Extensions of Expected Utility Theory and Some Limitations of Pairwise Comparisons*

M. J. SCHERVISH
Carnegie Mellon University, USA

T. SEIDENFELD
Carnegie Mellon University, USA

J. B. KADANE
Carnegie Mellon University, USA

I. LEVI
Columbia University, USA

Abstract

We contrast three decision rules that extend Expected Utility to contexts where a convex set of probabilities is used to depict uncertainty: Γ-Maximin, Maximality, and *E*-admissibility. The rules extend Expected Utility theory as they require that an option is inadmissible if there is another that carries greater expected utility for each probability in a (closed) convex set. If the convex set is a singleton, then each rule agrees with maximizing expected utility. We show that, even when the option set is convex, this pairwise comparison between acts may fail to identify those acts which are Bayes for some probability in a convex set that is not closed. This limitation affects two of the decision rules but not *E*-admissibility, which is not a pairwise decision rule. *E*-admissibility can be used to distinguish between two convex sets of probabilities that intersect all the same supporting hyperplanes.

1 Introduction

This paper offers a comparison among three decision rules for use when uncertainty is depicted by a non-trivial, convex set of probability functions $\mathcal{P}$. This setting for uncertainty is different from the canonical Bayesian decision theory of expected utility, which uses a singleton set, just one probability function, to represent a decision maker's uncertainty. Justifications for using a non-trivial set of probabilities to depict uncertainty date back at least a half century [4] and a

*The research of the first three authors is supported by NSF grant DMS-0139911

foreshadowing of that idea can be found even in [7], where he allows that not all hypotheses may be comparable by qualitative probability – in accord with, e.g., the situation where the respective intervals of probabilities for two events merely overlap with no further (joint) constraints, so that neither of the two events is more, or less, or equally probable compared with the other.

We study decision rules that are extensions of canonical Subjective Expected Utility [SEU] theory, using sets of probabilities, in the following sense. The decision rules we consider all satisfy the following pairwise comparison between two options.

Criterion 1 *For a pair of options f and g, if for each probability $P \in \mathcal{P}$, f has greater expected utility than g, then g is inadmissible whenever f is available.*

This pairwise comparison itself creates a strict partial order. It (or a similar relation) has been the subject of representation theorems by, e.g., [3, 14, 15]. Note that when $\mathcal{P}$ is a singleton set, then the partial order is a weak order that satisfies SEU theory. In this sense, a decision rule that embeds this partial order extends SEU theory.

Here, we avail ourselves of four simplifying assumptions:

1. The decision maker's values for outcomes are determinate and are depicted by a (cardinal) utility function.

   Reason: We use circumstances under which convexity of $\mathcal{P}$ is not controversial. [1]

2. The algebra of uncertainty is finite, with finite state space $\Omega = \{\omega_1, \omega_2, ..., \omega_k\}$.

   Reason: We avoid the controversies surrounding countable versus finite additivity, which arise with infinite algebras.

3. Acts (or options) are *gambles*, i.e. functions from states to utilities, $f : \Omega \longrightarrow \mathbb{R}$.

   Reason: This assumption is commonplace and affords us an opportunity to contrast a variety of decision rules.

4. Each decision problem presents the decision maker a *uniformly bounded* choice set $\mathcal{A}$ of *gambles.*

   Reason: We avoid complications with unbounded utilities. Moreover, by considering the convex hull of a family of such gambles, we are assured of achieving the infimum and supremum operations with respect to expected utilities calculated with respect to the set $\mathcal{P}$.

[1] The issue of convexity of $\mathcal{P}$ is controversial. See [14] for a representation of partially ordered strict preferences that does not require convexity unless the decision maker has a determinate (cardinal) utility for outcomes. Rebuttal is presented in Section 7 of [11].

Of the three decision rules we discuss, perhaps the most familiar one is Γ-Maximin[2]. This rule requires that the decision maker ranks a gamble by its lower expected value, taken with respect to a closed, convex set of probabilities, $\mathcal{P}$, and then to choose an option from $\mathcal{A}$ whose lower expected value is maximum. This decision rule (as simplified by the assumptions, above) was given a representation in terms of a binary preference relation over Anscombe-Aumann horse lotteries [2], has been discussed by, e.g., Section 4.7.6 of [1] and recently by [5], who defend it as a form of Robust Bayesian decision theory. The Γ-Maximin decision rule creates a preference ranking of options independent of the alternatives available in $\mathcal{A}$: it is context independent in that sense. But Γ-Maximin corresponds to a preference ranking that fails the so-called (von Neumann-Morgenstern's) "Independence" or (Savage's) "Sure-thing" postulate of SEU theory. In Section 2 of [15], we note that such theories suffer from *sequential incoherence* in particular sequential decision problems.

The second decision rule that we consider, called *E*-admissibility ('*E*' for "expectation"), was formulated in [8, 9]. *E*-admissibility constrains the decision maker's admissible choices to those gambles in $\mathcal{A}$ that are Bayes for at least one probability $P \in \mathcal{P}$. That is, given a choice set $\mathcal{A}$, the gamble f is *E*-admissible on the condition that, for at least one $P \in \mathcal{P}$, f maximizes subjective expected utility with respect to the options in $\mathcal{A}$.[3] Section 7.2 of [12][4] defends a precursor to this decision rule in connection with cooperative group decision making. *E*-admissibility does not support an ordering of options, real-valued or otherwise, so that it is inappropriate to characterize *E*-admissibility by a ranking of gambles independent of the set $\mathcal{A}$ of feasible options. However, the distinction between options that are and are not *E*-admissible does support the "Independence" postulate. For example, if neither option f nor g is *E*-admissible in a given decision problem $\mathcal{A}$, then the convex combination, the mixed option $h = \alpha f \oplus (1\text{-}\alpha)g$ $(0 \leq \alpha \leq 1)$ likewise is *E*-inadmissible when added to $\mathcal{A}$. This is evident from the basic SEU property: the expected utility of a convex combination of two gambles is the corresponding weighted average of their separate expected utilities; hence, for a given $P \in \mathcal{P}$ the expected utility of the mixture of two gambles is bounded above by the maximum of the two expected utilities. The assumption that neither of two gambles is *E*-admissible entails that their mixture has *P*-expected utility less than some *E*-admissible option in $\mathcal{A}$.

The third decision rule we consider is called *Maximality* by Walley in [17][5],

[2]When outcomes are cast in terms of a (statistical) loss function, the rule is then Γ-Minimax: rank options by their maximum expected risk and choose an option whose maximum expected risk is minimum.

[3]Levi's decision theory is lexicographic, in which the first consideration is *E*-admissibility, followed by other considerations, e.g. what he calls a Security index. Here, we attend solely to *E*-admissibility.

[4]Savage's analysis of the decision problem depicted by his Figure 1, p. 123, and his rejection of option *b*, p. 124 is the key point.

[5]There is, for our discussion here, a minor difference with Walley's formulation of Maximality

who appears to endorse it (p. 166). *Maximality* uses the strict partial order (above) to fix the admissible gambles from $\mathcal{A}$ to be those that are not strictly preferred by any other member of $\mathcal{A}$. That is, f is a *Maximal* choice from $\mathcal{A}$ provided that there is no other element $g \in \mathcal{A}$ that, for each $P \in \mathcal{P}$, carries greater expected utility than f does. *Maximality* (under different names) has been studied, for example, in [6, 8, 10, 13, 16]. Evidently, the E-admissible gambles in a decision problem are a subset of the Maximally admissible ones.

The three rules have different sets of admissible options. Here is a heuristic illustration of that difference.

Example 1 *Consider a binary-state decision problem,* $\Omega = \{\omega_1, \omega_2\}$*, with three feasible options. Option f yields an outcome worth 1 utile if state* ω_1 *obtains and an outcome worth 0 utiles if* ω_2 *obtains. Option g is the mirror image of f and yields an outcome worth 1 utile if* ω_2 *obtains and an outcomes worth 0 utiles if* ω_1 *obtains. Option h is constant in value, yielding an outcome worth 0.4 utiles regardless whether* ω_1 *or* ω_2 *obtains. Figure 1 graphs the expected utilities for these three acts. Let* $\mathcal{P} = \{P\colon 0.25 \leqslant P(\omega_1) \leqslant 0.75\}$*. The surface of Bayes solutions is highlighted. The expected utility for options f and g each has the interval of values [0.25, 0.75], whereas h of course has constant expected utility of 0.4. From the choice set of these three options* $\mathcal{A} = \{f, g, h\}$ *the* Γ*-Maximin decision rule determines that h is (uniquely) best, assigning it a value of 0.4, whereas f and g each has a* Γ*-Maximin value of 0.25. By contrast, under E-admissibility, only the option h is E-*inadmissible *from the trio. Either of f or g is E-admissible. And, as no option is strictly preferred to any other by expectations with respect to* $\mathcal{P}$*, all three gambles are admissible under* Maximality.

What normative considerations can be offered to distinguish among these three rules? For example, all three rules are immune to a Dutch Book, in the following sense:

Definition 1 *Call an option* favorable *if it is uniquely admissible in a pairwise choice against the status-quo of "no bet," which we represent as the constant 0.*

Proposition 1 *For each of the three decision rules above, no finite combination of favorable options can result in a Dutch Book, i.e., a sure loss.*

Proof. Reason indirectly. Suppose that the sum of a finite set of favorable gambles is negative in each state ω. Choose an element P from $\mathcal{P}$. The probability P is a convex combination of the extreme (0-1) probabilities, corresponding to a convex combination of the finite partition by states. The expectation of a convex

involving null-events. Walley's notion of Maximality requires, also, that an admissible gamble be classically admissible, i.e., not weakly dominated with respect to state-payoffs. This means that, e.g., our Theorem 1(i) is slightly different in content from Walley's corresponding result.

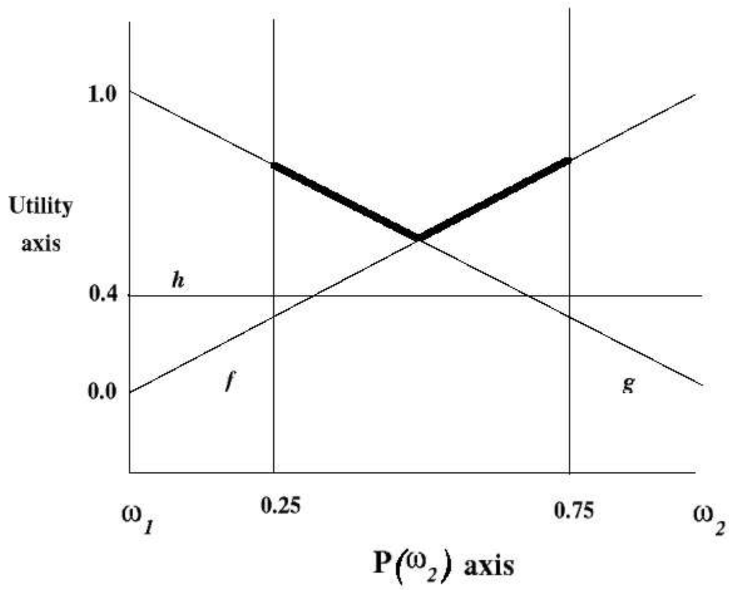

Figure 1: Expected utilities for three acts in Example 1. The thicker line indicates the surface of Bayes solutions.

combination of probabilities is the convex combination of the individual expectations. This makes the P-expectation of the sum of the finite set of favorable options negative. But the P-expectation of the sum cannot be negative unless at least one of the finitely many gambles has a negative P-expectation. Then that gamble cannot be favorable under any of the three decision rules. Thus, none of these three decision rules is subject to sure loss. □

In this paper, we develop an additional criterion for contrasting these decision rules. In Section 2 we address the question of what operational content the rules give to *distinguishing among different (convex) sets of probabilities*. That is, we are concerned to understand which convex sets of probabilities are treated as equivalent under a given decision rule. When do two convex sets of probabilities lead to all the same admissible options for a given decision rule? Γ-Maximin and Maximality are based solely on pairwise comparisons. Not so for E-admissibility. Even when the choice set $\mathcal{A}$ of feasible options is convex (e.g., closed under mixed strategies), these rules have distinct classes of admissible options.

2 Gambles and pairwise choice rules

It is evident that for Γ-Maximin generally to satisfy Criterion 1, the convex set of probabilities $\mathcal{P}$ must be closed. For an illustration why, if Example 1 is modified so that $\mathcal{P}' = \{P : 0.4 < P(\omega_1) \leqslant 0.75\}$ then, even though f and h both have the same infimum, 0.4, of expectations with respect to $\mathcal{P}'$, for each $P \in \mathcal{P}'$ f has greater expected utility than does h. Thus, from the perspective of operational content, the Γ-Maximin rule fails to distinguish between different convex sets of probabilities that differ with respect to Criterion 1, although each of Maximality and E-admissibility does distinguish the two sets $\mathcal{P}$ and $\mathcal{P}'$.

In order to contrast Maximality and E-admissibility, first we ask when do they lead to the same choices? Walley's Theorem 3.9.5 of [17] shows that, when the option space $\mathcal{A}$ is convex and the convex set of probabilities $\mathcal{P}$ is closed, the two rules are equivalent, i.e. both E-admissibility and Maximality reduce to a pairwise comparison of options according to Criterion 1. In this circumstance, an option is admissible, under either rule, just in case there is no other option that makes it inadmissible under Criterion 1. Then, with decision problems using convex sets of options, the two rules are capable of distinguishing between any two closed convex sets of probabilities, since distinct closed convex sets have distinct sets of supporting hyperplanes.

In Corollary 1 we re-establish Walley's result, and we extend the equivalence to decision problems in which $\mathcal{P}$ is open and $\mathcal{A}$ is finitely generated. The example following Theorem 1 establishes that for part (ii), the restriction to a finite (or finitely generated) option set, $\mathcal{A}$, is necessary. More important, however, we think is the second example following Theorem 1. That example is of a finite decision problem with a convex set of probabilities $\mathcal{P}$ (neither closed nor open) where, even though the option set is made convex, some Maximal options are not Bayes with respect to $\mathcal{P}$. Hence, even when the option space is convex, E-admissibility does not in general reduce to pairwise comparisons.

We preface Theorem 1 with a restatement of the structural assumptions for decision problems that we use in this paper. Let Ω be a finite state space with k states. Let $\mathcal{A}$ be a uniformly bounded collection of acts or gambles (real-valued functions from Ω). Let $\mathcal{C}$ be the convex hull of $\mathcal{A}$. For each probability vector P = $(p_1,\ldots,p_k) \in \mathcal{P}$ and each $f \in \mathcal{C}$ there is a point $(p_1,\ldots,p_{k-1},E_p(f)) \in \mathbb{R}^k$, where $E_p(f) = \sum_{j=1}^{k} p_j f(\omega_j)$. For each $f \in \mathcal{C}$ there is a hyperplane that contains all of the points of the form $(p_1,\ldots,p_{k-1},E_p(f))$. For each $f \in \mathcal{C}$, the halfspace at or above its corresponding hyperplane is

$$\{x \in \mathbb{R}^k : \alpha_f^\top x \geq c_f\},$$

where

$$\alpha_f = (f(\omega_k) - f(\omega_1),\ldots,f(\omega_k) - f(\omega_{k-1}),1),$$

and $c_f = f(\omega_k)$.

Definition 2 *Let $\mathcal{P}$ be a convex set of probability vectors. We say that $f \in \mathcal{A}$ is* Bayes with respect to $\mathcal{P}$ *if there exists $p \in \mathcal{P}$ such that $E_p(f) \geq E_p(g)$ for all $g \in \mathcal{A}$.*

Theorem 1 *Let $\mathcal{B}$ be the set of all $f \in \mathcal{A}$ such that f is Bayes with respect to $\mathcal{P}$. Suppose that $g \in \mathcal{A} \setminus \mathcal{B}$. Assume either*

(i) that $\mathcal{P}$ is closed, or

(ii) that $\mathcal{A}$ is finite and that $\mathcal{P}$ is open. That is,

$$\{(p_1, \ldots, p_{k-1}) : (p_1, \ldots, p_k) \in \mathcal{P}\}$$

is an open subset of $\mathbb{R}^{k-1}$.

Then there exists h in the convex hull of $\mathcal{B}$ such that $E_p(h) > E_p(g)$ for all $p \in \mathcal{P}$.

Corollary 1 *Assume that $\mathcal{A}$ is closed and convex. Let $\mathcal{B}$ be the set of all $f \in \mathcal{A}$ such that f is Bayes with respect to $\mathcal{P}$. Suppose that $g \in \mathcal{A} \setminus \mathcal{B}$ is not Bayes with respect to $\mathcal{P}$. Assume either*

(i) that $\mathcal{P}$ is closed, or

(ii) that $\mathcal{A}$ is the convex hull of finitely many acts and that $\mathcal{P}$ is open.

Then there exists $h \in \mathcal{B}$ such that $E_p(h) > E_p(g)$ for all $p \in \mathcal{P}$.

The proofs of Theorem 1 and Corollary 1 rely on a series of results about convex sets and are given in Appendix A.

Example 2 *The following example illustrates that Theorem 1(ii) does not hold if $\mathcal{A}$ is allowed to be infinite. Let Ω have only $k = 2$ states. Let $\mathcal{A}$ consist of the gambles $\{f_\theta : 0 \leq \theta \leq \pi/4\}$ where*

$$f_\theta = (0.4 + 0.8\tan(\theta) - 0.2\sec(\theta), 0.4 - 0.2\tan(\theta) - 0.2\sec(\theta)).$$

Notice that $f_0 = (0.2, 0.2)$. Let

$$\mathcal{P} = \{(p_1, p_2) : p_1 > 0.2\}.$$

For each $p_1 \in (0.2, 0.3)$, the act f_θ is Bayes with respect to $\mathcal{P}$ when $\theta = 0.5\sin^{-1}(10[p_1 - 0.2])$. For $p_1 \geq 0.3$, $f_{\pi/4}$ is Bayes with respect to $\mathcal{P}$. Let $g = f_0$, which is not Bayes with respect to $\mathcal{P}$. Notice that, for every θ,

$$E_p(f_\theta) = (p_1 - 0.2)\tan(\theta) + 0.4 - 0.2\sec(\theta).$$

So, $E_p(f_\theta) < 0.2$ when $p_1 = 0.2$. Since $E_p(f_\theta)$ is a continuous function of p, $E_p(f_\theta) < 0.2$ for p in an open set around $(0.2, 0.8)$, which includes part of $\mathcal{P}$. It follows that every convex combination h of f_θ's has $E_p(h) < 0.2$ somewhere inside of $\mathcal{P}$.

Example 3 *This example illustrates why we assume that $\mathcal{P}$ is closed in Theorem 1(i). Let Ω consist of three states. Let*

$$\begin{aligned}\mathcal{P} &= \{(p_1, p_2, p_3) : p_2 < 2p_1 \text{ for } p_1 \le 0.2\} \\ &\quad \cup \{(p_1, p_2, p_3) : p_2 \le 2p_1 \text{ for } 0.2 < p_1 \le 1/3\}.\end{aligned}$$

The set of acts $\mathcal{A}$ contains only the following three acts (each expressed as a vector of its payoffs in the three states):

$$\begin{aligned} f_1 &= (0.2, 0.2, 0.2), \\ f_2 &= (1, 0, 0), \\ g &= (-1.8, 1.2, .2). \end{aligned}$$

Notice that $E_p(f_2)$ is the highest of the three whenever $p_1 \ge 0.2$, $E_p(f_1)$ is the highest whenever $p_1 \le 0.2$, and $E_p(g)$ is never the highest. So, $\mathcal{B} = \{f_1, f_2\}$ and g is not Bayes with respect to $\mathcal{A}$. For each $0 \le \alpha \le 1$, we compute

$$\begin{aligned} E_p(\alpha f_1 + (1-\alpha) f_2) &= p_1(1-\alpha) + 0.2\alpha, \\ E_p(g) &= -2p_1 + p_2 + 0.2. \end{aligned}$$

Notice that $E_p(\alpha f_1 + (1-\alpha) f_2)$ is strictly greater than $E_p(g)$ if and only if $p_2 < (3-\alpha)p_1 - 0.2(1-\alpha)$. There is no α such that this inequality holds for all $p \in \mathcal{P}$.

Remark 1 *Note that is it irrelevant to this example that $p_2 = 0$ for some $p \in \mathcal{P}$.*

Definition 3 *Say that two convex sets* intersect all the same supporting hyperplanes *if they have the same closure and a supporting hyperplane intersects one convex set if and only if it intersects the other.*

In addition to showing that *E*-admissibility does not reduce to pairwise comparisons even when the option set is convex, this example also brings out the important point the *E*-admissibility (but not Maximality) can distinguish between some convex sets that intersect all the same supporting hyperplanes. As we noted some years ago (Section III of [15]), the strict preference relation induced by Criterion 1 cannot distinguish between pairs of convex sets that intersect all the same supporting hyperplanes. Of course, Γ-Maximin does even worse than Maximality, as it cannot distinguish open convex sets from their closure.

Figure 2 illustrates Example 3 and that the presence or absence of probability point $D = (0.2, 0.2, 0.4)$ determines whether or not act g is Bayes from the choice set $\mathcal{A} = \{f_1, f_2, g\}$. The closure of the convex set $\mathcal{P}$ is the triangle with extreme points $A = (1/3, 0, 2/3)$, $B = (1/3, 2/3, 0)$, and $C = (0, 0, 1)$. In Example 3, set $\mathcal{P}$ is the result of removing the closed line segment $[C, D]$ from the left face $[B, C]$ of the triangle ABC, leaving the half-open line segment $[B, D)$ along that face. The convex set $\mathcal{P}^*$ is the set of probabilities that results by adding point D to

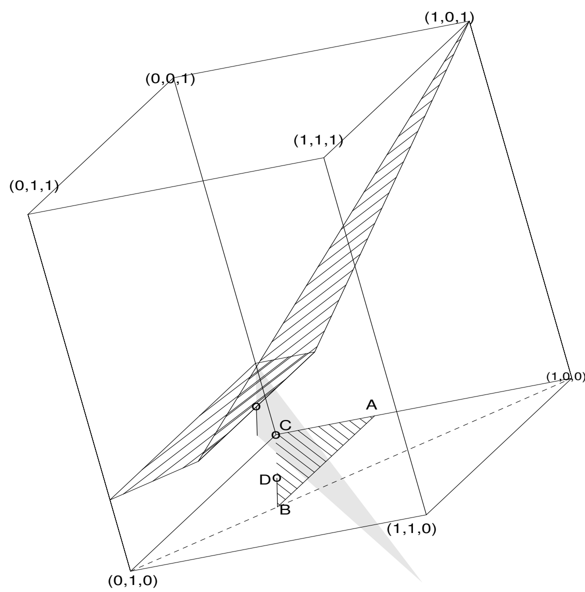

Figure 2: Illustration for Example 3. The set of (p_1, p_2) such that $(p_1, p_2, 1-p_1-p_2) \in \mathcal{P}$ is the diagonally shaded set inside the probability triangle at the bottom of the figure with the points *A*, *B*, *C*, and *D* that are discussed in the text labeled. The diagonally shaded surface is the surface of Bayes solutions for all probabilities (not just those in $\mathcal{P}$). The solid shaded set is $\{(p_1, p_2, E_p(g)) : p \in \mathcal{P}\}$. The points $(0.2, 0.4)$, $(0,0)$, and $(0.2, 0.4, E_p(g))$ are indicated by open circles.

set $\mathcal{P}$. Point *D* then is an extreme but not exposed point in $\mathcal{P}^*$. Evidently, $\mathcal{P}$ and $\mathcal{P}^*$ intersect all the same supporting hyperplanes. Next, we indicate how to use *E*-admissibility to distinguish between these two convex sets of probabilities.

For this exercise, we bypass the details of what can easily be done with pairwise comparisons to fix the common boundaries of $\mathcal{P}$ and $\mathcal{P}^*$. Specifically, binary comparisons suffice to fix the closed interval [*A*,*B*] belongs to both sets, as the upper probability $\overline{P}(\omega_1) = 1/3$; they suffice to fix that point *C* does *not* belong to either set, as the lower probability $\underline{P}(\omega_1) > 0$; they suffice to fix the half-open interval $[A, C)$ belongs to both sets, as the lower probability $\underline{P}(\omega_2) = 0$, and they suffice to fix the half open interval $[B, C)$ as a boundary for both sets, as the upper called-off (conditional) odds ratio $\overline{P}(\omega_1|\{\omega_1, \omega_2\}) > 1/3$. But pairwise comparisons according to Criterion 1, alone, cannot determine *how much* of the half-open interval $[B, C)$ belongs to either set $\mathcal{P}$ or $\mathcal{P}^*$. For that, we use non-binary choice problems and *E*-admissibility.

In order to establish that the half open line segment [*C*, *D*) does not belong to either set $\mathcal{P}$ or $\mathcal{P}^*$, consider the family of decision problems defined by the three-way choices: $\mathcal{A}_{-\varepsilon} = \{f_1, f_2, g_{-\varepsilon}\}$, where $g_{-\varepsilon}$ is the act with payoffs $(1.8, 1.2 - \varepsilon, 0.2)$. For each $\varepsilon > 0$, only the pair $\{f_1, f_2\}$ is *E*-admissible from such a three-way choice, with respect to each of the two convex sets of probabilities.

Likewise, in order to establish that the half-open line segment $(D, B]$ belongs to both sets, $\mathcal{P}$ and $\mathcal{P}^*$, consider the family of decision problems defined by the three-way choices: $\mathcal{A}_{+\varepsilon} = \{f_1, f_2, g_{+\varepsilon}\}$, where $g_{+\varepsilon}$ is the act with payoffs $(1.8, 1.2 + \varepsilon, .2)$. For each $\varepsilon > 0$, all three options are E-admissible with respect to each of the two convex sets of probabilities.

However, in the decision problem with options $\mathcal{A} = \{f_1, f_2, g\}$, as shown above, only the pair $\{f_1, f_2\}$ is *E*-admissible with respect to the convex set $\mathcal{P}$, whereas all three options are *E*-admissible with respect to the convex set $\mathcal{P}^*$.

By contrast, given a choice set, Maximality makes the same ruling about which options are admissible from that choice set, regardless whether convex set $\mathcal{P}$ or convex set $\mathcal{P}^*$ is used. That is, Maximality cannot distinguish between these two convex sets of probabilities in terms of admissibility of choices, as the two convex sets of probabilities intersect all the same supporting hyperplanes.

3 Summary

The discussion here contrasts three decision rules that extend Expected Utility and which apply when uncertainty is represented by a convex set of probabilities, $\mathcal{P}$, rather than when uncertainty is represented only by a single probability distribution. The decision rules are: Γ-Maximin, Maximality, and *E*-admissibility. We show that these decision rules have different operational content in terms of their ability to distinguish different convex sets of probabilities. When do the admissible choices differ for different convex sets of probabilities? Γ-Maximin is least sensitive among the three in this regard. We show that, even when the option set is convex, one decision rule (*E*-admissibility) distinguishes among more convex sets than the other two. This is because it alone among these three is not based on pairwise comparisons among options. The upshot it that it, but neither of the other two rules, can distinguish between two convex sets of probabilities that intersect all the same supporting hyperplanes.

A Proofs of Theorem 1 and Corollary 1

The proofs rely on some lemmas about convex sets.

Lemma 1 *Let k be a positive integer. Let C be a closed convex subset of $\mathbb{R}^k$ that contains the origin. There exists a unique closed convex subset D of $\mathbb{R}^{k+1}$ with the following properties:*

- $C = \{x \in \mathbb{R}^k : \alpha^\top x \geq c,$ *for all* $(\alpha, c) \in D\}$.
- $(\alpha, c) \in D$ *implies* $(a\alpha, ac) \in D$ *for all* $a \geq 0$,
- $(\alpha, c) \in D$ *implies* $(\alpha, c - a) \in D$ *for all* $a > 0$,

Also, for each $(\alpha, c) \in D$, $c \leq 0$.

Proof. To see that $(\alpha, c) \in D$ implies $c \leq 0$, let $\mathbf{0}$ be the origin. Then $\alpha^\top \mathbf{0} = 0 \geq c$. Define the following set

$$D_0 = \{(\alpha, c) : \alpha^\top x \geq c, \text{ for all } x \in C\}. \tag{1}$$

To see that D_0 is convex, let (γ_1, d_1) and (γ_2, d_2) be in D_0 and $0 \leq \beta \leq 1$. Then, for all $x \in C$,

$$(\beta\gamma_1 + [1 - \beta]\gamma_1)^\top x \geq \beta d_1 + (1 - \beta) d_2.$$

This means that $\beta(\gamma_1, d_1) + [1 - \beta](\gamma_2, d_2) \in D_0$, and D_0 is convex. To see that D_0 is closed, notice that $D_0 = \bigcap_{x \in C} D_x$, where $D_x = \{(\alpha, c) : \alpha^\top c \geq c\}$ and each D_x is closed. It is clear that D_0 has the last two properties in the itemized list. For the first condition, let E be the set defined in the first condition. It is clear that $C \subseteq E$. Suppose that there is $x_0 \in E$ such that $x_0 \notin C$. Then there is a hyperplane that separates $\{x_0\}$ from C. That is, there is $\gamma \in \mathbb{R}^k$ and d such that $\gamma^\top x \geq d$ for all $x \in C$ and $\gamma^\top x_0 < d$. It follows that $(\gamma, d) \in D_0$, but then $x_0 \notin E$, a contradiction.

To see that the set that satisfies the conditions is unique, suppose that D and F are both sets satisfying the listed conditions. If $F \neq D$, then there is (α, c) either in $D \setminus F$ or in $F \setminus D$. We will show, by way of contradiction, that neither of these cases can occur. The two cases are handled the same way. We will do only the first. In the first case, there is a hyperplane separating $\{(\alpha, c)\}$ from F. That is, there is (γ, d, f) with $\gamma \in \mathbb{R}^k$ and $d, f \in \mathbb{R}$ such that

$$\gamma^\top \delta + dg \geq f, \text{ for all } (\delta, g) \in F, \tag{2}$$

and $\gamma^\top \alpha + dc < f$. It follows that $a\gamma^\top \delta + da(g - b) \geq f$ for all $(\delta, g) \in F$ and all $a, b > 0$. As $a \to 0$, we see that $f \leq 0$ is required. As $b \to \infty$, we see that $d \leq 0$ is required. As $a \to \infty$ we see that $\gamma^\top \delta + dg \geq 0$ for all $(\delta, g) \in F$, hence we can assume that $f = 0$. Because $d, c \leq 0$ and $\gamma^\top \alpha + dc < 0$ it follows that $\gamma^\top \alpha < 0$ and there exists $d_0 < 0$ such that $\gamma^\top \alpha + d_0 c < 0$. Because $g \leq 0$ for all $(\delta, g) \in F$, we see that, even if $d = 0$, $\gamma^\top \delta + d_0 g \geq 0$ for all $(\delta, g) \in F$. Hence, we can assume that the separating hyperplane has the form $(\gamma, d_0, 0)$ with $d_0 < 0$. Define $\gamma_0 = \gamma/(-d_0)$. It follows from (2) that $\delta^\top \gamma_0 \geq g$ for all $(\delta, g) \in F$ and so $\gamma_0 \in C$. Hence $\alpha^\top \gamma_0 \geq c$ which contradicts $\gamma^\top \alpha + d_0 c < 0$. □

Lemma 2 *Let* V *be a closed convex subset of* $\mathbb{R}^{k+1}$, *and express elements of* V *as* (α, d) *where* $\alpha \in \mathbb{R}^k$ *and* d *is real. Define*

$$A = \{x \in \mathbb{R}^k : \alpha^\top x \geq d, \textit{ for all } (\alpha, d) \in V\}.$$

Assume that A is nonempty. Define D to be the set of all vectors in $\mathbb{R}^{k+1}$ *of the form* $(a\alpha, ad-b)$ *with* $a,b \geq 0$ *and* $(\alpha,d) \in V$*. Then* $D = \{(\alpha,d) \in \mathbb{R}^{k+1} : \alpha^\top x \geq d,$ *for all* $x \in A\}$.

Proof. Let $x_0 \in A$, and define

$$\begin{aligned} C &= \{x - x_0 : x \in A\}, \\ V' &= \{(\alpha, d - \alpha^\top x_0) : (\alpha,d) \in V\}. \end{aligned}$$

It follows that

$$C = \{x \in \mathbb{R}^k : \alpha^\top x \geq c, \text{ for all } (\alpha,c) \in V'\}, \tag{3}$$

and C contains the origin and is a closed convex set. Define $D_1 = \{(\alpha, d - \alpha^\top x_0) : (\alpha,d) \in D\}$. In other words, D_1 is the convex closed convex set of all vectors in $\mathbb{R}^{k+1}$ of the form $(a\alpha, ac-b)$ with $a,b \geq 0$ and $(\alpha,c) \in V'$. The definitions of D and D_1 were rigged so that D_1 satisfies all the conditions required of the set called D in Lemma 1 except possibly the first condition in the itemized list. To verify this condition, define

$$C' = \{x \in \mathbb{R}^k : \alpha^\top x \geq c, \text{ for all } (\alpha,c) \in D_1\}.$$

To see that $C \subseteq C'$, let $x \in C$. Then $a\alpha^\top x \geq ac - b$ for all $(\alpha,c) \in V'$ and all $a,b \geq 0$. Hence, $\alpha^\top x \geq c$ for all $(\alpha,c) \in D_1$. To see that $C' \subseteq C$, let $x \in C'$. Since $(\alpha,c) \in V'$ implies $(\alpha,c) \in D_1$, we have $\alpha^\top x \geq c$ for all $(\alpha,c) \in D_1$ and $x \in C$. It follows from Lemma 1 that D_1 is the set D_0 defined in (1) and D is the claimed set as well. □

Proof of Theorem 1. (i) Let

$$U = \left\{ x \in \mathbb{R}^k : \left(x_1, \ldots, x_{k-1}, 1 - \sum_{j=1}^{k-1} x_j \right) \in \mathcal{P} \right\}.$$

Let $\mathcal{C}'$ be the convex hull of $\mathcal{B}$. Let V consist of all points of the form (α_f, c_f) where $f \in \mathcal{C}'$. Let A be as defined in the statement of Lemma 2. Since g is not Bayes with respect to $\mathcal{P}$, the set $H_g = \{x \in U : \alpha_g^\top x = c_g\}$ does not intersect A. Now, notice that H_g and A are disjoint closed convex sets, hence there is a separating hyperplane. That is, there exists a nonzero $\gamma \in \mathbb{R}^k$ and c such that $\gamma^\top x \geq c$ for all $x \in A$ and $\gamma^\top y < c$ for all $y \in H_g$. Because $\gamma^\top x \geq c$ for all $x \in A$, it follows from Lemma 2 that (γ, c) is in the set D defined in the statement of Lemma 2. Hence, $\gamma = a\alpha$ and $c = ad - b$ for some $(\alpha,c) \in V$ and some $a,b \geq 0$. Because γ is nonzero, we have $a > 0$ and we can assume without loss of generality that $a = 1$ and $(\gamma, d-b) \in V$. So, $\gamma = \alpha_h$ for some $h \in \mathcal{C}'$ and $c = c_h - b$, and we can assume without loss of generality that $b = 0$ and $c = c_h$. Now, for all real t,

$$\alpha_h^\top (p_1, \ldots, p_{k-1}, t) = h(\omega_k) - E_p(h) + t = c_h - E_p(h) + t.$$

So, for all $x \in H_g$,

$$c_h > \alpha_h^\top x = c_h - E_p(h) + E_p(g).$$

It follows that, for all $p \in \mathcal{P}$, $E_p(h) > E_p(g)$.

(ii) Define U, $\mathcal{C}'$, V, A, and H_g exactly as in the proof of part (i). It is still true that H_g and A are convex, that A is closed, and that H_g does not intersect A. But H_g is now relatively open. That is, it is the intersection of the hyperplane $H'_g = \{x : \alpha_g^\top x = c_g\}$ with an open set. For this reason, H'_g is the unique hyperplane that contains H_g. Of course, if the closure of H_g fails to intersect A, the rest of the proof of part (i) continues to work. So, suppose that the closure $\overline{H_g}$ of H_g intersects A. Even so, there is a weakly separating hyperplane (γ, c), i.e., there is a $\gamma \in \mathbb{R}^k$ and c such that $\gamma^\top x \geq c$ for all $x \in A$ and $\gamma^\top y \leq c$ for all $y \in H_g$. We need to show that among all such separating hyperplanes, there is at least one such that the second inequality is strict, i.e., at least one of the separating hyperplanes fails to intersect H_g. Then the rest of the proof of part (i) will finish the proof.

Because $\mathcal{A}$ is finite, A is the intersection of finitely many closed halfspaces, and each of these halfspaces is of the form $\{x : \alpha_f^\top x \geq c_f\}$ for some $f \in \mathcal{B}$. Now, $\overline{H_g}$ intersects A in some convex subset of the union of the hyperplanes that determine these halfspaces. No subset of the union of finitely many distinct hyperplanes can be convex unless it is contained in the intersection of one or more of the hyperplanes. (Just check that $\alpha x + (1-\alpha)y$ is in the same hyperplane with x if and only if y is as well.) Hence, $A \cap \overline{H_g}$ is a subset of the intersection of one or more of the hyperplanes of the form $H'_f = \{x : \alpha_f^\top x = c_f\}$ for some $f \in \mathcal{B}$. Define

$$W = \{f \in \mathcal{B} : A \cap \overline{H_g} \subset H'_f\}.$$

If $W = \mathcal{B}$, then $H_g \subset A$, a contradiction. Let $f_0 \in \mathcal{B} \setminus W$ be such that H'_{f_0} is closest to $A \cap \overline{H_g}$. Such f_0 exists because $\mathcal{B}$ is finite. Let ε be one-half of the distance from H'_{f_0} to $A \cap \overline{H_g}$, and define

$$O = \{x : \|x - A \cap \overline{H_g}\| < \varepsilon\}.$$

Then

$$T = O \cap \left(\bigcap_{f \in W} \{x : \alpha_f^\top x \geq c_f\} \right) \subset A.$$

For each $f \in W$, define $M_f = \{x \in H_g : \alpha_f^\top x \geq c_f\}$. If at least one $M_f = \emptyset$, then H'_f fails to intersect H_g, and the proof is complete. So assume, to the contrary, that every $M_f \neq \emptyset$. Then for each f, the closure $\overline{M_f}$ of M_f contains $A \cap \overline{H_g}$. It follows that each M_f contains points in every neighborhood of $A \cap \overline{H_g}$, including O. Hence, for each f, there exists $x \in T \cap M_f$. Each such $x \in H_g \cap A$, a contradiction. □

Proof of Corollary 1. Let $\mathcal{C}$ be the convex hull of $\mathcal{B}$. Either assumption (i) or (ii) is strong enough to imply that Theorem 1 applies, hence there is $h' \in \mathcal{C}$ such that $E_p(h') > E_p(g)$ for all $p \in \mathcal{P}$. If $h' \notin \mathcal{B}$ let $h_1 = h'$, and apply Theorem 1 repeatedly in a transfinite induction as follows. At each successor ordinal $\gamma + 1$, find $h_{\gamma+1} \in \mathcal{C}$ such that $E_p(h_{\gamma+1}) > E_p(h_\gamma)$ for all $p \in \mathcal{P}$. At a countable limit ordinal γ choose any countable sequence $\{\gamma_n\}_{n=1}^{\infty}$ of ordinals that is cofinal with γ. By the induction hypothesis, $E_p(h_{\gamma_i}) < E_p(h_{\gamma_j})$ for all $p \in \mathcal{P}$ if $i < j$. The sequence $\{h_{\gamma_n}\}_{n=1}^{\infty}$ belongs to the closed bounded set $\mathcal{A}$, hence it has a limit h_γ and

$$E_p(h_\gamma) = \lim_{n \to \infty} E_p(h_{\gamma_n}) = \sup_n E_p(h_{\gamma_n}),$$

for all p, and hence does not depend on which limit point we take. Also, $\sup_n E_p(h_\gamma) > E_p(h_\alpha)$ for all $\alpha < \gamma$, so we continue to satisfy the induction hypothesis. Since $\mathcal{A}$ is bounded, there cannot exist an uncountable increasing sequence of $E_p(h_\gamma)$ values, hence the transfinite induction terminates at some countable ordinal γ_0 with $h_{\gamma_0} \in \mathcal{B}$.

References

[1] Berger, J. O. *Statistical Decision Theory and Bayesian Analysis*, 2 ed. Springer-Verlag, New York, 1985.

[2] Gilboa, I., and Schmeidler, D. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18* (1989), 141–153.

[3] Giron, F. J., and Rios, S. Quasi-Bayesian behavior: A more realistic approach to decision making? In *Bayesian Statistics*, e. a. J. M. Bernardo, Ed. University of Valencia Press, Valencia, 1980, pp. 17–38.

[4] Good, I. J. Rational decisions. *Journal of the Royal Statistical Society 14* (1952), 117–114.

[5] Grunwald, P. D., and Dawid, A. P. Game theory, maximum entropy, minimum discrepancy, and robust bayesian decision theory. Tech. Rep. 223, University College London, 2002.

[6] Herzberger, H. Ordinal preferences and rational choice. *Econometrica 41* (1973).

[7] Keynes, J. M. *A Treatise on Probability*. MacMillan and Co., London, 1921.

[8] Levi, I. On indeterminate probabilities. *Journal of Philosophy 71* (1974), 391–418.

[9] Levi, I. *The Enterprise of Knowledge*. MIT Press, Cambridge, MA, 1980.

[10] Levi, I. Imprecise and indeterminate probabilities. In *Proceedings of the 1st International Symposium on Imprecise Probabilities and Their Applications*. University of Ghent, Ghent, Belgium, 1999.

[11] Levi, I. Value commitments, value conflict, and the separability of belief and value. *Philosophy of Science 66* (1999), 509–533.

[12] Savage, L. J. *The Foundations of Statistics*. Wiley, New York, 1954.

[13] Seidenfeld, T. Comment on "Probability, evidence, and judgment" by A. P. Dempster. In *Bayesian Statistics 2: Proceedings of the Second Valencia International Meeting, 1983*, e. a. J. M. Bernardo, Ed. North-Holland, Amsterdam, 1985, pp. 127–129.

[14] Seidenfeld, T., and Schervish, M. J.and Kadane, J. A representation of partially ordered preferences. *Annals of Statistics 23* (1995), 2168–2217.

[15] Seidenfeld, T., Schervish, M. J., and Kadane, J. Decisions without ordering. In *Acting and Reflecting*, W. Sieg, Ed. Kluwer Academic Publishers, Dordrecht, 1990, pp. 143–170.

[16] Sen, A. K. Social choice theory: a re-examination. *Econometrica 45* (1977), 53–89.

[17] Walley, P. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1990.

M. J. Schervish is with Carnegie Mellon University, USA

T. Seidenfeld is with Carnegie Mellon University, USA

J. B. Kadane is with Carnegie Mellon University, USA

I. Levi is with Columbia University, USA