Bayesian consistent prior selection

Christopher P. Chambers[*] and Takashi Hayashi[†‡§]

August 2005

Abstract

A subjective expected utility agent is given information about the state of the world in the form of a set of possible priors. She is allowed to condition her prior on this information. A set of priors may be updated according to Bayes' rule, prior-by-prior, upon learning that some state of the world has not obtained. We show that there exists no decision maker who obeys Bayes' rule, conditions her prior on the available information (by selecting a prior in the announced set), and who updates the information prior-by-prior using Bayes' rule. The result implies that at least

[*]Assistant Professor of Economics, Division of the Humanities and Social Sciences, Mail Code 228-77, California Institute of Technology, Pasadena, CA 91125. Email: chambers@hss.caltech.edu. Phone: (626) 395-3559.

[†]Assistant Professor of Economics, Department of Economics, University of Texas at Austin, BRB 1.116, Austin, TX 78712. Email: th925@eco.utexas.edu. Phone: (512) 475-8543.

[‡]Corresponding author.

[§]We would like to thank Kim Border, Federico Echenique, Larry Epstein, Max Stinchcombe, John Quiggin, and Bill Zame for helpful discussions and comments. All errors are our own.

one of several familiar decision theoretic “paradoxes” is a mathematical necessity.

1 Introduction

The classical theories of Savage [16] and Anscombe and Aumann [2] serve to provide behavioral foundations in environments of subjective uncertainty justifying the hypothesis of subjective expected utility. An agent whose behavior conforms to the axioms of either model can be viewed as if she perceives uncertainty in a probabilistic sense. Thus, she attributes some probability measure to the set of states of the world, and evaluates state-contingent payoffs in an expected utility fashion, with respect to this probability measure. The theories are very elegant and the behavioral conditions posited are quite intuitive. However, a significant gap in the theory is that it provides no method of specifying how such a probability measure should be formed.

Recently, several models have been proposed which attempt to fill this gap. One example is the case-based decision theory of Gilboa and Schmeidler [12]. In this model, past experiences of the decision maker are explicitly modelled and incorporated in the formation of a prior. Other models assume the decision maker is given some information about the true state of the world, in the form of a *set of priors* (for example, see Ahn [1], Damiano [3], Gajdos et al. [8, 9], Hayashi [13], Klibanoff et al. [14], and Stinchcombe [18]). The decision-maker is given the information that the “true” prior lies in some possible set. She is then allowed to form a subjective probability which depends on this information.

This is the approach we follow.

To understand this approach, recall the classical Ellsberg paradox [4]. An urn contains balls of many colors. The decision maker is told that the composition (relative proportions of colors) of the balls in the urn lies within some set, but is told nothing else. A ball is drawn, and the state of the world is the color of ball drawn. A composition of balls in an urn is identified with a probability measure over the states. The idea is that the decision maker's subjective probability over the states of the world can depend on the information she is given.

Our primary result in this note demonstrates a fundamental conflict in such a model. Informally speaking, we establish that any subjective expected utility agent who obeys Bayes' rule cannot reasonably condition her subjective probability on the information given to her. More specifically, we take "reasonable" to mean that if the set of priors revealed to the agent is both closed and convex, then the agent is required to select a subjective probability from this set.

To understand the result more concretely, return to the Ellsberg urn example. Suppose there are three colors of balls in an urn, say, red, blue and, green. The decision maker is told that the composition of balls lies within some set, and she constructs a subjective probability over the states (colors) as a function of this information. Now, suppose that it is revealed to the decision maker that all green balls have been removed from the urn. This is equivalent to being told that the state of the world is not green, so the decision maker should naturally update her subjective probability according to Bayes' rule (conditioning on the event that red or blue occurs). Moreover, the decision maker's objective information has also clearly changed. For any possible composition of balls that the decision maker was originally given, the composition has changed so that there are no green balls. Viewing a composition as a probability, the new compositon is simply the Bayes' update of the original probability conditional on the event that the green state did not obtain. Thus, the decision maker possesses an information set now which is the prior-by-prior update of the original information set (the new set consists of the set of Bayesian updates of the priors in the original set).

It is important to stress here that the set of priors models *objective information*. Prior-by-prior updating is totally objective in this environment. This stands in stark contrast to models which feature a set of subjective priors, where different subjective updating rules have been proposed (see, for example, Gilboa and Schmeidler [11]).

We require that the Bayesian update of the decision maker's original prior in this environment is exactly the prior she would have selected had she originally been told the updated set of priors. A decision maker who behaves in such a way always looks ahead, considering what she would do if one or more states does not obtain. A decision maker who does not behave in this way faces two situations with identical information, and treats them differently depending on how she arrived at the information.

The impossibility result demonstrates that such a decision maker does not exist. It is impossible for a subjective expected utility decision maker to obey

Bayes' rule, always select a prior from the information set given to her, and always treat situations independently of how she arrived at them. Thus, we establish that *all* decision makers must violate one of the principles specified above. We believe this provides a mathematical argument explaining the prevalence of various decision-theoretic paradoxes (the Ellsberg paradox, non-Bayesian updating). *All* decision makers *must* exhibit behavior that can be described as "paradoxical," at least according to a subjective expected utilty standpoint.

Section 2 discusses and proves the main theorem. Section 3 discusses the implications of the result and concludes. An Appendix extends the analysis to an environment in which decision makers are allowed to make set-valued selections. The results in this environment are surprisingly not much more promising.

2 The model and primary result

Let Ω be a finite set of states of the world. For any nonempty subset $E \subset \Omega$, the set of probability measures over E is denoted $\Delta(E)$. The set of all convex and compact subsets of $\Delta(E)$ is denoted $\mathcal{P}(E)$. The set of all convex and compact subsets of $\Delta(E)$ consisting only of measures having full support on E is denoted $\mathcal{P}^{fs}(E)$.

A **prior selection problem** consists of a nonempty subset of Ω, say, E, and a set $P \in \mathcal{P}(E)$. The domain of all prior selection problems will be written $\mathcal{X}$, and the domain of all prior selection problems (E, P) for which $P \in \mathcal{P}^{fs}(E)$

is denoted $\mathcal{X}^{fs}$.

A **prior selection rule** is a function $\psi : \mathcal{X} \to \bigcup_{E \in 2^{\Omega} \setminus \varnothing} \Delta(E)$, such that for all $(E, P) \in \mathcal{X}$, $\psi(E, P) \in P$. A **full support prior selection rule** takes as domain $\mathcal{X}^{fs}$. We generalize the notion of a prior selection rule in the appendix, in order to accommodate the possibility of an agent who selects a set of priors.

Our main interest is in studying the prior-by-prior updating rule for sets of priors, and in understanding when a prior selection rule is consistent with respect to Bayesian updating. To this end, we will be concerned only with full support prior selection rules (so as not to worry about the case in which one must update on a set of probability zero). Thus, for all nonempty $E \subset \Omega$, and for all nonempty $F \subset E$, for all $P \subset \mathcal{P}^{fs}(E)$, the set $P^F \subset \mathcal{P}^{fs}(F)$ is given by

$$P^F \equiv \left\{ \frac{p}{p(F)}|_{2^F} : p \in P \right\}.$$

Hence, P^F is simply that set of probabilities that results from updating P prior-by-prior.

A full support prior selection rule is **Bayesian consistent** if for all $(E, P) \in \mathcal{X}^{fs}$, and for all nonempty $F \subset E$, $\psi\left(F, P^F\right)(\cdot) = \psi(E, P)(\cdot|F)$. In other words, the prior selected when information about the state is revealed is simply the Bayesian update of the originally selected prior.

The primary purpose of this note is to establish that there exists no full support prior selection rule satisfying Bayesian consistency.

Theorem: Suppose that $|\Omega| \geq 3$. Then there exists no Bayesian consistent full support prior selection rule.

Proof. Let $E \subset \Omega$ such that $|E| = 3$. Without loss of generality, label $E \equiv \{a, b, c\}$. We will construct four elements of $\mathcal{P}^{fs}(E)$ These four sets are illustrated in Figure 1 1. Here, conv denotes the convex hull). To do so, we need to define some preliminary elements of $\Delta(E)$. Define $\{p^i\}_{i=1}^6$ as follows:

$$\begin{array}{cccc} & a & b & c \\ p^1 & 3/13 & 9/13 & 1/13 \\ p^2 & 3/7 & 3/7 & 1/7 \\ p^3 & 1/5 & 3/5 & 1/5 \\ p^4 & 1/3 & 1/3 & 1/3 \\ p^5 & 3/5 & 1/5 & 1/5 \\ p^6 & 3/7 & 1/7 & 3/7 \end{array} \cdot$$

We define the following elements of $\mathcal{P}(E)$.

$$\begin{aligned} P_1 &\equiv \text{conv}\left\{p^1, p^2, p^3\right\}, \\ P_2 &\equiv \text{conv}\left\{p^2, p^3, p^4\right\}, \\ P_3 &\equiv \text{conv}\left\{p^2, p^4, p^5\right\}, \\ P_4 &\equiv \text{conv}\left\{p^4, p^5, p^6\right\}. \end{aligned}$$

Consider the problems $(E, P_1), (E, P_2), (E, P_3), (E, P_4) \subset \mathcal{X}^{fs}$. We claim that $P_1 \cap P_4 = \varnothing$. This is obvious; for all $p \in P_1$, $p(b) \geq 3p(c)$. However, for all $p \in P_4$, $p(c) \geq p(b)$.

We will now establish that $\psi(E, P_1) = \psi(E, P_4)$, which is a contradiction.

The argument is very geometric, so we will illustrate the first step in Figure

2 2.

Let $p^* = \psi(E, P_1)$. Clearly, $P_1^{\{a,b\}} = \operatorname{conv}\left\{p^1(\cdot|\{a,b\}), p^2(\cdot|\{a,b\}), p^3(\cdot|\{a,b\})\right\}$ and $P_1^{\{a,c\}} = \operatorname{conv}\left\{p^1(\cdot|\{a,c\}), p^2(\cdot|\{a,c\}), p^3(\cdot|\{a,c\})\right\}$. By Bayesian consistency, $\psi\left(\{a,b\}, P_1^{\{a,b\}}\right)(\cdot) = p^*(\cdot|\{a,b\})$ and $\psi\left(\{a,c\}, P_1^{\{a,c\}}\right)(\cdot) = p^*(\cdot|\{a,c\})$.

The shaded lines on the faces of the simplex on Figure 2 are $P_1^{\{a,b\}}$ and $P_1^{\{a,c\}}$. In particular, this illustrates geometrically how to find the Bayesian update of an information set. It is a projection of the information set onto the corresponding face of the simplex from the opposite vertex.

But note that $P_2^{\{a,b\}} = P_1^{\{a,b\}}$ and that $P_2^{\{a,c\}} = P_1^{\{a,c\}}$. (This is also cleary visible from Figure 2). Therefore, $\psi\left(\{a,b\}, P_2^{\{a,b\}}\right)(\cdot) = p^*(\cdot|\{a,b\})$ and $\psi\left(\{a,c\}, P_2^{\{a,c\}}\right)(\cdot) = p^*(\cdot|\{a,c\})$. Let $p^{**} = \psi(E, P_2)$. By Bayesian consistency, $p^{**}(\cdot|\{a,b\}) = p^*(\cdot|\{a,b\})$ and $p^{**}(\cdot|\{a,c\}) = p^*(\cdot|\{a,c\})$. But this is only possible if $p^{**} = p^*$. Hence, we conclude that $\psi(E, P_2) = \psi(E, P_1)$, and it has to lie on the facet on which P_1 and P_2 intersect.

The remainder of the proof lies in establishing that $\psi(E, P_3) = \psi(E, P_2)$ and that $\psi(E, P_4) = \psi(E, P_3)$. That $\psi(E, P_3) = \psi(E, P_2)$ follows from the fact that $P_3^{\{a,c\}} = P_2^{\{a,c\}}$ and $P_3^{\{b,c\}} = P_2^{\{b,c\}}$, and an identical argument using Bayesian consistency, and it has to be the point at which P_1, P_2, P_3 intersect.

That $\psi(E, P_4) = \psi(E, P_3)$ follows from the fact that $P_4^{\{a,b\}} = P_3^{\{a,b\}}$ and $P_4^{\{a,c\}} = P_3^{\{a,c\}}$, and an identical argument using Bayesian consistency. Hence, $\psi(E, P_4) = \psi(E, P_1)$, which demonstrates the existence of $p^* \in P_1 \cap P_4$, a contradiction. ■

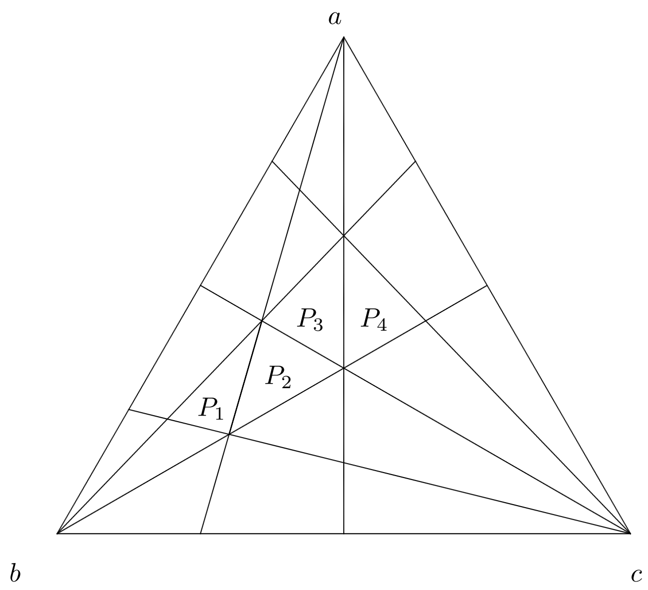

Figure 1: Information sets

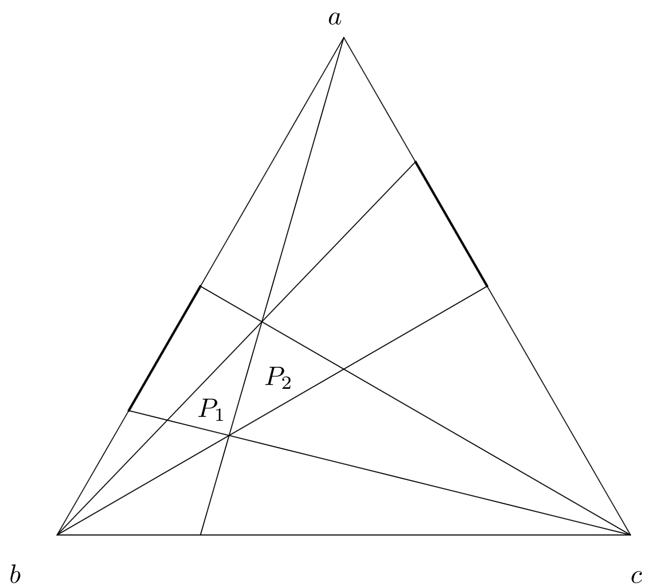

Figure 2: Constructing Bayesian updates of information sets

3 Discussion and conclusion

Our impossibility result demonstrates that there can be no Bayesian consistent prior selection rule. But what does this really mean for the theory? As we see it, it demonstrates the impossibility of the conjunction of several principles taken all at once. The primary point of this note is to demonstrate that one of these principles must be violated.

What are these principles? The first is the subjective expected utility hypothesis. The second is the selection hypothesis. The third is the dynamic consistency hypothesis. The fourth we shall refer to as the history independence hypothesis. Our claim is that at least one of these principles must be violated for any decision maker. We discuss each of these principles in turn.

3.1 The subjective expected utility hypothesis

Our model relies on the assumption that the decision maker obeys the axioms of subjective expected utility. There are many models which have been developed generalizing subjective expected utility (for example, see Gilboa and Schmeidler [10] and Schmeidler [17]). These models are originally developed to accommodate the "Ellsberg paradox." The Ellsberg paradox demonstrates a type of nonseparability in beliefs. One version of the paradox works like this: An urn is filled with three colors of balls. They are colored either blue, green, or red. Two-thirds of the balls are either blue or green, and one third is red. A decision maker is offered a choice between a bet on the draw of a blue or green ball or a bet on the draw of a red or green ball. Many decision makers prefer the first bet, as the total proportion of blue and green balls is known. However, when offered a choice between a bet on the draw of a blue ball or on the bet on the draw of a red ball, many decision makers would prefer the second bet But such a decision maker cannot conform to the subjective expected utility axioms. Of course, an agent who does not even select a prior violates the most basic of our assumptions. Hence, such a decision maker is not ruled out by our other axioms. However, in the appendix, we show that an agent whose behavior conforms to the multiple priors model [10] and who uses a prior-by-prior updating rule also cannot exist, at least when she always makes a nontrivial selection from the information set given to her.

3.2 The selection hypothesis

The next possibility we mention is the idea that a decision maker need not select a prior belonging to the set of priors that is revealed to her. While this is of course possible, for a "rational" decision maker, some very mild requirements on behavior will rule this out. Suppose that the utility index of a decision maker is affine (one can do this in the Anscombe-Aumann theory) and that the decision maker's behavior satisfies the following. For every possible prior in the information set, the expectation of a given bet is greater than zero. If she would always accept such a bet under her subjective prior, then a simple separation argument establishes that she must select a prior from the information given to her. Nevertheless, a decision maker who does not always select a prior from the information set offered to her is not ruled out by our remaining assumptions.

3.3 The dynamic consistency hypothesis (Bayesian updating)

What about an expected utility agent who does not update her prior in a Bayesian fashion? Indeed, this is again observed often in the experimental literature, and a new theoretical literature has risen up to explain this type of behavior (see Epstein [5] and Epstein et al. [7]). Bayesian updating and dynamic consistency are implicit in many axiomatizations of subjective expected utility (see for example, Epstein and LeBreton [6]) Nevertheless, an agent who does not update her prior according to Bayes' rule escapes the impossibility result.

3.4 The history independence hypothesis

We lastly discuss the idea of history independence built into our model. Starting from a particular set of states, when a certain state is ruled out, the decision maker is not allowed to take into consideration the original problem being faced. Thus, two information sets whose Bayesian updates induce the same new information set are treated equivalently after updating. Independence of this type is implicit in the decision maker. Non-expected utility models suggest that it may be normatively sound to allow decisions to depend on the original information set before updating (see, for example, Machina [15], who expounds upon the argument in an objective setting). Thus, the only remaining way to avoid our impossibility result is to allow the choice of a prior to depend nontrivially on any original information set faced in the past. As mentioned in the introduction, a decision maker of this type may face two situations which are identical in their information, but behave differently depending on how she arrived at this information.

3.5 Conclusion

All of the preceding types of behavior have been discussed before in the literature. Each of them has been recognized by decision theorists, and argued both experimentally and normatively for many years. What is new here is the mathematical necessity of behavior of one of the preceding types. In the Appendix, we show that the situation does not change much when moving to the multiple priors model with prior-by-prior updating. While multiple priors decision makers exist that are Bayesian consistent, such decision makers exhibit pathological behavior. For a large class of information sets (including those which are the convex hulls of two priors), such a decision maker is not allowed to make any kind of subjective judgment. Her set of subjective priors must coincide with the objective information given to her. This is a powerful statement about the behavior of decision makers facing objective information in the form of a set of priors. Such a decision maker, when facing certain sets, is required to use the objective information as subjective information. Note that the setup of the standard Ellsberg paradox mentioned above is the convex hull of two priors, so this is in the realm of our corollary (up to some modification–we have worked throughout with probability measures with full support in order to make our results more powerful. The classical Ellsberg paradox uses an information set that includes probability measures without full support. However, the argument is easily extended to such information sets).

4 Appendix: Set-valued selections

When set-valued selections are allowed, obviously the identity mapping is Bayesian consistent. However, below we show that there is no 'nontrivial' selection which is Bayesian consistent. Set-valued selections are interesting as there is a large literature devoted to the "multiple priors" model, initiated axiomatically by Gilboa and Schmeidler [10]. In this model, the decision maker forms a set of priors as her subjective belief. The *subjective* prior-by-prior updating rule is often advocated as a method of updating ambiguous beliefs (though there is not nearly as much consensus on this issue as there is in the subjective expected utility case).

Let us redefine the problem so as to include set-valued selections.

A **prior selection problem** consists of a nonempty subset of Ω, say, E, and a set $P \in \mathcal{P}(E)$. The domain of all prior selection problems will be written $\mathcal{X}$, and the domain of all prior selection problems (E, P) for which $P \in \mathcal{P}^{fs}(E)$ is denoted $\mathcal{X}^{fs}$.

A **prior selection rule** is a function $\psi : \mathcal{X} \to \bigcup_{E \in 2^{\Omega} \setminus \varnothing} \mathcal{P}(E)$, such that for all $(E, P) \in \mathcal{X}$, $\psi(E, P) \subset P$. A **full support prior selection rule** takes as domain $\mathcal{X}^{fs}$.

Similarly as before, for all nonempty $E \subset \Omega$, and for all nonempty $F \subset E$, for all $P \subset \mathcal{P}^{fs}(E)$, the set $P^F \subset \mathcal{P}^{fs}(F)$ is given by

$$P^F \equiv \left\{ \frac{p}{p(F)}|_{2^F} : p \in P \right\}.$$

Hence, P^F is simply that set of probabilities that results from updating P prior-by-prior.

A full support prior selection rule is **Bayesian consistent** if for all $(E, P) \in \mathcal{X}^{fs}$, and for all nonempty $F \subset E$, $\psi\left(F, P^F\right) = \psi(E, P)^F$.Also, a selection rule ψ is **proper** if $\psi(E, P)$ is a proper subset of P whenever P is non-singleton.

Corollary 1: Suppose that $|\Omega| \geq 3$. Then there is no full support prior selection rule that is proper and Bayesian consistent.

Proof. Suppose, by means of contradiction, that a full support prior selection rule exists that is both proper and Bayesian consistent, say, ψ. We will derive a contradiction.

Order the set of all Bayesian consistent rules by $\varphi \preceq \varphi'$ if for all $(E, P) \in \mathcal{X}^{fs}$

$$\varphi'(E, P) \subseteq \varphi(E, P).$$

Note that for all $(E, P) \in \mathcal{X}^{fs}$ and for all Bayesian consistent rules ϕ,

$$\psi\left(F, \varphi\left(F, P^F\right)\right) = \psi\left(F, \varphi(E, P)^F\right) = \psi(E, \varphi(E, P))^F,$$

by a double application of Bayesian consistency. Hence, φ' defined by

$$\varphi'(E, P) = \psi(E, \varphi(E, P))$$

is also a Bayesian consistent rule.

Now let $\{\varphi^\lambda\}_{\lambda\in\Lambda}$ be a chain according to $\preceq$. We claim that there exists a maximal element of this chain, say φ^*. Define, for all $(E,P)\in\mathcal{X}^{fs}$,

$$\varphi^*(E,P)=\bigcap_{\lambda\in\Lambda}\varphi^\lambda(E,P).$$

Since $\varphi^\lambda(E,P)$ is a nested, decreasing chain of compact sets, its intersection is nonempty and compact. Clearly, for all $(E,P)\in\mathcal{X}^{fs}$, and all $\lambda\in\Lambda$, $\varphi^*(E,P)\subseteq\varphi^\lambda(E,P)$. It remains to show that φ^* is itself Bayesian consistent. For all $(E,P)\in\mathcal{X}^{fs}$,

$$\begin{aligned}
&\varphi^*(E,P)^F\\
=\;&\left[\bigcap_{\lambda\in\Lambda}\varphi^\lambda(E,P)\right]^F\\
=\;&\bigcap_{\lambda\in\Lambda}\varphi^\lambda(E,P)^F\\
=\;&\bigcap_{\lambda\in\Lambda}\varphi^\lambda\left(F,P^F\right)\\
=\;&\varphi^*\left(F,P^F\right).
\end{aligned}$$

To see that the second equality is true, suppose that $p\in\left[\bigcap_{\lambda\in\Lambda}\varphi^\lambda(E,P)\right]^F$. Then there exists $p^*\in\bigcap_{\lambda\in\Lambda}\varphi^\lambda(E,P)$ for which $p^*(\cdot|F)=p$. Hence, $p\in\varphi^\lambda(E,P)^F$ for all λ, or $p\in\bigcap_{\lambda\in\Lambda}\varphi^\lambda(E,P)^F$. Conversely, suppose that $p\in\bigcap_{\lambda\in\Lambda}\varphi^\lambda(E,P)^F$. Then for all $\lambda\in\Lambda$, there exists $p^\lambda\in\varphi^\lambda(E,P)$ for which $p^\lambda(\cdot|F)=p$. Let p^* be a limit point of p^λ; clearly, $p^*\in\bigcap_{\lambda\in\Lambda}\varphi^\lambda(E,P)$; and moreover, $p^*(\cdot|F)=p$. Hence $p\in\bigcap_{\lambda\in\Lambda}\varphi^\lambda(E,P)$. Therefore, for all $\lambda\in\Lambda$, $\varphi^*\succeq\varphi^\lambda$.

As each chain has a maximal element, Zorn's Lemma implies that the set of all Bayesian consistent prior selection rules has a maximal element, say, ψ^*.

We claim that this maximal element must be singleton-valued. If it is not, then there exists some $(E, P) \in \mathcal{X}^{fs}$ for which $\psi^*(E, P)$ is nonsingleton. As ψ is proper, it follows that $\psi(E, \psi^*(E, P))$ is a proper subset of $\psi^*(E, P)$. Defining ψ^{**} so that for all $(E', P') \in \mathcal{X}^{fs}$,

$$\psi^{**}(E', P') = \psi(E, \psi^*(E', P')),$$

we know that ψ^{**} is both Bayesian consistent and that $\psi^{**} \succeq \psi^*$ (where $\psi^* \succeq \psi^{**}$ is false). This contradicts the maximality of ψ^*, so that ψ^* must be singleton-valued. However, we know that there cannot exist a singleton-valued Bayesian consistent prior selection rule. This is a contradiction. ■

The preceding corollary indicates that the most natural selection rule for a multiple prior decision maker who uses the prior-by-prior updating rule is to select the entire information set. In fact, the corollary can be used to show that for all information sets P which are the convex hull of at most two points, any Bayesian consistent full support prior selection rule must coincide with the identity.

Corollary 2: Suppose that $|\Omega| \geq 3$. Let ψ be a full support Bayesian consistent prior selection rule, and let $(E, P) \in \mathcal{X}^{fs}$ such that $P = \{\lambda p + (1 - \lambda) q : p, q \in \Delta(E), \lambda \in [0, 1]\}$. Then $\psi(E, P) = P$.

Proof. By Corollary 1, any full support Bayesian consistent prior selection rule ψ cannot be proper; hence, there exists $(E, P) \in \mathcal{X}^{fs}$ for which P is nonsingleton, and for which $\psi(E, P) = P$. Since P is nonsingleton, $|E| \geq 2$. If $|E| > 2$, then there exist $\omega_1, \omega_2 \in E$ and $p_1, p_2 \in P$ for which $\frac{p_1(\omega_2)}{p_1(\omega_1)} < \frac{p_2(\omega_2)}{p_2(\omega_1)}$.

Let $F = \{\omega_1, \omega_2\}$. Clearly, P^F is nonsingleton. By Bayesian consistency, $\psi\left(F, P^F\right) = \psi\left(E, P\right)^F = P^F$. This argument establishes that we may without loss of generality assume that $|E| = 2$.

Let $\omega_3 \in \Omega \backslash \{\omega_1, \omega_2\}$, and let $E' = \{\omega_1, \omega_3\}$. Let $P' \in \mathcal{P}^{fs}(E')$ be arbitrary. We claim that $\psi(E', P') = P'$. If P' is a singleton, this is obvious. Otherwise, let $p, q \in P$ be extreme points of P for which $p(\omega_1) > q(\omega_1)$, and let $p', q' \in P'$ be extreme points of P' for which $p'(\omega_1) > q'(\omega_1)$. Let $G = \{\omega_1, \omega_2, \omega_3\}$, and define $p^*, q^* \in \Delta^{fs}(G)$ so that

$$\begin{aligned}
\frac{p^*(\omega_1)}{p^*(\omega_2)} &= \frac{p(\omega_1)}{p(\omega_2)}, \\
\frac{p^*(\omega_1)}{p^*(\omega_3)} &= \frac{p'(\omega_1)}{p'(\omega_3)}, \\
\frac{q^*(\omega_1)}{q^*(\omega_2)} &= \frac{q(\omega_1)}{q(\omega_2)},
\end{aligned}$$

and

$$\frac{q^*(\omega_1)}{q^*(\omega_3)} = \frac{q'(\omega_1)}{q'(\omega_3)}.$$

Note that these linear inequalities in addition to the fact that p^* and q^* are probability measures, uniquely determine p^* and q^*. Define $Q \in \mathcal{P}^{fs}(G)$ as $Q = \{\lambda p^* + (1-\lambda) q^* : \lambda \in [0,1]\}$. Clearly, $Q^E = P$ and $Q^{E'} = P'$. Moreover, as $p \in \psi(E, P)$ and as $\psi(E, P) = \psi(G, Q)^E$, it must be that $p^* \in \psi(G, Q)$ (as p^* is the unique element of Q for which $p^*(\cdot|G) = p$). A similar argument establishes that $q^* \in \psi(G, Q)$. Hence, $\psi(G, Q) = Q$, and by Bayesian consistency, we conclude

$$\begin{aligned}
&P' \\
&= Q^{E'}
\end{aligned}$$

$$
\begin{aligned}
&= \psi\left(G,Q\right)^{E'} \\
&= \psi\left(E',Q^{E'}\right) \\
&= \psi\left(E',P'\right).
\end{aligned}
$$

It is now trivial to extend the argument to all problems $(E,P) \in \mathcal{X}^{fs}$ for which $|E| = 2$ and P is of the required form. Now, let $(E,P) \in \mathcal{X}^{fs}$ be arbitrary, where P is of the required form. It is clear that $\psi\left(E,P\right) = P$; otherwise, there exists some $F \subset E$ such that $|F| = 2$ and $\psi\left(E,P\right)^{F}$ is a proper subset of P^{F}. ■

This statement is quite strong. It says that for the special case in which an information set is the convex hull of two priors, *no subjective judgment is permitted*. In fact, the argument is easy to generalize to certain classes of simplicies, but we will not provide such a statement.

References

[1] D.S. Ahn, Ambiguity without a state space, Working paper, Stanford University, 2003.

[2] F.J. Anscombe and R. Aumann, A definition of subjective probability, *Annals of Mathematical Statistics* 34 (1963), 199-205.

[3] E. Damiano, Choice under limited uncertainty, mimeo, Yale University, 1999.

[4] D. Ellsberg, Risk, ambiguity, and the Savage axioms, *Quarterly Journal of Economics* 75 (1961), 643-669.

[5] L. Epstein, An axiomatic model of non-Bayesian updating, mimeo, University of Rochester, 2003.

[6] L. Epstein and M. LeBreton, Dynamically consistent beliefs must be Bayesian, *Journal of Economic Theory* 61 (1993), 1-22.

[7] L. Epstein, J. Noor, and A. Sandroni, Non-Bayesian updating: a theoretical framework, University of Rochester, 2005.

[8] T. Gajdos, J.-M. Tallon, and J.-C. Vergnaud, Decision making with imprecise probabilistic information, *Journal of Mathematical Economics* 40 (2004), 647-681.

[9] T. Gajdos, J.-M. Tallon, and J.-C. Vergnaud, Coping with imprecise information: A decision theoretic approach,

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

[11] I. Gilboa and D. Schmeidler, Updating ambiguous beliefs, *Journal of Economic Theory* 59 (1993), 33-49.

[12] I. Gilboa and D. Schmeidler, Case-based decision theory, *Quarterly Journal of Economics* 110 (1995), 605-639.

[13] T. Hayashi, Information, subjective belief, and preference, mimeo, University of Texas, Austin, 2003.

[14] P. Klibanoff, M. Marinacci, and S. Mukerji, A smooth model of decision making under uncertainty, *Econometrica*, forthcoming.

[15] M. Machina, Dynamic consistency and non-expected utility in choice under uncertainty, *Journal of Economic Literature* 27 (1989), 1622-1668.

[16] L.J. Savage, "The Foundations of Statistics," 2nd edition, Dover Publications, Minneola, NY 1972.

[17] D. Schmeidler, Subjective probability and expected utility without additivity, *Econometrica* 57 (1989), 571-587.

[18] M. Stinchcombe, Choice and games with ambiguity as sets of probabilities, Working Paper, University of Texas, Austin, 2003.