



Savage for dummies and experts

Mohammed Abdellaoui^a, Peter P. Wakker^{b,*}

^a CNRS-GREGHEC & HEC-Paris, France

^b Erasmus School of Economics, Erasmus University Rotterdam, Rotterdam, the Netherlands

Received 22 February 2019; final version received 29 November 2019; accepted 7 January 2020

Available online 24 January 2020

Abstract

Savage's foundation of expected utility is considered to be the most convincing justification of Bayesian expected utility and the crowning glory of decision theory. It combines exceptionally appealing axioms with deep mathematics. Despite the wide influence and deep respect that Savage received in economics and statistics, virtually no one touched his mathematical tools. We provide an updated analysis that is more general and more accessible. Our derivations are self-contained. This helps to better appreciate the depth and beauty of Savage's work and the foundations of Bayesianism, to teach it more easily, and to develop non-Bayesian generalizations incorporating ambiguity more efficiently.

© 2020 Elsevier Inc. All rights reserved.

JEL classification: C02; C60

Keywords: Subjective expected utility; Behavioral foundation; Mixture spaces; Foundations of statistics

1. Introduction

More than 60 years after the publication of *Foundations of Statistics* (Savage, 1954), its subjective expected utility derivation remains the crowning glory of decision theory (Kreps, 1988 p. 120). Combining ideas of de Finetti (1937) and von Neumann and Morgenstern (1947) (vNM), Savage gave the first complete revealed preference axiomatization of Bayesian expected utility. He provided exceptionally intuitive and elegant axioms. At the same time his axiomati-

* Corresponding author.

E-mail addresses: abdellaoui@hec.fr (M. Abdellaoui), Wakker@ese.eur.nl (P.P. Wakker).

zation was mathematically deep. His construction has had a profound impact in many fields, being widely accepted, and sometimes criticized (Allais, 1953; Cerreia-Vioglio et al., 2013; Ellsberg, 1961) as the normative foundation of decision under uncertainty in economics and Bayesian statistics. It has received deep respect.

Besides intuitive preference axioms, Savage also used nonnecessary richness axioms, requiring a σ -algebra of events that constitute a continuum. Such axioms are complex and technical. Up to now, researchers used Savage's theorem as a black box. Apart from a few exceptions, discussed later, researchers did not look into Savage's internal mechanism and did not alter or generalize it. Instead, Anscombe and Aumann's (1963) framework with its rich and even linear outcome space is most commonly used in economics today to derive generalizations that reckon with ambiguity (surveyed by Gilboa and Marinacci, 2016). Our paper simplifies Savage's mechanism, using rich state spaces that are naturally available in many applications, so that they become accessible and suitable for generalizations.

The main inefficiency in Savage's derivation comes from his choice to go by vNM's, in itself appealing, mixture-based derivation of expected utility. To make this possible, Savage had to transform his domain of uncertainty (no probabilities available) into vNM's domain of risk where all probabilities are available. Fishburn (1970) gave a clear and well-organized presentation of Savage's proof.¹ Savage himself left many details to the readers. This route through vNM was followed in all later analyses of Savage's theorem that we know of, including Arrow (1971), Chateauneuf et al. (2006), Gilboa and Marinacci (2016 §2.10), Kopylov (2007), Kreps (1988), and Machina and Schmeidler (1992). As we show, mainly through Fig. 11, the route through vNM is roundabout. It is simpler to directly derive expected utility for uncertainty, which can be done in an elementary manner.²

Savage's detour through vNM not only complicated his proof but also led to a loss of generality. vNM assumed all probabilities available, and Savage had to assume a corresponding richness. Hence, he imposed a restrictive continuity condition P6, requiring a continuum of events and precluding finite state spaces. Our proof does not need such richness but only needs solvability and Archimedean axioms. Those axioms are more general and their empirical meaning is also clearer. Hence, we find them more appealing than P6, although some readers may disagree. Many authors discussed the problematic nature of continuity assumptions in preference axiomatizations (Fuhrken and Richter, 1991 p. 94; Ghirardato and Marinacci, 2001; Halpern, 1999; Khan and Uyan, 2018; Krantz et al., 1971 §9.10; Luce et al., 1990 p. 49; Pfanzagl, 1968 §6.6 and §9.5).

Because we do not need mixtures, we can also simplify the construction of subjective probabilities in the first stage³ that we base on the more flexible technique of Hölder's (1901) lemma. Unlike mixture approaches, Hölder's (1901) lemma does not need a multiplication operation or a continuum domain. Thus, we do not need to incorporate limits in σ -algebras and we can therefore weaken Savage's P6 axiom. We generalize his result in a structural sense, that is, by allowing for more general structures, including discrete cases without convex-rangedness of probability. We also generalize Savage's result in a logical sense, that is, by deriving his theorem as a corollary

¹ He specified the route through vNM in his Theorem 14.3 and related proofs.

² In particular, we do not need to derive probabilistic sophistication (defined and characterized by Machina and Schmeidler, 1992) as an intermediate step. It follows from our expected utility representation in one blow.

³ Savage (1954 p. 39, para preceding P6) explained that later requirements in his analysis complicated the derivation of subjective probabilities, leading to P6 there: "but in Chapter 5 a slightly stronger assumption will be needed that bears on acts generally, not only on those very special acts by which probability is defined."

of ours (Proposition 4). Our approach follows Savage in using richness of states, but does so in a more tractable and more general manner.

Savage's (1954) theorem requires a commitment to finite additivity, and to abandoning countable additivity, in agreement with de Finetti's views. Alternative derivations required a commitment to countable additivity (Arrow, 1971) and, thus, to abandoning (strict) finite additivity. Our main theorem gives general finite additivity.⁴ Proposition 5 specifies the additional condition that is necessary and sufficient for countable additivity. Thus, our approach can be implemented with either finite or countable additivity, and neither is excluded.

Whereas in most papers proofs are only to be read by specialists, we hope that our proof will be read and enjoyed by many readers. It is complete, much simpler than preceding ones, and more didactical. It can be understood by nonspecialists and readily be used for teaching purposes. Following Cozic and Hill's (2015 §7) principle of constructive proofs, it shows more clearly than before how expected utility is constructed from preferences. This facilitates normative defenses—and criticisms. For specialists, our result provides a useful starting point for developing non-Bayesian generalizations of Savage's model that can, for instance, incorporate ambiguity (surveyed by Karni et al., 2014), robustness, non-additive belief functions, nonexpected utility, and imprecise probabilities (Walley, 1991).

This paper is organized as follows. Section 2 gives basic definitions and Savage's axioms. Section 3 presents our new axioms, our main result, and shows that our result is logically and structurally more general than Savage's. Because one of our aims is to deliver an appealing proof, we present it in the main text (§4). We also show that Hölder's lemma provides a powerful technique to obtain preference axiomatizations, which is an alternative to mixture techniques as used in the Anscombe–Aumann (1963) framework. This provides an additional reason for presenting the proof in the main text. Section 5 discusses papers that used Savage's mechanisms, and §6 concludes. The appendix provides further details and shows that Savage's axioms imply ours.

2. Basic definitions and Savage's axioms

We begin by presenting the basic definitions and preference conditions of Savage (1954). For alternative ways to model uncertainty, see Battigalli et al. (2017) and Marinacci (2015 §2.2). S denotes a set of *states of nature*. Exactly one state is true, but it is unknown which one. States can describe tomorrow's weather conditions, the performance of a stock in a year from now, and so on. Whereas Savage required S to be infinite, our approach includes some cases of finite state spaces. \mathcal{E} denotes an *algebra* of subsets of S called *events*. That is, \mathcal{E} contains \emptyset and S and is closed under complement taking and finite unions and intersections. Savage's (1954) main text assumes that \mathcal{E} contains all subsets of S . Savage (1954 p. 43) pointed out that his analysis and results remain valid if \mathcal{E} is a σ -*algebra*. That is, \mathcal{E} is also closed under countable unions and intersections. For a long time it was an open question whether Savage's analysis remains valid on algebras of events. Kopylov (2007) gave an affirmative answer. Our analysis provides a simpler derivation for algebras.

X denotes a set of *consequences* and can be finite or infinite. Consequences are general and can be monetary or anything else. Acts map states to consequences. We avoid measure-theoretic complications and assume throughout that acts are simple, taking only finitely many conse-

⁴ Necessary and sufficient richness conditions are in Observation 15. These do not preclude countable additivity.

quences. The novelty of this paper concerns the derivation of expected utility for simple acts and, hence, we focus on these. Acts are denoted f, g, h ; $(E_1 : x_1, \dots, E_n : x_n)$ denotes an act assigning x_i to all $s \in E_i$. We assume throughout that the E_i s partition S and are contained in \mathcal{E} . Thus, acts are measurable finite-valued mappings from states to consequences. Consequences are identified with constant acts. By $\alpha_E f$ we denote the act resulting from f if all consequences for event E are replaced by α . Thus, for consequences α, β , $\alpha_E \beta$ denotes an act assigning α to E and β to E^c . Similarly, $\alpha_A \beta_B f$ yields α under A , β under B , and is identical to f otherwise; here A and B are assumed disjoint.

The preference relation \succsim is a binary relation on the acts; \succ denotes the asymmetric part of \succsim , \sim denotes the symmetric part, and \preccurlyeq and \prec denote reversed preferences. Preference symbols also designate the preference relations over consequences induced by constant acts. Thus, $\alpha \succsim \beta$ denotes both a preference between consequences and a preference between constant acts.

In the subjective expected utility model, we assume that there exists a probability measure P on \mathcal{E} . That is, $P : \mathcal{E} \rightarrow [0, 1]$ satisfies $P(A \cup B) = P(A) + P(B)$ whenever A and B are disjoint events ((finite) additivity) and $P(S) = 1$. Additivity implies $P(\emptyset) = 0$. P is countably additive if $P(\cup_{j=1}^{\infty} E_j) = \sum_{j=1}^{\infty} P(E_j)$ for countably many disjoint E_j . Following Savage (1954), we do not require countable additivity. Finite additivity can run into paradoxes (Kadane et al., 1996) but provides more flexibility.

For subjective expected utility we further assume a function $U : X \rightarrow \mathbb{R}$, called utility function, which throughout is assumed not to be constant. The subjective expected utility (SEU) of an act $(E_1 : x_1, \dots, E_n : x_n)$ is $\sum_{j=1}^n P(E_j)U(x_j)$. Subjective expected utility (SEU) holds if there exist P, U such that $f \succsim g$ if and only if $SEU(f) \geq SEU(g)$.

Event E is nonnull if $(E : \gamma, E^c : \beta) \succ \beta$ for some consequences γ (“good”) and β (“bad”) and it is null otherwise. Under SEU, event E is null if and only if $P(E) = 0$. P is convex-ranged if for each event A and $0 \leq \lambda \leq P(A)$ there exists an event $B \subset A$ with $P(B) = \lambda$. Under countable additivity, convex-rangedness is equivalent to atomlessness, but we consider finite additivity where convex-rangedness is stronger.

The following five intuitive preference axioms are all necessary for SEU representations. They were introduced by Savage (1954) and will also be used in our paper.

P1 [weak ordering]

\succsim is complete ($f \succsim g$ or $g \succsim f$ for all f, g) and transitive.

P2 [sure-thing principle]

$\alpha_E f \succsim \alpha_E g \Leftrightarrow \beta_E f \succsim \beta_E g$ for all acts $\alpha_E f, \alpha_E g, \beta_E f$, and $\beta_E g$.

P3 [monotonicity]

Whenever E is nonnull: $\gamma \succ \beta \Leftrightarrow \gamma_E f \succ \beta_E f$.

P4 [independence of beliefs from tastes]

Whenever $\gamma \succ \beta$ and $\gamma' \succ \beta'$: $\gamma_E \beta \succ \gamma_B \beta \Leftrightarrow \gamma'_E \beta' \succ \gamma'_B \beta'$.

By P4, we can define a more likely than relation, also denoted \succsim , on events: $A \succsim B$ if there exist consequences $\gamma \succ \beta$ such that $\gamma_A \beta \succ \gamma_B \beta$. Because of P4, this relation is independent of the particular consequences γ and β . Because of P1 and P4, it is a weak order. Under SEU, $A \succsim B$ if and only if $P(A) \geq P(B)$.

P5 [nontriviality]

$\gamma \succ \beta$ for some consequences γ, β .

Savage further used the following technical axiom.

P6 [event continuity]

If $f \succ g$, then for all consequences γ, β there exists a partition $\{E_1, \dots, E_n\}$ of S such that for all $i : \beta_{E_i} f \succ g$ and $f \succ \gamma_{E_i} g$.

Savage also used an axiom P7 that exclusively concerns nonsimple acts. We focus on simple acts because this is where our novelty lies. Hence, we will not consider P7. Savage showed:

Theorem 1. [Savage’s theorem] Assume event continuity (P6), with \mathcal{E} a σ -algebra. Then the following two statements are equivalent:

- (i) Subjective expected utility holds with probability convex-ranged.
- (ii) P1-P5 hold. \square

3. Savage’s subjective expected utility simplified and generalized

We only use the following two implications of Savage’s P6, derived in Appendix B. The first is necessary for the very existence of subjective probability because infinitely many disjoint events with the same positive probability cannot exist.

P6.1 [Archimedeanity]

No infinite sequence of disjoint equally likely nonnull events exists.

The next condition, the second implication of Savage’s P6, is our only nonnecessary condition, introducing richness in a less demanding manner than Savage did. Imagine that a first act is preferred less than a second. And imagine that we can improve a consequence of the first act on an event B that is so big that the act becomes better than the second. Then there exists an “equalizing” subpart E of B such that improving only on E gives indifference.

P6.2 [(event) solvability].

For all acts f, g , consequences $\gamma \succ \beta$, and events B , $[\beta_B f \prec g \prec \gamma_B f \Rightarrow g \sim \gamma_E \beta_{B-E} f]$ for some $\{E, B - E\}$ partitioning B . See Fig. 1.

Solvability is reminiscent of the intermediate value property of continuous real-valued functions. Without richness conditions such as solvability or Savage’s P6, axiomatizations become

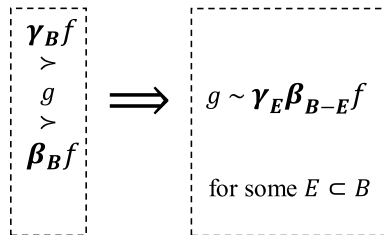


Fig. 1. Solvability.

considerably more complex and may even need infinite sequences of cancellation axioms (Alon and Lehrer, 2014). We now present the main result of our paper.

Theorem 2. [Main theorem] Assume solvability (P6.2), with \mathcal{E} an algebra. Then the following two statements are equivalent:

- (i) Subjective expected utility holds.
- (ii) P1-P5 and P6.1 hold.

In Statement (i), the probability measure is unique and utility is unique up to level and unit. \square

Uniqueness up to level and unit means that any real number can be added to utility and it can be multiplied by any positive real number.

Savage's axioms imply solvability, the only nonnecessary condition in our theorem. Therefore, a preference relation satisfying Savage's axioms also satisfies ours. In this sense, our model is structurally more general than Savage's. We do not only avoid mixtures and the vNM expected utility derivation so as to simplify the proof, but we also achieve bigger generality. This is illustrated by the following two examples, where the second is from Abdellaoui and Wakker (2005, Example 5.4.(ii)). Our axioms are satisfied in the examples, but Savage's are not. The examples have no continuums of probabilities so that the vNM technique to derive expected utility cannot be used.

Example 3.

- (i) [Comparative probability] The case of two (indifference classes of) consequences (say $\gamma \succ \beta$) is of special interest, because Theorem 2 then contributes to comparative probability theory. Here, every act can then be related one-to-one with the event of receiving the good consequence, similarly as indicator functions can be. Obtaining an expected utility representation is then equivalent to obtaining a probability measure representing the more-likely-than relation. Such representations are the topic of comparative probability theory (Fishburn, 1986). All our axioms are satisfied for a finite state space $S = \{s_1, \dots, s_n\}$ where \mathcal{E} contains all subsets of S , $U(\gamma) = 1$, $U(\beta) = 0$, and $P(s_j) = 1/n$ for all j . Savage's P6 and its implied atomlessness are violated. Extensions to finite state spaces with more than two equivalence classes of consequences are discussed in Appendix A.
- (ii) [Only rational-number probabilities] $S = [0, 1) \cap \mathbb{Q}$. \mathcal{E} contains all finite unions of intervals $[a, b)$ with a and b rational and the interval taken as a subset of S . P is the Lebesgue measure, determined by $P[a, b) = b - a$ for all $[a, b)$. $X = \mathbb{Q}$ and U is the identity function. All our axioms are satisfied, but Savage's assumption of a σ -algebra of events is violated as is his implied convex-rangedness of P . \square

Theorem 2 is also logically stronger than Savage's result in the sense that his result is a corollary of ours. The following proposition is proved in Appendix B.

Proposition 4. Savage's assumptions and Statement (ii) in Theorem 1 imply our assumptions and Statement (ii) in Theorem 2. \square

Whether one finds our axioms P6.1 and P6.2 or Savage's P6 more appealing is a matter of taste. Our analysis can equally well be used by readers who prefer Savage's axioms, because his axioms imply ours and his theorem is a corollary of ours, as shown in Proposition 4. We chose our axioms to maximally satisfy Cozic and Hill's (2015 §7) principle of constructive proofs, and maximally escape from Steven's (1968 p. 854) criticism of axiomatizations that are not related to empirical measurements. Our axioms clarify more directly how the theoretical constructs of utility and probability are related to the empirical primitive of preference (see, e.g., Figs. 4 and 7). Cozic and Hill argued that this is a pro for preference axiomatizations serving to evaluate decision models.

Savage (1954) only delivered finite additivity of probability and did not axiomatize countable additivity. His theorem assumes the powerset of events, and then countable additivity is not possible. In our Theorem 2 countable additivity is optional. We characterize it by the following well-known condition (Arrow, 1971; Fishburn, 1982, F7 in §10.3; Villegas, 1964, Theorem 4.2; Wakker, 1993 §4.1). *Set continuity* holds if, for all $\gamma > \beta$, nested decreasing sequences (A_j) with $(A_j) \downarrow \emptyset$, and $f > \beta$, there exists A_i such that $f > \gamma_{A_i}\beta$. The following result readily follows from substitution and from the references given.

Proposition 5. *In Theorem 2, P is countably additive if and only if set continuity holds.* \square

Observation 15 in Appendix A shows that range solvability of P , a condition defined there, is necessary and sufficient for solvability. This Observation thus provides an if-and-only-if version of Theorem 2. We present Theorem 2 in the main text because it is more accessible.

4. Proof of Theorem 2

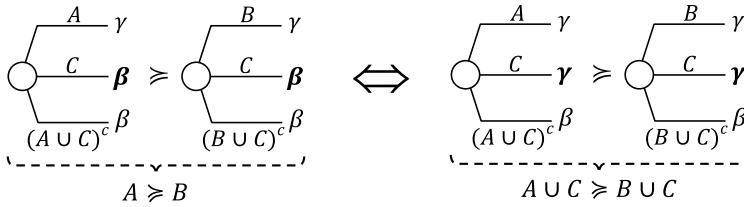
This section presents the proof of Theorem 2, where some details are deferred to Appendix A. All construction and approximation techniques used in our proofs are as in Hölder's lemma⁵ and its variations, explained in Krantz et al. (1971, Chs. 1-3) and Michell (1990 pp. 47-59). They only involve an addition operation and no scalar multiplication as in mixtures, making the results and analysis more accessible intuitively and more general mathematically.

Necessity of the conditions in Statement (i) of Theorem 2 readily follows from substitution and is left to the readers. The following two subsections prove sufficiency.

4.1. Subjective probabilities in the proof of Theorem 2

An *atom* is an event $A > \emptyset$ such that no event $B \subset A$ exists with $A > B > \emptyset$. Our theorem allows for some cases with atoms; see Appendix A. In the main text (§§4-5) we henceforth assume that no atoms exist. This is implied by Savage's (1954) axioms and hence this subsection is of most interest to readers primarily interested in his model. We start with some standard results for the more likely than relation. In figures, we denote acts using event nodes (circles). For instance, the left act in Fig. 2 depicts $(A : \gamma, C : \beta, (A \cup C)^c : \beta)$.

⁵ Although our proof will be complete and self-contained, we give the lemma here for completeness. A group with addition operation \circ , and with a binary relation \succcurlyeq , is embeddable in $(\mathbb{R}, +, \geq)$ if and only if: $a \succcurlyeq b \Leftrightarrow a \circ c \succcurlyeq b \circ c$ and, further, \circ is commutative, \succcurlyeq is an order, and for all a, b preferred to the identity element, there exists n such that $na \succcurlyeq b$. Moscati (2018) discusses the history of this lemma.



$\gamma > \beta$. By P2, changing the bold common consequence does not affect preference.

Fig. 2. Proof of (i).

Lemma 6. Assume P1-P5.

- (i) [common parts are immaterial in more-likely-than relations] $(A \cap C = B \cap C = \emptyset) \Rightarrow (A \succ B \Leftrightarrow A \cup C \succ B \cup C)$;
- (ii) [additivity of more-likely-than] $(A_1 \succ B_1, A_2 \succ B_2, A_1 \cap A_2 = B_1 \cap B_2 = \emptyset) \Rightarrow A_1 \cup A_2 \succ B_1 \cup B_2$, where the last preference is strict if one of the former two is;
- (iii) [additivity of equally likely] $(A_1 \sim B_1, A_2 \sim B_2, \text{ and } A_1 \cap A_2 = B_1 \cap B_2 = \emptyset) \Rightarrow A_1 \cup A_2 \sim B_1 \cup B_2$;
- (iv) [equally likely unions] $(A_1, \dots, A_n \text{ are disjoint, so are } B_1, \dots, B_n, \text{ and } A_j \sim B_j \text{ for all } j) \Rightarrow A_1 \cup \dots \cup A_n \sim B_1 \cup \dots \cup B_n$.

Proof.

- (i) See Fig. 2.
- (ii) Understanding the following visual proof takes little mental effort, unlike preceding proofs in the literature (Arrow, 1971; Fishburn, 1970). Define $A_0 := (A_1 \cup A_2)^c$ and $B_0 := (B_1 \cup B_2)^c$. Each square in Fig. 3 presents the state space S , being a union of nine rectangles $A_i \cap B_j$ (some may be empty). In each square, the union of rectangles with + is more likely than with -. Thus, the left two squares depict the two relations in the premise of (ii). By (i), common parts can be removed and added (indicated by bold and green \pm), giving the upper two implications “ \Rightarrow ”. Because the - event in the right upper square equals the + event in the square below, transitivity gives the lower left square. (i) finally gives the lower right square.
- (iii) follows from applying (ii) both with \succ and with \preceq .
- (iv) follows from repeated application of (iii). \square

To construct subjective probability, we use a (small) nonnull event A as a measuring rod. We recall that we assume no atoms in this part of the proof. By solvability, for all events $C \subset A^c$ with $C \succ A$ (i.e., $\gamma_C \beta \succ \gamma_A \beta$ for $\gamma > \beta$) we can find $B \subset C$ with $B \sim A$ (i.e., $\gamma_B \beta \sim \gamma_A \beta$). Applying this repeatedly, we obtain an almost uniform partition as in Fig. 4 and the following lemma. Because Savage’s almost uniform partitions are slightly more general than ours, we use a different formal term: a partition $\{A_1, \dots, A_n, R\}$ of S is close to uniform (CUP) if $A_1 \sim A_j$ for all j and $R \prec A_1$.

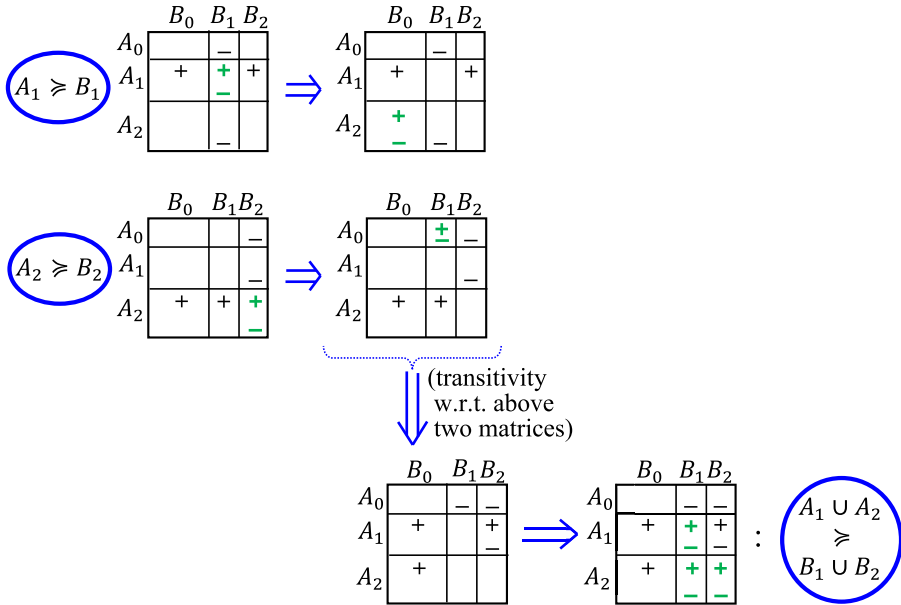


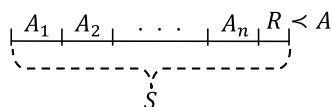
Fig. 3. Proof of (ii).

Lemma 7. Under P1-P6.2, for each event $A \succ \emptyset$ there is a CUP $\{A_1, \dots, A_n, R\}$ of S with $A \sim A_j$ for all j and $R \prec A$.

Proof. Inductively, define A_1, \dots, A_j . As long as the rest event $(A_1 \cup \dots \cup A_j)^c \succcurlyeq A$ we can by solvability define A_{j+1} as a subset of the rest event. By Archimedeanity, the process must stop after a finite number of steps. At the end, we must have $R \prec A$, and possibly $R = \emptyset$. \square

Uniform partitions result if there is no rest event, i.e., $R = \emptyset$. They need not exist in our approach, unlike in Savage's (1954). For every nonnull event E and CUP, the CUP suggests an upper and lower bound of the probability $P(E)$ by sandwiching E ; i.e., $A_1 \cup \dots \cup A_k \prec E \preccurlyeq A_1 \cup \dots \cup A_{k+1}$ (Fig. 5). As we will show, $P(E)$ then results as a limit, because we can take CUPs as refined as we want. As a corollary of our main theorem it will follow that we could have taken tighter bounds in Fig. 5, being $\frac{k}{n+1}$ and $\frac{k+1}{n}$, and we recommend those for empirical applications. However, at this stage the proof is easier with more conservative bounds.

Lemma 8. For each m , no matter how large, we can find a CUP with at least m elements.



$A_i \sim A$ for all i

Fig. 4. A close to uniform partition.

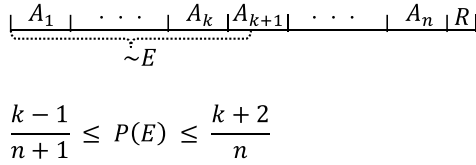


Fig. 5. Using a CUP to sandwich $P(E)$.

Proof. Take an arbitrary CUP, say the one of Lemma 7. We obtain one with at least $2n$ elements, as follows. Because there are no atoms, there exists $A' \subset A$ with $\emptyset \prec A' \prec A$. Define B as the \succcurlyeq smallest of A' and $A - A'$. By solvability, for each A_j we can find $B_j \subset A_j$ with $B_j \sim B$. By Lemma 6(ii), $A_j - B_j \sim A - B \succcurlyeq B \sim B_j$ so that, by solvability, each A_j contains at least two disjoint events indifferent to B . This delivers $2n$ disjoint equally likely events, which can be extended to a CUP with at least $2n$ elements, all except the rest event equally likely as B . \square

We next show that for all sufficiently refined CUPs we obtain upper and lower bounds that are consistent (no lower bound exceeds any upper bound) and, because their distance converges to 0, we obtain the following result. This result, and Lemma 13 presented later, are the only two places in the proof that require some nontrivial thinking, including the consideration of limits.

Observation 9. Under P1-P6.2, for each event E there exists a unique limit $P(E)$ of $k/(n+1)$ (see Fig. 5) over all CUPs as $n \rightarrow \infty$. \square

Proof. This proof is the only part where our analysis is more difficult than in mixture-approaches that have convex-rangedness. Under convex-rangedness, we can exactly double the refinement of partitions, and then the result below is straightforward. In Fig. 5, we can formally define $A_{n+1} = R$ for large events E that exceed $A_1 \cup \dots \cup A_n$, giving upper bound $(n+2)/n$. The following lemma shows consistency of the bounds.

Lemma 10. For each CUP $\{B_1, \dots, B_N, R'\}$ with N sufficiently large, the bounds $(\ell - 1)/(N + 1)$ and $(\ell + 2)/N$ are tighter than for the CUP above:

$$(k - 1)/(n + 1) < (\ell - 1)/(N + 1) < (\ell + 2)/N < (k + 2)/n. \tag{1}$$

The lemma is proved in Appendix A. The upper and lower bounds converge to each other and, hence, to one limit $P(E)$. \square

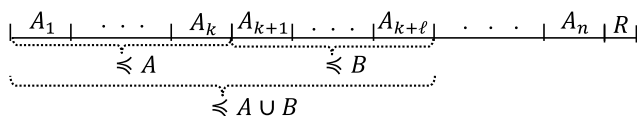
It is obvious that $P(\emptyset) = 0$ and $P(S) = 1$. We next show additivity of P .

Lemma 11. Under P1-P6.2, P is additive.

Proof. Consider disjoint A, B . The lower bound for $A \cup B$ (Lemma 10) in Fig. 6 is between $\frac{k+\ell-2}{n+1}$ and $\frac{k+\ell}{n+1}$. Its limit $P(A \cup B)$ is $P(A) + P(B)$. \square

Lemma 12. Under P1-P6.2, P represents \succcurlyeq on events.

Proof. If $A \succcurlyeq B$ then all estimations of $P(A)$ exceed the corresponding ones of $P(B)$. Hence, $P(A) \geq P(B)$. It implies $A \sim B \Rightarrow P(A) = P(B)$. We next assume $A \succ B$, and show $P(A) >$



$$\frac{k-1}{n+1} \leq P(A); \frac{\ell-1}{n+1} \leq P(B); \frac{k+\ell-2}{n+1} \leq P(A \cup B)$$

Fig. 6. Additivity of (lower bounds of) probability.

$P(B)$. By solvability, we get $B' \subset A$ with $B' \sim B$, so $P(B') = P(B)$. $A - B'$ is nonnull. Because of additivity of P , it suffices to show $P(E) > 0$ for any nonnull E . By Archimedeanity, the CUP generated by E (all elements except the rest event equivalent to E) is finite, say $\{E_1, \dots, E_n, R\}$. $P(E_1) = \dots = P(E_n) \geq P(R)$ and $P(S) = 1$ implies $P(E) \geq 1/(n + 1) > 0$. \square

The derivation of subjective probability is complete. Unlike preceding proofs in the literature (cf. Wakker, 1981), we did not derive convex-rangedness of probability because we do not need it in the following subsection. It indeed need not hold. We accordingly needed no σ -algebra structure. This is our main simplification in the derivation of subjective probability relative to preceding works.

4.2. Subjective utilities and expected utility in the proof of Theorem 2

The most appealing proof of the vNM expected utility theorem for risk is based on a substitution version of their independence axiom (Luce and Raiffa, 1957 pp. 27-28). In this proof, for each lottery (act in our case) the outcomes x_j are replaced by equivalent “standard gambles” (defined later) that all involve the same two outcomes $M \succ m$. After that, a lottery (act) results with only M and m , after which dominance trivially determines preference. This dominance relation is then shown to give the expected utility formula. This subsection shows that this proof technique can directly be used for uncertainty.

Throughout this subsection we assume P1-P6.2. We also continue to assume that no atom exists. We first fix two consequences $M \succ m$, and focus on consequences α in between ($m \preceq \alpha \preceq M$). We normalize $U(m) = 0$ and $U(M) = 1$. Fig. 7 shows how we obtain utility $U(\alpha)$. In empirical measurements of utility this method, using objective probabilities instead of events, is often employed and is called the standard gamble method (Gold et al., 1996; van Osch et al., 2004). We adapt it to uncertainty here. For each $\alpha \in [0, 1]$ we can, by solvability, find an event S^α giving the indifference in the figure. We define

$$U(\alpha) = P(S^\alpha). \tag{2}$$

We call any act as in the figure a *standard gamble* of α . Because of indifference, $U(\alpha)$ is independent of the particular S^α chosen.

We can also measure utility conditional on any nonnull event A instead of S , as in Fig. 8, where $A^\alpha \subset A$ exists by solvability. Here the act conditional on the complementary event A^c , f , is the same for both acts. Any right act in Fig. 8, conditional on A , is called a *conditional standard gamble* of α . The next lemma shows that we can obtain utility from conditional standard gambles and that it is consistent with the unconditional measurement. In other words, the lemma shows that expected utility adopts the same exchange rate between probability and utility everywhere.

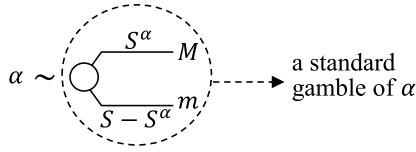


Fig. 7. Standard gamble measurement of U .

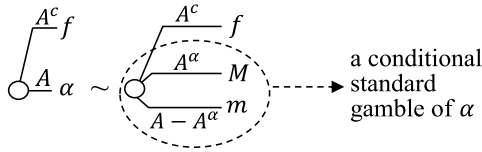


Fig. 8. Conditional measurement of U .

Lemma 13. For events A and A^α with $P(A) > 0$, as in Fig. 8,

$$P(A^\alpha)/P(A) = U(\alpha) (= P(S^\alpha)). \tag{3}$$

Proof. If $\alpha \sim m$ then $P(A^\alpha) = 0$ and Eq. (3) follows. Next assume $\alpha \succ m$. We often use monotonicity of A^α in the sense that $A \succcurlyeq B \Leftrightarrow A^\alpha \succcurlyeq B^\alpha$.

We first show that Eq. (3) approximately holds for an event A from a CUP $\{A_1, \dots, A_n, R\}$ with $A = A_1 \sim \dots \sim A_n \succ R$. See Fig. 9. The indifference between the left and right act in Fig. 9 is of the same nature as in Fig. 7, implying $\sum_{j=1}^n P(A_j^\alpha) + P(R^\alpha) = U(\alpha)$. In the reasoning that follows, it is instructive to first assume $R = \emptyset$.

We denote $[Ak] = \cup_{i=1}^k A_i$. All A_j are equally likely and have the same $P(A_j^\alpha)/P(A_j) = P(A^\alpha)/P(A)$. We have

$$U(\alpha) = \frac{\sum_{j=1}^n P(A_j^\alpha) + P(R^\alpha)}{\sum_{j=1}^n P(A_j) + P(R)} = \frac{nP(A^\alpha) + P(R^\alpha)}{nP(A) + P(R)}. \tag{4}$$

Because R is small and becomes negligible ($0 \leq P(R^\alpha) \leq P(R) < P(A) \leq 1/n$), it follows that $P(A^\alpha)/P(A)$ must tend to $U(\alpha)$ as $n \rightarrow \infty$. Then so does

$$P([Aj]^\alpha)/P[Aj] = P(\cup A_j^\alpha)/P(\cup A_j) = jP(A^\alpha)/jP(A) = P(A^\alpha)/P(A) \tag{5}$$

for every j . As $n \rightarrow \infty$, the ratios in Eq. (5) tend to $U(\alpha)$. Informally stated, Eq. (3) approximately holds for every union of elements of a CUP.

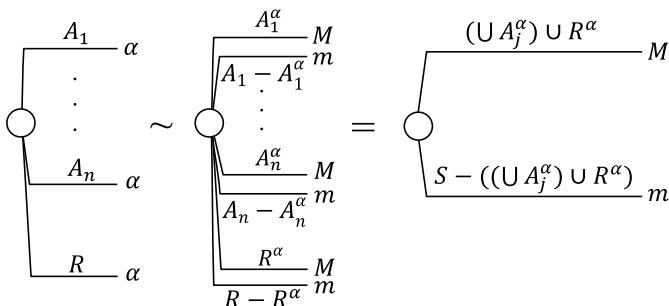


Fig. 9. Connecting conditional and unconditional U measurement using CUPs.

$$\frac{P([A(j-1)]^\alpha)}{P([A(j-1)])} \leq \frac{P(A^\alpha)}{P(A)} \leq \frac{P([A(j)]^\alpha)}{P([A(j)])}$$

} Numerators in ascending order; the two extremes converge to each other and, hence, to the middle.
 } Denominators in ascending order; the two extremes converge to the middle.

The two extreme fractions (identical to each other for each j) converge to $U(\alpha)$. This limit must be the middle fraction.

Fig. 10. Final step in proof of Lemma.

We now show that Eq. (3) holds exactly for any general event A , by using close approximations of CUPs with n larger and larger. Formally, take a CUP $\{A_1, \dots, A_n, R\}$ and take $[A(j-1)] \preccurlyeq A \preccurlyeq [A(j)]$. Then also $[A(j-1)]^\alpha \preccurlyeq A^\alpha \preccurlyeq [A(j)]^\alpha$. Let $n \rightarrow \infty$. Eq. (3) follows from Fig. 10. \square

Observation 14. By the sure-thing principle, f in Fig. 8 does not affect preference. \square

We can thus replace the subact $(A : \alpha)$ by the subact $(A^\alpha : M, A - A^\alpha : m)$ without affecting preference in every act f .⁶ We next derive SEU for a general act $(E_1 : x_1, \dots, E_n, x_n)$, depicted as the left act in Fig. 11. We here use the derivation of Luce and Raiffa (1957 p. 27/28). We immediately transplant their derivation to uncertainty. This is why we, unlike all preceding authors using Savage’s setup, do not need to derive a risk structure.

We, one by one, replace each outcome by a conditional standard gamble. By transitivity, the resulting act, depicted twice in the middle of Fig. 11, is indifferent. The equations in the figure show that this act yields M with probability SEU and m otherwise. This holds for any arbitrary left act. Hence, acts are preferred according to their SEU, and Statement (i) in Theorem 2 has been proved.

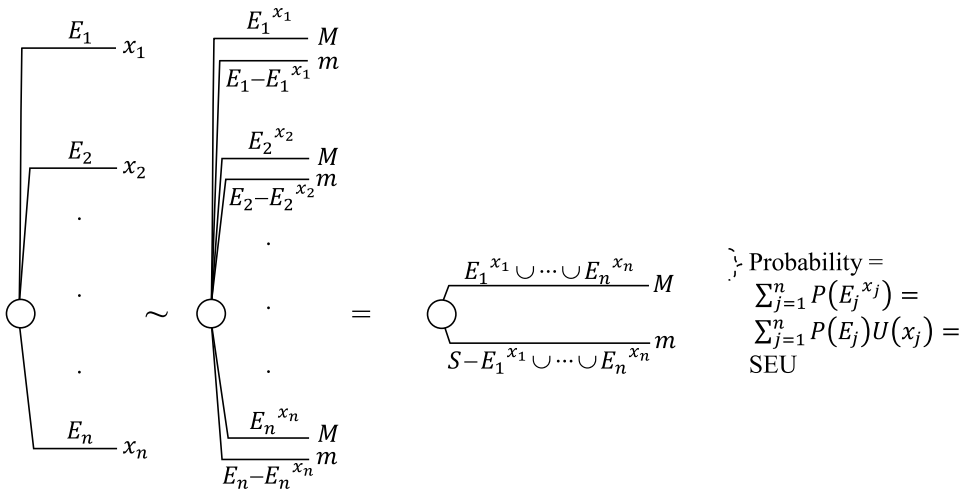


Fig. 11. SEU of a general act by substituting conditional standard gambles (using the sure-thing principle).

⁶ Cerreia-Vioglio et al. (2015) generalized expected utility by assuming that the above replacement always decreases the preference value of the act, and axiomatized it for risk in their cautious expected utility model.

UNIQUENESS RESULT As regards the uniqueness result, it is immediate that we can add any term to U , and can multiply it by any positive factor. After we chose two arbitrary values $U(M) > U(m)$, by Eq. (2), U is uniquely determined ($U(\alpha) = P(S^\alpha)U(M) + (1 - P(S^\alpha))U(m)$), so there is no further liberty.

We have now established the SEU representation and uniqueness result for the atomless case and consequences between m and M . Consequences beyond m and M can be covered by enlarging m' and M' more and more. For any $m' < m < M < M'$ we obtain an SEU model as we did for m and M . We rescale it to have U values 0 and 1 at m and M . Then, by the above uniqueness result, it must agree with the SEU representation derived above for m and M . Similarly, all SEU representations for all pairs m', M' agree on common domain, representing the same preference relation with the same scaling at m and M . All these SEU representations are part of the same overall SEU representation that captures every preference between every pair of acts by taking m' and M' extreme enough. The uniqueness result remains as before. This completes the proof of Theorem 2 for the atomless case.

5. Related literature

Because of its deep mathematics, Savage's analysis is a black box for most researchers. Hence, few papers have used his richness in events, even though it is natural. Kopylov (2007) is so far the only one that essentially generalized Savage's SEU derivation. Not only did he extend it to algebras of events, but further to a more general domain called mosaics. His analysis is, accordingly, more sophisticated and complex than Savage's. We focus on Savage's setup and leave the extension of our approach to mosaics to future research. Arrow (1971) added a preference axiom of Villegas' (1964) result on comparative probability that is equivalent to our set continuity, to imply countably additive probabilities rather than Savage's finitely additive probabilities. He could then replace Savage's P6 by a nonatomicity axiom that, under countable additivity, is equivalent to convex-rangedness. This is an appealing variation of Savage's theorem. Mackenzie (2019) generalized Villegas' analysis by allowing for atoms. Using MacKenzie's structure to obtain a generalization of Arrow (1971) is a topic for future research. Dietrich (2018) used Savage's theorem in an analysis of changing awareness.

Hartmann (2020) showed that P3 is redundant in Savage's (1954) main theorem. This holds only if nonsimple acts are included and Savage's P7, which only impacts nonsimple acts, is crucial for it. As Hartmann pointed out, the redundancy does not hold if only simple acts are considered, as in our case.

In analyses of ambiguity models and robust statistics, only a few papers used richness of events, similarly to Savage (1954), and virtually all used richness of consequences through the popular Anscombe-Aumann framework.⁷ Gilboa (1987) replaced Savage's P2 and P4 by a new intuitive axiom to obtain a generalization of SEU that uses nonadditive convex-ranged measures. He replaced Savage's P6 by solvability and Archimedean axioms similar to our axioms P6.1 and P6.2 in Theorem 2, although they do imply full convex-rangedness. Abdellaoui and Wakker (2005) used a similar approach. They replaced Savage's P2 and P4 by an intuitive axiom similar to Gilboa's and used technical conditions similar to P6.1 and P6.2. They additionally derived SEU from this result by reinforcing their intuitive axiom. Their resulting axiom is stronger than our P2 and P4. Their derivation essentially used the one for nonadditive measures, which entails

⁷ Ghirardato et al. (2003) did not assume this richness a priori, but derived it endogenously. Dean and Ortoleva (2017) similarly endogenized it.

a longer detour than the one in Savage (1954) and Fishburn (1970) (Abdellaoui and Wakker, 2005 p. 38 ll. 18-22).

Other papers using richness as in Savage include Ergin and Gul (2009), Grant (1995), Gul and Pesendorfer (2015), and Machina and Schmeidler (1992) for probabilistic sophistication, Abdellaoui (2002) and Nakamura (1995) for rank-dependent utility, Karni (1993) for state-dependent utility, Neilson (2010) for recursive expected utility, and Grant et al. (2016) in game theory. Chew and Sagi (2006, 2008) used it for probabilistic sophistication under ambiguity but, like our paper, can also handle finite state spaces with two utility levels. Their axioms imply convex-rangedness for the atomless case. Webb (2017) used Savage's setup in a rich two-stage model, but used solvability and Archimedeanity instead of Savage's P6, similarly as we did.

6. Conclusion

von Neumann and Morgenstern used a mixture technique to axiomatize a decision model and this has remained the most popular technique in decision theory. Luce and Tukey (1964) and Krantz et al. (1971) introduced an alternative technique based on Hölder's (1901) lemma, which is more powerful. We used it in this paper to improve the most famous theorem in decision theory: Savage's (1954) derivation of expected utility. The resulting derivation is more general and more appealing. Thus, we have made the most impressive intellectual achievement in decision theory, Savage (1954), accessible to a wide community of practitioners, researchers, and students. This also facilitates the development of non-expected utility and ambiguity models using Savage's appealing framework. A topic for future research is how to derive such models by using Hölder's lemma rather than mixture techniques as in the popular Anscombe-Aumann (1963) framework.

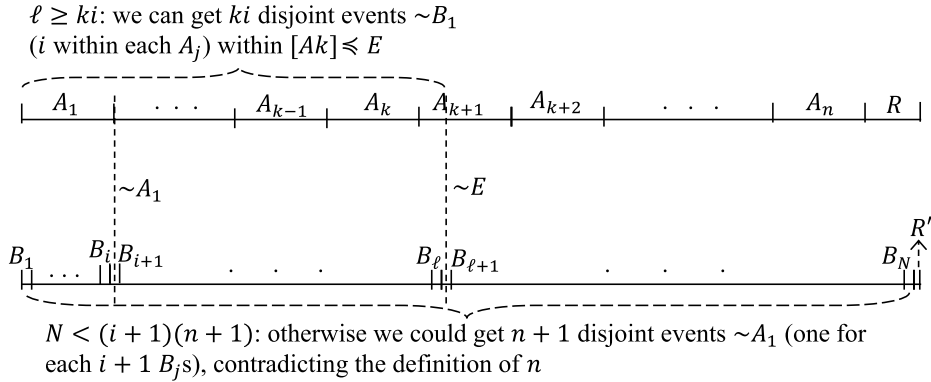
Acknowledgments

Jean Baccelli made helpful comments.

Appendix A. Details of the proof of Theorem 2

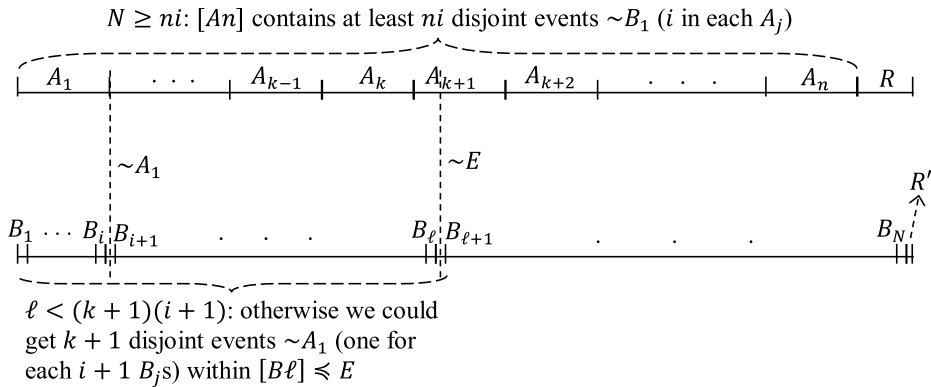
Proof of Lemma 10. See Fig. 12 for the first inequality and Fig. 13 for the second. In each figure, the state space S is a horizontal line partitioned by CUPs. Dashed vertical lines indicate the likelihood level of A_1 and E , respectively. Thus, $[Bi] \preceq A_1 \preceq [B(i+1)]$, where $[Bi]$ denotes $\cup_{j=1}^i B_j$ (as in Section 4.2). In Fig. 12, the upper brace gives one inequality, the lower brace a second, and these are then applied to numerator and denominator in the fraction below. The limiting result shows that for i large enough, i.e., N large enough, we get the first inequality in the lemma. Fig. 13 is organized in the same way. (It can be seen that $N > (n+1)^2$ gives the desired inequalities for $i \rightarrow \infty$ in the figures.) \square

Proof of Theorem 2 for atoms. We finally consider the case where an atom A exists. Lemmas 6 and 7, derived for the atomless case, hold for atoms without modification. There cannot exist a nonnull event less likely than A because then by solvability A would not be an atom after all. By repeated application of solvability, every event $C \succ A$ can be written as a union of events equally likely as A . These events must all be atoms. Say S is a union of n such atoms. By solvability, for every event, all of the n aforementioned atoms are either entirely contained in the event or disjoint, up to a null event. Hence the event must be equal to a union of some of these n atoms up



$$\frac{\ell - 1}{N + 1} \geq \frac{ki - 1}{(i + 1)(n + 1) + 1} = \frac{k - \frac{1}{i}}{n + 1 + \frac{n + 2}{i}} \xrightarrow{i \rightarrow \infty} \frac{k}{n + 1} > \frac{k - 1}{n + 1}$$

Fig. 12. Consistency of probability estimates for tentative lower bounds.



$$\frac{\ell + 2}{N} \leq \frac{(k + 1)(i + 1) + 2}{ni} = \frac{k + 1 + \frac{k + 3}{i}}{n} \xrightarrow{i \rightarrow \infty} \frac{k + 1}{n} < \frac{k + 2}{n}$$

Fig. 13. Consistency of probability estimates for tentative upper bounds.

to a null event. We define $P(A) = 1/n$, and each union of j atoms has probability j/n , as does every event differing by a null event.

Our solvability precludes the existence of three consequences $\gamma > \beta > \alpha$. To wit, then A being an atom, $\gamma_A \alpha > \beta_A \alpha > \alpha_A \alpha$, and (as required by solvability) $\gamma_B \alpha_{A-B} \sim \beta_A \alpha$ for some $B \subset A$ are contradictory. Hence we may assume that there are only two consequences $\gamma > \beta$ plus other consequences indifferent to one of these two. We define $U(\gamma) = 1$ and $U(\beta) = 0$. Preference is guided by the most likely receipt of a consequence indifferent to γ . Thus SEU holds. Any value $U(\gamma) > U(\beta)$ works and the uniqueness result follows. This completes the proof of Theorem 2 for the case of atoms. \square

We end this appendix with some further comments. Unlike Savage (1954), we allow for finite cases, but only in the limited sense of two consequences. This is due to the restrictive nature of the solvability axiom that we used. Richer cases with, for instance, any set of equally-spaced utilities can be obtained by adapting solvability axioms of Abdellaoui and Wakker (2005 p. 62) to our setup. The restriction of equally likely atoms for finite state spaces can generate every rational probability and, in this sense, is not very restrictive.

Theorem 2 contained solvability as a nonnecessary axiom. We now state the restrictions that this axiom imposes on probability and utility. P satisfies *range solvability* if: whenever $P(E_1)U(\gamma) + \sum_{j=2}^n P(E_j)U(x_j) > EU(g) > P(E_1)U(\beta) + \sum_{j=2}^n P(E_j)U(x_j)$, there exists a partition E_{11}, E_{12} of E_1 such that $P(E_{11})U(\gamma) + P(E_{12})U(\beta) + \sum_{j=2}^n P(E_j)U(x_j) = EU(g)$. The condition imposes richness on the probability measure P , which, roughly, should be able to match every EU difference, and should be at least as refined as U . This richness is in the spirit of Savage's convex-rangedness but is more general, as Example 3 demonstrated. It can be used to turn Theorem 2 into an if-and-only-if result, with necessary and sufficient preference conditions:

Observation 15. *In Theorem 2, solvability can be moved to Statement (i) if range solvability is added in Statement (ii). □*

We chose our solvability and Theorem 2 so as to maximize accessibility rather than generality.

Fig. 11 shows how, mainly due to the sure-thing principle, every subact (E_j, x_j) contributes to the value of the act (laid down in Fig. 8) independently of the rest of the act. This shows how the sure-thing principle leads to the additive aggregation over disjoint events that is most characteristic of the SEU formula.

Unlike Savage (1954), we only need an algebra of events and not a sigma-algebra. One reason is that all our limiting results are based on monotonic sandwiching between upper and lower bounds without requiring the existence of any limits.

Appendix B. Proof of Proposition 4

In the following proofs we only use results derived from P1-P5, such as Lemma 6, and we do not assume P6.1 or P6.2. The following lemma derives a stronger version of the Archimedean axiom that we need, in the absence of solvability, for Lemma 17.

Lemma 16. *Under P1-P5, Savage's P6 implies reinforced Archimedeanity: there exist no nonnull event E and infinite sequence of disjoint events E_j with $E_j \succcurlyeq E$ for all j .*

Proof. For contradiction, assume that the sequence E_1, \dots is infinite, and consider $\gamma_E \beta$ for $\gamma \succ \beta$. For the preference $\gamma_E \beta \succ \beta$, take the finite partition delivered by P6 that we denote $\{A_1, \dots, A_n\}$ instead of $\{E_1, \dots, E_n\}$. On each of the A_j s, improving β into γ is worse than improving it on E , so all A_j are smaller than E . But then $S = A_1 \cup \dots \cup A_n \preccurlyeq E_1 \cup \dots \cup E_n < E_1 \cup \dots \cup E_{n+1}$ by repeated application of Lemma 6(ii): contradiction. □

Savage assumed that \mathcal{E} is a σ -algebra of events, and so do we in the following lemma.

Lemma 17. *Assume that \mathcal{E} is a σ -algebra. Then Savage's P6 implies solvability.*

Proof. Assume $\beta_E f \prec g \prec \gamma_E f$ with $\gamma \succ \beta$. We may assume $f = \beta_E f$. We write $E = E^0$ and $f = f^0$, and have

$$\beta_{E^0} f^0 \prec g \prec \gamma_{E^0} f^0. \tag{6}$$

For the preference $\beta_{E^0} f^0 \prec g$, take the partition $\{E_1, \dots, E_n\}$ of P6. Define $E'_j = E_j \cap E$. Say E'_1 is the most likely of the (E'_j) s, implying that it is nonnull. Take j such that for $B^1 = E'_1 \cup \dots \cup E'_j$ and $E^1 = E'_{j+1}$ we have $\gamma_{B^1} f \prec g \prec \gamma_{B^1 \cup E^1} f$. By P6, $j \geq 1$. Write $D^1 = E'_{j+2} \cup \dots \cup E'_n$. We have partitioned $E = E^0$ into (B^1, E^1, D^1) , with E^1 smaller than B^1 (even smaller than E'_1), so, informally, smaller than half E^0 . Define $f^1 = \gamma_{B^1} f$. Now $\beta_{E^1} f^1 \prec g \prec \gamma_{E^1} f^1$. We can continue the process inductively, to get, for all natural j , f^j and a partition $\{B^j, E^j, D^j\}$ of E^{j-1} with $E^j \preceq E^{j-1} - E^j$ and

$$\beta_{E^j} f^j \prec g \prec \gamma_{E^j} f^j. \tag{7}$$

Because of reinforced Archimedeanity, $\cap E^j$ is null, and for the cumulative $C^j = \cup_{i=1}^j B^i$ we have

$$\gamma_{C^j} f \prec g \prec \gamma_{C^j \cup E^j} f. \tag{8}$$

Because Savage uses a σ -algebra of events, we can define $C = \cup_{j=1}^\infty C^j$. We claim

$$\gamma_C f \sim g. \tag{9}$$

First assume, for contradiction, $\gamma_C f \prec g$. Using P6 as above, there is a nonnull C' disjoint from C with $\gamma_{C \cup C'} f \prec g$, implying $E^j \succ C'$ (E^j is added to an event $C^j \preceq C$ but still turning β into γ on E^j switches preference) for all j , violating reinforced Archimedeanity.

Finally assume, for contradiction, $\gamma_C f \succ g$. Using P6 as above, there is a nonnull $C' \subset C$ with still $\gamma_{C-C'} f \succ g$, implying $E^j \succ C'$ (E^j is subtracted from an event $C^j \cup E^j \succ C$ but still turning γ back into β on E^j switches preference), violating reinforced Archimedeanity. Eq. (9) and solvability must hold. \square

References

Abdellaoui, Mohammed, 2002. A genuine rank-dependent generalization of the von Neumann-Morgenstern expected utility theorem. *Econometrica* 70, 717–736.

Abdellaoui, Mohammed, Wakker, Peter P., 2005. The likelihood method for decision under uncertainty. *Theory Decis.* 58, 3–76.

Allais, Maurice, 1953. *Fondements d'une Théorie Positive des Choix Comportant un Risque et Critique des Postulats et Axiomes de l'Ecole Américaine*. Colloques Internationaux du Centre National de la Recherche Scientifique (Econométrie) 40, 257–332. Paris: Centre National de la Recherche Scientifique. Translated into English, with additions, as the *Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School*, In: Allais, Maurice, Hagen, Ole (Eds.), *Expected Utility Hypotheses and the Allais Paradox*. Reidel, Dordrecht, the Netherlands, pp. 27–145, 1979.

Alon, Shiri, Lehrer, Ehud, 2014. Subjective multi-prior probability: a representation of a partial likelihood relation. *J. Econ. Theory* 151, 476–492.

Anscombe, Francis J., Aumann, Robert J., 1963. A definition of subjective probability. *Ann. Math. Stat.* 34, 199–205.

Arrow, Kenneth J., 1971. *Essays in the Theory of Risk-Bearing*. North-Holland, Amsterdam.

Battigalli, Pierpaolo, Cerreia-Vioglio, Simone, Maccheroni, Fabio, Marinacci, Massimo, 2017. Mixed extensions of decision problems under uncertainty. *Econ. Theory* 63, 827–866.

Cerreia-Vioglio, Simone, Dillenberger, David, Ortleva, Pietro, 2015. Cautious expected utility and the certainty effect. *Econometrica* 83, 693–728.

- Cerreia-Vioglio, Simone, Maccheroni, Fabio, Marinacci, Massimo, Montrucchio, Luigi, 2013. Classical subjective expected utility. *Proc. Natl. Acad. Sci.* 110, 6754–6759.
- Chateauneuf, Alain, Cohen, Michèle, Jaffray, Jean-Yves, 2006. Decision under uncertainty: the classical models. In: Bouyssou, Denis, Dubois, Didier, Prade, Henri, Pilot, Marc (Eds.), *Decision-Making Process: Concepts and Methods*. Wiley, New York, pp. 385–400.
- Chew, Soo Hong, Sagi, Jacob S., 2006. Event exchangeability: probabilistic sophistication without continuity or monotonicity. *Econometrica* 74, 771–786.
- Chew, Soo Hong, Sagi, Jacob S., 2008. Small worlds: modeling attitudes toward sources of uncertainty. *J. Econ. Theory* 139, 1–24.
- Cozic, Mikael, Hill, Brian, 2015. Representation theorems and the semantics of decision-theoretic concepts. *J. Econ. Methodol.* 22, 292–311.
- de Finetti, Bruno, 1937. La Prévision: Ses Lois Logiques, ses Sources Subjectives. *Ann. Inst. Henri Poincaré* 7, 1–68. Translated into English by Henry E. Kyburg, “Foresight: Its Logical Laws, its Subjective Sources, In: Kyburg, Henry E., Smokler, Howard E. (Eds.), *Studies in Subjective Probability*. Wiley, New York, 1964, pp. 53–118, 2nd edition. Krieger, New York, 1980.
- Dean, Mark, Ortoleva, Pietro, 2017. Allais, Ellsberg, and preferences for hedging. *Theor. Econ.* 12, 377–424.
- Dietrich, Franz, 2018. Savage’s theorem under changing awareness. *J. Econ. Theory* 176, 1–54.
- Ellsberg, Daniel, 1961. Risk, ambiguity and the Savage axioms. *Q. J. Econ.* 75, 643–669.
- Ergin, Haluk, Gul, Faruk, 2009. A theory of subjective compound lotteries. *J. Econ. Theory* 144, 899–929.
- Fishburn, Peter C., 1970. *Utility Theory for Decision Making*. Wiley, New York.
- Fishburn, Peter C., 1982. *The Foundations of Expected Utility*. Reidel, Dordrecht.
- Fishburn, Peter C., 1986. The axioms of subjective probability. *Stat. Sci.* 1, 335–358.
- Fuhrken, Gebhard, Richter, Marcel K., 1991. Additive utility. *Econ. Theory* 1, 83–105.
- Ghirardato, Paolo, Marinacci, Massimo, 2001. Range convexity and ambiguity averse preferences. *Econ. Theory* 17, 599–617.
- Gilboa, Itzhak, 1987. Expected utility with purely subjective non-additive probabilities. *J. Math. Econ.* 16, 65–88.
- Gilboa, Itzhak, Marinacci, Massimo, 2016. Ambiguity and the Bayesian paradigm. In: Arló-Costa, Horacio, Hendricks, Vincent F., van Benthem, Johan F.A.K. (Eds.), *Readings in Formal Epistemology*. Springer, Berlin, pp. 385–439.
- Ghirardato, Paolo, Maccheroni, Fabio, Marinacci, Massimo, Siniscalchi, Marciano, 2003. A subjective spin on roulette wheels. *Econometrica* 71, 1897–1908.
- Gold, Marthe R., Siegel, Joanna E., Russell, Louise B., Weinstein, Milton C., 1996. *Cost-Effectiveness in Health and Medicine*. Oxford University Press, New York.
- Grant, Simon, 1995. Subjective probability without eventwise monotonicity: or: how Machina’s Mom May also be probabilistically sophisticated. *Econometrica* 63, 159–189.
- Grant, Simon, Meneghel, Idione, Tourky, Rabee, 2016. Savage games. *Theor. Econ.* 11, 641–682.
- Gul, Faruk, Pesendorfer, Wolfgang, 2015. Hurwicz expected utility and subjective sources. *J. Econ. Theory* 159, 465–488.
- Halpern, Joseph Y., 1999. A counterexample to theorems of Cox and Fine. *J. Artif. Intell. Res.* 10, 67–85.
- Hartmann, Lorenz, 2020. Savage’s P3 is redundant. *Econometrica*. forthcoming.
- Hölder, Otto, 1901. Die Axiome der Quantität und die Lehre vom Mass. In: *Berichte Verhand. König. Sächs. Gesell. Wiss. (Leipzig)*. Math. Phys., Classe 53, 1–64;
Part I is translated into English by Michell, Joel, Ernst, Catherine, In: *The Axioms of Quantity and the Theory of Measurement*. *J. Math. Psychol.* 40, 1996 235–252;
Part II is translated into English by Michell, Joel, Ernst, Catherine. *J. Math. Psychol.* 41, 1997 345–356.
- Kadane, Joseph B., Schervish, Mark J., Seidenfeld, Teddy, 1996. Reasoning to a foregone conclusion. *J. Am. Stat. Assoc.* 91, 1228–1235.
- Khan, Ali M., Uyan, Metin, 2018. Topological Connectedness and Behavioral Assumptions on Preferences: a Two-Way Relationship. Working paper.
- Karni, Edi, 1993. Subjective expected utility with state-dependent preferences. *J. Econ. Theory* 60, 428–438.
- Karni, Edi, Maccheroni, Fabio, Marinacci, Massimo, 2014. Ambiguity and nonexpected utility. In: Young, Peyton H., Zamir, Shmuel (Eds.), *Handbook of Game Theory*, vol. 4. North-Holland, Amsterdam.
- Kopylov, Igor, 2007. Subjective probabilities on “small” domains. *J. Econ. Theory* 133, 236–265.
- Krantz, David H., Duncan Luce, R., Suppes, Patrick, Tversky, Amos, 1971. *Foundations of Measurement, Vol. I (Additive and Polynomial Representations)*. Academic Press, New York.
- Kreps, David M., 1988. *Notes on the Theory of Choice*. Westview Press, Boulder Colorado.
- Luce, Duncan R., Krantz, David H., Suppes, Patrick, Tversky, Amos, 1990. *Foundations of Measurement, Vol. III. (Representation, Axiomatization, and Invariance)*. Academic Press, New York.

- Luce, Duncan R., Raiffa, Howard, 1957. *Games and Decisions*. Wiley, New York.
- Luce, Duncan R., Tukey, John W., 1964. Simultaneous conjoint measurement: a new type of fundamental measurement. *J. Math. Psychol.* 1, 1–27.
- Machina, Mark J., Schmeidler, David, 1992. A more robust definition of subjective probability. *Econometrica* 60, 745–780.
- Mackenzie, Andrew, 2019. A foundation for probabilistic beliefs with or without atoms. *Theor. Econ.* 14, 709–778.
- Marinacci, Massimo, 2015. Model uncertainty. *J. Eur. Econ. Assoc.* 13, 1022–1100.
- Michell, Joel, 1990. *An Introduction to the Logic of Psychological Measurement*. Lawrence Erlbaum Associates, Hillsdale, NJ.
- Moscati, Ivan, 2018. *Measuring Utility: From the Marginal Revolution to Behavioral Economics*. Oxford University Press, Oxford, UK.
- Nakamura, Yutaka, 1995. Rank dependent utility for arbitrary consequence spaces. *Math. Soc. Sci.* 29, 103–129.
- Neilson, William S., 2010. A simplified axiomatic approach to ambiguity aversion. *J. Risk Uncertain.* 41, 113–124.
- Pfanzagl, Johann, 1968. *Theory of Measurement*. Physica-Verlag, Vienna.
- Savage, Leonard J., 1954. *The Foundations of Statistics*. Wiley, New York, second edition. Dover Publications, New York, 1972.
- Stevens, Stanley S., 1968. Measurement, statistics, and the schemapiric view. *Science* 161, 849–856.
- van Osch, Sylvie M.C., Wakker, Peter P., van den Hout, Wilbert B., Stiggelbout, Anne M., 2004. Correcting biases in standard gamble and time tradeoff utilities. *Med. Decis. Mak.* 24, 511–517.
- Villegas, Cesáreo, 1964. On quantitative probability σ -algebras. *Ann. Math. Stat.* 35, 1787–1796.
- von Neumann, John, Morgenstern, Oskar, 1947. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton NJ. 1953.
- Wakker, Peter P., 1981. Agreeing probability measures for comparative probability structures. *Ann. Stat.* 9, 658–662.
- Wakker, Peter P., 1993. Unbounded utility for Savage’s “foundations of statistics,” and other models. *Math. Oper. Res.* 18, 446–485.
- Walley, Peter, 1991. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London.
- Webb, Craig, 2017. Purely subjective variational preferences. *Econ. Theory* 64, 121–137.