RECURSIVE MULTIPLE-PRIORS*

Larry G. Epstein Martin Schneider

April, 2002

Abstract

This paper axiomatizes an intertemporal version of multiple-priors utility. A central axiom is dynamic consistency, which leads to a recursive structure for utility, to 'rectangular' sets of priors and to prior-by-prior Bayesian updating as the updating rule for such sets of priors. It is argued that dynamic consistency is intuitive in a wide range of situations and that the model is consistent with a rich set of possibilities for dynamic behavior under ambiguity.

1. INTRODUCTION

1.1. Outline

The Ellsberg Paradox [9] illustrates that aversion to ambiguity, as distinct from risk, is behaviorally meaningful. Motivated by subsequent related experimental evidence and by intuition that ambiguity aversion is important much more widely, particularly in market settings, this paper addresses the following question: "*Does there exist an axiomatically well-founded model of intertemporal utility that accommodates ambiguity aversion*?" We provide a positive response that builds on the atemporal multiple-priors model of Gilboa and Schmeidler [16]. Because

*Department of Economics, U. Rochester, Rochester, NY, 14627, lepn@troi.cc.rochester.edu, and Department of Economics, UCLA, California, 90095-1477, martin@econ.ucla.edu. Epstein gratefully acknowledges the financial support of the NSF (grant SES-9972442). We are grateful also for comments from Lars Hansen, Massimo Marinacci, Atsushi Nishimura, Marciano Siniscalchi and a referee and to Igor Kopylov for suggesting Figure 1. First version: February 2001.

intertemporal utility is also recursive, we refer to it as *recursive multiple-priors* utility.

We view intertemporal utility as a summary of *dynamic behavior* in settings where complete commitment to a future course of action is not possible. Accordingly, foundations are provided by axioms imposed on the entire utility (or preference) process, rather than merely on initial utility. Importantly, axioms do not simply apply to conditional preference after each history separately. To ensure that dynamic behavior is completely determined by preferences, a connection between conditional preferences is needed. This connection is provided by dynamic consistency.

There is another reason to assume dynamic consistency. In the Bayesian model, dynamic consistency delivers a compelling normative argument for Bayesian updating. In contrast, in nonprobabilistic models of beliefs there is no consensus about how to update (see Gilboa and Schmeidler [17] for some of the updating rules that have been studied). It is natural, therefore, to assume dynamic consistency in the multiple-priors framework to see if a unique updating rule is implied.

Our axiomatization is formulated in the domain of Anscombe-Aumann acts [2], suitably adapted to the multi-period setting, where we adopt a simple set of axioms. The essential axioms are roughly that (i) conditional preference at each time-event pair satisfies the Gilboa-Schmeidler axioms (appropriately translated to the intertemporal setting), and (ii) the process of conditional preferences is dynamically consistent.

The resulting representation for the utility of a consumption process $c = (c_t)$ is

$$V_t(c) \;=\; \min_{Q\in\mathcal{P}} E_Q\,\left[\Sigma_{s\geq t}\,\beta^{s-t}\,u(c_s)\;\mid\;\mathcal{F}_t\right], \tag{1.1}$$

where $\mathcal{P}$ is the agent's set of priors over the state space and the σ-algebra $\mathcal{F}_t$ represents information available at time t.[1] An essential feature is that $\mathcal{P}$ is restricted by the noted axioms to satisfy not only the regularity conditions for sets of priors in the atemporal model, but also a property that (following Chen and Epstein [7]) we call rectangularity. Because of rectangularity, utilities satisfy the recursive relation

$$V_t(c) \;=\; \min_{Q\in\mathcal{P}} E_Q\,\left[\Sigma_{s=t}^{\tau-1}\,\beta^{s-t}\,u(c_s)\;+\;\beta^{\tau-t}V_\tau(c)\;\mid\;\mathcal{F}_t\right] \tag{1.2}$$

[1] A limitation of our model is that each $\mathcal{F}_t$ is assumed to correspond to a finite partition; that is, we deal with event trees.

for all $\tau > t$, which in turn delivers dynamic consistency. As is apparent from these functional forms, the corresponding updating rule for sets of priors is Bayes' Rule applied prior by prior.

The close parallel between the foundations provided here for dynamic modeling with ambiguity and those that justify traditional expected utility modeling are sharper when specialized to the case where consumption takes place only at the (finite) terminal time. In that setting, we have the following results, where the first is well known and the second is a variant of our main theorem:

Bayesian result If conditional preferences at every time-event pair satisfy expected utility theory, then they are dynamically consistent if and only if each prior is updated by Bayes' Rule.

Multiple-priors result If conditional preferences at every time-event pair satisfy multiple-priors utility theory (suitably adapted), then they are dynamically consistent if and only if each set of priors is rectangular and it is updated by Bayes' Rule applied prior by prior.

Besides clarifying the nature of our analysis, the close parallel also supports our view that recursive multiple-priors utility is the counterpart of the Bayesian model for a setting with ambiguity. A similar parallel exists in the paper's setting of consumption streams, if one adds the usual assumptions of stationarity and intertemporal separability that lead to the additivity and geometric discounting in (1.1).

1.2. Related Literature

Conclude this introduction with mention of related literature. The model (1.1) is essentially that adopted in Epstein and Wang [13], though without axiomatic foundations; a continuous-time counterpart is formulated in Chen and Epstein [7].[2] A related nonaxiomatic model based on robust control theory has been proposed by Hansen and Sargent and several coauthors; see [1] and [18], for example. While these authors refer to 'model uncertainty' rather than 'ambiguity' as we do here, their model is also motivated in part by the Ellsberg Paradox and it is proposed as an intertemporal version of the Gilboa-Schmeidler model. In Section 5 we clarify

[2] A counterpart of (1.1) appears also in [3] in a social choice setting where t indexes individuals rather than time. However, no axiomatization is provided.

the behavioral content of the robust control model and draw comparisons with recursive multiple-priors.

There is a small literature on axiomatic models of intertemporal utility under risk or uncertainty. For the case of risky consumption processes, that is, where objects of choice are suitably defined lotteries (probability measures), recursive models are axiomatized in Kreps and Porteus [23], Epstein [10] and Chew and Epstein [8]. Skiadas [29] axiomatizes recursive utility when the domain consists of consumption processes, or Savage-style acts, rather than lotteries. However, his model is still restricted to choice between risky prospects; in general terms, it is related to the previously cited papers in the same way that Savage extends von-Neumann Morgenstern. Two papers that axiomatize intertemporal utility that admit a role for ambiguity are Klibanoff [21] and Wang [33]. They adopt different and more complicated preference domains and axioms. This permits them to derive a range of results that are not delivered here. Wang axiomatizes a representation similar to (1.1). However, our model provides a simpler axiomatization and hence also a clearer and sharper response to the question posed in the opening paragraph. (See Section 3.2 for further comparison.)

Sarin and Wakker [27, p. 94] observe (in their special setting) that a rectangular set of priors implies dynamic consistency. Finally, after completion of earlier versions of this paper, we learned of independent work by Wakai [31] who arrives at a characterization similar to our main result in the context of exploring conditions for the no-trade theorem to be valid when agents have multiple-priors preferences.

2. THE MODEL

2.1. Domain

Time is discrete and varies over $\mathcal{T} = \{0, 1, ..., T\}$. We focus on the finite horizon setting $T < \infty$ because of its relative simplicity. However, Appendix B considers the infinite horizon case and thus we adopt notation and formulations (of axioms, for example) that are compatible with both settings.

The state space is Ω. The information structure is represented by the filtration $\{\mathcal{F}_t\}_0^T$ that is given and fixed throughout. We assume that $\mathcal{F}_0$ is trivial and that for each finite t, $\mathcal{F}_t$ corresponds to a finite partition; $\mathcal{F}_t(\omega)$ denotes the partition component containing ω. Thus if ω is the true state, then at t the decision-maker knows that $\mathcal{F}_t(\omega)$ is true. One can think of this information structure also in terms of an event tree.

Consumption in any single period lies in the set C; for example, $C = \mathbb{R}^1_+$. Thus we are interested primarily in C-valued adapted consumption processes and how they are ranked. However, as is common in axiomatic work, we suppose that preference is defined on a larger domain, where the outcome in any period is a (simple) lottery over C, that is, a probability measure on C having finite support; the set of such lotteries is denoted $\Delta_s(C)$. Thus, adapting the Anscombe-Aumann formulation to our dynamic setting, we consider $\Delta_s(C)$-valued adapted processes, or acts of the form $h = (h_t)$, where each $h_t : \Omega \longrightarrow \Delta_s(C)$ is $\mathcal{F}_t$-measurable.[3] The set of all such acts, denoted $\mathcal{H}$, is a mixture space under the obvious mixture operation.[4]

An adapted consumption process $c = (c_t)$ can be identified with the act h such that for each ω and finite t, $h_t(\omega)$ assigns probability 1 to $c_t(\omega)$. In this way, the domain of ultimate interest can be viewed as a subspace of $\mathcal{H}$. Another important subset of $\mathcal{H}$ is $(\Delta_s(C))^{T+1}$, referred to as the subset of *lottery acts*. To elaborate, identify the act $h = (h_t)$ for which each h_t is constant at the lottery ℓ_t with $\ell = (\ell_t) \in (\Delta_s(C))^{T+1}$. Consumption levels delivered by any lottery act ℓ depend on time and on the realization of each lottery ℓ_t but not on the state ω. Thus lottery acts involve risk but not ambiguity.

The acts (ℓ_0, h_{-0}) and $(\ell_0, \ell_1, h_{-0,1})$ have the obvious meanings. Similarly, $(\ell_{-\tau,-(\tau+k)}, q, q')$ denotes the lottery act ℓ' in which $\ell'_t = \ell_t$ for $t \neq \tau, \tau+k$, $\ell'_\tau = q$ and $\ell'_{\tau+k} = q'$.

Finally, note that Gilboa and Schmeidler also adopt the Anscombe-Aumann framework. For axiomatizations of the atemporal multiple-priors model in a Savage framework see Klibanoff et al [5] and Ghirardato et al [15]. The latter provides a procedure for translating axiomatizations formulated in the Anscombe-Aumann domain into a Savage-style domain. We suspect that their procedure could be adapted to our setting.

[3] Alternatively, one might consider acts of the form $h : \Omega \longrightarrow \Delta_s(C^{T+1})$, which correspond precisely to Anscombe-Aumann acts where the deterministic outcome set is C^{T+1}. However, such a specification leaves open the question how to restrict h to respect the information structure.

[4] In the infinite horizon setting, we will deal with a suitable subset of $\mathcal{H}$ as explained in Appendix B.

2.2. Axioms

The decision maker has a preference ordering on $\mathcal{H}$ at any time-event pair represented by (t, ω). Denote by $\succeq_{t,\omega}$ the latter preference ordering, thought of as the ordering conditional on information prevailing at (t, ω). We impose axioms on the collection of preference orderings $\{\succeq_{t,\omega}\} \equiv \{\succeq_{t,\omega}: (t, \omega) \in \mathcal{T} \times \Omega\}$.

The first axiom formalizes what is usually meant by 'conditional preference.'

Axiom 1 (Conditional Preference - CP). *For each t and ω:*
(i) $\succeq_{t,\omega} = \succeq_{t,\omega^*}$ if $\mathcal{F}_t(\omega) = \mathcal{F}_t(\omega^*)$.
(ii) If $h'_\tau(\omega') = h_\tau(\omega')\ \forall \tau \geq t$ and $\omega' \in \mathcal{F}_t(\omega)$, then $h' \sim_{t,\omega} h$.

Part (i) ensures that the conditional preference ordering depends only on available information. Part (ii) reflects the fact that $\mathcal{F}_t(\omega)$ is known at t if ω is realized. Accordingly, (ii) states that at (t, ω) only the corresponding continuations of acts matter for preference. This rules out the possibility that the decision-maker, in evaluating h at (t, ω), cares about the nature of h on parts of the event tree that are inconsistent with her current information about which states are conceivable.

Next we assume that each conditional ordering $\succeq_{t,\omega}$ satisfies the appropriate versions of the Gilboa-Schmeidler axioms. We state these explicitly both for the convenience of the reader and also because our formal setup differs slightly from that in [16] as explained below in the proof of our theorem (Lemma A.1).

Axiom 2 (Multiple-Priors - MP). *For each t and ω: (i) $\succeq_{t,\omega}$ is complete and transitive. (ii) For all h, h' and lottery acts ℓ, and for all α in $(0,1)$, $h' \succ_{t,\omega} h$ if and only if $\alpha h' + (1-\alpha)\,\ell \succ_{t,\omega} \alpha h + (1-\alpha)\,\ell$. (iii) If $h'' \succ_{t,\omega} h' \succ_{t,\omega} h$, then $\alpha h'' + (1-\alpha)\,h \succ_{t,\omega} h' \succ_{t,\omega} \beta h'' + (1-\beta)\,h$ for some α and β in $(0,1)$. (iv) If $h'(\omega') \succeq_{t,\omega} h(\omega')$ for all ω', then $h' \succeq_{t,\omega} h$.[5] (v) If $h' \sim_{t,\omega} h$, then $\alpha h' + (1-\alpha)\,h \succeq_{t,\omega} h$ for all α in $(0,1)$. (vi) $h' \succ_{t,\omega} h$ for some h' and h.*

Gilboa and Schmeidler refer to their versions of the component axioms respectively as Weak Order, Certainty Independence, Continuity, Monotonicity, Uncertainty Aversion and Non-degeneracy, which names suggest interpretations. The motivation they offer applies here as well. We refer the reader to [16] and Appendix A for further discussion.

The next axiom restricts preferences only over (the purely risky) lottery acts.[6]

[5] For any given ω', $h'(\omega')$ and $h(\omega')$ are the lottery acts that deliver lotteries $h'_\tau(\omega')$ and $h_\tau(\omega')$ in every period τ and in every state.

[6] It would be unnecessary in a model where consumption occurs only at the terminal time.

Axiom 3 (Risk Preference - RP). *For any lottery act ℓ, for all p, p', q and q' in $\Delta_s(C)$, if*

$$\left(\ell_{-\tau,-(\tau+1)}, p, p'\right) \succeq_{t,\omega} \left(\ell_{-\tau,-(\tau+1)}, q, q'\right)$$

for some *ω, t and $\tau \geq t$, then it is true for* every *ω, t and $\tau \geq t$.*

Because beliefs about likelihoods are irrelevant to the evaluation of lottery acts, their ranking should not depend on the state. This property is imposed via the indicated invariance with respect to ω. Invariance with respect to τ imposes the following form of time stationarity in the ranking $\succeq_{t,\omega}$ of lottery acts $(\ell_0, ..., \ell_T)$: The ranking of (p, p') versus (q, q'), where these single-period lotteries are delivered at times τ and $\tau + 1$ respectively, (and where $\ell_{-\tau,-(\tau+1)}$ describes payoffs at other times in both prospects), does not depend on τ. Invariance with respect to t requires that $(\ell_0, ..., \ell_{t-1}, p, p', \ell_{t+1}...)$ is preferred to $(\ell_0, ..., \ell_{t-1}, q, q', \ell_{t+1}...)$ at time 0 if and only if the same ranking prevails at time t. If we assume CP, whereby only the time t continuations matter when ranking acts at t, then the ranking at t can be viewed as one between $(p, p', \ell_{t+1}...)$ and $(q, q', \ell_{t+1}...)$, and we arrive at a familiar form of stationarity (see Koopmans [22]).

The Risk Preference axiom is satisfied if the ranking of lottery acts induced by each $\succeq_{t,\omega}$ may be represented by a utility function of the form

$$U_t(\ell_0, ..., \ell_T; \omega) = \Sigma_{\tau \geq t}\, \beta^{\tau - t}\, u(\ell_\tau)$$

for some $\beta > 0$ and $u : \Delta_s(C) \longrightarrow \mathbb{R}^1$. Since this specification is common, indeed it is typically assumed further that u conforms to vNM theory, and since the axiom imposes no restrictions on how the decision-maker addresses ambiguity, which is our principal focus, we view RP as uncontentious in the present setting.

A central axiom is dynamic consistency. To state it, define nullity in the usual way. For any $\tau > t$, say that the event A in $\mathcal{F}_\tau$ is $\succeq_{t,\omega}$-null if

$$h'(\cdot) = h(\cdot) \text{ on } A^c \quad \Longrightarrow \quad h' \sim_{t,\omega} h.$$

Axiom 4 (Dynamic Consistency - DC). *For every t and ω and for all acts h' and h, if $h'_\tau(\cdot) = h_\tau(\cdot)$ for all $\tau \leq t$ and if $h' \succeq_{t+1,\omega'} h$ for all ω', then $h' \succeq_{t,\omega} h$; and the latter ranking is strict if the former ranking is strict at every ω' in a $\succeq_{t,\omega}$-nonnull event.*

According to the hypothesis, h' and h are identical for times up to t, while h' is ranked (weakly) better in every state at $t + 1$. 'Therefore', it should be ranked better also at (t,ω). A stronger and more customary version of the axiom would require the same conclusion given the weaker hypothesis that

$$h'_t(\omega) = h_t(\omega) \text{ and } h' \succeq_{t+1,\omega'} h \text{ for all } \omega' \in \mathcal{F}_t(\omega).$$

In fact, given CP, the two versions are equivalent.

Dynamic consistency may be limiting (see the example in Section 4.1). On the other hand, doing without leaves behavior unexplained unless one adds assumptions about how the conflict between different selves is resolved. Further motivation for assuming dynamic consistency was provided in the introduction.

The final axiom is adopted purely for simplicity.

Axiom 5 (Full Support - FS). *Each nonempty event in* $\cup_{t=0}^{T}\mathcal{F}_t$ *is* $\succeq_0$*-nonnull.*

More generally, if a component of the partition defined by some $\mathcal{F}_t$ were null according to $\succeq_0$, we could discard it and apply the preceding axioms to the smaller state space. In a general formulation without FS, the preceding axioms would be modified so as to apply only for a suitable subset of states rather than for all ω.

3. THE REPRESENTATION RESULT

3.1. Rectangularity

Dynamic consistency of the expected utility model is due to the law of iterated expectations and this, in turn, is due to the familiar decomposition of a probability measure in terms of its conditionals and marginals in the form:

$$p_t(\omega) = \int_{\Omega} p_{t+1}\, dp_t^{+1}(\omega). \tag{3.1}$$

Here, for any measure p on $(\Omega,\mathcal{F}_T)$, $p_t(\omega) = p(\cdot \mid \mathcal{F}_t)(\omega)$ is its $\mathcal{F}_t$-conditional and p_t^{+1} is the restriction of p_t to $\mathcal{F}_{t+1}$. A set of priors $\mathcal{P}$ on $(\Omega,\mathcal{F}_T)$ is rectangular if its induced sets of conditionals and marginals admit a corresponding decomposition.

To define rectangularity precisely, define the set of Bayesian updates by

$$\mathcal{P}_t(\omega) = \{p_t(\omega) : p \in \mathcal{P}\},$$

and define the set of conditional one-step-ahead measures by

$$\mathcal{P}_t^{+1}(\omega) = \{p_t^{+1}(\omega) : p \in \mathcal{P}\}.$$

These sets can be viewed as realizations of $\mathcal{F}_t$-measurable correspondences into $\Delta(\Omega, \mathcal{F}_T)$ and $\Delta(\Omega, \mathcal{F}_{t+1})$ respectively.[7]

Because the FS axiom will deliver measures having full support on $\mathcal{F}_T$, we formulate the following simpler definition that is appropriate for that case and avoiding thereby reference to '*a.e.*' qualifications.

Definition 3.1. *$\mathcal{P}$ is $\{\mathcal{F}_t\}$-rectangular if for all t and ω,*

$$\mathcal{P}_t(\omega) = \{\int_\Omega p_{t+1}(\omega')\, dm : p_{t+1}(\omega') \in \mathcal{P}_{t+1}(\omega')\ \forall \omega', m \in \mathcal{P}_t^{+1}(\omega)\}, \tag{3.2}$$

$$\textit{or} \quad \mathcal{P}_t(\omega) = \int \mathcal{P}_{t+1}\, d\mathcal{P}_t^{+1}(\omega).$$

When $\mathcal{P}$ is the singleton $\{p\}$, (3.2) reduces to (3.1). The key feature is that the decomposition on the right includes combinations of a marginal from $\mathcal{P}_t^{+1}(\omega)$ with *any* measurable selection of conditionals. This will typically involve 'foreign' conditionals, that is, combining the marginal of some p with the conditionals of measures other than p. Thus the essential content of rectangularity is '$\supset$', asserting that $\mathcal{P}$ is suitably large. Indeed, the inclusion '$\subset$' in (3.2) is true for any $\mathcal{P}$: for given p in $\mathcal{P}$, simply apply the decomposition (3.1). An additional observation is that rectangularity of $\mathcal{P}$ implies that of each $\mathcal{P}_t(\omega)$.

To illustrate, if $t = 0$ and if $\mathcal{F}_1$ corresponds to the binary partition $\{F_1, F_1'\}$, then the set on the right consists of all probability mixtures of the form

$$m(F_1)\, p(\cdot \mid F_1) + m(F_1')\, p'(\cdot \mid F_1'),$$

where m is a measure in $\mathcal{P}$ (restricted to $\mathcal{F}_1$) and where $p(\cdot \mid F_1)$ and $p'(\cdot \mid F_1')$, measures on $\mathcal{F}_T$, are eventwise conditionals of some measures p and p' in $\mathcal{P}$. If $p = p' = m$ above, then this mixture equals p and thus lies in $\mathcal{P}$. Rectangularity requires that the mixture lie in $\mathcal{P}$ even if the noted measures are distinct.

An important feature of rectangularity is that it implies that $\mathcal{P}$ is uniquely determined by the process of conditional 1-step-ahead correspondences $\mathcal{P}_t^{+1}$. More precisely, begin with an arbitrary set of correspondences[8]

$$P_t^{+1} : \Omega \rightsquigarrow \Delta(\Omega, \mathcal{F}_{t+1}), \tag{3.3}$$

[7] For each t, $\Delta(\Omega, \mathcal{F}_t)$ denotes the set of probability measures on the σ-algebra $\mathcal{F}_t$.

When we refer to p_t^{+1} as a selection from $\mathcal{P}_t^{+1}$, or when we write $p_t^{+1}(\omega) \in \mathcal{P}_t^{+1}(\omega)$ for all ω, then it is understood that $\omega \longmapsto p_t^{+1}(\omega)$ is $\mathcal{F}_t$-measurable; in other words, selections are understood to be measurable.

[8] Thus P_t^{+1} denotes a primitive correspondence while the calligraphic $\mathcal{P}_t^{+1}$ denotes a correspondence induced by a primitive set of priors $\mathcal{P}$. Similar notation is adopted throughout.

where P_t^{+1} is $\mathcal{F}_t$-measurable for each t. Because each measure in $P_t^{+1}(\omega)$ is a measure on $\mathcal{F}_{t+1}$, think of $P_t^{+1}(\omega)$ as the set of conditional 1-step-ahead measures describing beliefs about the 'next step'. Then there exists a unique rectangular set of priors $\mathcal{P}$ whose 1-step-ahead conditionals are given by the P_t^{+1}'s, that is,

$$\mathcal{P}_t^{+1}(\omega) \;=\; P_t^{+1}(\omega) \text{ for all } t \text{ and } \omega. \tag{3.4}$$

The asserted set $\mathcal{P}$ can be constructed by backward induction using the relation[9]

$$\mathcal{P}_t(\omega) \;=\; \{\int p_{t+1}\, dm : p_{t+1}(\omega') \in \mathcal{P}_{t+1}(\omega') \;\forall\omega',\; m \in P_t^{+1}(\omega)\}. \tag{3.5}$$

It is readily seen that the set $\mathcal{P}$ constructed in this way is the set of all measures p whose 1-step-ahead conditionals conform with the P_t^{+1}'s, that is,

$$\mathcal{P} \;=\; \{p \in \Delta(\Omega, \mathcal{F}_T) : \, p_t^{+1} \in P_t^{+1}(\omega) \text{ for all } t \text{ and } \omega\}.$$

Further, every rectangular set $\mathcal{P}$ can be described in this way; simply use (3.4) to define P_t^{+1}.

A Graphical Illustration of Rectangularity

Rectangularity can be illustrated geometrically in the probability simplex. Let $\Omega = \{R, B, G\}$, corresponding to the three colors of balls in an Ellsberg urn, and refer to Figure 1 for the corresponding probability simplex.[10] For the filtration, take $\{\mathcal{F}_t\}$ where all information is revealed at time 2, while

$$\mathcal{F}_1 \;=\; \{\{G\}, \{R, B\}\},$$

that is, the decision-maker learns at $t = 1$ whether or not the ball is green.

Every rectangular set of priors $\mathcal{P}$ is determined by the specification of 1-step-ahead conditional measures. Thus consider 1-step-ahead beliefs at time 0, that is, time 0 beliefs about the likelihood of G. Given ambiguity, these are naturally represented by a probability interval for G. Because the probability of

[9] At T, the decision-maker knows whether or not any given event in $\mathcal{F}_T$ has occurred. Thus $\mathcal{P}_T(\omega')$ consists of the single measure p, where $p(A) = 1_A(\omega')$ for A in $\mathcal{F}_T$. The proof that (3.5) implies (3.4) is similar to Step 3 in the proof of our theorem.

[10] The vertex R denotes red with probability 1. More generally, a point p in the simplex delivers red with probability given by the shortest distance between p and the face opposite R. Similarly for other colors.

G is constant along any line parallel to the face opposite G, the noted interval is defined by the region between the two negatively sloped lines shown. At time 1, conditional beliefs are trivial if G has been revealed to be true. Given $\{R, B\}$, conditional beliefs are described by an interval for the conditional probability of R. Because the conditional probability of R is constant along any ray emanating from G, an interval is determined by the region between the two rays shown. The collection of all probability measures satisfying both interval bounds is the rectangular set $\mathcal{P}$; and all rectangular sets in the simplex have this form.

Some features of rectangularity merit emphasis. First, rectangularity imposes no restrictions on 1-step-ahead conditionals - these can be specified arbitrarily. Moreover, the rectangular set $\mathcal{P}$ constructed as above from the 1-step-ahead conditionals, induces these same sets of conditionals and is the largest set of priors to do so. Third, any set of priors induces a smallest rectangular set containing it; for example, $\mathcal{P}$ is the smallest rectangular set containing $\mathcal{P}'$. More specifically, $\mathcal{P}'$ induces sets of 1-step-ahead conditionals and these generate $\mathcal{P}$ as described above. Because induced 1-step-ahead conditionals are precisely what one needs to compute utility by backward induction, we can view $\mathcal{P}$ as precisely the enlargement of $\mathcal{P}'$ needed in order to incorporate the logic of backward induction. Hence the connection between rectangularity and dynamic consistency. Finally, rectangularity is tied to the filtration. For example, if the information learned at time 1 is whether or not the color is R, then a rectangular set would have a similar geometric representation but from the perspective of the vertex R. In particular, while $\mathcal{P}'$ is rectangular relative to the new filtration, $\mathcal{P}$ is not.

3.2. The Theorem

We need some further terminology. Say that a measure p in $\Delta(\Omega, \mathcal{F}_T)$ has *full support* if

$$p(A) > 0 \text{ for every } \emptyset \neq A \in \mathcal{F}_T.$$

Say that $u : \Delta_s(C) \longrightarrow \mathbb{R}^1$ is *mixture linear* if $u(\alpha p + (1-\alpha)q) = \alpha u(p) + (1-\alpha)u(q)$ for all p and q in $\Delta_s(C)$ and $0 \leq \alpha \leq 1$.

We can now state our main result.

Theorem 3.2. *The following statements are equivalent:*

(a) $\{\succeq_{t,\omega}\}$ satisfy CP, MP, RP, DC and FS.

(b) There exists $\mathcal{P} \subset \Delta(\Omega, \mathcal{F}_T)$, closed, convex and $\{\mathcal{F}_t\}$-rectangular, with all measures in $\mathcal{P}$ having full support, $\beta > 0$ and a mixture linear and nonconstant

$u : \Delta_s(C) \longrightarrow \mathbb{R}^1$ such that: for every t and ω, $\succeq_{t,\omega}$ is represented by $V_t(\cdot, \omega)$, where

$$V_t(h,\omega) \;=\; \min_{m \in \mathcal{P}_t(\omega)} \int \Sigma_{\tau \geq t}\, \beta^{\tau-t}\, u(h_\tau)\, dm. \tag{3.6}$$

Moreover, β and $\mathcal{P}$ are unique and u is unique up to a positive linear transformation.

Because consumption processes form a subset of $\mathcal{H}$ in the way described in Section 2.1, the theorem delivers the representation promised in the introduction. In particular, in (b), rectangularity of $\mathcal{P}$ implies that utilities satisfy the recursive relation

$$V_t(h,\omega) \;=\; \min_{m \in \mathcal{P}_t^{+1}(\omega)} \int \left[u(h_t(\omega)) + \beta V_{t+1}(h)\right] dm, \tag{3.7}$$

which extends (1.2) to the domain $\mathcal{H}$.

Another point made in the introduction was the parallel with foundations for the Bayesian model. In that connection, note that one obtains an axiomatization of the subjective expected additive (geometric discounting) utility model, with Bayesian updating, if the multiple-priors axiom MP is strengthened to the appropriate versions of the Anscombe-Aumann axioms; more precisely, if MP (ii) and (v) are replaced by the independence axiom on the domain $\mathcal{H}$.

A representation analogous to that in the theorem is axiomatized in Wang [33, Theorems 5.3-5.4]. In addition to the greater complexity of that axiomatization, due in part to the more complicated domain assumed for preference, it delivers only the special case where each set of conditional 1-step-ahead measures $\mathcal{P}_t^{+1}(\omega)$ is the core of a convex capacity [28]; this restriction is not made explicit but it is clear from the proof.

To apply our model, one needs to begin with the specification of a rectangular set $\mathcal{P}$ (in the same way that to apply the Savage model, the modeler needs to select a prior). We showed in the previous section that this can be done by specifying 1-step-ahead correspondences $\{P_t^{+1}\}$. Moreover, *any* specification of $\{P_t^{+1}\}$ is admissible and generates, by backward recursion, a unique rectangular set of priors. Thus rectangularity is consistent with *any* specification of conditional beliefs about 'the next step'. Examples of such specifications are provided in the next section. The noted backward recursion underlies the dynamic consistency of preference, which in turn delivers tractability as demonstrated in [13], [7] and [11].[11]

[11] It is worth mentioning that frequently successful analysis does not require an explicit solution

Consider briefly some extensions of the theorem. It is straightforward to characterize the model in which β is restricted to be less than 1. For example, the following additional axiom on the ranking of lottery acts would characterize (3.6) with $\beta < 1$: For any p' and p in $\Delta_s(C)$, if $(p', p', ..., p') \succ_0 (p, p, ..., p)$, then $(p', p, p, ..., p) \succ_0 (p, p', p, ..., p)$. Two other extensions are discussed next.

An infinite horizon framework is desirable for the usual reasons and also because it would permit study of the long-run persistence of ambiguity. In our finite horizon model, the decision-maker knows at T the truth or falsity of any event in $\mathcal{F}_T$ and thus 'eventually' there is neither risk nor ambiguity. However, this need not be the case if we take $T = \infty$. Appendix B provides a representation result in an infinite horizon setting. Of particular note is that the set of measures $\mathcal{P}$ that it delivers is (mutually) *locally* absolutely continuous, that is, mutually absolutely continuous on $\cup_{t=0}^{\infty} \mathcal{F}_t$. However, measures in $\mathcal{P}$ need not be mutually absolutely continuous on the limiting σ-algebra $\mathcal{F}_\infty$, and thus they need not merge asymptotically to a single measure as in Blackwell and Dubins [4]. In that sense, the model permits ambiguity to persist even in the long run after repeated observations (see our paper [12] for more details).

A final extension is more speculative but we mention it in order to provide perspective on the theorem. A generalization of (3.7) is the recursive relation

$$V_t(h) = \min_{m \in \mathcal{P}_t^{+1}(\omega)} W\left(h_t, \int V_{t+1}(h)\, dm\right), \tag{3.8}$$

for a suitable aggregator function W (strictly increasing in its second argument). If $\mathcal{P}$ is a singleton, then this recursive relation is analogous to that axiomatized in Kreps and Porteus [23] and Skiadas [29] that is motivated in Epstein and Zin [14] by the desire to disentangle willingness to substitute intertemporally from attitudes towards risk. A continuous-time version of (3.8) is provided in [7], where it is argued that it permits a three-way separation between the two noted aspects of preference and attitudes towards ambiguity. As for an axiomatization of (3.8), the implied ordering $\succeq_{t,\omega}$ satisfies CP and DC and it weakens RP in ways that are well understood from studies of risk preference. In addition, it violates MP, but satisfies the Gilboa-Schmeidler axioms on the subdomain of acts that are $\mathcal{F}_{t+1}$-measurable. It seems clear from [23] and [29] that axiomatization of (3.8) would require a more complicated hierarchical domain for preference such

for the set $\mathcal{P}$ arising from the backward recursion (3.5). The recursive relation for utility, and hence the 1-step-ahead correspondences, often suffice, as shown in the cited papers.

as those adopted in Klibanoff [21] and Wang [33]. Indeed, these studies deliver related representations.

4. EXAMPLES

Our objective in this section is to cast further light on the scope of the theorem and on rectangularity. Section 4.1 shows, in the context of a dynamic version of the classic Ellsberg urn example, that dynamic consistency is problematic in some settings. However, there are many other settings, including those that are typical in dynamic modeling in macroeconomics and finance, where backward induction and hence dynamic consistency are natural. The remaining examples illustrate such settings. Sections 4.2 and 4.3 show how a rich set of models of dynamic behavior can easily be constructed by specifying the process of 1-step-ahead conditionals. These, in turn, lead naturally to a rectangular set of priors through the logic of backward induction. For concreteness, these examples specify relatively simple types of history dependence for these conditionals.[12] Much more general history dependence can be accommodated as explained further in Section 4.4. The examples, particularly the last one, also illustrate why the time zero set of priors $\mathcal{P}$ is *not* in general equal to the natural set of 'possible probability laws' or 'possible models of the environment' that the decision-maker may have in mind.

4.1. Ellsberg

Consider the 3-color Ellsberg urn experiment in which there are 30 balls that are red and 60 that are either blue or green. A ball is drawn at random from the urn at time 0. The goal is to model the decision-maker's preferences over acts that pay off according to the color of the ball that is drawn. A natural state space is $\Omega = \{R, B, G\}$. To introduce dynamics in a simple way, suppose that the color is revealed to the decision-maker at $t = 2$, leaving essentially a 3-period model. At the intermediate stage time 1, the decision-maker is told whether or not the color drawn is G. Thus the filtration is $\{\mathcal{F}_t\}$, where $\mathcal{F}_t$ is the power set for all $t \geq 2$ and

$$\mathcal{F}_1 = \{\{R, B\}, \{G\}\}.$$

[12] In the language of time series analysis, the models of Sections 4.2 and 4.3 permit the interpretation that the 'set of possible models' does not contain models with 'hidden state variables'.

To see that dynamic consistency may be problematic in this setting, consider the ranking of $(1,0,1)$ versus $(0,1,1)$, where the former denotes the act that pays 1 unit of consumption (or utils) at time 2 in the states R and G and where the latter is interpreted similarly. (There is no consumption in other periods.) The time 0 ranking

$$(1,0,1) \prec_0 (0,1,1)\,. \tag{4.1}$$

is intuitive in an atemporal setting and arguably also in the present dynamic setting. This ranking is supported by the set of priors $\mathcal{P}'$, where

$$\mathcal{P}' = \left\{ p = \left(\tfrac{1}{3}, p_B, \tfrac{2}{3} - p_B\right) : \tfrac{1}{6} \leq p_B \leq \tfrac{1}{2} \right\}. \tag{4.2}$$

Ambiguity about the number of blue versus green balls is reflected in the range of probabilities for p_B. Assuming for the moment that $\mathcal{P}'$ is indeed the initial set of priors, then the conditional rankings at time 1 depend on how $\mathcal{P}'$ is updated. Under prior by prior Bayesian updating, one concludes that

$$(1,0,1) \succ_{1,\{R,B\}} (0,1,1) \text{ and } (1,0,1) \sim_{1,\{G\}} (0,1,1)\,, \tag{4.3}$$

in contradiction to dynamic consistency.[13]

To clarify the connection to our theorem, note that $\mathcal{P}'$ is not $\{\mathcal{F}_t\}$-rectangular (see Figure 1 and recall its discussion in Section 3.1). Thus it is not surprising that $\mathcal{P}'$ leads to a violation of dynamic consistency. Our modeling approach would suggest replacing $\mathcal{P}'$ by the smallest $\{\mathcal{F}_t\}$-rectangular set containing $\mathcal{P}'$, which is readily seen to be given by the set $\mathcal{P}$,[14]

$$\mathcal{P} = \left\{ \left(\tfrac{1}{3} \frac{\tfrac{1}{3} + p'_B}{\tfrac{1}{3} + p_B}, \, p_B \frac{\tfrac{1}{3} + p'_B}{\tfrac{1}{3} + p_B}, \tfrac{2}{3} - p'_B \right) : \tfrac{1}{6} \leq p_B, p'_B \leq \tfrac{1}{2} \right\}.$$

Because $\mathcal{P}$ is $\{\mathcal{F}_t\}$-rectangular, it would ensure dynamic consistency. However, this would be at the cost of reversing the ranking (4.1). Therefore, the lesson we take from this is not that it is impossible to deliver dynamic consistency within the multiple-priors framework, but rather that in *some settings*, ambiguity may render dynamic consistency problematic.

The essence of these problematic settings seems clear. Begin with any specification of 1-step-ahead beliefs. These determine 1-step-ahead preferences, by

[13] $\succ_{1,\{R,B\}}$ denotes the common preference order $\succ_{1,R} = \succ_{1,B}$.

[14] As described in 3.1, $\mathcal{P}$ and $\mathcal{P}'$, induce the identical 1-step-ahead conditionals and thus generate the same rankings at any t of acts that are $\mathcal{F}_{t+1}$-measurable. In particular, they both lead to (4.3).

which we mean the collection of conditional preference orders at any (t, ω) over acts that are $\mathcal{F}_{t+1}$-measurable. Backward induction leads to a utility process over all acts satisfying dynamic consistency. In this construction of utility, 1-step-ahead beliefs or preferences are unrestricted. A difficulty arises only if there are intuitive conditional choices that are not expressible in terms of 1-step-ahead preferences; thus they involve acts that are not measurable with respect to the next period's information. The choice (4.1) is an example because the acts given there are not $\mathcal{F}_1$-measurable.

We turn now to examples, based on specifications that are common in applied dynamic modeling, where the appeal of dynamic consistency seems to us to be unqualified.

4.2. Ambiguous Random Walk

In many dynamic settings, the description of the environment is most naturally expressed in terms of 1-step-ahead correspondences and thus a rectangular specification of the set of priors. As a stylistic benchmark example, suppose that uncertainty is driven by an integer-valued state process W_t which begins at the value zero ($W_0 = 0$). All processes of interest are adapted to $\{\mathcal{F}_t\}$, where $\mathcal{F}_t = \sigma(W_s : s \leq t)$ defines a filtration on the state space $\Omega = \mathbf{N}^{T+1}$. The decision-maker's subjective view of the law of motion of W_t is that, given t and conditional on the realized value of W_t, then $W_{t+1} - W_t = \pm 1$. However, she is not completely confident about the transition probabilities. Thus she thinks in terms of a set of transition probability measures, or equivalently, in terms of an interval $[\frac{1-\kappa}{2}, \frac{1+\kappa}{2}]$ for the probability that the increment equal $+1$, where $0 \leq \kappa \leq 1$ parametrizes the extent of ambiguity.

As an initial specification, suppose that the same interval describes conditional beliefs at every realized W_t, reflecting the view that the increments $W_{t+1} - W_t$ are unaffected by current (or past) values of the state process (a type of IID assumption for increments) and also by the calendar time t (a form of stationarity). The conditional 1-step ahead correspondences P_t^{+1} are defined thereby and they in turn determine a rectangular set of priors $\mathcal{P}$, as described in (3.3)-(3.5). If there is no ambiguity ($\kappa = 0$), then $\mathcal{P}$ is a singleton and it describes a random walk. More generally, $\mathcal{P}$ describes an *ambiguous random walk*.

A rich range of generalizations of this model are possible, including the next example, in which P_t^{+1} depends on history reflecting learning. In all such cases, a rectangular set of priors emerges naturally and dynamic consistency is unproblematic.

4.3. Conditional Ambiguity

The ambiguous random walk features ambiguity about both the conditional mean and the conditional variance, parametrized by κ. The idea is easily generalized. As a further example of the rich dynamics that is compatible with rectangularity, consider the following 'autoregressive conditional ambiguity' model. The state process is now (y_t),

$$\begin{aligned} y_t &= a y_{t-1} + b\varepsilon_t + \sqrt{h_t} u_t \\ h_t &= \rho_0 + \rho_1 h_{t-1} + \left(y_{t-1} - \frac{\underline{a} + \overline{a}}{2} y_{t-2} \right)^2 \\ (a, \rho_1) &\in [\underline{a}, \overline{a}] \times [\underline{\rho}, \overline{\rho}] \subset (-1, 1) \times [0, 1) \\ \varepsilon_t &\in [-\overline{\varepsilon}_t, \overline{\varepsilon}_t],\ \overline{\varepsilon}_t = \phi \overline{\varepsilon}_{t-1} + \left(y_t - \frac{\underline{a} + \overline{a}}{2} y_{t-1} \right)^2 \end{aligned}$$

where u_t is white noise, and $\underline{a}$, $\overline{a}$, $\rho_0 > 0$, $\underline{\rho}$, $\overline{\rho}$, $\phi \in (0, 1)$ and a_0 are fixed parameters.

If $b = 0$, $\underline{a} = \overline{a}$ and $\underline{\rho} = \overline{\rho}$, the model reduces to a standard $AR(1)$ with zero mean and $GARCH(1, 1)$ errors. More generally, the decision-maker's beliefs reflect confidence that the next observation is generated by a density from this class, but there is ambiguity about the conditional mean and variance.[15] Since each admissible vector of these parameters determines a 1-step-ahead conditional measure, a set of such measures, and hence also a rectangular set of priors, are determined by the given specification which thus fits directly into our framework.[16]

This model captures *time-varying conditional ambiguity* that can depend both on the level of y_t and the 'surprises' that occur relative to the 'center' forecast $\frac{\overline{a}+\underline{a}}{2} y_{t-1}$. As one example of the former, if $b = 0$ and $\underline{a} = -\overline{a}$, then the interval for the conditional mean $[-\overline{a} y_{t-1}, \overline{a} y_{t-1}]$ is wider, the further away was the last observation from zero. Given recursive multiple-priors utility, such an observation would induce 'greater pessimism' for a decision-maker with a value function

[15] The somewhat unusual specification of the conditional variance equation reflects the fact that the 'forecast error' that feeds into the behavior of the conditional variance is measured with respect to $\frac{\overline{a}+\underline{a}}{2} y_{t-1}$, the center of the interval of means that was thought possible at $t - 1$.

[16] More accurately, it fits into our framework extended to accommodate filtrations that do not correspond to finite partitions.

increasing in y_t. As another example, if $1 > \overline{a} > \underline{a} > 0$, the decision maker is confident that there is mean reversion in y_t. Again assuming an increasing value function, we would now have *asymmetric* behavior, in that the decision-maker fears that bad times ($y_t < 0$) last longer ($a = \overline{a}$) than good times, in which mean reversion is expected to occur more quickly ($a = \underline{a}$).

If $\underline{\rho} < \overline{\rho}$, the interval for the conditional variance increases if there have been a lot of 'surprises' (relative to the forecast $\frac{\overline{a}+\underline{a}}{2} y_{t-1}$) in the recent past. Finally, if $b > 0$, the term ε_t provides a link between forecast errors and ambiguity *about the conditional mean.* Assuming for simplicity that $\overline{a} = \underline{a} = 0$, surprises widen the interval $[-b\overline{\varepsilon}_t, b\overline{\varepsilon}_t]$ for the mean. An increase in ambiguity caused by a large surprise is persistent as the ambiguity is resolved gradually.

4.4. An Entropy-Based Set of Priors

In the preceding two examples, the most natural description of the environment (or of the set of 'possible probability laws') is in terms of 1-step-ahead beliefs. The final example shows that dynamic consistency can be natural even where the primitive description does not have the 1-step-ahead form.

Suppose the set of probability models considered possible by the decision-maker is given by

$$\mathcal{P}^{rob} = \{ Q \in \mathcal{Q} : \, d\,(Q, P) \, \leq r \,\}, \tag{4.4}$$

where $\mathcal{Q} \subset \Delta\,(\Omega, \mathcal{F}_T)$ is a family of probability measures, $P \in \mathcal{Q}$ is a reference measure, d denotes relative entropy and r determines the size of the set. (Q and P are assumed mutually absolutely continuous and $d(Q, P) \equiv \Sigma_{\tau \geq \mathbf{0}} \beta^{\tau} \, E_Q \left[log \left(\frac{dQ_\tau}{dP_\tau} \right) \right]$, where Q_τ and P_τ denote the restrictions of Q and P to $\mathcal{F}_\tau$.) As described in the next section, such sets of priors have been adopted in the robust control approach, which explains the superscript attached to $\mathcal{P}$. The set $\mathcal{P}^{rob}$ is not rectangular and thus is not admissible in our model. Because this specification may seem natural, some readers may be concerned that our model limits unduly the dimensions of ambiguity that can be accommodated.

The key point concerns the interpretation of sets of priors. In particular, there is an important conceptual distinction between the set of probability laws that the decision-maker views as possible, such as $\mathcal{P}^{rob}$, and the set of priors $\mathcal{P}$ that is part of the representation of preference. Only the latter includes elements of reasoning or processing, backward induction for example, on the part of the decision-maker. Thus the description of the environment represented by $\mathcal{P}^{rob}$ is consistent with our model and the use of a rectangular set of priors in the following sense:[17] Determine the 1-step-ahead sets of conditionals P_t^{+1} implied by applying Bayes' Rule prior-by-prior to $\mathcal{P}^{rob}$. Then use the P_t^{+1}'s to construct, via the backward recursion (3.5), a new time 0 set of priors $\mathcal{P}$. This set is rectangular and, though larger than $\mathcal{P}^{rob}$, yields the identical 1-step-ahead conditionals. Indeed, because $\mathcal{P}$ is the smallest rectangular set containing $\mathcal{P}^{rob}$, it is the minimal enlargement of $\mathcal{P}^{rob}$ needed in order to accommodate the logic of backward induction. Thus $\mathcal{P}$ may be viewed as the natural vehicle for both capturing the set of possible models $\mathcal{P}^{rob}$ and simultaneously representing a dynamically consistent preference process.

The question remains whether the dynamic behavior implied by $\mathcal{P}$ is intuitive. In particular, is there intuitive choice behavior, paralleling (4.1) from the Ellsberg example, that is contradicted by $\mathcal{P}$? No such behavior is apparent to us, though admittedly, we cannot prove that all the behavioral implications of $\mathcal{P}$ are intuitive.

The 1-step-ahead conditionals constructed as above from (4.4) will involve relative entropy in their definitions. Because, entropy plays a large role in related statistical theory, as well as in the robust control modeling approach, we conclude by adding that there is a more direct way to build relative entropy into parametric specifications of sets of priors. For example, define the 1-step-ahead correspondence at any t and ω directly as a relative entropy neighborhood of the 1-step-ahead conditional of a reference measure, much as in (5.4) below, and then work with the corresponding rectangular set of time 0 priors.

5. COMPARISON WITH ROBUST CONTROL

In work with several coauthors, Hansen and Sargent have adapted and extended robust control theory to economic settings. Because there now exist a number of descriptive and normative applications of the robust control model,[18] we take this opportunity to compare their approach with ours. Hansen and Sargent [19] describe the utility specification that supports (or is implicit in) the robust control approach. We take this utility specification as the economic foundations for their approach and thus we use it as the basis for comparison. To permit a clearer com-

[17] The argument that follows applies equally to any time 0 set of measures and not just to that given by (4.4).

[18] For descriptive (e.g., asset pricing) applications, see [20], [1] and [18]. Normative applications are typically to optimal monetary policy in a setting where the monetary authority does not know precisely the true model describing the environment; see, for example, [18], [24] and [25].

parison of the two models, we translate the description in [19] into the framework of this paper, thereby modifying their model somewhat, but not in ways that are germane to the comparison.[19] Finally, the related model in Uppal and Wang [30] is also discussed.

The entropy-based model described in the preceding section can be viewed as a reformulation of the robust control approach that fits into our framework. Thus the reader may wish to refer back to Section 4.4 after reading the comparison that follows.

5.1. Utility Specification

We are interested in the collection $\{\succeq_{t,\omega}\}$ of conditional preferences, with representing utility functions $\{V_t(\cdot,\omega)\}$, implied by the robust control model. Fix a 'reference model' P, a measure in $\Delta(\Omega, \mathcal{F}_T)$, and a set of 'possible models' (or priors) $\mathcal{P}_0^{rob} \subset \Delta(\Omega, \mathcal{F}_T)$, containing P. Utility at time 0 is given by

$$V_0(h) = \min_{m \in \mathcal{P}_0^{rob}} \int \Sigma_{\tau \geq 0} \beta^\tau u(h_\tau)\, dm, \quad h \in \mathcal{H}, \tag{5.1}$$

for some β and u as in our theorem. Here the time 0 set of priors $\mathcal{P}_0^{rob}$ has the parametric form (4.4) for radius $r = r_0$.

To define subsequent utility functions $V_t(\cdot,\omega)$, specify an updating rule for the set of priors. This is done by first fixing an act h^*; for example, Hansen and Sargent take h^* to be optimal relative to $\succeq_0$ in a planning problem of interest, having time 0 feasible set Υ. Let Q^* be a minimizing measure in (5.1) when $h = h^*$, and let $r_t(\omega)$ denote the relative entropy between Q^* and the reference measure conditional on time t information, that is,[20]

$$r_t(\omega) = d\left(Q^*(\cdot \mid \mathcal{F}_t)(\omega),\, P(\cdot \mid \mathcal{F}_t)(\omega)\right). \tag{5.2}$$

Finally, define

$$V_t(h,\omega) = \min_{m \in \mathcal{P}_t^{rob}(\omega)} \int \Sigma_{\tau \geq t} \beta^{\tau - t} u(h_\tau)\, dm, \quad h \in \mathcal{H}, \tag{5.3}$$

[19] For example, we use discrete rather than continuous time, we exclude time nonseparabilities due to habit formation and we refer to the natural extension of their model from the domain of consumption processes to our domain $\mathcal{H}$.

[20] As in (4.4), d denotes relative entropy.

for the updated set of measures $\mathcal{P}_t^{rob}(\omega)$ given by

$$\mathcal{P}_t^{rob}(\omega) = \left\{ Q(\cdot \mid \mathcal{F}_t)(\omega) : Q \in \mathcal{P}_0^{rob}, \, d\left(Q(\cdot \mid \mathcal{F}_t)(\omega), P(\cdot \mid \mathcal{F}_t)(\omega)\right) \leq r_t(\omega) \right\}. \tag{5.4}$$

This completes our outline of the utility specification.[21]

At one level, the difference between the robust control and recursive multiple-priors models is a matter of alternative restrictions on initial sets of priors and on updating rules. Our model delivers rectangular sets of priors that are updated prior-by-prior, while robust control delivers sets of priors constrained by relative entropy and updated by (5.4). In what follows, we clarify the behavioral significance of these formal differences.

5.2. Discussion

For any given h^* or Υ, it is immediate that in common with recursive multiple-priors, axioms CP, MP and RP are satisfied. A difference between the models is that the robust control model violates DC. However, its construction delivers the following weaker form of dynamic consistency:

Axiom 6 (h^*-DC). *For every t and ω and for every act h, if $h_\tau(\omega) = h_\tau^*(\omega)$ for all $\tau \leq t$ and if $h \succeq_{t+1,\omega'} h^*$ for all ω', then $h \succeq_{t,\omega} h^*$; and the latter ranking is strict if the former ranking is strict at every ω' in a $\succeq_{t,\omega}$-nonnull event.*

The difference from DC is that here only comparisons with the given h^* are considered. Under h^*-DC, if h^* is optimal at time 0 in the feasible set Υ, then it will be carried out in (almost) all future contingencies. Under DC, the ranking of *any* two acts is time consistent.

Which set of assumptions on preferences is appropriate will typically depend on the application. In many descriptive modeling contexts, we want to describe an agent who solves a single intertemporal optimization problem. A typical example is consumption-savings decisions for given prices. Both axioms DC and h^*-DC permit the interpretation that a plan that the agent would choose ex ante under commitment would in fact be carried out ex post under discretion. One

[21] When $h = h^*$ in (5.1) and (5.3), the minimizations over measures can be characterized via Lagrangeans and deliver the multipliers θ_0 and $\theta_t(\omega)$. Under the specification described, $\theta_t(\omega) = \theta_0$ for all t and ω, a fact that plays an important role in the discussion and empirical implementation of the robust control model.

might argue that the stronger axiom DC is not needed if one is interested only in rationalizing h^* as an optimum in Υ.

However, rationalization of a single optimum cannot be the entire point. If it were, then there would be no need to deviate from the Bayesian model since, if the specification (5.1)-(5.3) rationalizes h^*, then so does the Bayesian model where the decision-maker uses the single prior Q^*. In fact, given alternative models that rationalize a given set of data, or here h^*, it is standard practice to evaluate them based also on how they accord with behavior in other settings (Ellsberg-type behavior, for example) or even with introspection, say about concern with model misspecification. These auxiliary criteria support the non-Bayesian alternative, whether robust control or recursive multiple-priors.

To distinguish between these two models, consider comparative statics predictions which provide another litmus test that extends beyond the framework of the particular planning problem of interest. To illustrate, consider consumption-savings models. Let the planning problem of prime interest be associated with the feasible set Υ corresponding to the time 0 budget constraint

$$E\left[\Sigma_0^T \beta^t \pi_t c_t\right] \leq y_0,$$

where the expectation is with respect to a reference measure P, (π_t) is the state-price density process, $\pi_0 = 1$ and where y_0 denotes initial wealth. Preferences are as in the robust control model with the added assumption that the utility index u is a power function, $u(c) = c^\alpha/\alpha$, $0 \neq \alpha < 1$. Consider now a change to feasible set Υ', where $\pi'_t = \pi_t$ for all $t > 1$, but where y'_0 and π'_1 may differ from their counterparts in Υ. To simplify, suppose that the filtration is such that $\mathcal{F}_1$ corresponds to the binary partition $\{F_{1a}, F_{1b}\}$, where each component has positive probability under P, and that state prices differ only in period 1 and then only in event F_{1b}:

$$\pi'_1 = (\pi'_{1a}, \pi'_{1b}), \ \pi_1 = (\pi_{1a}, \pi_{1b}), \ \pi'_{1a} = \pi_{1a} \text{ and } \pi'_{1b} \neq \pi_{1b}.$$

Let c and c' be the corresponding optimal plans. Suppose finally that y'_0 has been chosen so that[22]

$$y'_1 \equiv y'_0 - c'_0 = y_0 - c_0 \equiv y_1.$$

Then, as shown at the end of this subsection, the two optimal plans satisfy

$$P\left\{\omega \in F_{1a} : (c'_1(\omega), ..., c'_T(\omega)) \neq (c_1(\omega), ..., c_T(\omega))\right\} > 0, \tag{5.5}$$

[22] By the homotheticity of preference, c'_0 is a linearly homogeneous function of y'_0 and thus one can rescale y'_0 to ensure the equality.

that is, continuations from time 1 and event F_{1a} differ. Because in each case the time 0 optimal plan is carried out under discretion, the decision-maker when reoptimizing at time 1 and event F_{1a} will make different choices across the two situations. This is so in spite of the fact that the two time 1 optimization problems share common initial wealth levels ($y_1' = y_1$) and common state price processes for the relevant horizon, that is, they have identical feasible sets.

The formal reason for the differing behavior across the two continuation problems is that, according to the robust control model, the agent has different utility functions in these two situations; more particularly, the updated set of priors (5.4) differs at time 1 and event F_{1a} across the two situations. It remains to understand 'why' this is the case and ultimately 'why' choices differ even though it is the 'same decision-maker' in either case.[23] Admittedly, past consumption levels c_0' and c_0 differ, but time nonseparabilities (e.g., habit formation) are ruled out in the robust control model that we are employing. The other way in which the two histories at $t = 1$ differ is in the preceding time 0 plans contingent on the unrealized event F_{1b}, or alternatively, in the state prices that would have applied had the event F_{1b} been realized. The bottom line is that, *according to the robust control model, behavior at any time-event pair depends on what might have happened in unrealized parts of the tree.*

This feature of the robust control model is apparent directly from the specification (5.2) and (5.4) for updating at (t, ω). Because Q^* depends on the entire process h^* and not just on values of h^* realized along the path leading to (t, ω), conditional preference $\succeq_{t,\omega}$ depends also on what might have happened ex ante.

Finally, we sketch a proof of (5.5): Suppose the contrary and let Q^* and Q^{**} be minimizing measures given c and c' respectively. The key point is that even though c' and c agree in their continuations beyond F_{1a}, the corresponding supporting measures, conditioned on F_{1a}, differ. (This is because the Lagrange multipliers for the two minimizations are distinct.) However, the noted eventwise conditionals $Q^*(\cdot \mid F_{1a})$ and $Q^{**}(\cdot \mid F_{1a})$ are minimizing for the continuations of c' and c. Since the latter are identical by hypothesis, it follows that

$$Q^*(\cdot \mid F_{1a}) = Q^{**}(\cdot \mid F_{1a}),$$

which is a contradiction.

[23] It is not the case that the two situations feature different information and therefore 'naturally' different conditional beliefs. The same event tree applies in both cases.

5.3. A Related Model

Finally, consider the Uppal and Wang model. Suitably translated into our framework, their model of dynamic preferences $\{\succeq_{t,\omega}\}$ satisfies CP, RP and DC. It differs from recursive multiple-priors in its violation of MP; indeed, $\succeq_{t,\omega}$ violates MP even restricted to acts that are $\mathcal{F}_{t+1}$-measurable. A rough unifying perspective on both models is provided by Epstein and Zin [14] that provides a general approach for integrating alternative theories of atemporal decision-making into a dynamic setting. That approach was based on recursivity and the use of a 'certainty equivalent' functional as the vehicle for incorporating the relevant theory of one-shot choice. From this perspective, the models differ only in their underlying theories of one-shot choice - recursive multiple-priors adopts the Gilboa-Schmeidler model, while Uppal and Wang adopt an alternative specification that has not yet been axiomatized.

Though there are superficial similarities with the robust control approach, at the substantive behavioral level, Uppal-Wang differs from the robust control model *both* by violating MP *and* by satisfying DC.

6. CONCLUDING REMARKS

We have specified an axiomatic model of dynamic preference that extends the Gilboa-Schmeidler atemporal model. The model delivers dynamic consistency and permits a rich variety of dynamics and model uncertainty. Further, it permits the next logical step in modeling behavior under ambiguity, namely learning. Because rectangular sets of priors are equivalently specified through a process $\{P_t^{+1}\}$ of conditional 1-step-ahead correspondences, a theory of learning amounts to a specification of this process, that is, to a description of how histories are mapped into views about the next step and ultimately about the entire future trajectory. Our general model imposes no restrictions on how the decision-maker responds to data. However, intuitively plausible forms of response can be described, leading to a model of learning that is as well-founded as the Bayesian one [12].

A related normative application of recursive multiple-priors is to econometric estimation and forecasting. Chamberlain [6] describes a minimax approach and cites Gilboa and Schmeidler for foundations. When the statistical decision problem is sequential, however, one must rely on a dynamic model such as recursive multiple-priors for axiomatic foundations.

A. APPENDIX: Proof of Theorem

Only the direction (a) $\Longrightarrow$ (b) is nontrivial.

Lemma A.1. *There exist $0 < \beta$, $u : \Delta_s(C) \longrightarrow R^1$ mixture linear and nonconstant, and $P_t : \Omega \rightsquigarrow \Delta(\Omega, \mathcal{F}_T)$ that is convex-valued, closed-valued and $\mathcal{F}_t$-measurable, such that for each t and ω,*

$$p\left(\mathcal{F}_t(\omega)\right) = 1 \quad \text{for all } p \in P_t\left(\omega\right) \tag{A.1}$$

and $\succeq_{t,\omega}$ is represented by

$$V_t(h,\omega) = min_{m \in P_t(\omega)} \int \left(\Sigma_{\tau \geq t} \beta^{\tau - t} u(h_\tau)\right) dm. \tag{A.2}$$

Moreover, each $P_t(\omega)$ is unique and u is unique up to a positive linear transformation.

Proof. The Gilboa-Schmeidler theorem does not apply directly because our domain $\mathcal{H}$ is not formulated as the set of all measurable maps from Ω into the set of lotteries over some outcome set, which is the structure they assume for their domain. However, we can reformulate $\mathcal{H}$ in such a way as to make their theorem applicable.

Define $\mathcal{T} = \{0, 1, ..., T\}$. Each h in $\mathcal{H}$ can be viewed as the mapping from $\mathcal{T} \times \Omega$ into $\Delta_s\left(C\right)$ that takes (t, ω') into $h_t(\omega')$. Further the adapted nature of h corresponds to measurability of the above map with respect to Σ, where Σ is the σ-algebra on $\mathcal{T} \times \Omega$ generated by all sets of the form $\{t\} \times E$, where E lies in $\mathcal{F}_t$. Thus $\mathcal{H}$ consists of all Σ-measurable maps from the expanded state space $\mathcal{T} \times \Omega$ into $\Delta_s\left(C\right)$.

Moreover, by MP each $\succeq_{t,\omega}$ satisfies the Gilboa-Schmeidler axioms on this domain. Focus first on the ordering $\succeq_0$ at time 0. Then, by [16, Theorem 1], there exists $v : \Delta_s\left(C\right) \longrightarrow \mathbb{R}^1$, mixture linear and nonconstant, and a convex and closed set $\mathcal{Q} \subset \Delta\left(\mathcal{T} \times \Omega, \Sigma\right)$, such that $\succeq_0$ is represented by

$$V_0\left(h\right) = min_{q \in \mathcal{Q}} \int v\left(h(\tau, \omega')\right) dq(\tau, \omega'). \tag{A.3}$$

We argue now that $\mathcal{Q}$ has more structure than stated above. The point is that our axiom MP is stronger than what is required to deliver the preceding representation. The issue is the relevant notion of a 'constant act'. In the abstract framework with (expanded) state space $\mathcal{T} \times \Omega$, there is nothing that distinguishes between the two components of the state. Thus constant acts are maps h that are constant on $\mathcal{T} \times \Omega$. Consequently, direct translation of the Gilboa-Schmeidler analysis assumes MP(ii) only for lottery acts ℓ for which $\ell_\tau = \ell_0$ for all τ. Similarly, their analysis adopts the weakening of (iv) whereby: if for every (τ, ω'), the act that delivers $h'_\tau(\omega')$ in every time and state is weakly preferred to the corresponding act constructed from h, then $h' \succeq_0 h$. To clarify, our version of Monotonicity states, in contrast, that if for every ω', the lottery act $(h'_0(\omega'), ..., h'_T(\omega'))$ is weakly preferred to the corresponding act derived from h, then $h' \succeq_0 h$.

Our strengthening of these Gilboa-Schmeidler axioms is intuitive once one recognizes that there is a clear conceptual distinction between the two components of the expanded state (τ, ω'). For example, Gilboa and Schmeidler suggest that in a general mixture $\alpha h + (1 - \alpha) g$, g may hedge the variation across states in h, thus reducing ambiguity and leading to violations of Independence. However, no such hedging occurs if g is a constant act, which justifies Certainty Independence. In our setting, it is plausible to assume that *hedging across time* is not of value, which justifies our stronger axiom MP(ii).

Turn now to the added implications of our stronger axiom MP. By (ii), $\succeq_0$ satisfies the independence axiom on the set of lottery acts where it is represented by

$$V_0(\ell) = min_{q \in \mathcal{Q}} \int v(\ell_\tau(\omega'))\, dq(\tau, \omega') = min_{m \in mrg_{\mathcal{T}} \mathcal{Q}} \int v(\ell_\tau)\, dm(\tau),$$

where $mrg_{\mathcal{T}} \mathcal{Q}$ denotes the set of all $\mathcal{T}$-marginals of measures in $\mathcal{Q}$. Therefore, $mrg_{\mathcal{T}} \mathcal{Q}$ must be a singleton, that is, all measures in $\mathcal{Q}$ induce the identical probability measure, denoted λ, on $\mathcal{T}$. Consequently, for any h,

$$V_0(h) = min_{q \in \mathcal{Q}} \Sigma_\tau \left[\lambda_\tau \int_\Omega v(h(\tau, \omega'))\, dq(\omega' \mid \tau) \right]. \tag{A.4}$$

Monotonicity in the form MP(iv) implies that

$$V_0(h') = V_0(h) \text{ whenever } \Sigma_\tau\, \lambda_\tau\, v(h'(\tau, \cdot)) = \Sigma_\tau\, \lambda_\tau\, v(h(\tau, \cdot)). \tag{A.5}$$

Deduce that

$$V_0(h) = min_{p \in P_0} \int_\Omega [\Sigma_\tau\, \lambda_\tau\, v(h(\tau, \omega'))]\, dp(\omega'), \text{ for all } h, \tag{A.6}$$

for some closed and convex $P_0 \subset \Delta(\Omega, \mathcal{F}_T)$.

(Argue as follows:[24] Define $X = v(\Delta_s(C))$ and consider the domain D of all Anscombe-Aumann acts on $(\Omega, \mathcal{F}_T)$ with elementary outcomes in X. For generic element ψ, denote by $E\psi(\omega)$ the mean of the lottery $\psi(\omega)$ on X; thus $\omega \longmapsto E\psi(\omega)$ is a Savage-style act with outcomes in X. Define $U : D \longrightarrow \mathbb{R}^1$ by

$$U(\psi) \;=\; V_0(h), \text{ for any } h \text{ satisfying } E\psi(\cdot) = v\left(\Sigma_\tau \lambda_\tau \; h(\tau, \cdot)\right).$$

Then U is well-defined by (A.5), and its induced preference satisfies the axioms in [16, Theorem 1]. Thus, U admits a multiple-priors representation, perhaps after a monotonic transformation φ. Because risk linearity was built into U, deduce that

$$\varphi(V_0(h)) \;=\; \min_{p \in P_0} \int v\left(\Sigma_\tau \lambda_\tau \; h(\tau, \omega')\right) dp(\omega'),$$

for some $P_0 \subset \Delta(\Omega, \mathcal{F}_T)$.

Consider next the ranking of acts in D_t, the set of all acts $h \in \mathcal{H}$ such that $h(\tau, \cdot) \equiv m^* \in \Delta_s(C)$, for all $\tau \neq t$, where $v(m^*) = 0$; the existence of such m^* is wlog. Then on D_t,

$$\varphi(V_0(h)) \;=\; \min_{p \in P_0} \int_\Omega \lambda_t \; v\left(h(t, \omega')\right) dp(\omega'),$$

while from (A.4),

$$V_0(h) \;=\; min_{q \in \mathcal{Q}} \left[\int_\Omega \lambda_t \;\; v\left(h(t, \omega')\right) dq(\omega' \mid t)\right].$$

From the uniqueness of the set of priors in the multiple-priors representation, conclude that $\mathcal{Q}_t \equiv \{q(\cdot \mid t) : q \in \mathcal{Q}\}$ and P_0 coincide when viewed as measures on $\mathcal{F}_t$. This is true for any t. Finally, therefore, (A.4) implies (A.6).)

Argue similarly for each conditional ordering $\succeq_{t,\omega}$ to conclude that it is represented by

$$V_t(h, \omega) \;=\; min_{P_t(\omega)} \int U_t(h_0, ..., h_T; \omega)\, dp$$

where $U_t(\cdot, \omega) : (\Delta_s(C))^T \longrightarrow \mathbb{R}^1$ is mixture linear and has the form

$$U_t(\ell; \omega) \;=\; \Sigma_{\tau \geq t} \, \lambda_\tau(t, \omega) \; v_t\left(\ell_\tau; \omega\right) \;\equiv\; \Sigma_{\tau \geq t} \, v_{t,\tau}\left(\ell_\tau; \omega\right).$$

[24] We suspect that a shorter argument is possible, but we have not been able to find one.

Condition (A.1) follows from CP. By RP, $U_t(\cdot;\omega)$ and $U_t(\cdot;\omega')$ are ordinally equivalent for every ω and ω', with t fixed. Since both are mixture linear, they must be equal (after suitable affine transformations). Thus we can write

$$U_t(\ell) = \Sigma_{\tau \geq t}\, v_{t,\tau}(\ell_\tau). \tag{A.7}$$

RP implies further that the ordering on $\Delta_s(C) \times \Delta_s(C)$ that is represented by

$$(p, p') \longmapsto v_{t,\tau}(p) + v_{t,\tau+1}(p')$$

is the same for all t and τ such that $t \leq \tau \leq T-1$. In particular, for fixed t, the above ranking does not depend on τ in the indicated range. This is a form of stationarity and it implies, by familiar arguments (Koopmans [22] or Rader [26, pp. 162-3], for example) and after suitable cardinal transformations, that

$$v_{t,\tau} = (b_t)^{\tau-t}\, v_{t,t}$$

for some $b_t > 0$. Because the noted ordering is invariant also with respect to t, conclude that b_t is independent of t and hence that $v_{t,\tau} = \beta^{\tau-t}\, v_{t,t}$. Once again, the invariance yields (after suitable cardinal transformations) that $v_{t,t} = v_{0,0} \equiv u$ for all t. This establishes

$$U_t(\ell) = \Sigma_{\tau \geq t}\, \beta^{\tau-t} u(\ell_\tau) \tag{A.8}$$

and hence also (A.2). ■

From the Lemma,

$$V_t(h,\omega) = u(h_t) + \beta\, min_{m \in P_t(\omega)} \int \left(\Sigma_{\tau \geq t+1} \beta^{\tau-t-1}\, u(h_\tau)\right) dm.$$

$$\equiv u(h_t) + \beta\, W_t(h,\omega).$$

For each t, ω and lottery ℓ_t, define

$$D_{t,\omega,\ell_t} = \{(V_{t+1}(h,\omega'))_{\omega' \in \mathcal{F}_t(\omega)} : h \in \mathcal{H},\ h_t = \ell_t\}.$$

Then we can view D_{t,ω,ℓ_t} as a subset of $RV(\mathcal{F}_t(\omega),\ \mathcal{F}_{t+1})$, the set of $\mathcal{F}_{t+1}$-measurable (real-valued) random variables defined on $\mathcal{F}_t(\omega)$. Below, by $V_{t+1}(h,\cdot)$ we mean such a random variable, that is, the restriction of the second argument to $\mathcal{F}_t(\omega)$ is understood even where not stated explicitly.

Define $\Phi : D_{t,\omega,\ell_t} \longrightarrow R^1$ by

$$\Phi(V_{t+1}(h,\cdot)) = W_t(h,\omega). \tag{A.9}$$

DC implies that Φ is well-defined and increasing on D_{t,ω,ℓ_t} in the sense that

$$V_{t+1}(h',\cdot) \geq V_{t+1}(h,\cdot) \text{ on } \mathcal{F}_t(\omega) \Longrightarrow \Phi(V_{t+1}(h',\cdot)) \geq \Phi(V_{t+1}(h,\cdot)). \tag{A.10}$$

Lemma A.2. *There exists $Q \subset \Delta(\mathcal{F}_t(\omega), \mathcal{F}_{t+1})$, convex and closed, such that*

$$\Phi(x) = \min_{q \in Q} \int x\, dq, \quad \textit{for all } x \in D_{t,\omega,\ell_t}. \tag{A.11}$$

Proof. Adapt the arguments in [16, pp. 146-7].

(i) Φ is homogenous on D_{t,ω,ℓ_t}: Let $V_{t+1}(h',\cdot) = \alpha V_{t+1}(h,\cdot)$ on $\mathcal{F}_t(\omega)$ for $0 < \alpha \leq 1$. We need to show that

$$W_t(h',\omega) = \alpha\, W_t(h,\omega).$$

Let $h''_\tau(\cdot) = \alpha h_\tau(\cdot) + (1-\alpha)\ell^*$ for $\tau > t$ and $= \ell_t$ and defined arbitrarily for $\tau < t$. Then h'' lies in D_{t,ω,ℓ_t}, $u(h''_\tau(\cdot)) = \alpha\, u(h_\tau(\cdot)) + (1-\alpha)\, u(\ell^*) = \alpha u(h_\tau(\cdot))$ for $\tau > t$, $W_t(h'',\omega) = \alpha W_t(h,\omega)$ and

$$V_{t+1}(h'',\cdot) = \alpha V_{t+1}(h,\cdot) = V_{t+1}(h',\cdot) \text{ on } \mathcal{F}_t(\omega).$$

By DC, conclude that $W_t(h',\omega) = W_t(h'',\omega) = \alpha W_t(h,\omega)$.

(ii) Extend Φ by homogeneity to $RV_{simple}(\mathcal{F}_t(\omega), \mathcal{F}_{t+1})$: Because

$$\{u(h(\cdot)) : h \in \mathcal{F}_{t+1}\} \subset D_{t,\omega,\ell_t},$$

deduce that $RV_{simple}(\mathcal{F}_t(\omega), \mathcal{F}_{t+1}) = \cup_{\lambda \in R^1} (\lambda D_{t,\omega,\ell_t})$. (We can assume wlog that $\exists \ell_1$ and ℓ_2 with $u(\ell_1) < -1$ and $u(\ell_2) > 1$.) Thus a unique extension exists. Then Φ satisfies homogeneity there and the following form of monotonicity:

$$x'(\cdot) \geq x(\cdot) \Longrightarrow \Phi(x') \geq \Phi(x). \tag{A.12}$$

(iii) Φ satisfies Certainty Additivity: On D_{t,ω,ℓ_t}, argue as follows. For all lotteries $\ell \in (\Delta_s(C))^{T+1}$, $\Phi(\alpha V_{t+1}(h,\cdot) + (1-\alpha)V_{t+1}(\ell,\cdot)) = \Phi(V_{t+1}(\alpha h + (1-\alpha)\ell,\cdot)) =$

$W_t(\alpha h + (1-\alpha)\ell, \omega) = \alpha W_t(h,\omega) + (1-\alpha) W_t(\ell,\omega) = \alpha\Phi(V_{t+1}(h,\cdot)) + (1-\alpha)\Phi(V_{t+1}(\ell,\cdot))$, that is,

$$\Phi(\alpha V_{t+1}(h,\cdot) + (1-\alpha)V_{t+1}(\ell,\cdot)) = \alpha\Phi(V_{t+1}(h,\cdot)) + (1-\alpha)\Phi(V_{t+1}(\ell,\cdot)).$$

On $RV_{simple}(\mathcal{F}_t(\omega), \mathcal{F}_{t+1})$, argue as in [16, pp. 146-7].

(iv) Φ is superadditive: Prove first that

$$\begin{aligned}\Phi\left(\tfrac{1}{2}V_{t+1}(h',\cdot) + \tfrac{1}{2}V_{t+1}(h,\cdot)\right) &\geq \tfrac{1}{2}\Phi\left(V_{t+1}(h',\cdot)\right) + \tfrac{1}{2}\Phi\left(V_{t+1}(h,\cdot)\right) \\ &= \tfrac{1}{2}W_t(h',\omega) + \tfrac{1}{2}W_t(h,\omega).\end{aligned}$$

Suppose that $W_t(h',\omega) = W_t(h,\omega)$. Then because of the definition of $W_t(\cdot,\omega)$,

$$W_t\left(\tfrac{1}{2}h' + \tfrac{1}{2}h, \omega\right) \geq W_t(h,\omega) = \tfrac{1}{2}W_t(h',\omega) + \tfrac{1}{2}W_t(h,\omega).$$

Proceed as in [16, p. 147]. ■

The remainder of the proof is subdivided into a sequence of steps.
Step 1: Show that $Q = P_t^{+1}(\omega)$, the set of restrictions to $\mathcal{F}_{t+1}$ of measures in $P_t(\omega)$, that is, the set of 1-period ahead marginals. Apply (A.9) and the preceding lemmas to conclude that for any $\mathcal{F}_{t+1}$-measurable h in D_{t,ω,ℓ_t},

$$\begin{aligned}\min_{p_t \in P_t^{+1}(\omega)} \int \left(\Sigma_{\tau \geq t+1}\,\beta^{\tau-t-1}u(h_\tau)\right) dp_t &= \min_{p \in P_t(\omega)} \int \left(\Sigma_{\tau\geq t+1}\,\beta^{\tau-t-1}u(h_\tau)\right) dp \\ &\equiv W_t(h,\omega) = \min_{q\in Q} \int V_{t+1}(h,\cdot)\, dq \\ &= \min_{q \in Q} \int \left(\Sigma_{\tau\geq t+1}\,\beta^{\tau-t-1}u(h_\tau(\cdot))\right) dq.\end{aligned}$$

Thus uniqueness of the representing set of priors [16, Theorem 1] delivers the desired result since both $P_t^{+1}(\omega)$ and Q are convex and closed.

Step 2: The measures in $P_t^{+1}(\omega)$ are mutually absolutely continuous. The strict ranking component in DC implies that, for any h' and h in D_{t,ω,ℓ_t}, if $V_{t+1}(h',\cdot) \geq V_{t+1}(h,\cdot)$ and if $E = \{\omega' \in \mathcal{F}_t(\omega) : V_1(h',\omega') > V_1(h,\omega')\}$ is $V_t(\cdot,\omega)$-nonnull, then $V_t(h',\omega) > V_t(h,\omega)$, or equivalently, $W_t(h',\omega) > W_t(h,\omega)$. Because $V_t(\cdot,\omega)$ satisfies MP, the noted nonnullity is equivalent to

$$p(E) > 0 \text{ for some } p \text{ in } P_t(\omega).$$

Because E is in $\mathcal{F}_{t+1}$, there is a further equivalence with

$$q(E) > 0 \text{ for some } q \text{ in } P_t^{+1}(\omega).$$

From Step 1 and (A.9), conclude that

$$\min_{q \in P_t^{+1}(\omega)} \int V_{t+1}(h', \cdot)\, dq > \min_{q \in P_t^{+1}(\omega)} \int V_{t+1}(h, \cdot)\, dq$$

if $V_{t+1}(h', \cdot) \geq V_{t+1}(h, \cdot)$ with strict inequality on an event having positive q-probability for some q in $P_t^{+1}(\omega)$. In particular, for any $\mathcal{F}_{t+1}$-measurable h' and h in D_{t,ω,ℓ_t},

$$\min_{q \in P_t^{+1}(\omega)} \int \left(\Sigma_{\tau \geq t+1}\, \beta^{\tau-t-1} u(h'_\tau(\cdot))\right) dq > \min_{q \in P_t^{+1}(\omega)} \int \left(\Sigma_{\tau \geq t+1}\, \beta^{\tau-t-1} u(h_\tau(\cdot))\right) dq$$

if $\left(\Sigma_{\tau \geq t+1}\, \beta^{\tau-t-1} u(h'_\tau(\cdot))\right) \geq \left(\Sigma_{\tau \geq t+1}\, \beta^{\tau-t-1} u(h_\tau(\cdot))\right)$ with strict inequality on an event having positive q-probability for some q in $P_t^{+1}(\omega)$. Apply the preceding to the acts having, for $\tau \geq t$,

$$h'_\tau = (\ell' \text{ if } E; \ell^* \text{ if } E^c) \text{ and } h_\tau = (\ell \text{ if } E; \ell^* \text{ if } E^c),$$

where the lotteries ℓ' and ℓ are such that $u(\ell') > u(\ell) > u(\ell^*)$. Conclude that $\max_{P_t^{+1}(\omega)} m(E) > 0 \Longrightarrow \min_{P_t^{+1}(\omega)} m(E) > 0$.

Step 3: If $p(\cdot) = \int p_{t+1}(\omega')(\cdot)\, dm(\omega')$ for some measurable $p_{t+1} : (\Omega, \mathcal{F}_{t+1}) \longrightarrow \Delta(\Omega, \mathcal{F}_{t+2})$ such that $p_{t+1}(\cdot) \in P_{t+1}(\cdot)$ and $m \in P_t^{+1}(\omega)$, then

$$p(\cdot) = m(\cdot) \text{ on } \mathcal{F}_{t+1}, \text{ and} \tag{A.13}$$

$$p_{t+1}(\omega')(\cdot) = p(\cdot \mid \mathcal{F}_{t+1})(\omega') \qquad a.e.[p]. \tag{A.14}$$

Because $\mathcal{F}_{t+1}$ corresponds to a finite partition, then by (A.1),

$$p_{t+1}(\omega')(E) = \begin{cases} 0 & \text{if } E \cap \mathcal{F}_{t+1}(\omega') = \emptyset \\ 1 & \text{if } E \supset \mathcal{F}_{t+1}(\omega'). \end{cases}$$

In particular, if $E \in \mathcal{F}_{t+1}$, then the above two cases are exhaustive and

$$p(E) = \int_\Omega p_{t+1}(\omega')(E)\, dm(\omega') = m\left(\cup\{\mathcal{F}_{t+1}(\omega') : \mathcal{F}_1(\omega') \subset E\}\right) = m(E),$$

proving (A.13). Further,

$$p(E) \;=\; m(\mathcal{F}_{t+1}(\omega'))\, p_{t+1}(\omega')(E) \;\;\text{ for any } E \subset \mathcal{F}_{t+1}(\omega'),$$

E not necessarily in $\mathcal{F}_{t+1}$. Take also $E = \mathcal{F}_{t+1}(\omega)$. Then $p_{t+1}(\omega')(\mathcal{F}_{t+1}(\omega')) \;=\; 1$ and hence

$$p(\mathcal{F}_{t+1}(\omega')) \;=\; m(\mathcal{F}_{t+1}(\omega')).$$

Thus if $p(\mathcal{F}_{t+1}(\omega')) \neq 0$, or equivalently if $m(\mathcal{F}_{t+1}(\omega')) \neq 0$, then

$$p\left(E \mid \mathcal{F}_{t+1}\right)(\omega') \;=\; \frac{p(E \cap \mathcal{F}_{t+1}(\omega'))}{p(\mathcal{F}_{t+1}(\omega'))} \;=\; p_{t+1}(\omega')(E).$$

This proves (A.14).

Step 4: From (A.9) and Step 1,

$$\min_{p \in P_t(\omega)} \int \left(\Sigma_{\tau \geq t+1}\, \beta^{\tau-t-1} u(h_\tau)\right) dp \;\equiv\; W_t\left(h, \omega\right) \;=$$

$$\min_{m \in P_t^{+1}(\omega)} \int V_{t+1}(h, \omega')\, dm(\omega') \;=$$

$$\min_{m \in P_t^{+1}(\omega)} \int \left[\min_{p_{t+1} \in P_{t+1}(\omega')} \int \left(\Sigma_{\tau \geq t+1}\, \beta^{\tau-t-1} u(h_\tau)\right) dp_{t+1} \right] dm(\omega')$$

$$= \min_{p \in \overline{P_t}(\omega)} \int \left(\Sigma_{\tau \geq t+1}\, \beta^{\tau-t-1} u(h_\tau)\right) dp, \qquad \text{where}$$

$$\overline{P_t}(\omega) \;\equiv \left\{ p(\cdot) = \int p_{t+1}(\omega')(\cdot)\; dm : m \in P_t^{+1}(\omega),\, p_{t+1}(\cdot) \in P_{t+1}(\cdot) \right\}. \qquad \text{(A.15)}$$

Thus $P_t(\omega)$ and $\overline{P_t}(\omega)$ represent the same preference order. They must coincide because each is convex and closed.

It is immediate that $\overline{P_t}(\omega)$ is closed. To see that it is convex, let $\int p_{t+1}(\omega')(\cdot)\; dm$ and $\int p'_{t+1}(\omega')(\cdot)\; dm'$ lie in $\overline{P_t}(\omega)$. Then the $\frac{1}{2}/\frac{1}{2}$ mixture equals $\int p''_{t+1}(\omega')(\cdot)\; dm''$, where

$$p''_{t+1}(\omega')(\cdot) \;=\; \frac{\frac{1}{2}\, m(\omega')\, p_{t+1}(\omega')(\cdot) \;+\; \frac{1}{2}\, m'(\omega')\, p'_{t+1}(\omega')(\cdot)}{\frac{1}{2}\, m(\omega') \;+\; \frac{1}{2}\, m'(\omega')}$$

if the denominator is positive and equal to any measure in $P_{t+1}(\omega')$ otherwise, (in such way that $p''_1(\omega')(\cdot)$ is the same for all ω''s in the same component of the $\mathcal{F}_{t+1}$ partition), and where

$$m''(\cdot) \;=\; \tfrac{1}{2}\, m(\cdot) \;+\; \tfrac{1}{2}\, m'(\cdot).$$

For each ω, $p''_{t+1}(\omega')(\cdot)$ lies in $P_{t+1}(\omega')$ because the latter is convex; convexity of $mrgP_t(\omega)$ implies that it contains $m''(\cdot)$. Thus $\int p''_{t+1}(\omega')(\cdot)\ dm''$ lies in $\overline{P_t}(\omega)$. Similarly for other mixtures.

Step 5: Axiom FS implies that every measure in P_0 has full support on $\mathcal{F}_T$. From Step 4,

$$P_t(\omega) \equiv \left\{ p(\cdot) = \int p_{t+1}(\omega')(\cdot)\ dm : m \in P_t^{+1}(\omega),\ p_{t+1}(\cdot) \in P_{t+1}(\cdot) \right\},$$

for every t and ω. Use the full support observation and Step 3, particularly the appropriate version of (A.14), to prove by induction that for every t and ω: (i) $P_t(\omega)$ equals the set of all Bayesian $\mathcal{F}_t$-updates of measures in P_0; and (ii) each measure in $P_t(\omega)$ has full support on $\mathcal{F}_t(\omega)$.

Finally, define $\mathcal{P} = P_0$. ■

B. APPENDIX: Infinite Horizon

For reasons given in Section 3.2, this appendix axiomatizes an infinite horizon version of recursive multiple-priors. Thus set $T = \infty$ and interpret the formalism surrounding our axioms and the definition of rectangularity in the obvious way. Assume that $\mathcal{F}_T = \mathcal{F}_\infty = \sigma\left(\cup_1^\infty \mathcal{F}_t\right)$. Though we continue to assume that each $\mathcal{F}_t$ corresponds to a finite partition of Ω, that is not the case for the limiting σ-algebra $\mathcal{F}_\infty$. Measures in $\Delta\left(\Omega, \mathcal{F}_\infty\right)$ are required to be finitely (but not necessarily countably) additive. On $\Delta\left(\Omega, \mathcal{F}_\infty\right)$, adopt the weak topology induced by the set of all bounded measurable real-valued functions. Say that a measure p in $\Delta\left(\Omega, \mathcal{F}_\infty\right)$ has *full local support* if

$$p\left(A\right) > 0 \quad \text{for every } \emptyset \neq A \in \cup_{t=0}^\infty \mathcal{F}_t.$$

We are given preferences $\{\succeq_{t,\omega}\} \equiv \{\succeq_{t,\omega}\colon (t,\omega) \in \mathcal{T} \times \Omega\}$ on the domain $\mathcal{H}$, defined as above. Continue to adopt axioms CP, MP, RP, DC and FS.

Though the range of any h_t is finite for any act h in $\mathcal{H}$, the range of h, viewed as a mapping from $\mathcal{T} \times \Omega$ into $\Delta_s\left(C\right)$, need not be finite given that $T = \infty$. To handle the complications caused by this infinity, assume the existence of best and worst lotteries in the following sense.[25]

[25] Relaxation of BW is possible by more efficient application of [16, Proposition 4.1] than below.

Axiom 7 (Best-Worst - BW). *For each* $(t,\omega) \in \mathcal{T} \times \Omega$, *there exist lotteries* p^* *and* p^{**} *in* $\Delta_s(C)$ *such that* $(p^*)_0^\infty \preceq_{t,\omega} (p)_0^\infty \preceq_{t,\omega} (p^{**})_0^\infty$ *for all* p *in* $\Delta_s(C)$.

In addition, impose a form of impatience whereby the distant future receives little weight in each conditional preference order.

Axiom 8 (Impatience -IMP). *For any* (t,ω) *in* $\mathcal{T} \times \Omega$, p *in* $\Delta_s(C)$ *and acts* h, h^* *and* h^{**} *in* $\mathcal{H}$, *if* $h^* \prec_{t,\omega} h \prec_{t,\omega} h^{**}$ *and* $h^n = (h_0, ..., h_n, p, p, ...)$, *then* $h^* \prec_{t,\omega} h^n \prec_{t,\omega} h^{**}$ *for all sufficiently large* n.

Before stating the theorem, we point out a change in the uniqueness property of the representing set of priors $\mathcal{P}$ due to the infinite horizon setting. It is apparent from (3.6) that utilities depend only on the probabilities assigned to events in $\cup_1^\infty \mathcal{F}_\tau$. Thus uniqueness on $\mathcal{F}_\infty = \sigma(\cup_1^\infty \mathcal{F}_\tau)$ is not to be expected because one could change arbitrarily probabilities assigned to events in $\mathcal{F}_\infty \setminus (\cup_1^\infty \mathcal{F}_\tau)$ without affecting utilities. Even in the case of a singleton prior, the latter is uniquely determined by its values on $\cup_1^\infty \mathcal{F}_\tau$ only if the prior is countably additive, but countable additivity is not implied by our axioms (nor by those in [16]). Thus the following theorem refers only to the set of priors $\mathcal{P}$ being *unique on* $\cup_1^\infty \mathcal{F}_\tau$, by which we mean that if any other set $\mathcal{P}'$ also satisfies the conditions in part (b), then the set of all restrictions to $\cup_1^\infty \mathcal{F}_\tau$ of measures in $\mathcal{P}$ coincides with the set constructed from $\mathcal{P}'$.

To clarify further, the uniqueness assertion in [16, Theorem 1], translated into our setting, does yield uniqueness of the appropriate set of priors on the expanded state space $\mathcal{T} \times \Omega$. However, this does not deliver uniqueness of the set of priors on Ω because, roughly speaking, the $\mathcal{T}$-marginal is given by the discount factor β^t with $\beta < 1$ and this washes out effects of probabilities assigned to events in $\mathcal{F}_\infty \setminus (\cup_1^\infty \mathcal{F}_\tau)$.

Theorem B.1. *Let* $T = \infty$ *and let* $\{\succeq_{t,\omega}\}$ *be a collection of binary relations on* $\mathcal{H}$. *The following statements are equivalent:*

(a) $\{\succeq_{t,\omega}\}$ *satisfy CP, MP, RP, DC, FS, BW and IMP on* $\mathcal{H}$.

(b) There exists $\mathcal{P} \subset \Delta(\Omega, \mathcal{F}_\infty)$, *closed, convex and* $\{\mathcal{F}_t\}$*-rectangular, with all measures in* $\mathcal{P}$ *having full local support,* $0 < \beta < 1$ *and a mixture linear and nonconstant* $u : \Delta_s(C) \longrightarrow \mathbb{R}^1$, *where* $\max_{\Delta_s(C)} u$ *and* $\min_{\Delta_s(C)} u$ *exist, such that: for every* t *and* ω, $\succeq_{t,\omega}$ *is represented on* $\mathcal{H}$ *by* $V_t(\cdot, \omega)$, *where*

$$V_t(h,\omega) = \min_{m \in \mathcal{P}_t(\omega)} \int \Sigma_{\tau \geq t} \beta^{\tau - t} u(h_\tau) \, dm.$$

Moreover, β is unique, $\mathcal{P}$ is unique on $\cup_1^\infty \mathcal{F}_\tau$ and u is unique up to a positive linear transformation.

Proof. Necessity of the axioms is routine. To verify IMP, note that

$$\Sigma_{\tau \geq t}\, \beta^{\tau - t}\, u(h^n_\tau) \underset{n \longrightarrow \infty}{\longrightarrow} \Sigma_{\tau \geq t}\, \beta^{\tau - t}\, u(h_\tau)$$

in the sup norm topology, while the Maximum Theorem implies that the mapping $X \longmapsto \min_{m \in \mathcal{P}} \int X\, dm$, from the space of bounded $\mathcal{F}_\infty$-measurable functions into the reals, is sup-norm continuous.

To prove sufficiency, adapt the proof of Theorem 3.2 above. Gilboa and Schmeidler's central representation result (Theorem 1) does not apply directly to $\mathcal{H}$ as in the finite horizon case. That is because it deals only with the domain of finite-ranged acts, which in our setting equals the proper subset of $\mathcal{H}$ consisting of acts h that have finite range when viewed as mappings from $\mathcal{T} \times \Omega$ to $\Delta_s(C)$. However, because of BW, their extension result Proposition 4.1 delivers a multiple-priors representation on $\mathcal{H}$ for $\succeq_{t,\omega}$.

Now proceed as in the proof of Theorem 3.2 to deliver the asserted representation in terms of $\mathcal{P}$, β and u. Use IMP to complete the counterpart of Lemma A.1, for example (A.8). BW implies that $V_0(\cdot)$ is bounded above and below. Because u is not constant, conclude that $\beta < 1$. Existence of the noted maximum and minimum for u follows from BW.

Finally, the asserted uniqueness of $\mathcal{P}$ follows from [16], just as in the finite horizon setting, by restricting time 0 preference to acts $h = (h_\tau)$ such that every h_τ is $\mathcal{F}_t$-measurable for some fixed t. ■

References

[1] E. Anderson, L. Hansen and T. Sargent, Robustness, detection and the price of risk, 2000.

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

[3] E. Ben-Porath, I. Gilboa and D. Schmeidler, On the measurement of inequality under uncertainty, *J. Econ. Theory* 75 (1997), 194-204.

[4] D. Blackwell and L. Dubins, Merging of opinions with increasing information, *Ann. of Statistics* 38 (1962), 882-886.

[5] R. Casadesus-Masanell, P. Klibanoff and E. Ozdenoren, Maxmin expected utility over Savage acts with a set of priors, *J. Econ. Theory* 92 (2000), 35-65.

[6] G. Chamberlain, Minimax estimation and forecasting in a stationary autoregression model, *Amer. Econ. Rev. Papers and Proceedings*, 91 (2001), 55-59.

[7] Z. Chen and L.G. Epstein, Ambiguity, risk and asset returns in continuous time, *Econometrica*, forthcoming.

[8] S.H. Chew and L.G. Epstein, Recursive utility under uncertainty, pp. 352-369 in A. Khan and N. Yannelis eds., *Equilibrium Theory with an Infinite Number of Commodities*, Springer, 1991.

[9] D. Ellsberg, Risk, ambiguity and the Savage axioms, *Quart. J. Econ.* 75 (1961), 643-669.

[10] L.G. Epstein, Stationary cardinal utility and optimal growth under uncertainty, *J. Econ. Theory* 31 (1983), 133-152.

[11] L.G. Epstein and J. Miao, A two-person dynamic equilibrium under ambiguity, *J. Econ. Dyn. and Control*, forthcoming.

[12] L.G. Epstein and M. Schneider, Learning under ambiguity, 2002.

[13] L.G. Epstein and T. Wang, Intertemporal asset pricing under Knightian uncertainty, *Econometrica* 62 (1994), 283-322.

[14] L.G. Epstein and S. Zin, Substitution, risk aversion and the temporal behavior of consumption and asset returns: a theoretical framework, *Econometrica* 57 (1989), 937-969.

[15] P. Ghirardato, F. Maccheroni, M. Marinacci and M. Siniscalchi, A subjective spin on roulette wheels, Cal Tech 31127, August 2001.

[16] I. Gilboa and D. Schmeidler, Maxmin expected utility with non-unique priors, *J. Math. Econ.* 18 (1989), 141-153.

[17] I. Gilboa and D. Schmeidler, Updating ambiguous beliefs, *J. Econ. Theory* 59 (1993), 33-49.

[18] L. Hansen and T. Sargent, Acknowledging misspecification in macroeconomic theory, Sargent's Frisch Lecture at the 2000 World Congress of the Econometric Society.

[19] L. Hansen and T. Sargent, Robust control and model uncertainty, *Amer. Econ. Rev. Papers and Proceedings* 91 (2001), 60-66.

[20] L. Hansen, T. Sargent and T. Talliarini, Robust permanent income and pricing, *Rev. Ec. Studies* 66 (1999), 873-907.

[21] P. Klibanoff, Dynamic choice with uncertainty aversion, mimeo, 1993 (revised 1995).

[22] T.C. Koopmans, Representation of preference orderings over time, pp. 79-100 in C.B. McGuire and R. Radner eds., *Decision and Organization: A Volume in Honor of Jacob Marschak*, North Holland, 1972.

[23] D. Kreps and E.L. Porteus, Temporal resolution of uncertainty and dynamic choice theory, *Econometrica* 46 (1978), 185-200.

[24] A. Onatski, Robust monetary policy under model uncertainty: incorporating rational expectations, Harvard, 2000.

[25] A. Onatski and J. Stock, Robust monetary policy under model uncertainty in a small open economy, *Macroeconomic Dynamics* 6 (2002), 85-110.

[26] T. Rader, *Theory of Microeconomics*, Academic Press, 1972.

[27] R. Sarin and P. Wakker, Dynamic choice and nonexpected utility, *J. Risk and Uncertainty* 17 (1998), 87-119.

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

[29] C. Skiadas, Recursive utility and preferences for information, *Econ. Theory* 12 (1998), 293-312.

[30] R. Uppal and T. Wang, Model misspecification and limited diversification, U.B.C., 2001.

[31] K. Wakai, Conditions for dynamic consistency and no speculation under multiple-priors, Yale, 2001.

[32] P. Walley, *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall, 1991.

[33] T. Wang, Conditional preferences and updating, *J. Econ. Theory,* forthcoming.