Expected Utility in 3D

Jean Baccelli*

Munich Center for Mathematical Philosophy

July 29, 2021

Abstract

Consider a subjective expected utility preference relation. It is usually held that the representations which this relation admits differ only in one respect, namely, the possible scales for the measurement of utility. In this paper, I discuss the fact that there are, metaphorically speaking, two additional dimensions along which infinitely many more admissible representations can be found. The first additional dimension is that of state-dependence. The second—and, in this context, much lesser-known—additional dimension is that of act-dependence. The simplest implication of their usually neglected existence is that the standard axiomatizations of subjective expected utility fail to provide the measurement of subjective probability with satisfactory behavioral foundations.

*I am grateful to Thomas Augustin, Fabio Cozman, and Gregory Wheeler for inviting this contribution and offering me the opportunity to celebrate Teddy Seidenfeld's work. I have learnt immensely from Teddy's writings, and I have also come to owe him a lot for his time and advice. I am especially thankful for the many inspiring conversations he had the kindness to have with me during the year 2019-2020, while I was visiting Pittsburgh.

For helpful comments on earlier drafts of the present paper, I am indebted to two anonymous reviewers, Richard Bradley, Florian Brandl, Franz Dietrich, Raphaël Giraud, Brian Hill, Jay Lu, Marcus Pivato, Rush Stewart, Jiji Zhang, Fanyin Zheng, as well as the late and regretted Arthur Merin and Philippe Mongin. All errors and omissions are mine.

Introduction

In this paper, I discuss Savage's axiomatization of subjective expected utility (Savage, 1954; 1972). I will assume without discussion that the main function of this representation theorem is to provide behavioral foundations for subjective probability measures—e.g., the statisticians' priors appearing in Bayesian statistics. Throughout, I refer to Savage merely because his result is the prime example of its kind in decision theory; mutatis mutandis, all of the following applies to Ramsey, de Finetti, Anscombe and Aumann, and anyone in the business of "explicat[ing] [beliefs] in the language of bets" (Seidenfeld et al., 1990b, p. 523).

I will be concerned here exclusively with the uniqueness clause of Savage's theorem. Accordingly, unless stated otherwise, I will assume throughout the paper that the existence conditions of Savage's representation, i.e., the Savage axioms, are satisfied. Sticking to Savage's own terminology, this assumption can be expressed as follows. Let $\succcurlyeq$ denote the binary preference of a decision-maker over a set F of acts, with $\sim$ and $\succ$ denoting indifference and strict preference, respectively. The set F has the structure $F \equiv X^S$ with S, the set of all possible states of nature and X, the set of all possible consequences of the decision-maker's acts. Although the Savage theorem famously requires S to be infinite, throughout this paper, I will discuss it as if it allowed for a finite S, with $S = \{s_1, \ldots, s_n\}$, and I will take the maximal algebra of events 2^S. This is to simplify exposition and, for discussing the uniqueness of the representation, without loss of generality.[1] Accordingly, within this paper, I will say that $\succcurlyeq$ satisfies the Savage axioms if and only if there exists a (non-constant) utility function $u : X \to \mathbb{R}$ and a probability

[1]Besides, although the Savage theorem can be invoked only when S is infinite, reasonably close variants of this theorem can be invoked when S is finite (see, e.g., Köbberling and Wakker, 2003 and the references therein). In these variants, unlike in the Savage theorem, instead of being imposed on S, richness requirements are typically imposed on X.

measure $p : 2^S \to \mathbb{R}$ such that, for any f and $g \in F$, (1) holds:

$$f \succcurlyeq g \quad \Leftrightarrow \quad \sum_{i=1}^{n} p(s_i) \cdot u\big(f(s_i)\big) \geq \sum_{i=1}^{n} p(s_i) \cdot u\big(g(s_i)\big). \tag{1}$$

The standard uniqueness clause of the Savage theorem is that in (1), p is absolutely unique and u is unique up to an increasing affine transformation, i.e., u can be replaced by v if and only if $v = au + b$, for some $a \in \mathbb{R}_{>0}$, $b \in \mathbb{R}$.

What is said in the standard uniqueness clause of the Savage theorem is, of course, correct; but it is also significantly incomplete. As I will explain and discuss, there are—metaphorically speaking—two additional dimensions along which infinitely many other admissible representations of $\succcurlyeq$ can be found and across which the uniqueness of p, in particular, will be lost. The first additional dimension is that of *state-dependence*. The second additional dimension is that of *act-dependence* and it is by far, in the present state of the literature, the less familiar of the two dimensions. The most important fact in what follows will be that preferences satisfying the Savage axioms are compatible not only with the kind of representation given in (1), i.e., a state-independent and act-independent representation, but also with an infinity of alternative state-dependent and/or act-dependent representations. One major implication will be that the Savage theorem—once again: taken as the prime example of its kind in decision theory—fails to provide subjective probability measures with compelling behavioral foundations. I will be concerned here almost exclusively with mathematically establishing the above fact, because it is only incompletely understood in the current literature. Although this fact is of clear philosophical significance, I will not be able to elaborate here on the conceptual or the methodological implications.

The paper is organized as follows. Section 1 examines state-dependence. This dimension is now relatively well understood in the literature, thanks in particular to classic work by Teddy Seidenfeld and co-authors (Seidenfeld et al., 1990b; Schervish et al., 1990; 2013). I will only need to repeat their statement of the main problem, with two minor clarifications which I will contribute. Section 2 examines act-dependence. This dimension is currently not well known. To explore act-dependence and understand the symmetries or asymmetries with state-dependence, I will build on recent work by David Dillenberger and co-authors (Dillenberger et al., 2017). Their work will enable me to offer novel insights on act-dependence. A brief conclusion ensues.

1 State-Dependence

Consider a preference relation representable as in (1). For simplicity, assume, to start with, that p in (1) has full support over S, i.e., p gives strictly positive probability to all $s \in S$. Notice that, for any alternative full support q, with the definition $v_i(x) = \frac{p(s_i)}{q(s_i)} u(x)$, for $i = 1, \ldots, n$, equality (2) holds:

$$\sum_{i=1}^{n} p(s_i) \cdot u\big(f(s_i)\big) = \sum_{i=1}^{n} q(s_i) \cdot \frac{p(s_i)}{q(s_i)} u\big(f(s_i)\big) = \sum_{i=1}^{n} q(s_i) \cdot v_i\big(f(s_i)\big). \quad (2)$$

Therefore, with $\succcurlyeq$ the same preference relation as in (1), we have that, for any f and $g \in F$, (3) holds:

$$f \succcurlyeq g \quad \Leftrightarrow \quad \sum_{i=1}^{n} q(s_i) \cdot v_i\big(f(s_i)\big) \geq \sum_{i=1}^{n} q(s_i) \cdot v_i\big(g(s_i)\big). \quad (3)$$

The key point in (2) is that the v_i correspond to *state-dependent* increasing affine transformations of u from (1), and that such transformations, too, prove compatible with the respect of the Savage axioms—including P3 and P4, i.e., the axioms supposed to enforce state-independence in the Savage axiomatics.[2] Indeed, if q is coupled with the collection of v_i as constructed

[2] In general, seeing u as a function from X to $\mathbb{R}^{|S|}$, a state-dependent increasing affine transformation is of the form $\alpha u + \beta$, with $\alpha \in \mathbb{R}^{|S|}_{>0}$ and $\beta \in \mathbb{R}^{|S|}$. For simplicity, we here take $\beta = 0$ and a suitably restricted domain for α; but the problem applies more generally, with an order-preserving renormalization. Notice the parallel with the case of cardinal non-comparable utilities in social choice theory (e.g., d'Aspremont and Gevers, 2002, p. 485).

above, then, not only will their composition respect the Savage axioms; it will also be behaviorally indistinguishable from the product of the functions p and u initially given in (1). To this extent, Savage's approach cannot distinguish between p and q or, indeed, any pair of full support priors. Notice in particular that, following (2), one could associate any preference relation representable as in (1) with the same prior—say, the uniform prior over S. Accordingly, although, by assumption, their behavior differs, all the Savage decision-makers can be shown to share the same prior. This vividly illustrates the failure of the Savage theorem—once again: taken as the prime example of its kind in decision theory—to provide subjective probability measures with compelling behavioral foundations.

The argument just sketched is now standard in the literature, thanks to its being stressed by many over the years, and particularly forcefully by Teddy Seidenfeld and co-authors (see Seidenfeld et al., 1990b, p. 521; Schervish et al., 1990, p. 840, 2013, p. 508; see also, most importantly and in order of historical precedence, Drèze, 1961, Savage and Aumann, 1987, Arrow, 1974, and Karni, 1996). However, I would like to add two minor qualifications to the way Teddy Seidenfeld and co-authors typically present the issue. This is not to criticize these authors, whom I take to be aware of the necessary qualifications; this is merely to help their readers in directly appreciating the full extent of the problem.

First, I want to dispel the impression according to which, in (1), there is a substantial asymmetry between the states to which the Savage distribution p gives probability value zero and the states to which it gives a strictly positive probability value. Following Savage's terminology, I will call the former null states and the latter, non-null states. Second, I want to dispel the impression according to which, in the identification issues raised by (2), there is a substantial symmetry between the implications on the measurement of probability and the implications on the measurement of utility. For convenience, I will also informally refer to the former as the measurement of beliefs and the latter, the measurement of desires.

I discuss null states first. Teddy Seidenfeld and co-authors comment on (2) as follows: "for each coherent system of preferences $\prec$, the family of possible (...) probability/utility pairs that agree with $\prec$ according to expected utility is constrained solely by probability-0 (null) states" (Seidenfeld et al., 1990b, p. 521).[3] As I now show, the constraint is even weaker than what this statement suggests.

Assume now that, in (1), $p(s_{i^*}) = 0$ for some $i^* \in \{1, \dots, n\}$. Then, (2) applies all the same. The only difference is that by definition, for all $x \in X$, $v_{i^*}\big(x\big) = 0$. More generally, with c_{i^*} any constant from $\mathbb{R}$, the representation of $\succcurlyeq$ could be preserved with, for all $x \in X$, $v_{i^*}\big(x\big) = c_{i^*}$.[4] The key implication of this simple observation is that, even though p in (1) does not have full support over S, one can, following (2), construct behaviorally indistinguishable state-dependent representations featuring an alternative probability q that has full support over S. This is because, in a nutshell, zero probability values and constant state-dependent utilities are not behaviorally distinguishable from one another. The upshot is that even the support of the subjective probability measure is not behaviorally identified.

Of course, given the properties of expected utility, the support of p in (1) places constraints on the support of any alternative function q in (3). However, the only property that is common to all the priors q suitably compatible with the preferences given in (1) is the following weak property. For any

[3] I will eventually return (on p. 20) to the ellipsis in the above quotation.

[4] First, apply (2), thus inducing (3) with $v_{i^*}\big(x\big) = 0$ for all $x \in X$. All expected utility levels being preserved, so is the representation. Next, add the constant $q(s_{i^*})c_{i^*}$ to both sides of the inequality in (3). The inequality being preserved, so is the representation.

$s \in S$, if s is non-null, then, $q(s) > 0$; if s is null, then, $q(s) < 1$.[5] Therefore, with p the prior featured in the state-independent representation (1), the constraint characterizing the set of all priors q featured in some state-dependent alternative representation (3) is as follows: $\text{supp}(q) \supseteq \text{supp}(p)$. One notable implication of this characterization is that—however interesting it may be to conventionally restrict one's attention to such sets—the set P of all the priors suitably compatible with the preferences given in (1) is *not* a set of mutually absolutely continuous probability measures.[6] That is to say, it is not such that for any pair $q, q' \in P$ and any event E in the event algebra 2^S, if $q(E) = 0$, then, $q'(E) = 0$.

Second, Teddy Seidenfeld and co-authors make this other comment on (2): "it is difficult to justify treating probability values (...) as if they had meaning independently of the utility (...) function. Similarly, the utility function values do not have meaning independently of the probability of the events" (Schervish et al., 2013, p. 509). What follows focuses on the second half of the comment, more specifically, the parallel with the first. From one angle at least, this parallel is mathematically unquestionable; but it could lead to misappreciating one thought-provoking asymmetry, which I now highlight.

Consider (1), assuming again, to start with, that there is no null state. The state-dependent increasing affine transformation in (2) shows that the beliefs are essentially unidentified. But, in one important respect at least, it proves inconsequential as far as the identification of desires is concerned.

[5] In the first case, if $q(s_i) = 0$, then for any state-dependent utility function v_i it holds that $q(s_i)v_i(x) = 0$ for all $x \in X$, while to preserve the representation of the relation given in (1) with $p(s_i) > 0$, it must be that $q(s_i)v_i(x) \neq q(s_i)v_i(y)$ for some $x, y \in X$. In the second case, if $q(s_i) = 1$, then for any state-dependent utility function v_i it holds that $q(s_i)v_i(x) \neq q(s_i)v_i(y)$ for some $x, y \in X$, while to preserve the representation of the relation given in (1) with $p(s_i) = 0$, it must be that $q(s_i)v_i(x) = q(s_i)v_i(y)$ for all $x, y \in X$. Clearly, the converse of the two implications thus established do not hold; to see this, consider the case of a constant v_i and that of p and q having the same support, respectively.

[6] This is contrary to what is suggested by some passages in Schervish et al., 1990; 2013, such as the one quoted at the beginning of this development. The authors are nonetheless clearly aware of the necessary qualifications (see, e.g., Seidenfeld et al., 1995, p. 2174: "only when $p(s_i) \neq 0$ is it worth restricting U_j in a decomposition of a linear utility V").

Indeed, whatever the collection of state-dependent increasing affine transformations, the consequences and their utility differences (of first or higher order) will be ranked the same way in all states. Likewise, whenever defined, the so-called Arrow-Pratt index of risk aversion $I(x) \equiv -u''(x)/u'(x)$ (see Arrow, 1963, Pratt, 1964, and, e.g., Eeckhoudt et al., 2005) will be the same in all states. But, for most decision theorists, this is simply all there is, in itself, to the measurement of utility. Let me clarify that, in saying so, I do not mean to endorse a conceptual claim about desires, to the effect that it would not make sense to even aim at measuring, e.g., utility ratios or, beyond, absolutely unique utility values. I am merely making a practical observation about decision theory, according to which, absent any interest for the joint measurement of subjective probability (or, in a different context, interpersonal comparisons of utility values), the measurement of utility does not need to go beyond what is already achieved in (1) and, in fact, preserved by (2) in (3). Accordingly, in a nutshell, one can claim that, although the choice data in (1) leaves the underlying beliefs essentially unidentified, it essentially identifies the underlying desires. This is the promised thought-provoking asymmetry.[7]

If there are null states, the claim made above must be qualified by adding that the contrast which I am highlighting arises only insofar as non-null states are concerned. This is because, when there are null states, one can certainly construct state-dependent alternatives to (1) in which the null-state-dependent transformations will be non-increasing and/or non-affine—the previously mentioned case of conditionally constant utilities being the simplest example of all. Accordingly, as far as null states are concerned, the utility values are entirely unidentified, too. This last point is made clearly by Teddy Seidenfeld and co-authors (see, e.g., Schervish et al., 1990, p. 840).

[7] On which see Nau, 2001. Nau draws many other interesting implications from the key underlying fact here, viz. that up to a standard affine transformation of u, the compound summands or state-values in (1) are uniquely identified, irrespective of how each of them is to be further decomposed, state by state, in probability and utility values, respectively.

Once again, the foregoing qualifications are mere clarifications. Furthermore, overall, the main lesson to draw from (2) is exactly the one which Teddy Seidenfeld and co-authors forcefully emphasized.[8] To wit, the Savage axioms fail to provide subjective probability measures with satisfactory behavioral foundations because they exclude only some, not all forms of state-dependent utility. Demonstrably, this is not an accidental failure that could be fixed by more restrictive variants of these axioms. What (2) really illustrates is that the assumption of state-independent utility is, at the end of the day, "ineffable in Savage's language of preference over acts" (Seidenfeld et al., 1995, p. 2168–2169). Inasmuch as choosing between Savagian acts is akin to betting on the state of nature, the same point can also be put by saying that "state-independent utility (...) cannot be explicated in the language of bets" (Seidenfeld et al., 1990b, p. 523). Such general formulations are also meant to suggest that the problem of state-dependent utility, i.e., the identification issues detailed in the present section, is relevant not only when the existence conditions of representation (1) are satisfied, but more generally than that. To some extent (which I cannot make precise here),[9] the problem challenges any approach—be it a non-expected utility one—that would seek to explicate beliefs in the language of bets.

2 Act-Dependence

I now turn to the lesser known dimension of *act-dependence*. I start by recalling how this notion is most often understood. Next, I motivate a generalized conception of act-dependence, which is the one I am interested in discussing.

Traditionally (see, e.g., Seidenfeld et al., 1990a, p. 143; Seidenfeld et al., 1990b, p. 521), act-dependence is said to obtain when a representation like (1) must be replaced by another representation in which, unlike in (1),

[8] A detailed discussion of this lesson (including concrete interpretations for the forms of state-dependent utility associated with (2) and the like) is offered in Baccelli, 2017.

[9] See Baccelli, 2021b for a partial answer and a discussion.

the probability values assigned to the states depend on the acts under which the states are envisaged. That is to say, the act-independent probability measure $p : 2^S \to \mathbb{R}$ of (1) is replaced by a family of act-dependent probability measures, $p^f : 2^S \to \mathbb{R}$, $p^g : 2^S \to \mathbb{R}$, and the like, so that, for any f and $g \in F$, (4) holds:

$$f \succcurlyeq g \quad \Leftrightarrow \quad \sum_{i=1}^{n} p^f(s_i) \cdot u\big(f(s_i)\big) \geq \sum_{i=1}^{n} p^g(s_i) \cdot u\big(g(s_i)\big). \tag{4}$$

As Jacques Drèze may have been the first to notice (Drèze, 1961) and Teddy Seidenfeld and co-authors, among others, subsequently highlighted (see, e.g., Seidenfeld et al., 1990a, p. 143), (4) can be naturally interpreted with reference to the situations of "moral hazard" studied in information economics and contract theory. For instance, the decision-maker is the agent in a principal–agent situation, and the probability of the uncertain events of interest—say, the various levels of profit for the company which she is paid to manage—depends on the various courses of action which she can take (see, e.g., Hart and Holmström, 1987; Laffont and Martimort, 2002). More generally, in decision theory, "moral hazard" refers to a varieties of situations in which the resolution of uncertainty is not exogenous to the agent's decision-making (or so she believes), so that the probability of any given state generally depends on her chosen course of action—i.e., it is act-dependent.

Clearly, if act-dependence is understood as in (4), then, it must fall out of the scope of this paper, i.e., the Savage axioms are violated. Indeed, although neither its surface structure nor the above interpretation is the standard one, (4) falls, observationally and literally speaking, within the class of the "multi-prior" models, understood generically as the class of all models of individual decision-making featuring not one, but multiple priors. It is well known that the multi-prior models, starting with the seminal "maxmin expected utility" model (Gilboa and Schmeidler, 1989) or, in an independent tradition, Isaac

Levi's decision theory (Levi, 1974; 1980), violate the Savage axioms.[10]

However, conceptually as well as behaviorally speaking, there is in general no good reason for limiting act-dependence to probability, i.e., for not extending it to utility, too. Intuitively, there are even good reasons *against* thus limiting the scope of act-dependence. In a principal–agent situation, for instance, the agent can influence the probability value of various events, but at a utility cost—say, the disutility of the various levels of effort which she will put in her job, depending on the incentive scheme set up by the principal—that will vary with the courses of action open to her. This is why one finds at the core of such strategic situations a problem of "incentive compatibility", as economists put it. Pushing this line of thought one step further, one can argue that a genuine act-dependent generalization of (1) will feature not only a family of act-dependent probability measures, as in (4), but also a family of act-dependent utility functions, $u^f : X \to \mathbb{R}$, $u^g : X \to \mathbb{R}$, and the like, so that, for any f and $g \in F$, (5) holds:

$$f \succcurlyeq g \;\;\Leftrightarrow\;\; \sum_{i=1}^{n} p^f(s_i) \cdot u^f\big(f(s_i)\big) \geq \sum_{i=1}^{n} p^g(s_i) \cdot u^g\big(g(s_i)\big). \qquad (5)$$

Therefore, by analogy with (2) and (3), the following question arises. Can (1) and (5) ever be equivalent, i.e., can some forms of act-dependence be behaviorally indistinguishable from the traditional subjective expected

[10] The above analysis could be sharpened based on the following fact. Most non-expected utility models can be construed as departing from expected utility exactly in that they operate with act-dependent, rather than act-independent, probability values (see, e.g., Chambers and Echenique, 2016, p. 126; and more fundamentally Cerreia-Vioglio et al., 2011, Cor. 3). For example, maxmin expected utility is the particular case of (4) where $p^f = \arg\min_{p \in \Pi} \Big(\sum_{i=1}^{n} p(s_i) \cdot u\big(f(s_i)\big) \Big)$ for some state-independent u over X, and Π a closed, convex (act-independent) set of priors over 2^S.

utility model axiomatized by Savage?[11] The answer is: yes.[12] Intuitively, this happens when, in (5), the act-dependence of probability and the act-dependence of utility cancel out, in the following sense. Although these two forms of act-dependence generally entail—be it when they are considered separately from one another, or when they are considered together—that the existence conditions of act-independent representation (1) are violated, they can exceptionally entail, when considered together, that these conditions are satisfied.

A first case may be found shocking, but should be almost immediate at this stage. Assume that the existence conditions of representation (1) are satisfied. Implicit in (2) is the choice of one same alternative full support measure q for all acts f and g. However, one might as well take an act-specific full support measure q^f for each f. As (6) makes explicit, this will simply induce an act-dependent state-dependent utility function, like v_s^f:

$$\sum_{i=1}^{n} p(s_i) \cdot u\big(f(s_i)\big) = \sum_{i=1}^{n} q^f(s_i) \cdot \frac{p(s_i)}{q^f(s_i)} u\big(f(s_i)\big) = \sum_{i=1}^{n} q^f(s_i) \cdot v_i^f\big(f(s_i)\big). \quad (6)$$

The equality in (6) directly establishes that the preferences satisfying the Savage axioms are compatible not only with the state-independent, act-independent representation in (1), but also with an infinity of alternative act- *and* state-dependent representations. Specifically, the existence conditions of (1) do imply that there must exist an act- and state-independent representation; but the equality in (6) shows that alternative act- and state-

[11] I am aware of the fact that, considering decision-making from a first-person and a prescriptive point of view, rather than from the third-person and the descriptive point of view adopted here, act-dependence raises a different set of questions (see, e.g., Joyce, 1999). There is no reason to consider these questions as exclusive from the one focused on here.

[12] To the best of my knowledge, the idea that (4), (5), or the like can be compatible with (1) has appeared only once before—and rather indirectly—in the literature, namely, in Drèze and Rustichini, 1999 (Section 5). Drèze and Rustichini's motivations, assumptions, and conclusions, which I cannot present here, are sufficiently different from mine to justify that I offer the discussion to follow.

dependent representations exist, too.

Noteworthily, in the alternative representations induced by an equality like the one in (6), the concavity properties of the act-dependent utility functions cannot vary across acts, and a form of state-dependent utility must obtain. To highlight the properties specific to act-dependence, rather than state-dependence, I now establish that the previously announced positive answer remains correct even when the concavity properties of the act-dependent utility functions can vary across acts, and no form of state-dependent utility obtains. To show this, I will build on work recently done, under different interpretations and with different motivations, by David Dillenberger and co-authors (Dillenberger et al., 2017).[13] Abusing notation as usual, let x stand for the constant act associating consequence x to all states of nature. Assume from now on that X is (for simplicity) a bounded non-empty real interval.[14] Assume further that, in (1), $\succcurlyeq$ is increasing in X (endowed with the usual order) and such that, for any $f \in F$, there exists one and only $x_f \in X$, called the certainty equivalent of f, such that $f \sim x_f$. Notice that this is true if and only if $x_f = u^{-1}\left(\sum_{i=1}^{n} p(s_i) \cdot u\big(f(s_i)\big)\right)$. Thus defined, the certainty equivalent function u^{-1} represents $\succcurlyeq$ in (1), i.e., for any f and $g \in F$, (7) holds:

$$f \succcurlyeq g \quad \Leftrightarrow \quad u^{-1}\left(\sum_{i=1}^{n} p(s_i) \cdot u\big(f(s_i)\big)\right) \geq u^{-1}\left(\sum_{i=1}^{n} p(s_i) \cdot u\big(g(s_i)\big)\right). \tag{7}$$

Therefore, to answer positively the question raised in the previous paragraph, it suffices that, for at least one $f \in F$, some act-dependent pair

[13]David Dillenberger and co-authors do not discuss moral hazard, a particular form of which is behaviorally indistinguishable from what they call "optimism". Besides, because their focus is different, they provide an analysis of act-dependence that is less detailed than the one offered in the present paper. Finally, as the end of the present paper will make clear, their suggested analysis of the links between act-dependence and state-dependence (Dillenberger et al., 2017, fn. 14, p. 1171) is, in several important respects, incomplete.

[14]This is typically the case in the axiomatizations of subjective expected utility over a finite state space (see fn. 1), so as to make the uniqueness of the representation tractable.

$(p^f, u^f) \neq (p, u)$ be found such that the constraint given in (8) is respected:

$$u^{f^{-1}}\left(\sum_{i=1}^{n} p^f(s_i) \cdot u^f\big(f(s_i)\big)\right) = u^{-1}\left(\sum_{i=1}^{n} p(s_i) \cdot u\big(f(s_i)\big)\right). \tag{8}$$

Two preference relations characterized by the same certainty equivalent function being one and the same, it will then follow that, when it is thus constructed, (5) is equivalent to (1).

I start by merely stating how to find, for any $f \in F$, a pair (p^f, u^f) satisfying (8). For non-triviality, I will focus throughout on non-constant acts.[15] For concreteness, I will focus on finding a pair (p^f, u^f) such that p^f displays the following remarkable property: The lottery induced by p^f over the consequences of f puts more probability weight on the more preferred consequences than the lottery similarly induced by p. That is to say, calling $\widehat{p}^f$ and $\widehat{p}$ the former and the latter lottery, respectively, $\widehat{p}^f$ first-order stochastically dominates $\widehat{p}$. Such cases are of interest because, as first-order stochastic dominance is sufficient, ceteris paribus, for a greater expected utility, they most immediately fit (through the idea of a utility cost) the intuitions coming from the moral hazard cases studied in economics. Now, pick any increasing concave transformation h^f, and define u^f by $u^f \equiv h^f \circ u$. Then, it can be shown that, for any u^f thus defined, there exists at least one p^f satisfying (8) and the above first-order stochastic dominance requirement. It is the probability measure p^f that, under the constraint given in (8), minimizes the Euclidean distance to p, i.e., the expression in (9):

$$d(p, p^f) = \sqrt{\sum_{i=1}^{n}\Big(p(s_i) - p^f(s_i)\Big)^2}. \tag{9}$$

To the best of my present understanding, minimizing the Euclidean distance to p has no particularly straightforward decision-theoretic interpretation in

[15] Notice that, by the defining properties of certainty equivalent functions, if f is constant, then, whatever the function u^f, there is no constraint on p^f in (8).

our context, and it may be considered as playing a merely technical role in my analysis. The concave transform u^f of u, by contrast, can be rather naturally interpreted as integrating the utility cost paid by the decision-maker for her efforts resulting in the initial distribution $\widehat{p}$ being replaced by the more advantageous distribution $\widehat{p}^f$. In any case, the special twist of the situation thus constructed is that the two changes offset one another, in the sense that (8) holds.

When $|S| = 2$, the proof of the existence claim just made is almost immediate. Specifically, when $f(s_1) \neq f(s_2)$, for a given increasing transformation h^f, only one probability measure p^f satisfies (8). The solution being unique, it is also, trivially, the distance-minimizing solution. Besides, the concavity of h^f directly imposes that $\widehat{p}^f$ first-order stochastically dominates $\widehat{p}$. Call this Proposition 1, a proof of which is provided in the Appendix.

When $|S| \geq 3$, by contrast, the situation is of necessity more complex. For a given h^f, there is typically not just one distribution, but a convex set of measures satisfying (8). By the continuity in probability of the expected utility functional and the fact that any probability measure is in principle eligible to be p^f, this set is not empty. I will be content here with analyzing the case where $|S| = 3$ and f is such that $f(s_1) \neq f(s_2) \neq f(s_3)$, with, say, $f(s_1) > f(s_2) > f(s_3)$. This is an enlightening case to focus on because the main intuitions, that apply beyond this setting, can then be conveyed geometrically in the so-called Marschak-Machina triangle. Presented as pertains to the present analysis, this is the simplex representing all the lotteries over the outcome set $\{f(s_1), f(s_2), f(s_3)\}$, wherein each lottery q is located thanks to its identifying coordinates $q\Big(f(s_1)\Big)$ and $q\Big(f(s_3)\Big)$, i.e., the probability weight it places on outcomes $f(s_1)$ and $f(s_3)$, respectively (see Marschak, 1950, Machina, 1982, and, e.g., the survey in Sugden, 2004). Assuming some familiarity with this probability triangle and the geometric properties of expected utility within it, I illustrate the situation next, in Fig. 1.

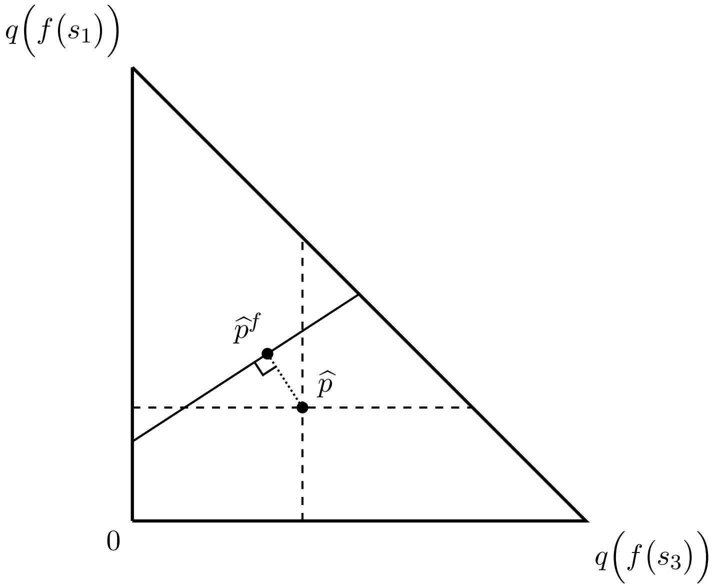

Figure 1: Act-dependence in the Marschak-Machina triangle

As illustrated in Fig. 1, $\widehat{p}$ corresponds to some point within the triangle (or on one of its boundaries). Given our assumptions and the properties of expected utility, the convex set of probability measures satisfying (8)—i.e., yielding with u^f x_f, the observed certainty equivalent of f—must appear in the triangle as a line with some positive slope. By concavity, this line cannot pass by $\widehat{p}$ or any point located southeast of $\widehat{p}$, i.e., any lottery that is first-order stochastically dominated by $\widehat{p}$. Therefore, this line must pass somewhere above $\widehat{p}$. Because this line has positive slope, its intercept with the perpendicular passing by $\widehat{p}$ (or, if this intercept lies strictly out of the triangle, the closest intercept of the line with one of the boundaries of the triangle) must lie somewhere northwest of $\widehat{p}$, i.e., correspond to a lottery that first-order stochastically dominates $\widehat{p}$. In other words, given a concave transform u^f of u, if p^f satisfies (8) and minimizes the distance to p as defined in (9), then, $\widehat{p}^f$ first-order stochastically dominates $\widehat{p}$. Call this Proposition 2, the algebraic proof of which is sketched in the Appendix.[16]

Here is an illustration of what the preceding analysis implies. Suppose for concreteness that $|S| = 3$, $X = [0, 100]$, and consider the two acts in Table 1.

	s_1	s_2	s_3
f	49	25	1
g	2	2	32

Table 1: Two Savagian acts

With reference to f and g in Table 1, consider a decision-maker whose preferences over f, g, and the like can be represented as in (1). Further assume that, in (1), the elicited functions are $u(x) = x$ and $p = \Big(p(s_1), p(s_2), p(s_3)\Big) = (1/5, 3/5, 1/5)$. Thus, we have that $f \succ g$ and the certainty equivalents of f and g are 25 and 8, respectively. However, notice that the same certainty equivalents are observed with $u^f(x) = \sqrt{x}$ together with $p^f = \Big(p^f(s_1), p^f(s_2), p^f(s_3)\Big) = (12/35, 17/35, 6/35)$ and $u^g(x) = \log_2 x$ together with $p^g = \Big(p^g(s_1), p^g(s_2), p^g(s_3)\Big) = (0, 1/2, 1/2)$, respectively. The two act-dependent probability measures have been computed following the construction geometrically presented in Fig. 1. Together with the act-dependent utility functions based on which the computations have been made, they provide a concrete illustration of the possible equivalence between (1) and (5).

Clearly, the compatibility between (1) and (5) generalizes to any number of states and any number of acts. It also generalizes to convex transformations of u together with act-dependent probability measures inducing distributions that are first-order stochastically dominated by $\widehat{p}$, to pairs (u^f, p^f) such that u^f is neither more concave nor more convex than u and $\widehat{p}^f$ is neither first-order stochastically dominating nor first-order stochastically dominated

[16] David Dillenberger and co-authors sketch the geometric analysis of Fig. 1, but they do not provide the complementary algebraic analysis presented for the case $|S| = 3$ in the Appendix. (The general case $|S| = n$ can be proved following a constrained optimization analysis.)

by $\widehat{p}$, and to still other cases. These other cases might less immediately fit the intuitions coming from the moral hazard literature, than the case in Fig. 1. However, importantly, they are equally compelling illustrations of the fact that some forms of act-dependence are behaviorally indistinguishable from the subjective expected utility model axiomatized by Savage. Moreover, their sheer variety, especially as regards the concavity properties of the act-dependent utility functions across all acts, is informative. It makes transparent that—whatever either of the following means at this level of intricacy—neither the beliefs nor the desires of the decision-maker are interestingly identified across all the admissible act-dependent representations of a given subjective expected utility relation. Notice that the second half of this conclusion contrasts with what has been previously explained to hold across all the admissible state-dependent representations of that relation (see p. 7).

Two general lessons can be drawn from the fact that (1) and (5) thus prove compatible with one another. The first and, perhaps, main lesson can be phrased so as to echo how Teddy Seidenfeld and co-authors phrase (in the quotations given on p. 8) the main lesson to draw from the equivalence between (1) and (3). Preferences satisfying the Savage axioms are compatible not only with the traditional expected utility representation in (1), but also with an infinity of "multi-prior expected multi-utility" representations displaying the special structure in (5), and (in at least some cases) rather naturally associated with the moral hazard interpretation introduced at the beginning of this section.[17] This is because the Savage axioms—starting with P2, that is the main axiom supposed to enforce act-independence in the Savage axiomatics—exclude only some, not all forms of act-dependence.

[17] Although its usage is typically more specialized, I freely borrow the phrase "multi-prior expected multi-utility representation" from the literature on incomplete preferences (see, e.g., Galaabaatar and Karni, 2013 and the references therein, including Seidenfeld et al., 1995). Observationally and literally speaking, this is justified inasmuch as (5) features multiple probability measures and multiple utility functions.

And this, in turn, is because the requirement that representation (1) be act-independent is partly "ineffable in Savage's language of preference over acts". In other words, it cannot be fully "explicated in the language of bets". Such general language is also meant to suggest that the problem of act-dependence, i.e., the identification issues explored in the present section, is relevant not only when the existence conditions of representation (1) are satisfied, but more generally than that.

Second, and perhaps most interestingly, the fact that (1) and (5) prove compatible has implications on the problem of state-dependent utility, i.e., the identification issues presented in Section 1, and its relation to the problem of act-dependence. Remarkably, it has been suggested in the literature that, for the problem of state-dependent utility to be solved, it suffices that the decision-maker presumes that she has some influence on the realization of the events, the likelihood of which she is betting upon (see, especially, Drèze, 1987; Karni, 2011). However, the possible equivalence between (1) and (5) makes it clear that, in general, this will not suffice.[18] Indeed, this equivalence means that there will be some cases of the following kind. The decision-maker does presume that she has such capacity of influence, which is one natural interpretation of (at least some instances of) (5); but the observer of her betting behavior is brought back to square one, i.e., representation (1), that is, based on equality (2), behaviorally indistinguishable from representation (3).

Moreover, it is not just indirectly, i.e., through its equivalence with (1), that (5) leads back to the identification issues raised by state-dependent utility. Indeed (even without once again mentioning the cases immediately covered by equality (6)), notice that whatever the act f, for any full sup-

[18]For an elaboration on this non-sufficiency, see Baccelli, 2021a, Sec. 3. Admittedly, both Drèze and Karni are well aware of the fact that, with actions that are unequally costly to the agent, moral hazard may not suffice to overcome the challenges posed by state-dependent utility to the behavioral identification of beliefs. However, neither points out let alone elaborates on the fact that this is even compatible with the respect of the Savage axioms.

port probability measure q^f, with the definition $v_i^f(x) = \frac{p^f(s_i)}{q^f(s_i)} u^f(x)$, for $i = 1, \ldots, n$, equality (10) holds:

$$\sum_{i=1}^{n} p^f(s_i) \cdot u^f\big(f(s_i)\big) = \sum_{i=1}^{n} q^f(s_i) \cdot \frac{p^f(s_i)}{q^f(s_i)} u^f\big(f(s_i)\big) = \sum_{i=1}^{n} q^f(s_i) \cdot v_i^f\big(f(s_i)\big). \tag{10}$$

Therefore, when (1) and (5) are equivalent (with potentially different concavity properties, across all acts, for the various act-dependent utility functions), we also have through (10) that, for any f and $g \in F$, (11) holds:

$$f \succcurlyeq g \;\; \Leftrightarrow \;\; \sum_{i=1}^{n} q^f(s_i) \cdot v_i^f\big(f(s_i)\big) \geq \sum_{i=1}^{n} q^g(s_i) \cdot v_i^g\big(f(s_i)\big). \tag{11}$$

However obvious they should be at this stage of the analysis, equality (10) and equivalence (11) are worth expliciting because they make the following fact clear. Even holding fixed the act-dependent concavity properties of the utility side of an act-dependent representation, the probability side is not behaviorally identified. More generally, what equality (10) shows is that the possible equivalence between (1) and (3) generalizes to any act-dependent variant of (1) and the induced variant of (3). Likewise, the possible equivalence between (1) and (5) generalizes to any state-dependent variant of (1) and the induced variant of (5). Thus (and even more so given the cases already immediately covered by (6)), the state-dependence and act-dependence identification issues examined in this paper do not cancel out, but combine with one another. Consequently, the respect of the Savage axioms proves compatible with an infinity of state-dependent *and* act-dependent representations. Metaphorically speaking, state-dependence and act-dependence form two dimensions along which one can find infinitely many admissible representations of the same Savage preference relation.

Incidentally, the above analysis offers a full clarification for what Teddy

Seidenfeld and co-authors likely refer to when they mention—now without the ellipsis, unlike in the first occurrence of this quotation on p. 5—"the family of possible *(act-independent)* probability/utility pairs that agree with $\prec$ according to expected utility" (Seidenfeld et al., 1990b, p. 521; my emphasis).

Conclusion

In this paper, I have discussed the uniqueness clause of the Savage theorem. I have kept referring to Savage's result simply because it is the prime example of its kind in axiomatic decision theory; for my purposes, all the essential conclusions reached by examining this theorem are equally relevant to Ramsey's, de Finetti's, or Anscombe and Aumann's alternative approaches, to mention but a few. Throughout the paper, I have assumed that the existence conditions of Savage's theorem are satisfied. I have also, in effect, held fixed one arbitrary set of choice data meeting this requirement. I have shown that, contrary to what is usually held, there is not just one (the possible scales for the measurement of utility), but three dimensions along which the non-uniqueness of the representation of the data must be appreciated. The two usually neglected dimensions are that of state-dependence and act-dependence. The latter dimension is, by far, the most unexplored of the two in the current literature; this justifies the greater detail in which it has been presented, as well as the greater caution with which it has been discussed. The existence of these two additional dimensions has especially damaging consequences on the behavioral identification of subjective probability, which the Savage theorem is supposed to deliver.

I have particularly emphasized the symmetries, asymmetries, and connections between state-dependence and act-dependence. But upon finishing, I want to specifically highlight that, although state-dependence and act-dependence are often—which to some extent includes, by its very construction, the present paper—put on a par, they are profoundly different from one another. This would become even more apparent if one stopped assuming that the existence conditions of the classical subjective expected utility representation are satisfied and started investigating which of the Savage axioms are violated by the forms of state-dependence and act-dependence that are incompatible with that representation. The question is especially interesting when both forms of dependence are allowed to obtain at the same time. Given the limited work currently available on act-dependence in axiomatic, behavioral, Savage-style decision theory, such investigation will require further research.

Appendix

PROPOSITION 1: *Assume that* $|S| = 2$ *and* $f(s_1) \neq f(s_2)$*. For any concave transform* u^f *of* u*, there is one and only one* p^f *that satisfies constraint* (8)*, and it is such that* $\widehat{p}^f$ *first-order stochastically dominates* $\widehat{p}$*.*

PROOF OF PROPOSITION 1

Given the definition $u^f \equiv h^f \circ u$ and the properties of function composition, (8) can be uniquely solved as detailed next. For brevity, in this proof, (u_1, u_2) stands for $\Big(u\big(f(s_1)\big), u\big(f(s_2)\big)\Big)$, (u_1^f, u_2^f) for $\Big(u^f\big(f(s_1)\big), u^f\big(f(s_2)\big)\Big)$, and $\Big(h^f(u_1), h^f(u_2)\Big)$ for $\Big(h^f \circ u\big(f(s_1)\big), h^f \circ u\big(f(s_2)\big)\Big)$.

$$
\begin{aligned}
&{u^f}^{-1}\Big(p^f(s_1)u_1^f + \big(1-p^f(s_1)\big)u_2^f\Big) = u^{-1}\Big(p(s_1)u_1 + \big(1-p(s_1)\big)u_2\Big) \\
&\quad \Leftrightarrow \quad p^f(s_1)u_1^f + \big(1-p^f(s_1)\big)u_2^f = h^f\Big(p(s_1)u_1 + \big(1-p(s_1)\big)u_2\Big) \\
&\quad \Leftrightarrow \quad p^f(s_1)h^f(u_1) + \big(1-p^f(s_1)\big)h^f(u_2) = h^f\Big(p(s_1)u_1 + \big(1-p(s_1)\big)u_2\Big) \\
&\quad \Leftrightarrow \quad p^f(s_1) = \frac{h^f\Big(p(s_1)u_1 + \big(1-p(s_1)\big)u_2\Big) - h^f(u_2)}{h^f(u_1) - h^f(u_2)}.
\end{aligned}
$$

It is readily checked that $p^f(s_1) \in [0,1]$, so that the solution defines a probability measure. Next, assume without loss of generality that $f(s_1) > f(s_2)$.

Jensen's inequality and the solution given above for $p^f(s_1)$ imply that $\widehat{p}^f$ first-order stochastically dominates $\widehat{p}$, for h^f is (strictly) concave if and only if:

$$\begin{aligned}
&h^f\Big(p(s_1)u_1 + \big(1-p(s_1)\big)u_2\Big) > p(s_1)h^f(u_1) + \big(1-p(s_1)\big)h^f(u_2) \\
&\Leftrightarrow \quad \frac{h^f\Big(p(s_1)u_1 + \big(1-p(s_1)\big)u_2\Big)}{h^f(u_1)-h^f(u_2)} > \frac{p(s_1)h^f(u_1) + \big(1-p(s_1)\big)h^f(u_2)}{h^f(u_1)-h^f(u_2)} \\
&\Leftrightarrow \quad p^f(s_1) + \frac{h^f(u_2)}{h^f(u_1)-h^f(u_2)} > p(s_1) + \frac{h^f(u_2)}{h^f(u_1)-h^f(u_2)} \checkmark \\
&\Leftrightarrow \quad p^f(s_1) > p(s_1).
\end{aligned}$$

□

Proposition 2: *Assume that* $|S| = 3$ *and* $f(s_1) \neq f(s_2) \neq f(s_3)$. *For any concave transform* u^f *of* u, *if* p^f *minimizes* (9) *under constraint* (8), *then,* p^f *is such that* $\widehat{p}^f$ *first-order stochastically dominates* $\widehat{p}$.

Proof of Proposition 2

Henceforth, for brevity, let (u_1, u_2, u_3) stand for $\Big(u\big(f(s_1)\big), u\big(f(s_2)\big), u\big(f(s_3)\big)\Big)$, $\Big(h^f(u_1), h^f(u_2), h^f(u_3)\Big)$ for $\Big(h^f \circ u\big(f(s_1)\big), h^f \circ u\big(f(s_2)\big), h^f \circ u\big(f(s_3)\big)\Big)$, $(\widehat{p}_1, \widehat{p}_2, \widehat{p}_3) \equiv \widehat{p}$ for $\Big(\widehat{p}\big(f(s_1)\big), \widehat{p}\big(f(s_2)\big), \widehat{p}\big(f(s_3)\big)\Big)$, and $(\widehat{p}^f_1, \widehat{p}^f_2, \widehat{p}^f_3) \equiv \widehat{p}^f$ for $\Big(\widehat{p}^f\big(f(s_1)\big), \widehat{p}^f\big(f(s_2)\big), \widehat{p}^f\big(f(s_3)\big)\Big)$.

I start with the following preliminary observation. Given a (strictly) concave h^f, $\widehat{p}^f$ cannot satisfy (8) and be first-order stochastically dominated by $\widehat{p}$. Assume, by way of contradiction, that such is the case. Then, by the properties of expected utility, we have $\widehat{p}_1 h^f(u_1) + \widehat{p}_2 h^f(u_2) + \widehat{p}_3 h^f(u_3) > \widehat{p}^f_1 h^f(u_1) + \widehat{p}^f_2 h^f(u_2) + \widehat{p}^f_3 h^f(u_3)$. By the concavity of h^f, we also have $h^f(\widehat{p}_1 u_1 + \widehat{p}_2 u_2 + \widehat{p}_3 u_3) > \widehat{p}_1 h^f(u_1) + \widehat{p}_2 h^f(u_2) + \widehat{p}_3 h^f(u_3)$. Therefore, we have $h^f(\widehat{p}_1 u_1 + \widehat{p}_2 u_2 + \widehat{p}_3 u_3) > \widehat{p}^f_1 h^f(u_1) + \widehat{p}^f_2 h^f(u_2) + \widehat{p}^f_3 h^f(u_3)$, thus contradicting (8). Similarly, in Fig. 1, the line on which $\widehat{p}^f$ is to be found cannot pass by $\widehat{p}$ itself. This is because (8) would then require that $\widehat{p}_1 h^f(u_1) + \widehat{p}_2 h^f(u_2) + \widehat{p}_3 h^f(u_3) = h^f(\widehat{p}_1 u_1 + \widehat{p}_2 u_2 + \widehat{p}_3 u_3)$, while concavity requires that

$h^f(\widehat{p}_1 u_1 + \widehat{p}_2 u_2 + \widehat{p}_3 u_3) > \widehat{p}_1 h^f(u_1) + \widehat{p}_2 h^f(u_2) \widehat{p}_3 h^f(u_3)$—a contradiction.

Next, without loss of generality, assume that $u_1 > u_2 > u_3$, as in Fig. 1. Then, $\widehat{p}^f$ first-order stochastically dominates $\widehat{p}$ if and only if $\widehat{p}_1^f \geq \widehat{p}_1$ and $\widehat{p}_1^f + \widehat{p}_2^f \geq \widehat{p}_1 + \widehat{p}_2$, with one of these inequalities being strict. Because, as explained in the preliminary observation, it is excluded that $\widehat{p}$ first-order stochastically dominates $\widehat{p}^f$, it remains to be shown that if $\widehat{p}_1^f \geq \widehat{p}_1$ (respectively, $\widehat{p}_1^f + \widehat{p}_2^f \geq \widehat{p}_1 + \widehat{p}_2$) and the minimal Euclidean distance condition is satisfied, then, $\widehat{p}_1^f + \widehat{p}_2^f \geq \widehat{p}_1 + \widehat{p}_2$ (respectively, $\widehat{p}_1^f \geq \widehat{p}_1$). I now show, by contraposition, that if $\widehat{p}_1^f \geq \widehat{p}_1$ but $\widehat{p}_1^f + \widehat{p}_2^f < \widehat{p}_1 + \widehat{p}_2$, then, the minimal Euclidean distance condition is not satisfied. (The other case is similar.) With reference to Fig. 1, this amounts to showing that, if one picks a point r, corresponding to $\widehat{p}^f$, that is on the line and northeast of $\widehat{p}$, then, one can always find another point r' that is also on the line, but closer—as measured by the Euclidean distance (adapted from measures over the algebra of events to lotteries over the set of consequences in the obvious way)—to $\widehat{p}$.

I now show how to construct r' from r. As is clear from Fig. 1 and as I now detail algebraically, in general, this can be done by transferring to $f(s_2)$ appropriately small probability weights ϵ_1, ϵ_3 from $f(s_1)$ and $f(s_3)$, respectively. First, notice that since *(i)* $\widehat{p}_1^f \geq \widehat{p}_1$ and $\widehat{p}_1^f + \widehat{p}_2^f < \widehat{p}_1 + \widehat{p}_2$ and *(ii)* as explained in the preliminary observation, $\widehat{p}$ cannot first-order stochastically dominate $\widehat{p}^f$, it must be not only that $\widehat{p}_1^f \geq \widehat{p}_1$, but more specifically that $\widehat{p}_1^f > \widehat{p}_1$; hence, that $\widehat{p}_1^f > 0$. Second, notice that if it also holds that $\widehat{p}_3^f > 0$, points r and r' will both satisfy (8) if and only if $\widehat{p}_1^f h^f(u_1) + \widehat{p}_2^f h^f(u_2) + \widehat{p}_3^f h^f(u_3) = (\widehat{p}_1^f - \epsilon_1) h^f(u_1) + (\widehat{p}_2^f + \epsilon_1 + \epsilon_3) h^f(u_2) + (\widehat{p}_3^f - \epsilon_3) h^f(u_3)$, which is true if and only if $\epsilon_3 = \Big(\Big(h^f(u_1) - h^f(u_2)\Big) / \Big(h^f(u_2) - h^f(u_3)\Big)\Big) \epsilon_1$. Next, assuming $\widehat{p}_3^f > 0$ still, consider the Euclidean distance between $\widehat{p}$ and r and r', respectively, defining it like in (9). For commodity, examine more specifically the squared Euclidean distances, denoting them by d_r and $d_{r'}$, respectively. Notice that, by (9), we have that:

$$
\begin{aligned}
d_{r'} - d_r &= \epsilon_1^2 + 2\epsilon_1(\widehat{p}_1 - \widehat{p}_1^f) + (\epsilon_1+\epsilon_3)^2 - 2(\epsilon_1+\epsilon_3)(\widehat{p}_2 - \widehat{p}_2^f) + \epsilon_3^2 + 2\epsilon_3\big((\widehat{p}_1+\widehat{p}_2) - (\widehat{p}_1^f + \widehat{p}_2^f)\big) \\
&= \epsilon_1\big(\epsilon_1 + 2(\widehat{p}_1 - \widehat{p}_1^f)\big) + (\epsilon_1+\epsilon_3)\big((\epsilon_1+\epsilon_3) - 2(\widehat{p}_2 - \widehat{p}_2^f)\big) + \epsilon_3\Big(\epsilon_3 - 2\big((\widehat{p}_1+\widehat{p}_2) - (\widehat{p}_1^f + \widehat{p}_2^f)\big)\Big) \\
&= \epsilon_1\big(\epsilon_1 + 2(\widehat{p}_1 - \widehat{p}_1^f)\big) + \left(\frac{h^f\big(u_1\big) - h^f\big(u_3\big)}{h^f\big(u_2\big) - h^f\big(u_3\big)}\right)\epsilon_1\left(\left(\frac{h^f\big(u_1\big) - h^f\big(u_3\big)\big)}{h^f\big(u_2\big) - h^f\big(u_3\big)}\right)\epsilon_1 - 2(\widehat{p}_2 - \widehat{p}_2^f)\right) \\
&\quad + \left(\frac{h^f\big(u_1\big) - h^f\big(u_2\big)}{h^f\big(u_2\big) - h^f\big(u_3\big)}\right)\epsilon_1\left(\left(\frac{h^f\big(u_1\big) - h^f\big(u_2\big)}{h^f\big(u_2\big) - h^f\big(u_3\big)}\right)\epsilon_1 - 2\big((\widehat{p}_1+\widehat{p}_2) - (\widehat{p}_1^f + \widehat{p}_2^f)\big)\right).
\end{aligned}
$$

Therefore, to have $d_{r'} - d_r < 0$, i.e., to find a point r' on the line but at a lesser Euclidean distance to $\widehat{p}$ than point r, it suffices to pick any ϵ_1 such that the following two conditions are satisfied:

1. $\epsilon_1 < 2(\widehat{p}_1^f - \widehat{p}_1) \equiv \alpha$;
2. $\epsilon_1 < 2\left(\frac{h^f\big(u_2\big) - h^f\big(u_3\big)}{h^f\big(u_1\big) - h^f\big(u_2\big)}\right)\Big((\widehat{p}_1+\widehat{p}_2) - (\widehat{p}_1^f + \widehat{p}_2^f)\Big) \equiv \beta$.

It is the case that $\alpha, \beta > 0$ since: *(i)* $\widehat{p}_1^f \geq \widehat{p}_1$ and $\widehat{p}_1^f + \widehat{p}_2^f < \widehat{p}_1 + \widehat{p}_2$ hold by assumption; *(ii)* for the reasons previously detailed, not only $\widehat{p}_1^f \geq \widehat{p}_1$, but more specifically $\widehat{p}_1^f > \widehat{p}_1$ must hold; *(iii)* $h^f\big(u_1\big) > h^f\big(u_2\big) > h^f\big(u_3\big)$ holds by assumption. Thus, if $\widehat{p}_3^f > 0$, any $\epsilon_1 \in (0, \min\{\alpha; \beta\})$ will define a point r' that, like r, satisfies (8), while generating, by (9), a lesser distance than r. If $\widehat{p}_3^f = 0$, there is no need to preliminarily express ϵ_3 in terms of ϵ_1, and one may directly compare the relevant squared Euclidean distances simply by setting $\epsilon_3 = 0$ in the preceding equalities; then, with α as defined above, any $\epsilon_1 \in (0, \alpha)$ has the desired properties. This establishes in all cases that under constraint (8), if $\widehat{p}_1^f \geq \widehat{p}_1$ but $\widehat{p}_1^f + \widehat{p}_2^f < \widehat{p}_1 + \widehat{p}_2$, then, the minimal Euclidean distance condition is not satisfied. By contraposition, under constraint (8), if $\widehat{p}_1^f \geq \widehat{p}_1$ and the minimal Euclidean distance condition is satisfied, then, $\widehat{p}_1^f + \widehat{p}_2^f \geq \widehat{p}_1 + \widehat{p}_2$, i.e., $\widehat{p}^f$ first-order stochastically dominates $\widehat{p}$.

□

References

Arrow, K. (1963): "Uncertainty and the Welfare Economics of Medical Care", *American Economic Review*, 53(5), 941–973.

——— (1974): "Optimal Insurance and Generalized Deductibles", *Scandinavian Actuarial Journal*, 1974(1), 1–42.

Baccelli, J. (2017): "Do Bets Reveal Beliefs?", *Synthese*, 194(9), 3393–3419.

——— (2021a): "Moral Hazard, the Savage Framework, and State-Dependent Utility", *Erkenntnis*, 86, 367–387.

——— (2021b): "The Problem of State-Dependent Utility: A Reappraisal", *British Journal for the Philosophy of Science*, 72(2), 617–634.

Cerreia-Vioglio, S., P. Ghirardato, F. Maccheroni, M. Marinacci, and M. Siniscalchi (2011): "Rational Preferences under Ambiguity", *Economic Theory*, 48(2-3), 341–375.

Chambers, C., and F. Echenique (2016): *Revealed Preference Theory.* New York: Cambridge University Press.

d'Aspremont, C., and L. Gevers (2002): "Social Welfare Functionals and Interpersonal Comparability", in *Handbook of Social Choice and Welfare*, ed. by K. Arrow, A. Sen, and K. Suzumura, vol. 1, pp. 459–541. Amsterdam: North-Holland.

Dillenberger, D., A. Postlewaite, and K. Rozen (2017): "Optimism and Pessimism with Expected Utility", *Journal of the European Economic Association*, 15(5), 1158–1175.

Drèze, J. (1961): "Les fondements logiques de la probabilité subjectives et de l'utilité", *La décision. Colloques Internationaux du Centre National de la Recherche Scientifique*, pp. 73–87.

——— (1987): "Decision Theory with Moral Hazard and State-Dependent Preferences", in *Essays on Economic Decisions under Uncertainty*, ed. by J. Drèze, pp. 23–89. Cambridge: Cambridge University Press.

Drèze, J., and A. Rustichini (1999): "Moral Hazard and Conditional Preferences", *Journal of Mathematical Economics*, 31(2), 159–181.

Eeckhoudt, L., C. Gollier, and H. Schlesinger (2005): *Economic and Financial Decisions under Risk*. Princeton: Princeton University Press.

Galaabaatar, T., and E. Karni (2013): "Subjective Expected Utility with Incomplete Preferences", *Econometrica*, 81(1), 255–284.

Gilboa, I., and D. Schmeidler (1989): "Maxmin Expected Utility with Non-Unique Prior", *Journal of Mathematical Economics*, 18(2), 141–153.

Hart, O., and B. Holmström (1987): "The Theory of Contracts", in *Advances in Economic Theory*, ed. by T. Bewley, pp. 71–155. Cambridge: Cambridge University Press.

Joyce, J. (1999): *The Foundations of Causal Decision Theory*. New York: Cambridge University Press.

Karni, E. (1996): "Probabilities and Beliefs", *Journal of Risk and Uncertainty*, 13(3), 249–262.

——— (2011): "Subjective Probabilities on a State Space", *American Economic Journal: Microeconomics*, 3(4), 172–185.

Kobberling, V., and P. Wakker (2003): "Preference Foundations for Nonexpected Utility: A Generalized and Simplified Technique", *Mathematics of Operations Research*, 28(3), 395–423.

Laffont, J.-J., and D. Martimort (2002): *The Theory of Incentives: The Principal-Agent Model*. Princeton: Princeton University Press.

Levi, I. (1974): "On Indeterminate Probabilities", *The Journal of Philosophy*, 71(13), 391–418.

——— (1980): *The Enterprise of Knowledge*. Cambridge, Massachusetts: MIT Press.

Machina, M. (1982): "'Expected Utility' Analysis without the Independence Axiom", *Econometrica*, 50(2), 277–323.

Marschak, J. (1950): "Rational Behavior, Uncertain Prospects, and Measurable Utility", *Econometrica*, 18(2), 111–141.

Nau, R. (2001): "de Finetti Was Right: Probability Does Not Exist", *Theory and Decision*, 51(2-4), 89–124.

Pratt, J. (1964): "Risk Aversion in the Small and in the Large", *Econometrica*, 32(1), 122–136.

Savage, L. (1954): *The Foundations of Statistics*. New York: Wiley (First Edition).

——— (1972): *The Foundations of Statistics*. New York: Dover (Second Edition).

Savage, L., and R. Aumann (1987): "Letters between Leonard Savage and Robert Aumann [January 1971]", in *Essays on Economic Decisions under Uncertainty*, ed. by J. Dreze, pp. 76–81. Cambridge: Cambridge University Press.

Schervish, M., T. Seidenfeld, and J. Kadane (1990): "State-Dependent Utilities", *Journal of the American Statistical Association*, 85(411), 840–847.

——— (2013): "The Effect of Exchange Rates on Statistical Decisions", *Philosophy of Science*, 80(4), 504–532.

Seidenfeld, T., M. Schervish, and J. Kadane (1990a): "Decisions without Ordering", in *Acting and Reflecting*, ed. by W. Sieg, pp. 143–170. Boston: Kluwer.

——— (1990b): "When Fair Betting Odds Are Not Degrees of Belief", *Philosophy of Science*, 1990(1), 517–524.

——— (1995): "A Representation of Partially Ordered Preferences", *The Annals of Statistics*, 23(6), 2168–2217.

Sugden, R. (2004): "Alternatives to Expected Utility: Foundations", in *Handbook of Utility Theory, Volume II: Extensions*, ed. by S. Barbera, P. Hammond, and C. Seidl, pp. 685–755. Boston: Kluwer Academic Press.