JOURNAL OF ECONOMIC THEORY 61, 1–22 (1993)

Dynamically Consistent Beliefs Must Be Bayesian*

LARRY G. EPSTEIN

Department of Economics, University of Toronto, 150 St. George Street, Toronto, Ontario, Canada M5S 1A1

AND

MICHEL LE BRETON

GREQE, Université d'Aix-Marseille 2, 2, Rue de la Charité, 13002 Marseille, France

Received April 17, 1992; revised August 23, 1992

Experimental evidence such as the Ellsberg Paradox contradicts the Savage model of decision making under uncertainty, since the representation of beliefs underlying preferces by a single probability measure leaves no room for the degree of imprecision in information to affect decisions. Proceeding axiomatically, this paper shows that the existence of a Bayesain prior is implied, even if Savage's Sure-Thing Principle is deleted, if preferences (i) are "based on beliefs" and (ii) admit dynamically consistent updating in response to new information. The result raises questions about the appeal of models of preference that feature a separation of tastes and beliefs. *Journal of Economic Literature* Classification Number: D81.

1. INTRODUCTION

1.1. *Motivation and Outline*

The standard way to model a decision maker's belief in a situation of uncertainty is by means of a subjective probability distribution or Bayesian prior. Axiomatic justification for the assumption that there exists a prior is provided by Savage [31]. However, behaviour such as that exhibited in the Ellsberg Paradox [10] is inconsistent with the above Bayesian approach.

* Epstein gratefully acknowledges the financial support of the Social Scienes and Humanities Research Council of Canada. Le Breton thanks the Department of Economics at the University of Toronto for its hospitality during the autumn of 1991 when this paper was conceived. We have benefitted from discussions with Mike Peters, Dale Poirier, Zvi Safra, and David Schmeidler, and from comments by Costis Skiadas, a referee, and participants in a number of seminar presentations.

0022-0531/93 $5.00

The difficulty is that the "vagueness" or "ambiguity" that is common in situations of uncertainty cannot be captured in the representation of beliefs by a single prior. Such motivation underlies recent generalizations of the Bayesian paradigm such as the non-additive probability model [32, 16] and the multiple priors model [17]. (See also [5, 6, 35, 37].)

In this paper we follow Savage in adopting a behaviourial or choice-theoretic approach to modeling beliefs. Thus our primitive is a preference ordering $\succsim$ over uncertain prospects (or acts). Beliefs are represented by a binary ordering $\succsim_l$ over events, where $A \succsim_l B$ indicates that A is believed to be at least as likely as B. Refer to $\succsim$ as being *based on beliefs* if there exists a relation $\succsim_l$ such that the decision maker would rather bet on the event A than on event B if and only if $A \succsim_l B$. The restriction that $\succsim$ be based on beliefs is Savage's axiom P4; it imposes in a minimal way the separation of beliefs regarding the likely realization of alternative events from the valuation of the consequences associated with the realization of these events.

Our principal contribution is to prove the following: Suppose that $\succsim$ is based on beliefs and that the same is true for all preference orderings derived from $\succsim$ by conditioning on available information in such a way as to ensure dynamic consistency. Assume also that $\succsim$ satisfies the relatively uncontroversial axioms in Savage, i.e., the Savage axioms minus the Sure-Thing Principle. Finally, suppose that there exist at least three outcomes that are pairwise nonindifferent. Then beliefs, in the form of $\succsim_l$, can be represented by a unique probability measure μ over events. Moreover, the preference ordering $\succsim$ is probabilistically sophisticated in the sense of [29]; that is, two acts that induce (through μ) identical probability distributions over outcomes must be indifferent according to $\succsim$. Note that though Savage's model of beliefs (a single prior) is implied, the other component of his model, namely the expected utility form for valuing probability distributions, is *not* implied.

The importance of the theorem stems from the observation that most choice problems are sequential and require the updating of preferences and beliefs as new information arrives. Thus a satisfactory treatment of updating is a prerequisite for fruitful applications of models of non-Bayesian beliefs or probabilistically non-sophisticated preferences, whether to intertemporal problems, game theory or statistical theory. Conditional upon acceptance of our treatment of updating, there are a number of possible reactions to our theorem. First, those committed to the Bayesian model of beliefs may find it reassuring. On the other hand, for modelers who think that vagueness in beliefs is an important component of decision-making under uncertainty, our theorem may represent a distressing negative result. At the very least, it indicates some "costs" to the modeler, in terms of appealing axioms that must be relaxed, involved in modeling vagueness in beliefs. We turn now to a brief discussion of the central feature of our approach to updating, namely "dynamic consistency."

Dynamic consistency is assumed first at a meta level. Our view is that for a model to be coherent, (nontechnical) axioms imposed on the initial preference ordering $\succsim$ should be satisfied also by subsequent updated orderings. In that sense, the *model* should be dynamically consistent. Such consistency is reflected in our assumption that updated preferences are based on beliefs. A separate justification for the latter assumption is that it is clearly necessary for the issue of "updating beliefs" to be meaningful.

The second way in which dynamic consistency is assumed is through our supposition that preferences are updated in response to new information in such a way as to ensure the dynamic consistency of preferences. The literature on "changing tastes," beginning with the seminal article by Strotz [36], has provided a number of approaches to describing choice behaviour given dynamically inconsistent preferences (for example, see [30] and for more recent illustrations in the context of choice under risk see [4, 24]). Thus we acknowledge that for descriptive purposes it is an empirical question whether dynamic consistency or inconsistency is more accurate. At the same time, the assumption of dynamic consistency is clearly advantageous in terms of theoretical elegance and simplicity and also analytical tractability, and so is "natural" for these reasons. From a normative point of view, it is difficult to imagine adopting or recommending a dynamically inconsistent updating rule for use in statistical decision problems. Walley [37, Chap. 6] reflects this view in his theory of statistical reasoning; he adopts as a fundamental principle the conglomerative principle, which is the counterpart of dynamic consistency for his slightly different formal famework. In addition, Hammond [21] has shown that generally, the choice behaviour implied by a set of dynamically inconsistent preferences cannot be rationalized by a preference ordering. In Section 2.2 below, we indicate for our specific famework some implications of dynamic inconsistency that seem undesirable in both normative and positive modeling contexts.

The rest of the paper proceeds as follows: The Ellsberg Paradox is described next, followed by a review of related literature. Then Section 2 contains our contribution. Section 3 reconsiders some of our assumptions and Section 4 concludes. Proofs are collected in an Appendix.

1.2. *The Ellsberg Paradox*

For the convenience of the reader, we describe a variation of the Ellsberg Paradox: There are 90 balls in an urn, 30 red ones and the rest either black of yellow, in unknown proportions. One ball is to be drawn at random. The following preferences over acts are typical,

$$\begin{bmatrix} \$100 & \text{if} & s \in R \\ 0 & \text{if} & s \in B \\ 0 & \text{if} & s \notin R \cup B \end{bmatrix} \succ \begin{bmatrix} 0 & \text{if} & s \in R \\ \$100 & \text{if} & s \in B \\ 0 & \text{if} & s \notin R \cup B \end{bmatrix} \quad \text{and}$$

$$\begin{bmatrix} \$100 & \text{if} & s \in R \\ 0 & \text{if} & s \in B \\ \$100 & \text{if} & s \notin R \cup B \end{bmatrix} \prec \begin{bmatrix} 0 & \text{if} & s \in R \\ \$100 & \text{if} & s \in B \\ \$100 & \text{if} & s \notin R \cup B \end{bmatrix},$$

where R and B denote the events corresponding to the chosen ball being red or black and s refers to the ball that is drawn. Since the substitution of the outcome \$100 for 0 in the event of a yellow ball being drawn reverses the ranking, these choices contradict the Sure-Thing Principle. (A statement of the latter is provided in Section 2.4.) Note that the above choices seem sensible at a normative level, since they correspond to an aversion to imprecise information. Therefore, they raise doubts about both the positive and prescriptive appeal of the Sure-Thing Principle.

For later reference note that, if preferences are "based on beliefs," then the first ranking indicates that R is viewed as "more likely than" B, while the second ranking indicates that the complement of R is viewed as "more likely than" the complement of B. Therefore, the above choices contradict not only the Sure-Thing Principle, but also the hypothesis that underlying beliefs can be represented by a probability measure.

1.3 *Related Literature*

Though the formal Savage framework is atemporal, the Sure-Thing Principle and its analogue for the case of risk, the independence axiom, have been widely interpreted as dynamic consistency requirements. Correspondingly, it has been thought that the dynamic consistency of preference essentially implies the Savage expected utility model. Recently, the basis for such an interpretation has been clarified. In particular, it has been shown to rest in part on the implicit assumption of "consequentialism" or "independence from unrealized alternatives" (see [22, 28], for example). Machina argues against consequentialism in part by drawing an analogy with the implications of nonseparability for consumer choice under certainty. Nonconsequentialism can arise also from a concern with the process of choice; for instance, for reasons of *ex ante* fairness in social choice situations (see [28, Sect. 4.2] and [12]), or as in orthodox (non-Bayesian) statistics violating the likelihood principle, where, for example, inference is based on the method of sampling as well as on the realized data.[1] In this paper we show that if preferences are based on beliefs, then dynamic

[1] For another nonconsequentialist model see [20]. For some difficulties with nonconsequentialism see [19, 2].

consistency implies Savage's model of beliefs (a single prior), *even without the assumption of consequentialism.* See Section 3.4 for further discussion.

This paper is also closely related to Machina and Schmeidler [29]. They disentangle axiomatically the property of probabilistic sophistication (roughly, that beliefs are represented by a single prior) from Savage's joint model of beliefs (a single prior) and valuation (expected utility). They accomplish this by dropping the Sure-Thing Principle and simultaneously strengthening Savage's P4 to what they call P4*. We adopt an alternative strengthening of P4, denoted $P4^c$, that we show is equivalent to P4*, and we then invoke Theorem 2 of Machina–Schmeidler to draw our conclusions. Thus, methematically our result is a corollary of theirs. The value added by this paper rests primarily on the superior, or at least substantially different, intuitive appeal and economic significance that we claim for our central axiom $P4^c$ as opposed to their P4*. (See Sections 2.2 and 2.4 for further discussion.)

There is a large literature dealing with updating rules for vague beliefs, frequently modeled by a set of priors. Walley [37, p. 279] argues that the Dempster–Shafer updating rule can lead the decision-maker to accept a "sure loss," essentially because of the associated dynamic inconsistency. Adopting a decision-theoretic approach, Gilboa and Schmeidler [18] impose some requirements on plausible updating rules in the context of the nonadditive-prior and multiple-priors models of beliefs. Further, on the intersection of these two models of beliefs, they cast light on the Dempster–Shafer updating rule. Dynamic consistency is not one of the requirements they impose; indeed, it is at least implicit in their paper that there do not exist any updating rules that are compatible with dynamic consistency, by which we mean that the models of preference they consider violate our $P4^c$. It is straightforward to demonstrate this, as we do below, and thus the models they consider possess the "undesirable" properties described in Section 2.2. Our analysis shows that this negative conclusion is not restricted to the models considered in [18]; rather, it holds for a large class of models of beliefs. In particular, our analysis reflects also on the updating rules in [14, 23, 37], where explicit decision-theoretic frameworks are not completely specified. This paper shows that there does not exist a "satisfactory" decision-theoretic foundation for *any* rule for updating vague (i.e., non-probabilistic) beliefs.

Finally, it is well-known (for example, see [26, Chap. 10] and [38]) that *given* the Savage model of preference, dynamic consistency requires that probability distributions be updated according to Bayes Rule as new information is received. We also provide a consistency-based justification for Bayes Rule, but with much weaker maintained assumptions regarding the nature of utility and beliefs.

2. Dynamically Consistent Beliefs

2.1. *Axioms and Theorem*

Adopt the Savage framework, which consists of

$S = \{..., s, ...\}$ a set of *states*

$\mathscr{E} = \{..., A, B, ..., E, ...\}$ the set of *events*, i.e., all subsets of S

$X = \{..., x, y, z, ...\}$ the set of outcomes or *consequenes*, and

$\mathscr{F} = \{..., f, g, ...\}$ the set of *finite-outcomes acts* on S, i.e,

$f \in \mathscr{F}$ if $f: S \to X$ and if $f(S)$ is a finite set.

Given a binary relation $\succsim$ on $\mathscr{F}$, say that the event E is *null*, if any pair of acts differing only on E are indifferent. Write $x \succsim y$ whenever the constant act yielding x for all $s \in S$ is weakly preferred to the constant act yielding y.

The following variation of Savage's axioms are assumed for $\succsim$. The numbering is Savage's and the names are adapted from those proposed by Machina–Schmeidler. Where the axioms are self-explanatory or familiar from Savage, little or no discussion is offered. See, however, Section 3 for further discussion.

P1 (Ordering). *The binary relation $\succsim$ is complete, reflexive and transitive.*

P3 (Eventwise Monotonicity). *For all outcomes x and y, non-null events E and acts g*

$$\begin{bmatrix} x & \text{if} & s \in E \\ g(s) & \text{if} & s \notin E \end{bmatrix} \succsim \begin{bmatrix} y & \text{if} & s \in E \\ g(s) & \text{if} & s \notin E \end{bmatrix} \Leftrightarrow x \succsim y.$$

P4 (Weak Comparative Probability). *For all events A, B and outcomes $x^* \succ x$ and $y^* \succ y$*

$$\begin{bmatrix} x^* & \text{if} & s \in A \\ x & \text{if} & s \notin A \end{bmatrix} \succsim \begin{bmatrix} x^* & \text{if} & s \in B \\ x & \text{if} & s \notin B \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} y^* & \text{if} & s \in A \\ y & \text{if} & s \notin A \end{bmatrix} \succsim \begin{bmatrix} y^* & \text{if} & s \in B \\ y & \text{if} & s \notin B \end{bmatrix}.$$

Axiom P4 induces a complete, reflexive and transitive "comparative likelihood" relation $\succsim_l$, where: for all events A and B, $A \succsim_l B$ if $\exists$ outcomes $x^* \succ x$ such that

$$\begin{bmatrix} x^* & \text{if} & s \in A \\ x & \text{if} & s \notin A \end{bmatrix} \succsim \begin{bmatrix} x^* & \text{if} & s \in B \\ x & \text{if} & s \notin B \end{bmatrix}.$$

The relation $\succsim_l$ represents the beliefs about comparative likelihoods that are implicit in $\succsim$.

Savage adopts

P5 (Nondegeneracy). *There exist outcomes x and y such that $x \succ y$.*

We strengthen P5 slightly to

P5* (Strong Nondegeneracy). *There exist outcomes x, y, and z such that $x \succ y$ and $y \succ z$.*

A "technical" axiom concerning the richness of the state space S is

P6 (Small Event Continuity). *For any acts $f \succ g$ and outcome x, there exists a finite partition $\{A_1, ..., A_n\}$ of S such that*

$$f \succ \begin{bmatrix} x & \text{if} & s \in A_i \\ g(s) & \text{if} & s \notin A_i \end{bmatrix}$$

and

$$\begin{bmatrix} x & \text{if} & s \in A_j \\ f(s) & \text{if} & s \notin A_j \end{bmatrix} \succ g, \qquad \text{for all} \quad i, j \in \{1, ..., n\}.$$

The above axioms, with the exception of P5*, are due to Savage. His axiom P2, the Sure-Thing Principle, is excluded, since it conflicts with a role for vagueness as demonstrated in the context of the Ellsberg Paradox. In its place, we propose the following strengthening of P4:

P4^c (Conditional Weak Comparative Probability). *For all events T, A, and B, $A \cup B \subseteq T$, outcomes x^*, x, y^*, and y, and acts g,*

$$\text{if} \quad \begin{bmatrix} x^* & \text{if} & s \in T \\ g(s) & \text{if} & s \notin T \end{bmatrix} \succ \begin{bmatrix} x & \text{if} & s \in T \\ g(s) & \text{if} & s \notin T \end{bmatrix}$$

and

$$\begin{bmatrix} y^* & \text{if} & s \in T \\ g(s) & \text{if} & s \notin T \end{bmatrix} \succ \begin{bmatrix} y & \text{if} & s \in T \\ g(s) & \text{if} & s \notin T \end{bmatrix},$$

then

$$\begin{bmatrix} x^* & \text{if} & s \in A \\ x & \text{if} & s \in T\setminus A \\ g(s) & \text{if} & s \notin T \end{bmatrix} \succsim \begin{bmatrix} x^* & \text{if} & s \in B \\ x & \text{if} & s \in T\setminus B \\ g(s) & \text{if} & s \notin T \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} y^* & \text{if} & s \in A \\ y & \text{if} & s \in T\setminus A \\ g(s) & \text{if} & s \notin T \end{bmatrix} \succsim \begin{bmatrix} y^* & \text{if} & s \in B \\ y & \text{if} & s \in T\setminus B \\ g(s) & \text{if} & s \notin T \end{bmatrix}.$$

By taking $T = S$, we see that P4^c implies Savage's P4 and thus the existence of the comparative likelihood relation $\succsim_l$. More generally, and as we justify further below, P4^c amounts to applying P4 "conditionally" on any event T, which explains the name we have chosen for the axiom.

Our central result is[2]

THEOREM. *If the preference relation $\succsim$ over $\mathscr{F}$ satisfies the axioms* P1, P3, P4^c, P5*, *and* P6, *then there exists a unique finitely additive non-atomic probability measure μ on $\mathscr{E}$ such that:* (a) *μ represents the comparative likelihood relation $\succsim_l$, that is*

$$A \succsim_l B \Leftrightarrow \mu(A) \geq \mu(B);$$

and (b) *for any pair of acts f and g in $\mathscr{F}$ with respective outcome sets contained in $\{x_1, ..., x_n\}$:*

If $\mu(f^{-1}(z)) = \mu(g^{-1}(z))$ $\forall z \in \{x_1, ..., x_n\}$, then $f \sim g$.

Under the conditions of the theorem, beliefs must be Bayesian in that $\succsim_l$ is representable by a prior probability distribution, contradicting typical behaviour in the Ellsberg Paradox. Moreover, by part (b), any preference ordering $\succsim$ satisfying the stated axioms must be *probabilistically sophisticated*—two acts that imply the identical probability distribution over outcomes are indifferent.[3]

Before turning to the interpretation of P4^c, we remark firstly that it is vacuously satisfied if X contains only two distinct indifference classes, which explains our strengthening of P5 to P5*. Secondly, note that, given P3, P4^c is equivalent to the following axiom, which is employed in the

[2] A probability measure μ is non-atomic if for any event A with $\mu(A) > 0$ and any α in $(0, 1)$, $\exists A^* \subset A$ such that $\mu(A^*) = \alpha \cdot \mu(A)$.

[3] As in [29, Th. 2], we can identify the exhaustive implications of the hypothesized axioms for the structure of the utility function representation of $\succsim$. It is the associated strengthening of (b) that Machina and Schmeidler refer to as probabilistic sophistication, but the difference from (b) is not important in this paper.

proof of our theorem and is more directly comparable with the Machina-Schmeidler axiom described in Section 2.4

P4′. *For all events* A, B, T *such that* $A \cup B \subseteq T$, *outcomes* $x^* \succ x$ *and* $y^* \succ y$ *and acts* g,

$$\begin{bmatrix} x^* & \textit{if} & s \in A \\ x & \textit{if} & s \in T \setminus A \\ g(s) & \textit{if} & s \notin T \end{bmatrix} \succsim \begin{bmatrix} x^* & \textit{if} & s \in B \\ x & \textit{if} & s \in T \setminus B \\ g(s) & \textit{if} & s \notin T \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} y^* & \textit{if} & s \in A \\ y & \textit{if} & s \in T \setminus A \\ g(s) & \textit{if} & s \notin T \end{bmatrix} \succsim \begin{bmatrix} y^* & \textit{if} & s \in B \\ y & \textit{if} & s \in T \setminus B \\ g(s) & \textit{if} & s \notin T \end{bmatrix}.$$

To see the equivalence, note that both P4^c and P4′ are satisfied vacuously if T is null. On the other hand, if T is non-null then by P3, the hypotheses concerning x^*, x, y^*, and y in the statement of P4^c are equivalent to $x^* \succ x$ and $y^* \succ y$.

2.2 *Dynamic Consistency*

To interpret P4^c, think of two "times" at which choices are made. The *ex ante* stage, $t = 0$, refers to the point in time we have been discussing where choices are dictated by $\succsim$. Fix an arbitrary event $T \subseteq S$. Suppose that by the intermediate stage, $t = 1$, the decision maker will have learned whether the true state of the world lies in the event $T \subseteq S$ or its complement $S \setminus T$. Suppose further that at $t = 1$ it is learned that in fact T contains the true state and let the decision maker reconsider her options. Relevant to deliberations at this stage are the new primitives: the set of states T, the set $\mathscr{E}_T$ of all subsets of T, the set of outcomes X (unchanged) and the set $\mathscr{F}_T$ of finite-outcome acts on T. Suppose that acts in $\mathscr{F}_T$ are ranked by a complete, reflexive and transitive relation $\succsim^*$. It is natural to *permit* that relation to depend on past choices. For example, suppose that at $t = 0$ the act

$$\begin{bmatrix} f(s) & \text{if} & s \in T \\ g(s) & \text{if} & s \notin T \end{bmatrix}$$

was chosen out of some feasible set. Then at $t = 1$ the preference ordering $\succsim^*$ over $\mathscr{F}_T$ may very well depend on g due, for example, to feelings of elation or disappointment at the realization of event T and thus at the preclusion of the outcomes associated with g. Since $\succsim^*$ obviously depends on T, we will write $\succsim_{T,g}$ for $\succsim^*$. Dynamic consistency requires that choice made at $t=0$ be respected at $t=1$ and thus may be formalized as follows: for all f, f' in $\mathscr{F}_T$ and g an act on $S \backslash T$,

$$f \succsim_{T,g} f' \quad \textit{iff} \begin{bmatrix} f(s) & \text{if} & s \in T \\ g(s) & \text{if} & s \notin T \end{bmatrix} \succsim \begin{bmatrix} f'(s) & \text{if} & s \in T \\ g(s) & \text{if} & s \notin T \end{bmatrix}. \tag{2.1}$$

(Note that if we had also required that $\succsim_{T,g}$ be independent of g, then dynamic consistency would imply that $\succsim$ satisfy the Sure-Thing Principle.)

The above consistency condition defines a *unique* rule for updating $\succsim$ to $\succsim_{T,g}$ given the observation T and the "unrealized alternative" g. There are two noteworthy implications of this rule for updating preferences. First, it provides a simple interpretation for our central axiom: $P4^c$ is simply the requirement that the updated ordering $\succsim_{T,g}$ be "based on beliefs" in the sense of satisfying the appropriate form of Savage's P4. Consequently $P4^c$ is a natural assumption if the theory is to be coherent in settings where updating is called for. *It makes little sense to impose P4 on $\succsim$ but not on its updates; indeed, $\succsim$ itself is presumably an updated version of some "earlier" preference ordering.* Note that the appropriate forms of P1 and P3 are automatically satisfied by $\succsim_{T,g}$ given that $\succsim$ satisfies them, while P5* is also inherited if T is non-null. Moreover, $P4^c$ (and not merely P4) is satisfied by the updated ordering if it is satisfied by $\succsim$. Thus, with the possible exception of the "technical" axiom P6, our set of axioms or model is "dynamically consistent" (recall the discussion in the introduction).

A second implication of the consistency requirement on updating preference is the following: Under the conditions of our theorem, the comparative likelihood relation $\succsim_l$ can be represented by the probability measure μ. Denote by $\succsim_{T,g,l}$ the updated comparative likelihood relation, i.e., that implicit in the updated preference ordering $\succsim_{T,g}$. It follows from Lemma 3 in the appendix that for all events A, B, and T with $A \cup B \subseteq T$ and T non-null,

$$A \succsim_{T,g,l} B \quad \text{iff} \quad A \succsim_l B.$$

Therefore, the Bayesian updated probability measure $\mu(\cdot)/\mu(T)$ represents beliefs about likelihoods of subevents of T. As noted in the introduction, this link between dynamic consistency and the necessity of Bayes Rule for updating beliefs generalizes the justification for the latter that is well known within the framework of the Savage model.

As the preceding has made abundantly clear, the justification for $P4^c$, and hence also for our theorem, rests in large part on the assumption that preferences are updated so as to ensure dynamic consistency in the standard sense of (2.1). To conclude this subsection, we elaborate upon the supporting arguments provided in the introduction for this assumption.

A similar issue arises in the context of behaviour under risk (i.e., objective probability distributions) given that the independence axiom is violated. The reader is referred to Machina [28, Sect. 6.4] for a detailed discussion which we here abbreviate and adapt to our setting.

The most common objection to the specification (2.1) for updating preference is that *behaviour* can be consistent even if *preferences* are not, and thus (2.1) is "unnecessary": In the two-stage framework described above, suppose that at $t=1$ preference over $\mathscr{F}_T$ is $\succsim_T^*$ and preference over $\mathscr{F}_{S\setminus T}$, the set of acts over $S\setminus T$, is $\succsim_{S\setminus T}^*$. The asterisk is used to distinguish these conditional orderings from those defined by (2.1). For simplicity, and in conformity with *all* of the earlier literature on updating cited in the introduction, we suppose that the conditional orderings do not depend on previously chosen acts. As an extreme form of dynamic inconsistency of the preference orderings $\succsim$, $\succsim_T$, and $\succsim_{S\setminus T}$, suppose that there exist acts satisfying

$$f' \succ_T^* f, \qquad g' \succ_{S\setminus T}^* g \qquad \text{and} \qquad \begin{bmatrix} f(s) & s\in T \\ g(s) & s\notin T \end{bmatrix} \succ \begin{bmatrix} f'(s) & s\in T \\ g'(s) & s\notin T \end{bmatrix}. \tag{2.2}$$

Consider the choice between the above two composite acts over S, i.e., between a and b in Fig. 1. It is argued that the inconsistency in (2.2) is not an issue, since the choice problem would be solved recursively: f' is chosen given T, g' is chosen given $S\setminus T$, and these choices are taken as constraints at time 0. But now consider the choice between the same two acts in the case where no new information, in the form of $s\in T$ or $s\notin T$, is forthcoming at $t=1$, i.e., c versus d in Fig. 1. Then clearly the act (f on T; g on $S\setminus T$) would be chosen. It follows that the actual choice between (f on T; g on $S\setminus T$) and (f' on T; g' on $S\setminus T$) depends on the prevailing information structure and so cannot be rationalized by any preference ordering over $\mathscr{F}$ [21]. While such dependence is sensible in some settings (see Section 3.3), it seems to us difficult to justify at a prescriptive level in sequential problems such as statistical decision problems.

Next suppose that at time 0 the decision maker can choose between the partition $\{T, S\setminus T\}$ and the trivial partition $\{S, \phi\}$ to determine the information forthcoming at $t=1$, i.e., she can choose the "experiment" to be conducted. Then the recursive approach leads to the choice of the trivial partition, i.e., information will be rejected, even if it is costless. To highlight the counterintuitive nature of this implication, consider a statistical decision problem in which at $t=1$ one chooses an element from $D\{\delta, \delta'\}$. That choice and the true state of the world jointly determine the outcome via the reward function $r\colon D\times S\to X$. Define $f'(\cdot)\equiv r(\delta',\cdot)$ and $f(\cdot)\equiv r(\delta,\cdot)$ on T, $g'(\cdot)\equiv r(\delta,\cdot)$ and $g(\cdot)\equiv r(\delta',\cdot)$ on $S\setminus T$. Then the two conditional rankings in (2.2) represent a situation where one would expect information to have

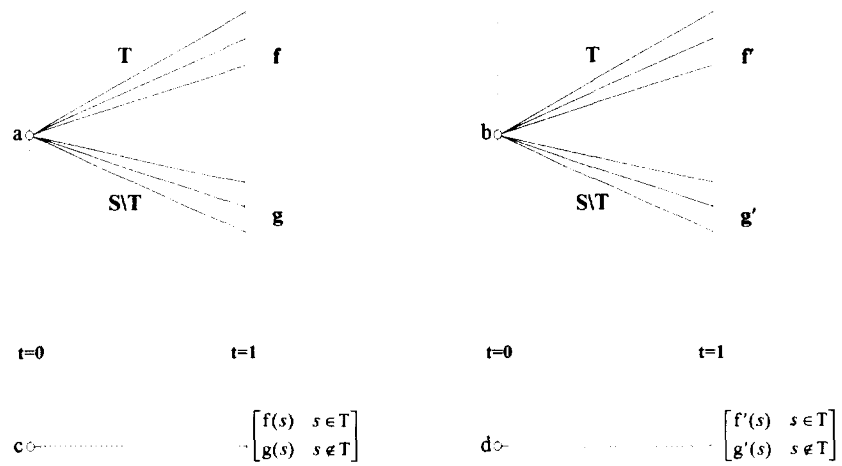

FIG. 1. Two-stage acts.

strictly positive value, namely δ' is better given T and δ is better given $S \setminus T$. But if the preference orderings violate dynamic consistency as in (2.2) and if the problem is solved recursively, then at $t=0$ the decision maker would strictly prefer to have no information available to guide the later choice of an element from D. Once again, though such an aversion to information is plausible in intertemporal settings where the "real" time separating $t=0$ and $t=1$ is substantial, we find it normatively unappealing in sequential settings such as statistical decision problems. Moreover, on a positive level, introspection and casual observation suggest to us that such aversion to information is uncommon in sequential contexts.

2.3. *An Example*

We illustrate our theorem and the surrounding discussion by considering, as in [18], the intersection of the nonadditive probability and multiple priors models. According to this model, one possible representation for the ordering $\succsim$ over $\mathscr{F}$ consists of a utility index $u: X \to \mathbb{R}$ and a set Δ of additive probability measures on S. (The set Δ is convex and closed in the weak* topology.) The utility function $U: \mathscr{F} \to \mathbb{R}$ representing $\succsim$ is given by

$$U(f) \equiv \min\left\{\int u(f(s))\, dp(s) \colon p \in \Delta\right\}. \tag{2.3}$$

The Savage expected utility model is obtained if Δ is a singleton. More generally, $\succsim$ satisfies P1 and P4, with the implied comparative likelihood relation $\succsim_l$ represented by $v(\cdot)$, where

$$v(A) \equiv \min\{p(A): p \in \Delta\}, \qquad \text{for all events } A.$$

P5* is satisfied if $u(X)$ contains at least three elements, while P6 is satisfied if v has convex range, i.e., if $A \subset B$ and $l \in (v(A), v(B))$ imply $\exists C$, $A \subset C \subset B$ such that $v(C) = l$. Axiom P3 is satisfied under additional restrictions on Δ, e.g., if all elements in Δ assign zero measure to the same events.

However, P4^c is generally violated. To determine the circumstances under which $\succsim_{T,g}$ satisfies P4, note that it is represented by the utility function $U_{T,g}: \mathscr{F}_T \to \mathbb{R}$, where for each $f \in \mathscr{F}_T$,

$$U_{T,g}(f) = \min\left\{\int_T u(f(s))\,dp(s) + \int_{S\setminus T} u(g(s))\,dp(s): p \in \Delta\right\}. \quad (2.4)$$

Let $x^* \succ z \succ x$ and suppose the unrealized alternative g is such that $g(s) \equiv z$ on $S\setminus T$. Further, let A and B be arbitrary subsets of T, and consider comparison with respect to $\succsim_{T,g}$ of the two bets

$$\begin{bmatrix} y^* & \text{if} & s \in A \\ y & \text{if} & s \in A\setminus T \end{bmatrix} \quad \text{vs.} \quad \begin{bmatrix} y^* & \text{if} & s \in B \\ y & \text{if} & s \in T\setminus B \end{bmatrix}.$$

Axiom P4 for $\succsim_{T,g}$ requires that the ranking of these bets be the same for all specifications of $y^* \succ y$. Taking $(y^*, y) = (x^*, z)$ and then (z, x), we conclude immediately that v must satisfy

$$v(B) \leqslant v(A) \qquad \text{iff} \quad v(B \cup (S\setminus T)) \leqslant v(A \cup (S\setminus T)).$$

Since this must be true for all events T, it follows that v is a qualitative probability and, given P6, that v can be represented by a probability measure μ ([15, Theorem 14.2]). In other words, v is a distortion of probability measure μ in the sense that

$$v(\cdot) = \phi(\mu(\cdot)) \quad (2.5)$$

for some strictly increasing function ϕ from the unit interval onto itself. Since, by assumption, $\succsim$ can be represented also by a Choquet integral with respect to v ([32]), it follows that $\succsim$ is probabilistically sophisticated, confirming our theorem.

To summarize, adopt the model of preference (2.3). Then the updating rule defined by (2.4) delivers dynamic consistency, but the updated ordering is generally not "based on beliefs"; in particular, it cannot be represented by a minimum of expected utilities over some set of updated probability measures over T. Alternatively, if $p(T)>0$ $\forall p \in \Delta$, one could apply Bayes Rule to each element of Δ to generate a set Δ_T of probability measures over T, and then represent the preference ordering conditional on T by the suitable minimum of expected utilities. This procedure for updating sets of measures element by element has been widely discussed (see [14, 23, 37], for example). But it generally leads to the dynamic inconsistency of preferences. Only in the event of (2.5) and hence that $\succsim$ is probabilistically sophisticated, are both desiderata—dynamic consistency and updated preference "based on beliefs"—satisfied; and then the above two procedures for updating coincide.[4]

2.4 *Machina–Schmeidler*

As indicated earlier (Section 1.3) and as evidenced by the proof of our theorem provided in the appendix, we rely heavily on Theorem 2 of Machina and Schmeidler [29]. Since they also conclude that preference must be probabilistically sophisticated, some elaboration on the value added by this paper is in order.

Their Theorem 2 assumes P1, P3, P5 (rather than P5*), P6, and P4* (rather than P4^c), where

P4* (Strong Comparative Probability). *For all pairs of disjoint events A and B, outcomes $x^* \succ x$ and $y^* \succ y$, and acts g and h,*

$$\begin{bmatrix} x^* & \text{if} & s \in A \\ x & \text{if} & s \in B \\ g(s) & \text{if} & s \notin A \cup B \end{bmatrix} \succsim \begin{bmatrix} x & \text{if} & s \in A \\ x^* & \text{if} & s \in B \\ g(s) & \text{if} & s \notin A \cup B \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} y^* & \text{if} & s \in A \\ y & \text{if} & s \in B \\ h(s) & \text{if} & s \notin A \cup B \end{bmatrix} \succsim \begin{bmatrix} y & \text{if} & s \in A \\ y^* & \text{if} & s \in B \\ h(s) & \text{if} & s \notin A \cup B \end{bmatrix}.$$

This axiom plays a central role in achieving the important objective these authors define for themselves—to demonstrate that the notion of probabilistic sophistication can be separated axiomatically from the expected utility valuation component of the Savage model. Though Machina and Schmeidler do not offer their Theorem 2 as an argument

[4] The fact that the model (2.3) does not admit dynamically consistent updating has played a key role in some recent applications to sequential settings. For example, [7, 9] adopt the Dempster–Shafer updating rule. These analyses can be understood as implicitly assuming that the decision maker behaves naively in the face of the resulting dynamic inconsistency of preferences. It is this naive behaviour, rather than the vagueness of beliefs, that underlies the derived results. In contrast, the applications to portfolio choice in [8] deal with one-shot choice problems and so updating is not an issue.

against non-Bayesian models of beliefs, one might be tempted to interpret it as such. But in that role, we feel that the theorem is unconvincing because of the limited force of P4*.

The axiom P4* seems questionable to us because it requires that the indicated ranking of acts be invariant if the act g on $S \backslash (A \cup B)$ is replaced by any other act h. If in P4* we restrict $(y^*, y) = (x^*, x)$, then such invariance is implied by the Sure-Thing Principle which is restated here for the convenience of the reader.

P2 (Sure-Thing Principle). *For all events T and acts f, f', g, and h,*

$$\begin{bmatrix} f(s) & \text{if} & s \in T \\ g(s) & \text{if} & s \notin T \end{bmatrix} \succsim \begin{bmatrix} f'(s) & \text{if} & s \in T \\ g(s) & \text{if} & s \notin T \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} f(s) & \text{if} & s \in T \\ h(s) & \text{if} & s \notin T \end{bmatrix} \succsim \begin{bmatrix} f'(s) & \text{if} & s \in T \\ h(s) & \text{if} & s \notin T \end{bmatrix}.$$

The invariance in P4* corresponds to the special case $T = A \cup B$, $f = (x^* \text{ on } A;\, x \text{ on } B)$ and $f' = (x \text{ on } A;\, x^* \text{ on } B)$. One can make the case, as Machina and Schmeidler do, that this represents a *plausible* weakening or specialization of the Sure-Thing Principle. However, once the latter is rejected, it seems to us impossible to argue that any form of invariance with respect to acts on complementary events is *compelling* at any level—either as a coherence requirement for the model, or at the level of the decision maker in a prescriptive or descriptive sense. In particular, we expect that most researchers who take the Ellsberg Paradox seriously and therefore admit violations of the Sure-Thing Principle, would be willing to drop P4* if that were the "price" to be paid for modeling vague beliefs.

A second difficulty with P4*, from the prespective of this paper, is that it is explicitly violated by the typical choices in the Ellsberg Paradox (Section 1.2).[5] Therefore, it hardly seems appropriate to assume P4* in arguing against models of beliefs that are motivated by the Ellsberg Paradox and similar evidence. Of course, Machina and Schmeidler do not do so, and they observe the above violation; once again, we are merely cautioning against a misinterpretation of their results.

In short, Machina and Schmeidler do not reveal any serious obstacles to modeling preferences that are not probabilistically sophisticated. In particular, the connection between probabilistic sophistication and dynamic consistency does not arise in their paper. That connection is the heart of our contribution.

[5] The conflict between $P4^c$ and Ellsberg-type behaviour is less direct (see Lemma 2 in the Appendix).

3. A Reexamination

While P4^c was discussed in detail above, the other assumptions in our theorem were not. We turn now to an examination of some of those assumptions, both explicit and implicit, to see if an "escape" can be found from what we interpret as the negative conclusion of the theorem. Three possibilities are considered briefly in turn, only the last of which seems to us to be promising.

3.1. *Incomplete Preferences*

It can be argued that incompleteness is a natural feature of preference in an environment where objective probabilities are unavailable (see [1, 37]). If P1 is weakened by dropping the completeness requirement, then P3, P4 and indeed even P4^c are satisfied by the following ordering: For a given utility index u and Δ, a set of additive probability measures on S,

$$f \succsim g \quad \text{if} \quad \int_S u(f(s))\, dp(s) \geq \int_S u(g(s))\, dp(s) \qquad \text{for } all \quad p \in \Delta. \tag{3.1}$$

On the other hand, as pointed out by Bewley, (3.1) cannot resolve the Ellsberg Paradox. More generally, even if we drop the completeness requirement from our set of axioms, Lemma 2 in the Appendix remains valid, contradicting Ellsberg-type behaviour.

This observation is relevant to the literature on updating sets of priors, including [14, 23]. Much of that literature does not specify an explicit decision-theoretic famework. Earlier we pointed out that a model of complete preference does not provide an appropriate foundation for any procedure for updating sets of priors. In contrast, the procedure of applying Bayes Rule to each prior in a set *can* be rationalized by a model of incomplete preference. For instance, given (3.1), when conditioning on an event T for which $p(T) > 0 \forall p \in \Delta$, the dynamic consistency of preference is implied if the updating of preference is accomplished by applying Bayes Rule to each element of Δ. However, we have just seen that any such rationalization is unsatisfactory in the sense of contradicting Ellsberg-type behaviour.

3.2. *Weak Dynamic Consistency*

One could argue that the following weakening of our dynamic consistency requirement would suffice:

$$f \succsim_{T,g} f' \Rightarrow \begin{bmatrix} f(s) & \text{if} & s \in T \\ g(s) & \text{if} & s \notin T \end{bmatrix} \succsim \begin{bmatrix} f'(s) & \text{if} & s \in T \\ g(s) & \text{if} & s \notin T \end{bmatrix}. \tag{3.2}$$

This weak form of dynamic consistency would permit $f \succsim_{T,g} f'$ and (f on T; g on $S \setminus T$) $\sim$ (f' on T; g on $S \setminus T$). If the latter were chosen at $t=0$, the choice would be overturned later, but the $t=0$ choice of (f on T; g on $S \setminus T$) would be respected at $t=1$. More, generally, condition (3.2) would ensure that in suitable sequential optimization problems, some but possibly not all acts that are optimal with respect to the ex ante preference ordering would remain optimal at future decision nodes with respect to updated preference.

The weakening of the dynamic consistency requirement from (2.1) to (3.2) is relevant to the problem of updating conditional on events that ex ante were not expected to occur [3]. If (2.1) is used to define updated preference and T is null with respect to $\succsim$, then all acts are indifferent according to $\succsim_{T,g}$. This problematic feature of an updating rule is avoided if only (3.2) is imposed.

Modify our theorem accordingly by assuming, instead of P4^c, that each $\succsim_{T,g}$ satisfies (3.2) and the appropriate form of P4. Then, by adapting the proof of Lemma 2, we can show that for all events E_1, E_2 and E, $(E_1 \cup E_2) \cap E = \phi$,

$$E_1 \succ_l E_2 \Rightarrow E_1 \cup E \succsim_l E_2 \cup E,$$

which is still incompatible with the typical choices in the Ellsberg problem described above (take $E_1 = B$, $E_2 = R$ and $E = S \setminus (R \cup B)$). In the context of the example of the last section, the above weaker set of axioms is satisfied if the distortion (2.5) is such that ϕ is nondecreasing and not necessarily strictly increasing. We have not yet succeeded in deducing further implications of the weakened axioms in the general case.

3.3. *Temporal Resolution of Uncertainty*

A final possibility for escaping the grasp of our theorem arises if it is recognized explicitly that arguments based on dynamic consistency invariably refer to a domain consisting of multistage decision trees. In contrast, our analysis assumes that the objects of preference are acts over S, and thus the domain does not permit one to encode the sequential resolution of uncertainty. For instance, the two-stage acts a and c from Fig. 1 are both identified with the same element of $\mathscr{F}$, namely (f on T; g on $S \setminus T$). The restriction to such a narrow domain is justified if the decision maker is indifferent to the way in which uncertainty is resolved through the tree [27], that is, if she is indifferent between a and c. But when real time is involved in passing between decision nodes, such indifference is not at all compelling. Explicit reconsideration of the issues of this paper in the richer domain implied by sequential decision making seems to us the most promising way to model dynamically consistent vague beliefs, at least for intertemporal problems. However, in cases where the interval between $t=0$ and $t=1$ is short or indeed only conceptual, indifference to the temporal resolution of uncertainty seems to us to be compelling. In particular, this "escape route" does not seem to apply to statistical decision problems.

This final suggestion draws attention to an interesting parallel with the axiomatic underpinnings of the von Neumann–Morgenstern utility model for the case of risk, i.e., objective probability distributions. Viewed from the prespective of a domain of multistage lotteries, the vNM model is implied by the assumptions of (a) dynamic consistency, (b) the axiom of reduction of compound lotteries, and (c) independence from unrealized alternatives or consequentialism. (See [28, 33, 11, 25], for elaboration.) Above we suggested that a counterpart to the axiom of reduction of compound lotteries is implicit in our theorem. In a sense, therefore, the theorem shows that in a framework of uncertainty the Bayesian beliefs component of the Savage model is implied by the counterparts to (a) and (b).

We are led to make two remarks: First, consequentialism is not assumed in our theorem in that we allow the updated ordering $\succsim_{T,g}$ to depend on g. In that sense, the argument that probabilistic non-sophistication entails dynamic inconsistency is "stronger" than the parallel argument in the case of risk that non-expected utility preferences necessarily lead to dynamic inconsistencies. In particular, Machina's [28] critique of the latter argument, since it is essentially a critique of consequentialism, is not relevant here. Second, in the case of risk, (a) and (c) are compatible with non-expected utility functions if reduction is dropped (see the discussion of recursive utility functions in [11]). Similarly, if reduction is dropped in our setting, a recursive construction of utility can deliver preferences that involve non-Bayesian beliefs and exhibit suitable forms of dynamic consistency and independence from unrealized alternatives. Such a recursive construction is used in [13] to define intertemporal utility in the presence of vague beliefs.

4. Concluding Remarks

We have examined the question "can preferences that are based on beliefs and admit dynamically consistent updating in response to new information be probabilistically non-sophisticated?" Our motivation has been the presumption that the effect of vagueness on behaviour that is exhibited in the context of the Ellsberg Paradox is present much more broadly in situations of decision making under uncertainty. We have provided a negative answer to the above question. Though an "escape" was suggested for intertemporal settings—admitting nonindifference to the temporal resolution of uncertainty—the negative answer seems to us to be significant, particularly from the perspective of normative choice theory. In particular, it casts doubt upon whether recently developed models based on non-additive probabilities or multiple priors can provide adequate foundations for statistics. More generally, this paper casts doubt upon the appeal of models of preference in which tastes and beliefs are separated in the sense of Savage's axiom P4. The investigation of alternatives to P4 seems to us to be an important subject for further research.[6]

Appendix: Proof of Theorem

In light of [29, Th. 2], we need only show that our axioms imply P4*. Recall from Section 2.1 that given P3, $P4^c$ is equivalent to P4′, which is more convenient below. The following lemma provides more than needed.

Lemma 1. (a) $P4^* \Rightarrow P4'$.

(b) *Given* P1 *and* P5*, *then* $P4' \Rightarrow P4^*$.

Proof. Consider A, B, T, x^*, x, y^*, y, and g as in the statement of P4′. The events $A \cap B$, $B \backslash A$, $A \backslash B$ and $T \backslash (A \cup B)$ form a partition of T. Therefore, the act

$$\begin{bmatrix} x^* & \text{if} & s \in A \\ x & \text{if} & s \in T \backslash A \\ g & \text{if} & s \notin T \end{bmatrix}$$

can be equivalently rewritten

$$\begin{bmatrix} x^* & \text{if} & s \in A \backslash B \\ x & \text{if} & s \in B \backslash A \\ x^* & \text{if} & s \in A \cap B \\ x & \text{if} & s \in T \backslash (A \cup B) \\ g(s) & \text{if} & s \notin T \end{bmatrix}.$$

A similar rewriting can be performed for the three other acts in the statement of P4′. Now, if P4* is satisfied, apply it to the pair of disjoint events $A \backslash B$ and $B \backslash A$, and in the "complementary act"

[6] Segal [34] describes a class of preferences, based on "probabilities over probabilities," that violate P4 and can resolve the Ellsberg Paradox. But they do not exhibit dynamic consistency in the sense of this paper.

$$\begin{bmatrix} x^* & \text{if} & A \cap B \\ x & \text{if} & T \backslash (A \cup B) \\ g(s) & \text{if} & s \notin T \end{bmatrix}$$

replace x^* by y^*, x by y, and g by g. Conclude that P4′ is satisfied.
Part (b) is an implication of the next two lemmas. ▮

LEMMA 2. *Assume* P4′ *and* P5*. *Then for all events* A, B, *and* C *such that* $C \cap A = C \cap B = \phi$,

$$A \succsim_l B \quad \textit{if and only if} \quad A \cup C \succsim_l B \cup C.$$

Proof. Consider outcomes x, y, and z such that $x \succ y$ and $y \succ z$. If $A \succsim_l B$, then

$$\begin{bmatrix} x & \text{if} & s \in A \\ y & \text{if} & s \notin A \cup C \\ y & \text{if} & s \in C \end{bmatrix} \succsim \begin{bmatrix} x & \text{if} & s \in B \\ y & \text{if} & s \notin B \cup C \\ y & \text{if} & s \in C \end{bmatrix}.$$

By P4′, therefore,

$$\begin{bmatrix} y & \text{if} & s \in A \\ z & \text{if} & s \notin A \cup C \\ y & \text{if} & s \in C \end{bmatrix} \succsim \begin{bmatrix} y & \text{if} & s \in B \\ z & \text{if} & s \notin B \cup C \\ y & \text{if} & s \in C \end{bmatrix},$$

which implies that $A \cup C \succsim_l B \cup C$. This argument can be reversed to prove the converse implication. ▮

LEMMA 3. *Assume* P1, P4′, *and* P5*. *Then for all events* A, B, *and* E *such that* $A \cup B \subseteq E$, $f \in \mathscr{F}$ *and outcomes* x^* *and* x *with* $x^* \succ x$,

$$A \succsim_l B \quad \textit{iff} \quad \begin{bmatrix} x^* & \textit{if} & s \in A \\ x & \textit{if} & s \in E \backslash A \\ f(s) & \textit{if} & s \notin E \end{bmatrix} \succsim \begin{bmatrix} x^* & \textit{if} & s \in B \\ x & \textit{if} & s \in E \backslash B \\ f(s) & \textit{if} & s \notin E \end{bmatrix}.$$

Proof. Suppose $A \succsim_l B$. We proceed by induction on $k \equiv \# f(S \backslash E)$, the cardinality of $f(S \backslash E)$. If $k = 1$, let $f(s) = y$ for $s \in S \backslash E$. By P1 and P5*, $\exists z$ in X such that either $y \succ z$ or $z \succ y$. In the first case the desired ranking is true if $x^* = y$ and $x = z$, since $A \cup (S \backslash E) \succsim_l B \cup (S \backslash E)$ by Lemma 2. By P4′, any $x^* \succ x$ will do. In the second case the desired ranking is true if $x^* = z$ and $x = y$ since $A \succsim_l B$. Again, P4′ implies that any $x^* \succ x$ will do.

Suppose now that $\# f(S\backslash E)=k+1$, $y \in f(S\backslash E)$, and $E_1 \equiv f^{-1}(y) \cap (S\backslash E)$. As before $\exists z \in X$ such that either $y \succ z$ or $z \succ y$. In the first case we obtain the desired ranking for $x^* = y$ and $x = z$ from the induction assumption, since $A \cup E_1 \succsim_I B \cup E_1$ be Lemma 2. In the second case the desired ranking is true for $x^* = z$ and $x = y$ because of the induction hypothesis and $A \succsim_I B$. In both cases, P4′ implies that the desired ranking holds for all $x^* \succ x$.

The preceding arguments can be reversed to prove the converse implication. ∎

References

1. T. Bewley, "Knightian Decision Theory," Cowles Paper 807, Yale University, 1986.
2. K. Border and U. Segal, Dynamic consistency implies expected utility (almost), *J. Econ. Theory*, in press.
3. L. Blume, A. Brandenburger, and E. Dekel, Lexicographic probabilities and choice under uncertainty, *Econometrica* **59** (1991), 61–79.
4. S. H. Chew and L. Epstein, Nonexpected utility preferences in a temporal framework with an application to consumption-savings behaviour, *J. Econ. Theory* **50** (1990), 54–81.
5. A. P. Dempster, Upper and lower probabilities induced by a mutlivalued map, *Ann. Math. Statist.* **38** (1967), 325–339.
6. A. P. Dempster, A generalization of Bayesian inference, *J. Roy. Statist. Soc. Ser. B* **30** (1968), 205–247.
7. J. Dow, V. Madrigal and S. Werlang, Preferences, common knowledge and speculative trade, mimeo, London Business School, 1990.
8. J. Dow and S. Werlang, Uncertainty aversion, risk aversion and the optimal choice of portfolio, *Econometrica* **60** (1992), 197–204.
9. J. Dow and S. Werlang, Excess volatility of stock prices and Knightian uncertainty, *European Econ. Rev.* **36** (1992), 631–638.
10. D. Ellsberg, Risk, ambiguity and the Savage axioms, *Quart. J. Econ.* **75** (1961), 643–669.
11. L. Epstein, Behaviour under risk: Recent developments in theory and applications, *in* "Advances in Economic Theory" (J. J. Laffont, Ed.), Cambridge Univ. Press, Cambridge, 1993, pp. 1–63.
12. L. Epstein and U. Segal, Quadratic social welfare functions, *J. Polit. Econ.* **100** (1992), 691–712.
13. L. Epstein and T. Wang, "Intertemporal Asset Pricing Under Knightian Uncertainty," Working Paper 9211, Univ. of Toronto, 1992.
14. R. Fagin and J. Halpern, A new approach to updating beliefs, mimeo, IBM, 1989.
15. P. Fishburn, "Utility Theory for Decision Making," Wiley, New York, 1970.
16. I. Gilboa, Expected utility with purely subjective non-additive probabilities, *J. Math. Econ.* **16** (1987), 65–88.
17. I. Gilboa and D. Schmeidler, Maximin expected utility with non-unique prior, *J. Math. Econ.* **18** (1989), 141–153.
18. I. Gilboa and D. Schmeidler, Updating ambiguous beliefs, *J. Econ. Theory* **59** (1993), 33–49.
19. F. Gul, Why do economists resist non-expected utility theories?, *J. Econ. Theory*, in press.
20. F. Gul and O. Lantto, Betweenness satisfying preferences and dynamic choice, *J. Econ. Theory* **52** (1990), 162–177.

21. P. HAMMOND, Changing tastes and coherent dynamic choice, *Rev. Econ. Stud.* **43** (1976), 159–173.
22. P. HAMMOND, Consequentialist foundations for expected utility, *Theory and Decision* **25** (1988), 25–78.
23. J. Y. JAFFRAY, Bayesian updating and belief functions, *IEEE Trans. on Systems, Man and Cybernetics* **22** (1992), 1144–1152.
24. E. KARNI AND Z. SAFRA, Behaviourally consistent optimal stopping rules, *J. Econ. Theory* **51** (1990), 391–401.
25. E. KARNI AND D. SCHMEIDLER, Atemporal dynamic consistency and expected utility, *J. Econ. Theory* **54** (1991), 401–408.
26. D. KREPS, "Notes on the Theory of Choice," Westview, Boulder, 1988.
27. D. KREPS AND E. PORTEUS, Temporal resolution of uncertainty and dynamic choice theory, *Econometrica* **46** (1978), 185–200.
28. M. MACHINA, Dynamic consistency and non-expected utility models of choice under uncertainty, *J. Econ. Lit.* **27** (1989), 1622–1668.
29. M. MACHINA AND D. SCHMEIDLER, A more robust definition of subjective probability, *Econometrica* **60** (1992), 745–780.
30. R. A. POLLAK, Consistent planning, *Rev. Econ. Stud.* **35** (1968), 201–208.
31. L. J. SAVAGE, "The Foundations of Statistics," Wiley, New York, 1954.
32. D. SCHMEIDLER, Subjective probability and expected utility without additivity, *Econometrica* **57** (1989), 571–587.
33. U. SEGAL, Two-stage lotteries without the reduction axiom, *Econometrica* **58** (1990), 349–377.
34. U. SEGAL, The Ellsberg paradox and risk aversion: An anticipated utility approach, *Internat. Econ. Rev.* **28** (1987), 175–202.
35. G. SHAFER, "A Mathematical Theory of Evidence," Princeton Univ. Press, Princeton, 1976.
36. R. STROTZ, Myopia and inconsistency in dynamic utility maximization, *Rev. Econ. Stud.* **23** (1956), 165–180.
37. P. WALLEY, "Statistical Reasoning with Imprecise Probabilities," Chapman and Hall, London, 1991.
38. P. WELLER, Consistent intertemporal decision making, *Rev. Econ. Stud.* **45** (1978), 263–266.