually run very large numbers of such trials and, having agreed on doing so, will reach the same measurement. Put simply, for such uncertainty, one can turn probability into something that is commonly observable. These are the typical examples of objective probabilities.

The Independence Axiom as Rationality Given Probability

In von Neumann and Morgenstern's foundation of EU, probabilities are given as data, so why mention it here when this should be about beliefs? Most importantly with regards to rationality, EU with given probabilities allows us to ask, without any ambiguity, the normative question of how these *given* probabilities should enter one's decision calculus. That is, if I accept the given probabilities as measuring the uncertainty I am facing, how should I use this measure in my decision-making? The behavioral foundation of EU offers a basis for why we might want to use the probabilities to calculate the expectation of some cardinal utility function over outcomes and then maximize this criterion.

A priori, this must be a hugely challenging question. However, it is one that illustrates the value of decision theorists' foundational work, because the EU theorem simplifies the question enormously; it boils this normative question down to whether I want my choices to respect the independence axiom.

For any two lotteries P and Q , and any $\alpha \in [0,1]$, let $\alpha P + (1 - \alpha)Q$ be the “mixed” lottery that assigns the probabilities $\alpha p(x) + (1 - \alpha)q(x)$ to various outcomes $x \in X$. The independence axiom says that if you prefer lottery P over lottery Q , then, for any third lottery R and any $\alpha \in (0,1)$, you shall prefer lottery $\alpha P + (1 - \alpha)R$ over lottery $\alpha Q + (1 - \alpha)R$. This axiom did not appear as such in the work of von Neumann and Morgenstern (Fishburn and Wakker 1995), and while it was being distilled, even Paul Samuelson initially took the position that it was rather arbitrary—he could not see how rationality would imply cardinal utility, and specifically, more than respecting first-order stochastic dominance. No other than Savage then convinced him in a letter of correspondence, by stating it along the following lines (Moscati 2016, 230): Given two mutually exclusive events E and F , and three outcomes $\{x, y, z\}$, suppose you prefer outcome x over outcome y and have to choose between the two contracts (i) “ x if E occurs and z if F occurs” and (ii) “ y if E occurs and z if F occurs.” Then, your choice is irrelevant if F occurs, so you should condition on the event E where it matters. This should lead you to choose (i) because it delivers the preferred outcome in that event. Note that Savage here makes no explicit reference to the nature of the uncertainty about events, nor to lotteries. Indeed, this is not really independence but actually the first statement of the sure-thing principle, the key axiom in Savage’s derivation of SEU (Moscati 2016, 230–231). However, upon replacing the three outcomes with lotteries $\{P, Q, R\}$ and the uncertainty about events with another lottery, yielding E with probability $\alpha \in (0,1)$, (i) and (ii) correspond to lotteries over lotteries aka compound lotteries that reduce to $\alpha P + (1 - \alpha)R$ and $\alpha Q + (1 - \alpha)R$, respectively.

Using this dynamic compounding interpretation, a common related way of arguing for the normative appeal of independence is dynamic consistency (Karni and Schmeidler 1991). If you preferred P over Q , but also $\alpha Q + (1 - \alpha)R$ over $\alpha P + (1 - \alpha)R$, this means you would like to conditionally choose Q over P instead, a preference reversal. This argument for independence relies on a reduction (of compound lotteries) axiom, in the ex ante evaluation, as well as a consequentialism axiom, so that ex post the fact you did not end up with lottery R , whatever it might be, would not alter your preference regarding P and Q .

Bounded Rationality and Prospect Theory

Yet, independence is famously violated by most people, which is known as the Allais paradox (Allais 1953) and was key for Kahneman and Tversky's development of prospect theory, which delivered the leading descriptive model of decision-making under risk (Kahneman and Tversky 1979). Consider their Problems 3 and 4, based on Allais, and then their Problem 10 (any residual chances in the lotteries below concern winning nothing, which I omit in the description):

- Problem 3 asks respondents to choose between (A) an 80% chance of winning 4000 and (B) winning 3000 for sure.
- Problem 4 asks respondents to choose between (C) a 20% chance of winning 4000 and (D) a 25% chance of winning 3000.
- Problem 10 has Problem 3 as a potential second stage preceded by a first stage, where there is a 75% chance that “the game” ends immediately without winning anything and a 25% that it moves to the second stage. It asks respondents for a conditional choice.

Note that Problem 4 is the reduced version of Problem 10, which features Problem 3 as a conditional choice. In particular, independence requires that in both Problems 3 and 4 the same “amount” is chosen (either both times the less likely 4000 or both times the more likely 3000), while reduction requires that this is true in both Problems 4 and 10. In their between-subjects study, while 80% chose the more likely 3000 in Problem 3, only 35% did so in Problem 4, and again 78% did so in Problem 10. Notably, essentially the same fraction chose the more likely 3000 in both Problems 3 and 10. This suggests that violations of independence are due to a cognitive difficulty with identifying common “sublotteries” in order to focus and condition on the “real” difference that matters. The dynamic compounding or description via events as in Problem 10 or Savage’s argument, respectively, make this transparent. Hence, while people rather clearly fail at reduction, they may well satisfy a weaker “compound independence” (Segal 1990), which experimental work has confirmed to a fair extent (Conlisk 1989).

Only very recently have researchers found a convincing way of directly eliciting how people would like to make choices for the (static) risk domain in an incentive-compatible manner (Nielsen and Rehbeck 2022). The authors operationalize the subjective notion of rationality suggested by Itzhak Gilboa: “a mode of behavior is *irrational for a decision-maker*, if, when the latter is exposed to the analysis of her choices, she would have liked to change her decision, or to make different choices in similar future circumstance. Note that this definition is based on a sense of regret, or embarrassment about one’s decisions, and not only to observed behavior” (Gilboa 2015, 316) A vast majority of 83% of the study participants (Ohio State University undergraduate students) select the independence axiom for their choices, compared to 14% that select an “opposite” control axiom. A majority of 75% of them end up violating it when left to their own devices, however. Notably, almost half of the

participants keep their choice inconsistent with the independence axiom when reminded of their prior selection of it and shown what the axiom would choose instead. For comparison, the transitivity axiom (and its opposite control axiom) was selected at very similar frequencies, but only 43% violated it, and of those only 17% left their choice and this axiom inconsistent when reminded. This further confirms that, while independence is generally normatively appealing, its application is cognitively difficult, and the fact that many accept the inconsistency suggests that people even lose sight of its appeal in concrete examples. Relative to EU, for static choice, the descriptive success of prospect theory stems from how given probabilities enter the decision calculus. The given probabilities are nonlinearly transformed via an inverse-S-shaped function into decision weights on outcome utility. Unlike probabilities, decision weights do not sum to the same constant/one. In particular, the utility of unlikely outcomes is therefore overweighted, while that of likely outcomes is underweighted, relative to EU. Very recent work interprets this as bounded rationality reflecting “cognitive uncertainty” (Enke and Graeber 2023). They elicit valuations for win-lose lotteries in a standard manner, varying the amount that may be won as well as the winning probability. However, they additionally elicit the participants’ confidence that their true valuation lies in a certain range around their earlier response. The valuations of those who report high confidence (low cognitive uncertainty) are closely in line with the risk-neutral EU prediction across the probability space, whereas those with low confidence (high cognitive uncertainty) show insensitivity in valuation to probability in line with the inverse-S-shape from prospect theory. Notably, the authors find that cognitive uncertainty is largely unrelated to response times or standard measures of cognitive ability. This opens an avenue for research to investigate the determinants of cognitive uncertainty, and their relevance for economic decisions in the field.

Most relevant to this chapter, however, it also raises the issue that the descriptive success of prospect theory might not be about preferences in the usual sense. It may be that people do not “accept” given and supposedly objective probabilities as the relevant measure of their uncertainty to begin with (Gollier and Muermann 2010); in particular, could we think of prospect theory’s decision weights as subjective non-probabilistic beliefs? Broadly speaking, even to this day, “the psychology of probability weighting is poorly understood” (O’Donoghue and Somerville 2018, 110).

Subjective Uncertainty and Bayesianism

In its extreme form, doing without any objectivity in assessing uncertainty, the subjectivist approach is most provocatively stated as “PROBABILITY DOES NOT EXIST” (de Finetti 1974, x). For most “naturally” occurring uncertainty (election outcomes, economic or financial conditions, health outcomes, relationship outcomes, etc.), the law of large numbers seems hopelessly inapplicable. How could one then give probability any meaning? Savage’s *Foundations of Statistics* (1954) offers an answer, behaviorally founding Bayesianism as SEU, which is considered the “crowning achievement of single-person decision theory” (Kreps 1988, 127). Savage followed Abraham Wald’s approach to statistics, whereby every statistical problem is best viewed as a decision problem (Giocoli 2013). There is a set of unknown states of the world S , fully describing all uncertainty, and a set of relevant outcomes X . The decision problem is to choose among “acts” $f: S \rightarrow X$, whose consequences/outcomes depend on the uncertainty. Choices between acts satisfy a system of axioms (including the sure-thing principle, see above), if and only if there exist (i) a unique probability measure over states and (ii) a cardinal utility function over outcomes such that those choices maximize the expectation of (ii) using (i).

Savage thus shows how preferences over acts allow to pin down *both* subjective constructs “probabilistic belief” and “utility” (the former uniquely, the latter up to positive affine transformations). Taking this model literally appears descriptively ridiculous, however—just think about making plans for retirement. Coming up with the state space, which here is assumed data, already seems daunting, and indeed, Savage requires an infinite state space. One is rather immediately led to accepting bounded rationality as the human condition (Simon 1955).

Bayesianism and the Ellsberg Paradox

Having already discussed the independence principle for choice given probability, including how humans violate it, I focus here on the Bayesian part of SEU—assigning precise subjective probabilities to various events (Savage speaks of “personal” probabilities)—regarding its descriptive validity and normative appeal, based on the “Ellsberg paradox,” a series of (thought) experiments that challenge it on both levels (Ellsberg 1961). As introduction to this discussion, it is worthwhile noting that Savage simply rejected the Allais paradox. He argued that any rational person, given probabilistic beliefs, when confronted with her violation of independence would revise her choice to be consistent with EU (Savage 1954, 102–103). Note that the aforementioned study of Nielsen and Rehbeck (2022) on undergraduate students directly relates to this argument, and Savage would have to declare a large fraction of those students “irrational.”

By contrast, Savage’s defense of the Bayesian approach’s assigning precise probabilities to various conceivable events, “before” evaluating different available acts conditionally on these events, appears more pragmatic: “Once a frequentist position is adopted, the most important uncertainties that affect science and other domains of application of statistics can no longer

be measured by probabilities. A frequentist can admit that he does not know whether whisky does more harm than good in the treatment of snake bite, but he can never, no matter how much evidence accumulates, join me in saying that it *probably* does more harm than good. (...) he is cut off from most applications of Bayes' theorem, the algorithm for calculating what the new probability of a proposition is on the basis of its original probability and new relevant evidence" (Savage 1961, 576).

This pragmatism is confirmed also by Savage's response to the Ellsberg paradox demonstrating the descriptive failure of SEU. Ellsberg himself subscribed to the subjective approach to uncertainty but criticized the reduction of all uncertainty to precise probabilities, as if it was risk. For instance, consider his single-urn experiment (Ellsberg 1961, 653ff.). An urn contains 90 balls, each of which is either red or black or yellow, and one ball will be drawn at random. You know that exactly 30 balls are red, but nothing about the color composition of the remaining 60 black and yellow balls. Consider bets that win you 100 if the underlying event concerning the color of the ball occurs, and nothing otherwise. Which bets will you choose in the following two problems?

1. Betting on "red" or betting on "black"?
2. Betting on "not red" or betting on "not black"?

Most people choose the "red" and "not red" bets, which entail only risk (where winning probabilities are given as $1/3$ and $2/3$, respectively). By contrast, the "black" and "not black" bets entail ambiguity. Anyone assigning a probability to "black" does so to "not black," however, as the probabilities of two complementary events add up to one. Betting on "red" reveals the probability assigned to "black" to be less than one third, whereas betting on "not red" reveals the probability assigned to "not black" to also be less than one third, an

inconsistency. As it turns out, various famous economists not only exhibited this paradox but also questioned whether this was irrational.

This sparked an intense debate in economics. Savage, who allegedly exhibited the paradox himself, remained largely absent from this debate, focusing instead on his impact on the practice of statistics (Zappia 2021, 172). Correspondence between him and Ellsberg, as well as other researchers like William Fellner or Bruno de Finetti shows a rather nuanced viewpoint, however, that stands in contrast to his response to the Allais paradox and his published works in statistics (Zappia 2021). Not only did Savage himself admit to “vagueness” (e.g., regarding “the likelihood of total war in the near future” (Zappia 2021, 186)) but he also appeared open to considering it even from a normative viewpoint, should it be developed axiomatically for a behavioral characterization. Today, sixty years after these exchanges, there is still no broad consensus among decision theorists regarding whether rationality implies Bayesianism (Gilboa 2015).

Bayesianism as Rationality given Objective Probability

Unfortunately, I found no record of whether or how Savage’s views were affected by Anscombe and Aumann’s (1963) alternative characterization of SEU, which I find to contain the most compelling argument for Bayesianism. Technically, their mathematical approach leverages the existence of objective probability and EU on this part of the domain, for behaviorally defining—or, better, eliciting—various subjective probabilities. This immensely simplifies the characterization in terms of axioms and proof, and it obtains for a finite state space as well. Later work on generalizations of SEU have therefore relied on it, and so does incentivized belief elicitation in economics.

I interpret Anscombe and Aumann as follows: Upon accepting that rationality implies EU for objective probabilities, one cannot but accept that it implies Bayesianism. Indeed, the much more transparent set of axioms shows the key role of the independence principle. In this framework, it is applied to acts that assign to every state a lottery (objective probabilities over outcomes) and where “mixing” of any two such acts is done state by state. This may be cognitively hard to perform but is still harder to object.

Normative reservations with respect to Bayesianism arise from unease with assigning precise probabilities to the truth of any proposition, even “meaningless” ones employing words that may just as well be made up. Why should it be irrational to say “I don’t know” (Gilboa 2015, 329–330)? However, SEU takes as given what the decision-maker knows. With regards to the outcome, “I don’t know” is the defining feature of decision-making under any kind of uncertainty. The fact that SEU does not give us guidance as to how we might come up with precise probabilistic beliefs given what we know or do not know in relation to the uncertainty can therefore hardly be a criticism. Moreover, SEU is not about any kind of judgment but about decision-making when something is at stake. While I surely find it difficult to assign a precise probability to the event that Donald Trump will win the next US presidential election, if I had to choose between a bet on that event and a bet on a fair coin toss, this choice would reveal something about this probability anyways. Put simply, my acceptance of the existence of objective probabilities and that my choices given such probabilities should satisfy independence implies that, in any situation of subjective uncertainty, I want to choose *as if* I were a Bayesian.

This difference between mere judgment and decision-making is reflected in the fact that experimental economics—in contrast to psychology—insists on incentivized elicitation,

including belief elicitation (whenever possible). To caricature this methodological approach, it would hold that “actions speak louder than words” and therefore make sure you “put your money where your mouth is.” This seems to assume that, without incentives, people would lie to us researchers. More seriously, it derives from the understanding that incentives induce respondents to take the task seriously, as they supposedly do in naturally occurring decision problems. It is worth noting that this is reflected also in the behavioral characterization of SEU: If the decision-maker is indifferent over all acts from which to choose, subjective probability is not defined (Gilboa 2009, 91).

My sense is that the normative implication arising from Anscombe and Aumann’s result is underappreciated in the literature, because Savage came first, of course, and with such an impact that the result’s main contribution seemed to be one of mere technical simplification. It therefore seems somewhat ironic that the decision-theoretic development of non-Bayesian models relied on the Anscombe-Aumann setting, including objective probability and hence a rather compelling normative case for Bayesianism. The seminal two such models here are Choquet EU (Schmeidler 1989), generalizing subjective probabilities to subjective “capacities” that need not be additive, and maxmin EU (Gilboa and Schmeidler 1989), allowing decision-makers to simultaneously maintain multiple subjective beliefs from which they would compare various acts according to their minimal EU. These generalizations of SEU were not only intended by the authors to accommodate the Ellsberg paradox with a notion of ambiguity aversion but also as something of a critique of Bayesianism (Gilboa 2015). If one agrees that rationality implies SEU, however, these theoretical approaches constitute positive decision theory (Dekel and Lipman 2010), helping us to better understand the cognitive reasons why being a Bayesian is difficult, i.e., the nature of our bounded rationality. Indeed, they all rely on weakening the independence axiom (as extended by

Anscombe and Aumann), and I am not aware of any normative argument with that (indeed, see the preface of Edwards (1992)).

There is one important limitation to SEU, however. The behavioral definition of subjective probability of an event cannot depend on the outcome that occurs in this event, so that probability and utility are separable (Gilboa 2009, 128). The definition of states and outcomes is a critical element in the applicability of SEU theory. In a letter to Aumann, who had posed problems with “state-dependent utility” (Drèze and Rustichini 2004), Savage would openly admit this general limitation, show how it may often be avoided by redefining outcomes, but then also that this could become ridiculous, by writing: “I should not mind being hung so long as it be done without damage to my health or reputation” (Gilboa, Minardi, et al. 2020, 382). Hence, from the behavioral viewpoint, subjective probabilities regarding adverse life events such as severe illness, divorce etc. are hardly defined, which is important to be aware of in the practice of belief elicitation (Schlag, Tremewan, and van der Weele 2015). Relatedly, even when subjective probabilistic beliefs about an event are well-defined, theoretically, no incentive scheme could truthfully elicit them if the respondent has a personal stake in the event—which may be hard to determine in practice (Kadane and Winkler 1988).

Bayesian Updating and Information

As Savage himself suggested, to motivate his Bayesian approach as a necessary replacement of frequentism, once the definition of probability is extended beyond the objective, frequentist domain, the logic of probability theory is broadly applicable to various problems. Scientific progress thus turns into Bayesian updating. Let $p(E)$ denote the prior probability of any event E (or a proposition’s truth), so for its complement E^C (or the truth of the

proposition's negation) we have that $p(E^c) = 1 - p(E)$. Bayes' Theorem says that upon learning event F that had positive prior probability $p(F) > 0$, the probability of any event E updates from prior $p(E)$ to posterior

$$p(E|F) = \frac{p(E \cap F)}{p(F)} = \frac{p(E \cap F)}{p(E \cap F) + p(E^c \cap F)} = \frac{p(F|E)p(E)}{p(F|E)p(E) + p(F|E^c)p(E^c)}.$$

Note that in order to apply it, the Bayesian decision-maker indeed has to have a prior probability for the event F and any joint event $E \cap F$ (or truth of both propositions). The prior thus defines the statistical meaning or information content of event F . This illustrates again the cognitive demands on beliefs as a subjective prior probability distribution.

In terms of behavior, given SEU, Bayesian updating is essentially equivalent to dynamic consistency of preferences, and some authors reverse-engineered updating under ambiguity from this property (Hanany and Klibanoff 2007). However, how people should and how they do update under ambiguity is still subject to much debate (Shishkin and Ortoleva 2023, see also this Handbook's Chapter 8 on "Contingent Expectations").

It should be noted that Bayesianism in the sense of behaviorally revealing well-defined subjective probabilities does not imply SEU, just as knowledge of outcome probabilities does not imply EU. Someone's behavior may define her subjective probabilities—separable from tastes—despite being inconsistent with EU; e.g., a prospect-theory agent may have well-defined subjective probabilities implicit in their decision weights. Machina and Schmeidler (1992) characterize such "probabilistic sophistication" behaviorally (in the purely subjective Savage framework), thus obtaining "a more robust definition of subjective probability." Of course, the Ellsberg paradox violates probabilistic sophistication, so it follows that there is indeed something to ambiguity beyond risk. Furthermore, probabilistic sophistication with Bayesian updating implies that dynamic consistency of preferences is essentially equivalent

to independence (recall the earlier illustration). In this stronger sense, SEU is indeed the behavioral definition of Bayesianism.

Bounded Rationality and Ambiguity Theory

When it comes to theory testing, to ensure separability between probability and utility under SEU, the practice in experimental economics is to use money and abstract states (e.g., colors of balls drawn from urns), as in Ellsberg's experiment above. For an extensive survey of such experimental evidence on ambiguity attitudes, see Trautmann and van de Kuilen (2015). In contrast to prospect theory for risk, there is no leading descriptive model for ambiguity. One notable reason is that the axiomatic approach in decision theory has modeled decision-makers with ambiguity by generalizing SEU such that for risk they boil down to EU, at odds with evidence from risk.

At the same time, capturing the key cognitive challenges that occur under *any* uncertainty separately from those that arise especially under ambiguity remains an important challenge in this literature (Wakker 2010). Here I mention only two recent works that relate to bounded rationality in a way that is relevant to both risk and ambiguity.

First, a recent study suggests that a major cognitive difficulty in decision-making under any kind of uncertainty concerns contingent thinking (Esponda and Vespa 2023). The authors frame several well-known decision problems—both the Allais and (single-urn) Ellsberg paradox, as well as second-price auctions and majority voting—either in the traditional (non-contingent) way or an alternative (contingent) way that essentially guides respondents through the relevant contingencies so that they become obvious (Li 2017). For instance, the choice between the “not red” and “not black” bets on the Ellsberg urn is described such that

one wins anyway with “yellow” and has to choose between winning with “black” but not “red” or winning with “red” but not “black.” They replicate earlier evidence in terms of mistakes and show that the contingent frame significantly helps decision makers to avoid them. For instance, violations of SEU on the Ellsberg urn halve from around 58% to around 28%. On the one hand, this idea is similar to and sits well with the aforementioned evidence that “compound independence” is more often satisfied than “mixture independence” under risk. On the other hand, it is more fundamental in that it extends to ambiguity and actually connects bounded rationality under risk and ambiguity. Moreover, the authors also cover strategic settings with a dominant strategy, which requires contingent thinking. (A dominant strategy dominates any alternative strategy in that it performs at least as well *in every possible contingency*.) Here, the contingent frame also reduces mistakes significantly, in line with another recent experiment’s finding that learning dominance in second-price auctions is greatly enhanced by dynamic framing (Breitmoser and Schweighofer-Kodritsch 2022).

In social behavior, recent co-authored work shows that the same person may reveal two contradictory beliefs, one held unconsciously and another formed consciously, depending on the complexity of the decision problem (Barron, Ditlmann, et al. 2022). We identify people betting on a female rather than a male candidate’s performance on a logic task whenever both candidates are equally qualified but reverse the “gender bet” whenever they are differently qualified. This is in line with an implicit/unconscious bias against women (stereotype) driving complex decisions where both qualification and gender differ, and which contradicts the same person’s explicit/conscious judgment about gender that is revealed when gender is the only difference between the candidates. “Belief” is thus not well-defined, which is relevant in settings when certain views are socially stigmatized, like “women are worse at logical reasoning than men,” but prevail (if only unconsciously). For a formal development of

implicit preferences, co-existing with opposing explicit preferences, which explains also related evidence on social behavior under risk (Exley 2016), see Cunningham and de Quidt (2023).

Psychology and Economics: The “Heuristics and Biases” Debate, and Some Recent Evidence

Psychology has for a long time studied human judgment, resulting in a huge body of evidence, most famously from Kahneman and Tversky (KT)’s “Heuristics and Biases” research program. KT’s research is nowadays well-known and influential among behavioral economists. It has been subject to intense debate within psychology, however, most notably in KT (1996) and Gigerenzer (1996), printed back-to-back in the *Psychological Review* in the very year of Tversky’s premature death. Both in terms of normative and methodological questions, this debate complements the decision theoretical discussion.

Kahneman-Tversky vs. Gigerenzer

Research in cognitive psychology on deviations from Bayesian reasoning with given probabilities began with the work of Ward Edwards as early as the 1950s. In abstract settings, he established a deviation known as conservatism bias, which has been found surprisingly robust in a recent meta-analysis (Benjamin 2019, Section 4). It means that the updated posterior is too close to the prior. Tversky, influenced as a PhD student by Edwards, gave a talk on this bias on the invitation of Kahneman, who found it at odds with the observation that people would often jump to conclusions. Their ensuing exchange marks the start of their “Heuristics and Biases” research in the late 1960s, investigating errors in human judgment more comprehensively and fundamentally.

KT went on to establish a wide range of judgmental biases and fallacies, which they explained as the consequence of a few basic heuristics. Most prominently, they propose the so-called “Representativeness Heuristic,” whereby in judging how likely it is that B is true when A is observed, “probabilities are evaluated by the degree to which A is representative of B, that is, by the degree to which A resembles B” (Tversky and Kahneman 1974, 1124). They argued that, e.g., the apparent opposite of conservatism bias, base-rate neglect—ignoring prior probabilities, as “jumping to conclusions”—would be a consequence of representativeness because judgments of resemblance ignore prior information.

For instance, consider the following famous cab problem: 85% of cabs in a city are blue and the remaining 15% are green, one was involved in a hit-and-run accident at night, and a witness identified this cab as green. Given that the witness, under night conditions, correctly identifies the color correctly 80% of the time, what are the chances that the errant cab was indeed green? Using Bayes’ Theorem, the correct answer is around 41%. (In the above, take E to be the event that it was green, so $p(E) = 0.15$, and F to be the event that the witness identified the car as green, so $p(F|E) = 0.8$.) The modal response, that given by most people, is 0.8, however, which completely neglects the “biased” prior, under which blue is much more likely ($p(E) < 0.5$).

Another equally famous example that KT also associate with representativeness is the so-called conjunction fallacy, demonstrated with the following problem: “Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.” Respondents are asked to judge which of the following is more probable:

- Linda is a bank teller.

- Linda is a bank teller and is active in the feminist movement.

A basic property of any probability is that $p(E \cap F) \leq p(E)$ for any two events E and F , so unless one believes all female bank tellers are active in the feminist movement, the answer must be the first, of which the second is a special case. However, consistent with the idea that the description is representative of a feminist, but not a bank teller, the vast majority of respondents in various studies indicated the opposite. This is taken to demonstrate that human intuitive (subjective) judgment need not obey even a basic law of probability.

Gigerenzer summarizes his critique of this finding as follows: “At issue is the imposition of unnecessarily narrow norms of sound reasoning that are used to diagnose so-called cognitive illusions and the continuing reliance on vague heuristics that explain everything and nothing. (...) The focus should be rather the construction of detailed models of cognitive processes that explain when and why they disappear” (Gigerenzer 1996, 592).

By unnecessarily narrow norms, Gigerenzer essentially refers to Bayesianism. He criticizes KT’s suggestion of human irrationality in judgment based on single-event problems like Linda, for which probability is not objectively defined. KT counter that the vast majority of biases in their 1974 *Science* article concern problems with objective probabilities, though there is apparent disagreement on this notion (Gigerenzer 1996, footnote 1). KT counter further that the “refusal to apply the concept of probability to unique events is a philosophical position that has some following among statisticians, but it is not generally shared by the public” (Kahneman and Tversky 1996, 585). Indeed, in everyday communication, we use the language of probability—or chances or odds—to express judgment/beliefs.

Here, Gigerenzer's second normative concern comes into play, namely that KT's Bayesian norms are "content-blind." Essentially, his argument concerns the role of framing and semantics in the presentation of a given abstract problem. One may read this also as KT's "tricking people" into mistakes. This is exemplified with the finding that frequency formats instead of probability formats (e.g., there are 200 cabs, 170 of which are blue and 30 of which are green, the witness would correctly identify 136 blue cars as blue and 24 green cars as green) strongly reduce biases in judgment.

This leads to Gigerenzer's third main criticism. He claims that KT's heuristics are so vague that they explain everything and hence nothing, and that progress will be achieved only with the development of models that are predictive of the extent of bias as a function of the problem's characteristics. KT countered that their proposed heuristics make predictions that can be experimentally studied, without any need to define them a priori.

I find this debate highly instructive, especially for behavioral economists. It highlights the difficulty of a normative debate regarding beliefs/judgment without connecting to observed behavior. In contrast, the celebration of Savage's result in decision theory stems from its *behavioral definition* of subjective probability. The normative disagreement between KT and Gigerenzer thus has to boil down to their respective judgments of the "ecological" relevance of the heuristics-induced biases that KT document (Gigerenzer and Todd 1999, 5). This should be familiar to experimental economists as concerns of external validity of laboratory experiments, which have been prominently raised against lab measures of social preferences and promoted the use of field experiments (Levitt and List 2007). Finally, and again related to this key point of disagreement, Gigerenzer's insistence on the development of precise models for explanation and prediction should resonate with any economist. It is the reason

why economists seek decision models with broad applicability in various richer economic environments (“markets”). In the face of practical constraints on data availability and collection, such applied theoretical analysis provides another route to external validity. Quite generally, this debate—including the fact that the models developed by Gigerenzer and his co-authors (Gigerenzer, Todd, and ABC Research Group 1999) are tailored toward judgment data and have not been useful for economists—serves as a useful reminder of the differences in methods and objectives between psychology and economics.

Selected Recent Evidence from Economics

The emphasis on behavior in economics has led to the methodological difference from psychology that experiments in economics generally employ monetary incentives. Sophisticated mechanisms have been developed to truthfully elicit a respondent’s probabilistic belief (Schlag, Tremewan, and van der Weele 2015). The theoretical concern here has been with incentives to distort one’s response away from the true belief due to non-EU risk preferences (e.g., prospect theory), and the leading incentive mechanism BSR (binarized scoring rule) achieves incentive compatibility also for these.

The use of incentives has resulted in qualifications regarding their incidence but by no means do the biases documented in psychology appear to be an artifact of a lack of incentives. This certainly includes all biases mentioned here (conservatism in updating, conjunction fallacy, base rate neglect). For a comprehensive survey I refer to Benjamin (2019).

This suggests that incentives appear to be unnecessary when noisiness is dealt with by a sufficiently large sample, which is good news for studies of beliefs that cannot be incentivized, such as macroeconomic expectations. In fact, recent evidence shows that

complex incentives may even backfire by introducing (additional) centrality bias, meaning reported beliefs are distorted towards 0.5 (Danz, Vesterlund, and Wilson 2022). The authors find that transparent quantitative information on incentives under the leading elicitation mechanism—known as the binarized scoring rule, BSR (Hossain and Okui 2013)—actually exacerbates centrality bias in participants’ responses, not only between but even within subjects (i.e., the same person responds “worse” under such transparency).

However, even when there is an objective probability like a Bayesian update, the psychology literature shows people typically do not know it, and recent evidence further suggests they know that they do not know, so from the respondent’s perspective, there is subjective “cognitive uncertainty” (Enke and Graeber 2023). The Ellsberg paradox already tells us that, when it comes to behavior, beliefs regarding subjective uncertainty are not probabilistic. The interpretation of elicited probabilities against an objective benchmark as behaviorally relevant bias (or also non-bias) anyway rests on strong assumptions that are probably violated for at least a significant share of respondents.

This suggests a cautious approach in the use and interpretation of probabilistic belief elicitation, as followed by the state-of-the-art in information provision experiments (Haaland, Roth, and Wohlfart 2023). Elicited beliefs are compared between different (information) treatments, thus between random samples of subjects, and further augmented with behavioral measures. Though I am not aware of any rigorous theoretical analysis, the assumption that differences in probabilistic responses between subjects tell us something relevant about differences in beliefs seems highly plausible, even when the latter are not probabilistic. Relatedly, in recent co-authored work, we exposed people to an objective lottery but randomly varied the time delay until uncertainty resolution, between subjects (Barron,

Bönisch, and Schweighofer-Kodritsch 2023). We separately elicited participants' beliefs regarding the event that they win (unincentivized) and their guess of the statistical chance of this event (incentivized). We found that longer resolution delay increases the reported subjective winning chances, in line with the leading explanation of optimism bias due to anticipatory utility, and these beliefs also correlate with information preferences. In contrast, participants' guesses of the statistical chance of winning are unaffected by resolution delay and unrelated to information preferences. The behaviorally relevant belief may therefore even consciously differ from an incentivized guess of the objective probability.

Another conclusion from this study is that anticipatory feelings matter, and accounting for these in terms of utility, belief bias can be rational. The same is true in the presence of other non-standard preferences. Indeed, in other recent work, we find evidence among students studying for an exam that they develop a belief bias to better achieve their study objectives, because it counteracts another well-documented and behavioral bias called present bias (Bönisch, König, et al. 2024). Study effort has immediate costs but delayed rewards, and present bias implies too little effort from the long-run perspective. Exploiting dynamic variation in the importance of study effort between two groups of students with exams at different dates, we find that students for whom self-control is more important due to the exam's imminence believe the returns to effort to be greater. Related work with field data also establishes that people deceive themselves in order to better deceive others (Schwardmann, Tripodi, and van der Weele 2022). In view of non-standard tastes or strategic interaction, belief bias can have behavioral/instrumental value, and human psychology appears able to deliver it.

Conclusion

Upon accepting the notion of probability from objective uncertainty, rationality implies Bayesian decision-making in the form of maximizing expected utility based on precise subjective probabilistic beliefs that are updated according to Bayes' Theorem. This standard model of decision-making is widely applicable and theoretically useful for economic analysis.

Evidence from psychology and experimental economics shows that this model is not an accurate description of how people think or act, however, due to their bounded rationality. In particular, subjective beliefs more often than not seem to differ from precise probabilities; even confronting objective uncertainty, many people act as if they did not accept the frequentist odds as measuring it. Moreover, elicited beliefs seem sensitive to framing effects, and researchers would be better off investing in larger sample sizes than monetary incentives. Practically any uncertainty is subjective, and framing belief elicitation is unavoidable. Interpreting bias on probabilistic belief measures against some absolute objective benchmark as indicative of irrationality, hence relies on strong assumptions. In view of other deviations from the standard model, belief bias may even be rational—whether by enhancing emotional well-being or serving instrumental purposes.

These issues more or less disappear when comparing identically measured beliefs between “comparable” groups; e.g., as part of understanding potential treatment effects in experiments, or when repeatedly studying beliefs within the same population over time in surveys. However, as a call to the profession, framing effects should be actively dealt with as part of a common methodology, by systematically studying them and agreeing on a standard approach to framing tasks/questions.

Establishing the ecological economic relevance of any particular psychological bias established in experimental research ultimately requires field evidence. The most fruitful approach towards this goal is then to model such biases as portable extensions of standard models and derive their distinct implications for “market” settings (Rabin 2013). This is only beginning to happen and should become a priority.

References

- Allais, Maurice. 1953. “Le comportement de l’homme rationnel devant le risque: Critique des postulats et axiomes de l’Ecole Americaine.” *Econometrica* 21, no. 4: 503–546. DOI: 10.2307/1907921.
- Anscombe, Francis J, and Robert J Aumann. 1963. “A Definition of Subjective Probability.” *The Annals of Mathematical Statistics* 34, no. 1: 199–205.
- Barron, Kai, Felix Bönisch, and Sebastian Schweighofer-Kodritsch. 2023. “How Anticipation Shapes Belief Formation and Information Preferences—Experimental Evidence.” Mimeo (in progress).
- Barron, Kai, Ruth Dittmann, Stefan Gehrig, and Sebastian Schweighofer-Kodritsch. 2022. “Explicit and Implicit Belief-Based Gender Discrimination: A Hiring Experiment.” Working Paper No. 9731, CESifo. SSRN: <https://ssrn.com/abstract=4097858>.
- Benjamin, Daniel J. 2019. “Errors in Probabilistic Reasoning and Judgment Biases.” In *Handbook of Behavioral Economics—Foundations and Applications* 2, edited by B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, 69–186. Amsterdam: North-Holland. DOI:10.1016/bs.hesbe.2018.11.002.
- Benjamin, Daniel J, Matthew Rabin, and Collin Raymond. 2016. “A Model of Nonbelief in the Law of Large Numbers.” *Journal of the European Economic Association* 14, no. 2: 515–544. DOI:10.1111/jeea.12139.
- Bönisch, Felix, Tobias König, Sebastian Schweighofer-Kodritsch, and Georg Weizsäcker. 2024. “Beliefs as a Means of Self-Control? Evidence from a Dynamic Student Survey.” Working Paper No. 10984, CESifo. SSRN: <https://ssrn.com/abstract=4756065>.
- Breitmoser, Yves, and Sebastian Schweighofer-Kodritsch. 2022. “Obviousness Around the Clock.” *Experimental Economics* 25, no. 2: 483–513. DOI: 10.1007/s10683-021-09720-z.
- Camerer, Colin F, George Loewenstein, and Matthew Rabin. 2004. *Advances in Behavioral Economics*. Russell Sage Foundation. New York: Princeton University Press.
- Conlisk, John. 1989. “Three Variants on the Allais Example.” *The American Economic Review* 79, no. 3: 392–407.
- Cunningham, Tom, and Jonathan de Quidt. 2023. “Implicit Preferences.” Working Paper (January 10). https://jondequidt.com/pdfs/paper_implicit.pdf.
- Danz, David, Lise Vesterlund, and Alistair J Wilson. 2022. “Belief Elicitation and Behavioral Incentive Compatibility.” *The American Economic Review* 112, no. 9: 2851–2883. DOI:10.1257/aer.20201248.
- de Finetti, Bruno. 1974. *Theory of Probability*. 2 volumes. New York: John Wiley & Sons.
- Dekel, Eddie, and Barton L Lipman. 2010. “How (Not) to Do Decision Theory.” *Annual Review of Economics* 2: 257–282. DOI:10.1146/annurev.economics.102308.124328.

- Diaconis, Persi, Susan Holmes, and Richard Montgomery. 2007. “Dynamical Bias in the Coin Toss.” *SIAM Review* 49, no. 2: 211–235. DOI: 10.1137/S0036144504446436.
- Drèze, Jacques H, and Aldo Rustichini. 2004. “State-Dependent Utility and Decision Theory.” In *Handbook of Utility Theory*, vol. 2, edited by Salvador Barberà, Peter J Hammond, and Christian Seidl, 839–892. New York: Springer Science+Business Media.
- Edwards, Ward. 1992. “Utility Theories: Measurements and Applications.” *Studies in Risk and Uncertainty*. New York: Springer Science+Business Media.
- Ellsberg, Daniel. 1961. “Risk, Ambiguity, and the Savage Axioms.” *The Quarterly Journal of Economics* 75, no. 4: 643–669. DOI: 10.2307/1884324.
- Enke, Benjamin, and Thomas Graeber. 2023. “Cognitive Uncertainty.” *The Quarterly Journal of Economics* 138, no. 4: 2021–2067. DOI:10.1093/qje/qjad025.
- Esponda, Ignacio, and Emanuel Vespa. Forthcoming. “Contingent Thinking and the Sure-Thing Principle: Revisiting Classic Anomalies in the Laboratory.” *The Review of Economic Studies*, forthcoming. DOI:10.1093/restud/rdad102.
- Exley, Christine L. 2016. “Excusing Selfishness in Charitable Giving: The Role of Risk.” *The Review of Economic Studies* 83, no. 2: 587–628. DOI:10.1093/restud/rdv051.
- Fishburn, Peter, and Peter Wakker. 1995. “The Invention of the Independence Condition.” *Management Science* 41, no. 7: 1130–1144. DOI:10.1287/mnsc.41.7.1130.
- Gabaix, Xavier. 2019. “Behavioral Inattention.” In *Handbook of Behavioral Economics: Applications and Foundations* 1, edited by B. Douglas Bernheim, Stefano DellaVigna, and David Laibson, 261–343. Amsterdam: North-Holland. DOI:10.1016/bs.hesbe.2018.11.001.
- Gigerenzer, Gerd. 1996. “On Narrow Norms and Vague Heuristics: A Reply to Kahneman and Tversky (1996).” *Psychological Review* 103, no. 3: 592–596. DOI:10.1037/0033-295X.103.3.592.
- Gigerenzer, Gerd, and Peter M Todd. 1999. “Fast and Frugal Heuristics: The Adaptive Toolbox.” In *Simple Heuristics That Make Us Smarter*, edited by Gerd Gigerenzer, Peter M. Todd, and The ABC Research Group, 3–34. New York: Oxford University Press.
- Gigerenzer, Gerd, and Reinhard Selten. 2002. *Bounded Rationality: The Adaptive Toolbox*. Cambridge MA: The MIT Press.
- Gigerenzer, Gerd, Peter M Todd, and the ABC Research Group. 1999. *Simple Heuristics That Make Us Smarter*. New York: Oxford University Press.
- Gilboa, Itzhak. 2015. “Rationality and the Bayesian Paradigm.” *Journal of Economic Methodology* 22, no. 3: 312–334. DOI:10.1080/1350178X.2015.1071505.
- Gilboa, Itzhak. 2009. *Theory of Decision under Uncertainty*. New York, NY: Cambridge University Press.
- Gilboa, Itzhak, and David Schmeidler. 1989. “Maxmin Expected Utility with Non-unique Prior.” *Journal of Mathematical Economics* 18, no. 2: 141–153. DOI: 10.1016/0304-4068(89)90018-9.

Gilboa, Itzhak, Stefania Minardi, Larry Samuelson, and David Schmeidler. 2020. “States and Contingencies: How to Understand Savage without Anyone Being Hanged.” *Revue économique* 71, no.2: 365–385.

Giocoli, Nicola. 2013. “From Wald to Savage: Homo Economicus Becomes a Bayesian Statistician.” *Journal of the History of Behavioral Sciences* 49, no. 1: 63–95. DOI:10.1002/jhbs.21579.

Gollier, Christian, and Alexander Muermann. 2010. “Optimal Choice and Beliefs with Ex Ante Savoring and Ex Post Disappointment.” *Management Science* 56, no. 8: 1272–1284. DOI:10.1287/mnsc.1100.1185.

Grether, David M. 1980. “Bayes Rule as a Descriptive Model: The Representativeness Heuristic.” *The Quarterly Journal of Economics* 95, no. 3: 537–557. DOI: 10.2307/1885092.

Gul, Faruk, and Wolfgang Pesendorfer. 2008. “The Case for Mindless Economics.” In *The Foundations of Positive and Normative Economics—A Handbook*, edited by Andrew Caplin and Andrew Schotter, 3–39. New York: Oxford University Press.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart. 2023. “Designing Information Provision Experiments.” *Journal of Economic Literature* 61, no. 1: 3–40. DOI:10.1257/jel.20211658.

Hanany, Eran, and Peter Klibanoff. 2007. “Updating Preferences with Multiple Priors.” *Theoretical Economics* 2, no. 3: 261–298.

Hossain, Tanjim, and Ryo Okui. 2013. “The Binarized Scoring Rule.” *The Review of Economic Studies* 80, no. 3: 984–1001. DOI:10.1093/restud/rdt006.

Kadane, Joseph B., and Robert L. Winkler. 1988. “Separating Probability Elicitation From Utilities.” *Journal of the American Statistical Association* 83, no. 402: 357–363.

Kahneman, Daniel. 2003. “Maps of Bounded Rationality: Psychology for Behavioral Economics.” *The American Economic Review* 93, no. 5: 1449–1475. DOI: 10.1257/000282803322655392.

Kahneman, Daniel, and Amos Tversky. 1996. “On the Reality of Cognitive Illusions.” *Psychological Review* 103, no. 3: 582–591. DOI:10.1037/0033-295X.103.3.582.

Kahneman, Daniel, and Amos Tversky. 1979. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica* 47, no. 2: 263–292. DOI: 10.2307/1914185.

Karni, Edi, and David Schmeidler. 1991. “Atemporal Dynamic Consistency and Expected Utility Theory.” *Journal of Economic Theory* 54, no. 2: 401–408. DOI: 10.1016/0022-0531(91)90130-V.

Kim, Jeongbin, Wooyoung Lim, and Sebastian Schweighofer-Kodritsch. 2022. “Patience Is Power: Bargaining and Payoff Delay.” Berlin School of Economics Discussion Papers 0015. DOI: 10.48462/opus4-4946.

Kreps, David M. 1988. *Notes on the Theory of Choice*. Boulder, CO: Westview Press.

- Levitt, Steven D, and John A List. 2007. “What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?” *Journal of Economic Perspectives* 21, no. 2: 153–174. DOI:10.1257/jep.21.2.153.
- Li, Shengwu. 2017. “Obviously Strategy-Proof Mechanisms.” *The American Economic Review* 107, no. 11: 3257–3287. DOI:10.1257/aer.20160425.
- Machina, Mark J, and David Schmeidler. 1992. “A More Robust Definition of Subjective Probability.” *Econometrica* 60, no. 4: 745–780. DOI: 10.2307/2951565.
- Manski, Charles F. 2004. “Measuring Expectations.” *Econometrica* 72, no. 5: 1329–1376. DOI:10.1111/j.1468-0262.2004.00537.x.
- Moscati, Ivan. 2016. “How Economists Came to Accept Expected Utility Theory: The Case of Samuelson and Savage.” *Journal of Economic Perspectives* 30, no. 2: 219–236. DOI:10.1257/jep.30.2.219.
- Nielsen, Kirby, and John Rehbeck. 2022. “When Choices Are Mistakes.” *The American Review* 112, no. 7: 2237–2268. DOI:10.1257/aer.20201550.
- O’Donoghue, Ted, and Jason Somerville. 2018. “Modeling Risk Aversion in Economics.” *Journal of Economic Perspectives* 32, no. 2: 91–114. DOI:10.1257/jep.32.2.91.
- Rabin, Matthew. 2013. “An Approach to Incorporating Psychology into Economics.” *The American Economic Review* 103, no. 3: 617–22. DOI:10.1257/aer.103.3.617.
- Savage, Leonard J. 1954. *The Foundations of Statistics*. New York: John Wiley & Sons.
- Savage, Leonard J. 1961. “The Foundations of Statistics Reconsidered.” *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*, edited by Jerzy Neyman, 575–586. Berkeley CA: University of California Press.
- Schlag, Karl H, James Tremewan, and Joël J. van der Weele. 2015. “A Penny for Your Thoughts: A Survey of Methods for Eliciting Beliefs.” *Experimental Economics* 18, no. 3: 457–490.
- Schmeidler, David. 1989. “Subjective Probability and Expected Utility without Additivity.” *Econometrica* 57, no. 3: 571–587. DOI:10.2307/1911053.
- Schwardmann, Peter, Egon Tripodi, and Joël van der Weele. 2022. “Self-persuasion: Evidence from International Debating Competitions.” *The American Economic Review* 112, no. 4: 1118–1146. DOI:10.1257/aer.20200372.
- Schweighofer-Kodritsch, Sebastian. 2022. “The Bargaining Trap.” *Games and Economic Behavior* 136: 249–254. DOI:10.1016/j.geb.2022.09.006.
- Schweighofer-Kodritsch, Sebastian. 2018. “Time Preferences and Bargaining.” *Econometrica* 86, no. 1: 173–217. DOI: 10.3982/ECTA14396.
- Segal, Uzi. 1990. “Two-Stage Lotteries without the Reduction Axiom.” *Econometrica* 58, no. 2: 349–377. DOI: 10.2307/2938207.
- Shishkin, Denis, and Pietro Ortoleva. 2023. “Ambiguous Information and Dilation: An Experiment.” *Journal of Economic Theory* 208: 105610. DOI:10.1016/j.jet.2023.105610.

Simon, Herbert A. 1955. “A Behavioral Model of Rational Choice.” *The Quarterly Journal of Economics* 69, no. 1: 99–118. DOI: 10.2307/1884852.

Trautmann, Stefan, and Gijs van de Kuilen. 2015. “Ambiguity Attitudes.” In *The Wiley Blackwell Handbook of Judgment and Decision Making*, edited by Gideon Keren and George Wu, 89–116. New York: John Wiley & Sons.

Tversky, Amos, and Daniel Kahneman. 1974. “Judgment under Uncertainty: Heuristics and Biases.” *Science* 185, no. 4157: 1124–1131. DOI:10.1126/science.185.4157.1124.

von Neumann, John, and Oskar Morgenstern. 1947. *Theory of Games and Economic Behavior*. 2nd edition. Princeton NJ: Princeton University Press.

Wakker, Peter P. 2010. *Prospect Theory: For Risk and Ambiguity*. Cambridge: Cambridge University Press.

Zappia, Carlo. 2021. “Leonard Savage, the Ellsberg Paradox and the Debate on Subjective Probabilities: Evidence from the Archives.” *Journal of the History of Economic Thought* 43, no. 2: 169–192. DOI: 10.1017/S1053837220000152.

¹ An important omission from this chapter is the literature on limited/costly attention, which may well be closest in spirit to Simon’s notion of bounded rationality, and also its interpretation by Gigerenzer and Selten (2002). For a recent survey of this literature see Gabaix (2019).

² There are some exceptions like surveys of professional forecasters (Manski 2004).

³ I retrieved citations data via the “Publish or Perish” software from google scholar on March 23, 2023, so it includes also citations outside of economics; the software is available at <https://harzing.com/resources/publish-or-perish>.

⁴ My initial work on bargaining theory maintained very strong assumptions on players’ beliefs, considering only non-standard time preferences (Schweighofer-Kodritsch (2018) or Schweighofer-Kodritsch (2022)), though I also experimentally tested some behavioral predictions in the laboratory under weaker such rationality assumptions (Kim, Lim, and Schweighofer-Kodritsch 2022).