CESALTINA PACHECO PIRES

A RULE FOR UPDATING AMBIGUOUS BELIEFS

ABSTRACT. When preferences are such that there is no unique additive prior, the issue of which updating rule to use is of extreme importance. This paper presents an axiomatization of the rule which requires updating of all the priors by Bayes rule. The decision maker has conditional preferences over acts. It is assumed that preferences over acts conditional on event E happening, do not depend on lotteries received on E^c, obey axioms which lead to maxmin expected utility representation with multiple priors, and have common induced preferences over lotteries. The paper shows that when all priors give positive probability to an event E, a certain coherence property between conditional and unconditional preferences is satisfied *if and only if* the set of subjective probability measures considered by the agent given E is obtained by updating all subjective prior probability measures using Bayes rule.

KEY WORDS: Ambiguous beliefs, Bayesian updating, Dynamic choice, Multiple priors, Uncertainty aversion

1. INTRODUCTION

The Savage (1954) model of decision making under uncertainty prescribes that a decision maker has a unique prior probability and a utility function such that decisions are made so as to maximize the *expected utility*. However, there exists a large body of empirical evidence which contradicts Savage's subjective expected utility model (see Camerer and Weber, 1992). In particular, behavior such as that exhibited in the Ellsberg Paradox (1961) is inconsistent with that approach. The main problem is that the *ambiguity* that is common in situations of uncertainty cannot be captured in the representation of beliefs by a single prior. This is the motivation behind several nonexpected utility models, which have been presented in the literature in the last two decades.

Among the generalizations of decision making under uncertainty we would like to refer two types of models: the non-additive probabilities and the multiple priors model. The first type of model was first introduced by Schmeidler (1982, 1989), but was also axiomatized,

Theory and Decision **53:** 137–152, 2002.

among others, by Gilboa (1987) and Wakker (1989). Non-additive probabilities or capacities, are monotone set-functions which may fail to satisfy additivity. Using the Choquet (1953) integral, the previous works axiomatize preferences which are representable by Choquet expected utility. The second model was used by Wald (1950) and axiomatized by Gilboa and Schmeidler (1989). The model assumes that the decision maker considers a set of priors as possible and, because he is uncertainty averse, he evaluates each alternative according to its minimal expected utility, where the minimum is taken over all priors in the set.

One issue which is not addressed in the previous models is the updating of beliefs and preferences as new information arrives. However, this problem is of enormous importance. In fact, most choice problems are sequential and, therefore, updating of beliefs must be specified. Thus application of these models to intertemporal decision making or to game theory, will necessarily require assumptions on how economic agents change their beliefs over time. The same problem is also present in statistical inference, whenever it is difficult to specify a unique prior and consequently a set of prior distributions is used (see, for example, Berger, 1990; Wasserman and Kadane, 1990; Lavine et al., 1991).

Savage's axiomatic derivation does not consider the problem of how to update beliefs. Nevertheless, the consensus view is that Bayes updating rule should be used (see, for example, Brown, 1976; Kreps 1988). Unfortunately, once we abandon the Savage framework, there is no single answer to the updating question. Among the rules that have been used there are two which have wider support. The first one is used when uncertainty can be described by a belief function, and it is the cornerstone of Dempster (1967, 1968) and Shafer (1976) belief functions theory – the *Dempster–Shafer updating rule*. The other rule, sometimes known as *full Bayesian updating*, consists in updating all the priors. This rule has been proposed, among others, by Fagin and Halpern (1990), Wasserman and Kadane (1990) and Jaffray (1992).

The Dempster–Shafer updating rule was axiomatized by Gilboa and Schmeidler (1993). Our paper provides a decision theoretic axiomatization of the rule that requires updating of all the priors. We then summarize Gilboa and Schmeidler (1993) work and compare their approach with ours.

In Savage's framework there is a natural way of defining conditional preferences from the unconditional preferences of the agent. Suppose we want to derive the conditional preferences *given* E between acts f and g. Then we only need to consider acts which coincide on E with f and g respectively, and coincide with act h over E^c, denoted by $f|^h_{E^c}$ and $g|^h_{E^c}$. The sure-thing principle says that preferences of the agent between these two acts, does not depend on h. Hence if $f|^h_{E^c} \succeq g|^h_{E^c}$ for some h, one knows that $f|^h_{E^c} \succeq g|^h_{E^c}$ for all h. Then if $f|^h_{E^c} \succeq g|^h_{E^c}$ for some h we say that $f \succeq g$, *given* E.

Such a natural way of defining conditional preferences does not exist in the non-additive expected utility framework because the sure-thing principle does not hold. Here the ordering of f and g derived from the unconditional preferences as above, depends on h. That is, the "induced" conditional preferences depend on h.

Gilboa and Schmeidler (1993) extend the idea of deducing conditional preferences from unconditional preferences. In the setup of non-additive expected utility, they define an *updating rule* of a set of preference relations. An *updating rule,* specifies for any preference relation and for any event of a measurable partition of the set of states of the world, what is the preference relation once that event is known to have occurred. The only two conditions the *updating rule* has to satisfy are that if the only event known to be true is the set of all states of the world, then the updating function is the identity and that the preference relation given that event E is known to have happened depends only on the outcomes on states of the world in E. The last property is known as consequentialism.

They then consider a particular set of *updating rules*, where the preference relation *given* E is *induced* from the unconditional preferences by choosing a common h on E^c; i.e., $f \succeq g$ *given* E if and only if $f|^h_{E^c} \succeq g|^h_{E^c}$. They call this the *h-Bayesian update rule*. Clearly, when preferences don't obey the sure-thing principle, the updated preferences depend on h.

Gilboa and Schmeidler argue that choosing $h = x^*$, where x^* is the best prize, is a pessimistic update rule which fits well with the pessimism of an uncertainty averse agent[1]. The decision maker implicitly assumes that if E had not occurred the best possible outcome would have happened. Hence, the rule reflects disappointment that E occurred. This rule for updating preference relations corresponds to the maximum likelihood principle. The agent would choose from the initial set of priors only the probability measures which assign the maximum probability to event E, and update these probability measures using Bayes rule. Moreover, for preferences which can be represented simultaneously by a non-additive probability measure and by multiple priors, the maximum likelihood update rule boils down to the Dempster–Shafer conditioning rule, an essential element of the Dempster and Shafer belief function theory.

In this work we do not follow the approach of deriving the conditional preferences from the unconditional preferences by specifying an update rule. Instead we characterize the conditional preferences and how they are related to unconditional preferences. We assume that preferences over acts conditional on event E happening, do not depend on the lotteries received on E^c and obey axioms which are equivalent, under that assumption, to the ones postulated by Gilboa and Schmeidler (1989) (henceforth GS). As a consequence one can derive a representation of the conditional preferences as in GS. That is, *given* E, the decision maker considers as possible a certain set of probability measures on E and evaluates each alternative according to its minimal expected utility.

There are two axioms which link the different preference orderings. One assumes that conditional preferences have common induced preferences over lotteries. The other one relates unconditional preferences to conditional preferences. It states that if the agent is indifferent *given* E, between a certain act f and a constant act $\bar{y}$, then the unconditional preferences should be such that the agent is also indifferent between an act which gives f if E happens and the constant act $\bar{y}$ if E^c happens and the constant act $\bar{y}$. We show that this axiom is satisfied if the set of probability measures which appears in GS representation of the conditional preferences, is the set of measures obtained by updating using Bayes rule the set of all prior probability measures derived in the GS representation of the unconditional preferences. In addition, we show that if the axiom is satisfied and $p(E) > 0$ for all prior probability measures, then the set of probability measures in the conditional preferences is the one obtained by *full Bayesian updating*.

The axiomatization presented in this paper is consistent with the existence of an *updating rule* of preferences as defined by Gilboa and Schmeidler (1993), as it is clear from the assumption that preferences *given E* only depend on the outcomes in states of the world in E. However the updating rule implicit in our paper is quite more complex than the *h-Bayesian update rule*. Such a rule is imbedded in the axiom which relates conditional and unconditional preferences.

In terms of updating a set of probability measures, the rule implicit in our work is that the agent updates all the priors using Bayes rule. Although it can be argued that this rule is extreme because it assumes that the agent continues to use probability measures which give very small probability to the events which occurred, we believe it is the most appropriate rule to use. The reason is that the agent who has GS type of preferences is likely to have a "conservative" attitude. He would not reject some probability measure unless he is sure that it cannot be the true probability measure.

As mentioned before, the idea of updating all priors without excluding any of them has been proposed by other authors. However, no previous work provides a complete decision-theoretic axiomatization of the rule. Fagin and Halpern (1989) and Jaffray (1992) used this rule in a context where a set of priors can be represented by its lower envelope and the lower envelope is a belief function as defined by Dempster (1967) and Shafer (1976). In that setup, they show that the lower envelope of the set of posterior measures is a belief function itself. However, Jaffray (1992) shows that, in general, the lower envelope does not characterize the set of posterior measures. Similar results hold for two-alternating Choquet capacities (belief functions are infinitely alternating capacities, hence they are also two-alternating), as proved by Walley (1981) and Wasserman and Kadane (1990).

The main contribution of this paper is to provide a decision-theoretic axiomatization of the *full Bayesian updating rule*. That is, the paper describes dynamic preferences which imply the updating of all priors by Bayes rule. The paper can also be interpreted as an extension of the multiple prior model axiomatized by Gilboa and Schmeidler (1989) to a dynamic setup. Given the importance of dynamic decision making under uncertainty this is an important value added.

In the remainder of the paper we describe the notation and the axioms satisfied by the conditional preferences and present the results.

2. FRAMEWORK

We will adopt the "lottery-acts" framework of Anscombe and Aumann. Let X be the set of consequences. Let Y be the set of distributions over X, with finite support. The elements of Y are called lotteries (roulette lotteries). Let S be the set of states of the world, and Σ be an algebra on S. The elements of Σ are events. L_0 is defined as the set of Σ-measurable finite step functions from S to Y. Let L_c be the set of constant functions in L_0. Let L be a convex subset of Y^S which includes L_c. Convex combinations in Y^S are performed pointwise; i.e., for f and g in Y^S and α in [0, 1], $\alpha f + (1-\alpha)g = h$ where $h(s) = \alpha f(s) + (1-\alpha)g(s)$ for all s in S.

Acts are functions from the set of states of the world to the set of lotteries ($S \to Y$). The set of acts is denoted by L. A constant act $\bar{y}$ is an act which gives the lottery y in any state of the world. We denote by $\overline{f(s)}$ the constant act which gives in every state of the world s', the same lottery that f gives in the state s, $f(s)$.

The decision-maker has conditional preferences over acts. We will denote by $\succeq^E$ the preferences *given* E. In other words if the agent knew that E happened, then $\succeq^E$ would be his preference ordering. Let $\succeq$ be the unconditional preference ordering ($\succeq = \succeq^S$). Let $\simeq^E$ and $\succ^E$ denote the symmetric and asymmetric parts, respectively, of $\succeq^E$.

Let $g|_E^f$ be the act which coincides with g on E^c and with f on E. In particular $g|_E^{\overline{f(s)}}$ is the act which gives the lottery $f(s)$ in any state belonging to E and coincides with g otherwise. Let L_c^E be the set of acts which are constant on E.

Given a preference ordering $\succeq^E$, an event $A \in \Sigma$ is $\succeq^E$-null, or in other words A is null with respect to $\succeq^E$ iff $\forall f, g \in L,\ s.t.\ \forall s \in A^c,\ \overline{f(s)} \simeq^E \overline{g(s)}$, it is true that $f \simeq^E g$. Otherwise we say that A is $\succeq^E$-non-null.

The conditional preferences are assumed to obey the following axioms:

$\mathcal{A}1$ (Weak Order) For all f and g in L, $f \succeq^E g$, or $g \succeq^E f$ or both. For all f, g and h in L if $f \succeq^E g$ and $g \succeq^E h$ then $f \succeq^E h$.

$\mathcal{A}2$ E^c is a null-event with respect to $\succeq^E$.

$\mathcal{A}3$ (State-Independence) $f \succeq^{\{s\}} g$ if and only if $h|_E^{\overline{f(s)}} \succeq^E h|_E^{\overline{g(s)}}, \forall h$ in L and for any E in Σ.

$\mathcal{A}4$ (Certainty Independence) For all f and g in L and h in L_c^E: $f \succ^E g$ if and only if $\alpha f + (1-\alpha)h \succ^E \alpha g + (1-\alpha)h$ for all $\alpha \in (0, 1)$.

$\mathcal{A}5$ (Continuity) For all f, g, h such that $f \succ^E g$ and $g \succ^E h$, there exist α^E and $\beta^E \in (0, 1)$ such that $\alpha^E f + (1-\alpha^E)h \succ^E g$ and $g \succ^E \beta^E f + (1-\beta^E)h$.

$\mathcal{A}6$ (Monotonicity) For all $f, g \in L$ such that $f(s) \succeq^{\{s\}} g(s)$ for all $s \in E$, $f \succeq^E g$.

$\mathcal{A}7$ (Uncertainty Aversion) For all $f, g \in L$ and $\alpha \in (0, 1)$, if $f \simeq^E g$ then $\alpha f + (1-\alpha)g \succeq^E f$.

$\mathcal{A}8$ (Non-degeneracy) For all E, not for all $f, g \in L$, $f \succeq^E g$.

$\mathcal{A}9$ For all $E \in \Sigma$ such that E is $\succeq$-non-null, if $f \simeq^E \bar{y}$ then $f|_{E^c}^{\bar{y}} \simeq \bar{y}$.

Axiom $\mathcal{A}2$ captures the idea that if the agent knew that E happened his (conditional) preferences over the acts on L should not depend on the lotteries received on E^c. This property is important in defining conditional preferences. It is the same condition – consequentialism – that is present in the definition of an "updating rule".

Axiom $\mathcal{A}3$ guarantees that preferences over lotteries are well defined and common to all conditional preferences. A constant act on E means that the agent knows which lottery he will receive if E happens. If we restrict $\succeq^E$ to constant acts on E, uncertainty will be irrelevant. Hence the ordering *given* E of the constant acts on E, should coincide with the ordering of the lotteries. $\mathcal{A}3$ states that this happens for any E. We notice that $\mathcal{A}3$ is implicit in GS. In fact they derive preferences over lotteries by considering the preference ordering over constant acts.

The axioms $\mathcal{A}1$, $\mathcal{A}4$, $\mathcal{A}5$, $\mathcal{A}6$, $\mathcal{A}7$, $\mathcal{A}8$ are GS axioms, adapted to the case of conditional preferences. Since by $\mathcal{A}2$, E^c is a null event with respect to $\succeq^E$, the monotonicity and certainty independence axioms impose conditions only on the lotteries on E.

Axiom $\mathcal{A}9$ relates conditional and unconditional preferences. In particular, it tells us how the conditional preferences restrict unconditional preferences. It says that if the agent is indifferent *given* E, between a given act f and a constant act $\bar{y}$, then the unconditional preferences should be such that the agent is also indifferent between an act which gives f if E happens and the constant act $\bar{y}$ if E^c happens and the constant act $\bar{y}$. With the act $f|_{E^c}^{\bar{y}}$ the agent receives something indifferent to the constant act $\bar{y}$ both on E and E^c. However this is not why the agent is indifferent between $\bar{y}$ and $f|^{\bar{y}}_{E^c}$. As it will become clear later, the reason why the agent whose preferences obey the previous axioms, is indifferent between these two acts is because the relative weight given to states in E in the evaluation of $f|^{\bar{y}}_{E^c}$ coincides with the relative weights used in evaluating the act f, *given* E.

Axiom $\mathcal{A}9$ imposes a certain coherence in the pessimism implicit in the conditional preferences. A decision maker with ambiguous subjective beliefs and who obeys to GS axioms, evaluates each act in such a way that it is as if he chooses among the set of possible subjective probability measures the one which gives less weight to the states with the best utility outcomes and more weight to the states with the worst utility outcomes and takes the expectation of his utility with respect to this probability measure.

Let us assume that the decision maker uses this pessimist evaluation criterion, *given* E and concludes $f \simeq^E \bar{y}$. Using also the pessimistic criterion in the evaluation of the act $f|^{\bar{y}}_{E^c}$, the agent considers the probability measure in the set of possible measures with the lowest expected utility. It is clear that no matter what are the weights given to E and E^c, provided the relative weights of the states in E are the same than the one used in the evaluation of f *given* E, the agent will be indifferent between $\bar{y}$ and $f|^{\bar{y}}_{E^c}$.

3. RESULTS

In this section we present the main result of the paper. We start by stating two lemmas which are useful to the proof of the main result.

LEMMA 1. *Under $\mathcal{A}1$ and $\mathcal{A}3$ the preference ordering over lotteries is the same for all the states of the world. I.e., if $f(s) = f(t)$, $g(s) = g(t)$ and $f \succ^{\{s\}} g$ then $f \succ^{\{t\}} g$.*

Proof. Suppose not, suppose $\sim (f \succ^{\{t\}} g)$. Then by $\mathcal{A}1$, $g \succeq^{\{t\}} f$. Applying $\mathcal{A}3$ for $E = S$, $\overline{g(t)} \succeq \overline{f(t)}$, which by assumption is equivalent to $\overline{g(s)} \succeq \overline{f(s)}$. However by $\mathcal{A}3$ this is equivalent to $g \succeq^{\{s\}} f$, contradicting the assumption that $f \succ^{\{s\}} g$. □

LEMMA 2. *If the conditional preferences $\succeq^E$ satisfy axioms $\mathcal{A}1$, $\mathcal{A}2$ and $\mathcal{A}4$-$\mathcal{A}7$ there exists a GS representation of the conditional preferences $\succeq^E$. That is, there exists an affine function $u : Y \to \Re$ and a non-empty, closed, convex set C^E of finitely additive probability measures on Σ, s.t. $\forall f, g \in L_0$, $f \succeq^E g$ iff $\min_{p \in C^E} \int u \circ f \, dp \geq \min_{p \in C^E} \int u \circ g \, dp$. In addition: (1) u is unique up to a positive linear transformation and (2) C^E is unique if and only if $\mathcal{A}8$ holds.*

Proof. Since, by $\mathcal{A}2$, preferences *given* E do not depend on the lotteries received in E^c, it is enough to verify that restricting the acts to E, all the axioms which are necessary and sufficient for GS representation hold. Axioms $\mathcal{A}1$, $\mathcal{A}4$–$\mathcal{A}8$ guarantee that Theorem 1 of GS applies to the conditional preferences (see Gilboa and Schmeidler, 1989). □

We denote by B the space of all bounded Σ-measurable real valued functions on S (this is denoted by $B(S, \Sigma)$ in Dunford and Schwartz (1957)). Let B_0 then denote the space of functions in B which assume finitely many values. Let $K = u(Y)$, and let $B_0(K)$ be the subset of functions in B_0 with values in K.

The previous lemma guarantees that under $\mathcal{A}1$, $\mathcal{A}2$ and $\mathcal{A}4$–$\mathcal{A}8$ the preference ordering *given* E has a GS representation. The function on L_0

$$J^E(f) = \min_{p \in C^E} \int u \circ f \, \mathrm{d}p = I^E(u \circ f)$$

where u is unique up to a positive linear transformation and C^E is unique, represents the preference ordering $\succeq^E$. $J^E : L_0 \longrightarrow \Re$ is the function which associates with each act the corresponding conditional utility and $I^E : B_0 \longrightarrow \Re$ is a functional such that $I^E(u \circ f) = J^E(f)$. The properties of I^E (derived from the axioms) guarantee that there exists a closed and convex set C^E of finitely additive probability measures on Σ, such that for all $b \in B$, $I(b) = \min_{p \in C^E} \int b dp$. Hence one can express the utility of an act as a function of the utility received in each state and the set of finitely additive probability measures[2].

Axiom $\mathcal{A}3$ implies that the function $u : Y \to \Re$ is the same, up to a positive linear transformation, to all preference orderings $\succeq^E$. The next proposition summarizes these results and establishes the link between the sets of additive probability measures C^E and $C^S = C$. In particular, the result relates C^E with the set of posterior measures obtained by *full Bayesian updating* of C. Since Bayes rule only applies if $p(E) > 0$, we assume that when C includes measures that assign probability zero to event E, those measures are excluded in computing the set of posterior measures of C given E, which we denote by C/E.

PROPOSITION 1. *Let $\succeq^E$ be a set of binary relations on L_0. Then the following conditions are equivalent:*

(i) The binary relations $\succeq^E$, for all $E \in \Sigma$ satisfy axioms $\mathcal{A}1$–$\mathcal{A}7$ for $L = L_0$.

(ii) There exists an affine function $u : Y \to \Re$, and non-empty closed and convex sets, C^E, of finitely additive measures on Σ such that:

$$f \succeq^E g \text{ iff } \min_{p \in C^E} \int u \circ f \, \mathrm{d}p \geq \min_{p \in C^E} \int u \circ g \, \mathrm{d}p$$

In addition:

(a) The function u is unique up to a positive linear transformation. The sets C^E are unique iff $\mathcal{A}8$ holds.

(b) If C^E is equal to the set of posterior probability measures using Bayes rule of C given E, $C^E = C/E$, then $\mathcal{A}9$ is satisfied. On the other hand, if $\mathcal{A}9$ is satisfied $C/E \subseteq C^E$. Moreover, if $p(E) > 0$ for all $p \in C$, and $\mathcal{A}9$ is satisfied then $C^E = C/E$.

Proof. The first part of the result is Lemma 2, hence only part (b) remains to be proved. By Lemma 2 one knows that:

$$I^E(u \circ f) = \min_{p^E \in C^E} \int u \circ f \, \mathrm{d}p^E \text{ and } \tag{1}$$

$$I(u \circ f|^{\bar{y}}_{E^c}) = \min_{p \in C} \left[\int_E u \circ f \, \mathrm{d}p + \int_{E^c} u \circ \bar{y} \, \mathrm{d}p \right] \tag{2}$$

where u is common in both functionals by Lemma 1.

Let us first show that if C^E is the set of posterior measures of C given E then $\mathcal{A}9$ holds. That is, if $C^E = C/E$ and $f \simeq^E \bar{y}$ then $f|^{\bar{y}}_{E^c} \simeq \bar{y}$. Or equivalently, using the GS representation, if $C^E = C/E$ and $I^E(u \circ f) = u(y)$ then $I(u \circ f|^{\bar{y}}_{E^c}) = u(y)$.

Assume p^{*E} is the minimand in problem (1) when $C^E = C/E$. Since $f \simeq^E \bar{y}$, $u(y)$ is the value of the functional in (1). Define the set C^* as the set of probability measures in C with posterior probability p^{*E}; i.e. $C^* = \{p \in C : p/E = p^{*E}\}$. If $p(E) = 0$ for some p in C, problem (2) can be written as:

$$\min\left[\min_{p\in C: p(E)>0}\left[p(E)\int_E \frac{u\circ f}{p(E)}dp+(1-p(E))u(y)\right], u(y)\right]. \tag{3}$$

The solution to this problem will be no greater than $u(y)$. If $p(E) > 0$ for all p in C, problem (2) can be written as:

$$\min_{p\in C}\left[\ p(E)\int_E \frac{u\circ f}{p(E)}\,dp + (1-p(E))u(y)\ \right]. \tag{4}$$

Since $C^* \subseteq C$ the solution to this problem will be no greater than

$$\min_{p\in C^*}\left[\ p(E)\int_E \frac{u\circ f}{p(E)}\,dp + (1-p(E))u(y)\ \right], \tag{5}$$

which is equal to $u(y)$. We need to prove that the solution to (3) or to (4) cannot be smaller than $u(y)$. Suppose it can, and assume $\exists\ \hat{p} \in C$ such that $\hat{p}$ is the solution to problem (3) or (4) and $I(u\circ f|^{\bar{y}}_{E^c}) < u(y)$. The only way this can happen is if $\hat{p}(E) > 0$ and $\int_E \frac{u\circ f}{\hat{p}(E)}\,d\hat{p} < u(y)$. But since the posterior of $\hat{p}$ belongs to $C^E = C/E$ this would mean that p^{*E} could not be the solution to (1), a contradiction.

Let us now show that if axiom $\mathcal{A}9$ holds then $C/E \subseteq C^E$. Supposing the statement is false, assume $\exists p' : p' \in C/E$ but $p' \notin C^E$, then we can show that $\mathcal{A}9$ fails, that is $\exists f \in L_0$ and $\bar{y} \in L_c$ such that $f \simeq^E \bar{y}$ but $\sim (f|^{\bar{y}}_{E^c} \simeq \bar{y})$.

If $\exists p' : p' \in C/E$ and $p' \notin C^E$ then, by a separation theorem (Dunford and Schwartz, 1957, V.2.10), there exists $a \in B$ such that $\int a\mathrm{d}p' < \min_{p\in C^E}\int a\mathrm{d}p$. Without loss of generality we may assume that $a \in B_0(K)$. Hence there exists $f \in L_0$, such that

$u \circ f = a$ and

$$\int u \circ f \, \mathrm{d}p' < \min_{p \in C^E} \int u \circ f \, \mathrm{d}p \Leftrightarrow$$
$$\int_E u \circ f \, \mathrm{d}p' < \min_{p \in C^E} \int_E u \circ f \, \mathrm{d}p. \tag{6}$$

Let $\bar{y} \in L_c$ be such that $f \simeq^E \bar{y}$ (by continuity $\bar{y}$ exists). Then the RHS of (6) is equal to $u(y)$. Multiplying the previous inequality by $0 < k \leq 1$ and adding up $(1-k)u(y)$ on both sides we obtain:

$$k \int_E u \circ f \, \mathrm{d}p' + (1-k)u(y) < u(y) \tag{7}$$

Inequality (7) holds for any $0 < k \leq 1$. But this implies that for any prior measure $p'_0 \in C$ with posterior p' we have $\int u \circ f|^{\bar{y}}_{E^c} \mathrm{d}p'_0 < u(y)$. Since $I(u \circ f|^{\bar{y}}_{E^c}) \leq \int u \circ f|^{\bar{y}}_{E^c} \mathrm{d}p'_0$, one concludes that $I(u \circ f|^{\bar{y}}_{E^c}) < u(y)$. But then $\mathcal{A}9$ does not hold.

Finally, we need to show that if $p(E) > 0$ for all $p \in C$ and $\mathcal{A}9$ holds then $C^E = C/E$. Since we already know that $C/E \subseteq C^E$, we only need to prove that there exist no $p'' : p'' \in C^E$ and $p'' \notin C/E$.

By contradiction, assume $\exists\, p'' : p'' \in C^E$ but $p'' \notin C/E$. Then, by a Dunford and Schwartz's separation theorem (1957, V.2.10), there exists an $a \in B$ such that $\int a \mathrm{d}p'' < \min_{p \in C/E} \int a \mathrm{d}p$. Without loss of generality we may assume that $a \in B_0(K)$. Hence there exists $f \in L_0$, such that:

$$\int u \circ f \, \mathrm{d}p'' < \min_{p \in C/E} \int u \circ f \, \mathrm{d}p$$

Let $\bar{y} \in L_c$ be such that $f \simeq^E \bar{y}$. Since $I^E(u \circ f) = u(y) \leq \int u \circ f \, \mathrm{d}p''$ the previous inequality implies

$$u(y) < \min_{p \in C/E} \int u \circ f \, \mathrm{d}p$$

which, by definition of C/E, is equivalent to

$$u(y) < \min_{p \in C} \int_E \frac{u \circ f}{p(E)} \mathrm{d}p.$$

However this implies that $u(y) < \int_E u \circ f \,/p(E)\mathrm{d}p$ for all $p \in C$ (recall that, by assumption, $p(E) > 0$ for all $p \in C$). Therefore,

$$u(y) < p(E) \int_E \frac{u \circ f}{p(E)} \mathrm{d}p + (1 - p(E))u(y)$$

for all $p \in C$. As a consequence

$$u(y) < \min_{p \in C} \left[p(E) \int_E \frac{u \circ f}{p(E)} \mathrm{d}p + (1 - p(E))u(y) \right] = I(u \circ f|^{\bar{y}}_{E^c}).$$

But this means that $\mathcal{A}9$ does not hold. □

Our result is the first one to give a complete decision-theoretic axiomatization of the *full Bayesian updating rule*. However, we would like to mention that one direction of our result is implicit in Proposition 1 of Jaffray (1994). In his work, Jaffray justifies the use of Hurwicz α–criteria with the axioms of rational decision making under mixed uncertainty (Cohen and Jaffray, 1985) and defends its use after conditioning. According to Hurwicz criterion when the decision maker considers the set of prior C as possible, the utility of an act is a convex combination of its minimal and its maximal expected utility. Hence, the Wald criterion is a special case of Hurwicz criterion, in which the decision maker gives zero weight to the maximal expected utility.

Following our assumptions, in particular that u is common to all preferences, and using our notation Proposition 1 of Jaffray (1994) could be stated as follows:

(1) *Let E be such that $p(E) > 0$ for all $p \in C$. The value of* $\min_{p \in C/E} \int u \circ f \ \mathrm{d}p$ *is the unique value $u(y)$ that solves the equation:*

$$\begin{aligned} u(y) &= \min_{p \in C} \left[\int_E u \circ f \, dp + (1 - p(E))u(y) \right] \\ &\Leftrightarrow \min_{p \in C} \left[\int u \circ f|^{\bar{y}}_{E^c} dp \right] - u(y) = 0. \end{aligned}$$

In other words, if $C^E = C/E$ then axiom A9 is satisfied.

Another interpretation of the previous property is that one can compute bounds on conditional expectations (in our case lower bounds, but a similar property holds for upper bounds) by computing bounds on prior expectations, one idea that has been used in *robust Bayesian inference*. Moreover, it is also related to the *linearization technique* first proposed by Lavine (1991a), and used, among others, by Wasserman et al. (1993). The technique also appears as a *generalized Bayes rule* in Walley (1991). The idea of this technique is to compute the bounds on conditional expectations, which is a nonlinear function of the prior, by evaluating as accurately as desired the bounds of a linear function of the prior.

In our framework the properties of the conditional preferences are primitives of the model. They are as important as the axioms on unconditional preferences. Our main result shows that when $p(E) > 0$ for all $p \in C$, the conditional preferences can be derived uniquely from the unconditional ones, since in this case $C^E = C/E$ and u is common to conditional and unconditional preferences. However, if $p(E) = 0$ for some $p \in C$ the conditional preferences are no longer fully determined by the unconditional ones. In this case, the conditional preference ordering $\succeq^E$ is needed to completely characterize the set of posterior measures. Note that in our framework it is possible to define the set of posterior measures, even if the event is a null event with respect to the unconditional preferences. This feature is also present in Walley (1991). This author does not use a decision theoretical framework, but his modelling strategy is similar to ours in many respects. He considers the *conditional previsions* as fundamental as the *unconditional previsions,* assumes that conditional previsions satisfy *separate coherence* and imposes a coherence condition between conditional and unconditional previsions.

4. CONCLUSION

In the context of maxmin expected utility with multiple prior this paper axiomatized a rule for updating ambiguous beliefs. The paper assumes that the decision maker has conditional preferences: he or she is able to order acts conditional on a given event happening. Conditional preferences have common induced preferences over lotteries and can be represented by a maxmin expected utility with multiple prior. The main result of the paper is related to a certain coherence property between conditional and unconditional preferences. It is assumed that if act f is indifferent, conditional on event E happening, to the constant act $\overline{y}$ then unconditional preferences are such that the constant act $\overline{y}$ and the act which gives f if E happens and $\overline{y}$ if E^c happens are also indifferent.

The paper shows that if the set of subjective posterior measures is obtained by full Bayesian updating of the set of subjective prior measures, then the previous coherence property holds. In addition, if all prior probability measures give positive probability to event E and the coherence property holds the set of measures in the conditional preferences given E is the one obtained by full Bayesian updating of the set of priors.

This paper makes two principal contributions to the economic literature. The first is a decision-theoretic axiomatization of the *full Bayesian updating rule*. The other contribution is an extension of the maxmin expected utility model to a variable information context.

NOTES

1. Gilboa (1989) has used another updating rule for capacities which corresponds to choosing $h = x_*$, where x_* is the worst prize.
2. Existence and uniqueness of J^E is proved in Lemma 3.2. of GS. Existence and properties of I^E (monotonicity, superadditivity, homogeneity of degree one and C-independence) are derived in Lemma 3.3 of GS. Existence of C^E is proved in Lemma 3.5 of GS.

REFERENCES

Berger, J. (1990), Robust Bayesian analysis: sensitivity to the prior, *Journal of Statistical Planning and Inference*, 25, 303–328.

Brown, P.M., (1976) Conditionalization and expected utility, *Philosophy of Science* 43, 415–419.

Camerer, C.F. and Weber M. (1941), Recent developments in modelling preferences: uncertainty and ambiguity, *Journal of Risk and Uncertainty*, 5, 325–370.

Choquet, G. (1953–1954), Theory of capacities, *Annals lÍnstitut Fourier*, 5, 131–295.

Cohen, M. and J.Y. Jaffray, (1985), Decision making in a case of mixed uncertainty: a normative model, *Journal of Mathematical Psychology* 29, 428–442.

Dempster, A.P. (1967), Upper and lower probabilities induced by a multivalued map, *Annals of Mathematical Statistics*, 38, 325–339.

Dempster, A.P. (1968), A generalization of Bayesian inference, *Journal of Royal Statistics Society, Ser. B*, 30, 205–247.
Dunford, N. and J.T. Schwartz, *Linear Operators, Part 1*, Interscience, New York, 1957.
Ellsberg, D. (1961), Risk, ambiguity and the Savage axioms, *Quarterly Journal of Economics*, 75, 643–669.
Fagin, R. and Halpern, J. (1989), A new approach to updating belief, in *Proc. 6th Conference Uncertainty and AI*, 1990. I.B.M. Res. Rep. RJ 7222.
Gilboa, I. (1987), Expected utility theory without purely subjective non-additive probabilities, *Journal of Mathematical Economics* 16, 65–88.
Gilboa, I. (1989), Additivizations of non-additive measures, *Mathematical Operations Research,* 14, 1–17.
Gilboa, I. and Schmeidler D. (1989), Maxmin expected utility with non-unique prior, *Journal of Mathematical Economics*, 18, 141–153.
Gilboa, I. and Schmeidler D. (1993), Updating ambiguous beliefs, *Journal of Economic Theory*, 59, 33–49.
Jaffray, J.Y. (1992), Bayesian updating and belief Functions, *IEEE Transactions on Systems, Man and Cybernetics*, 22, 1144–1152.
Jaffray, J.Y. (1994), Dynamic decision making with belief functions, in Yager et al. (ed.), *Advances in the Dempster-Shafer Theory of Evidence*, Wiley.
Kreps, D. (1998), *Notes on the Theory of Choice*. Westview, Boulder.
Savage, L.J. (1954), *The Foundations of Statistics*. Wiley, New York.
Schmeidler, D. (1982), Subjective probability without additivity, Working Paper, Foerder Institute for Economic Research, Tel Aviv University.
Schmeidler, D. (1989), Subjective probability and expected utility without additivity, *Econometrica*, 57, 571–587.
Shafer, G. (1976), *A Mathematical Theory of Evidence*, Princeton, NJ: Princeton University Press.
Wakker, P. (1989), Continuous subjective expected utility with non-additive probabilities, Journal of Mathematical Economics, 18, 1–17.
Wald, A. (1950), *Statistical Decision Functions*. Wiley, New York.
Walley, P. (1981), Coherent lower (and upper) probabilities, Statistics research report, University of Warwick, Coventry.
Walley, P. (1991), *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London.
Wasserman, L.A. and Kadane J. (1990), Bayes' theorem for Choquet capacities, *The Annals of Statistics*, 18, 1328–1339.
Wasserman, L., Lavine M., and Wolpert R. (1993), Linearization of Bayesian robustness problems, *J. Statistical Planning and Inference*, 37, 307–316.

Address for correspondence: Prof. Cesaltina Pacheco Pires, Departamento de Gestão de Empresas, Universidade de Évora, Largo dos Colegiais, no. 2, 7000-803 Évora, Portugal.
Phone: +351-266-740-892; Fax: +351-266-742-494;
E-mail: cpires@uevora.pt