ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics

http://www.econometricsociety.org/

Econometrica, Vol. 83, No. 1 (January, 2015), 239–262

TIME AND NO LOTTERIES: AN AXIOMATIZATION OF MAXMIN EXPECTED UTILITY

ASEN KOCHOV
University of Rochester, Rochester, NY 14627, U.S.A.

Econometrica, Vol. 83, No. 1 (January, 2015), 239–262

TIME AND NO LOTTERIES: AN AXIOMATIZATION OF MAXMIN EXPECTED UTILITY

BY ASEN KOCHOV[1]

This paper axiomatizes an intertemporal version of the maxmin expected-utility model. It employs two axioms specific to a dynamic setting. The first requires that smoothing consumption across states of the world is more beneficial to the individual than smoothing consumption across time. Such behavior is viewed as the intertemporal manifestation of ambiguity aversion. The second axiom extends Koopmans' notion of stationarity from deterministic to stochastic environments.

KEYWORDS: Intertemporal choice, ambiguity, stationary preferences.

1. INTRODUCTION

THIS PAPER AXIOMATIZES THE FOLLOWING INTERTEMPORAL VERSION of the maxmin expected-utility model

$$V(h_0, h_1, \dots) = \min_{p\in\mathcal{P}} \mathbb{E}_p\left[\sum_t \beta^t (u\circ h_t)\right], \tag{1.1}$$

where $(h_0, h_1, \dots)$ is a stochastic consumption stream and $\mathcal{P}$ is a set of probability measures over the state space. The set $\mathcal{P}$ reflects ambiguity or the fact that the individual cannot quantify the relevant uncertainty in terms of a single prior belief. To obtain the utility representation in (1.1), the paper employs two axioms specific to a dynamic setting. The first requires that smoothing consumption across states of the world is more beneficial to the individual than smoothing consumption across time. Such behavior is viewed as the intertemporal manifestation of ambiguity aversion. The second axiom extends Koopmans' (1972) notion of stationarity from deterministic to stochastic environments.

The maxmin model was introduced in a seminal paper by Gilboa and Schmeidler (1989). In the tradition of Savage (1954) and Anscombe and Aumann (1963), they adopted an atemporal or static domain of choice that focuses on uncertainty, but that abstracts from issues pertaining to time. The axiomatization in this paper is complementary. Its primary objective is to generate new testable implications by focusing on intertemporal behavior. It is also argued that the dynamic framework and the choice-theoretic comparisons it permits are useful in discriminating among alternative models of ambiguity aversion. In this regard, the paper compares the intertemporal predictions of the maxmin model to those of the variational model introduced by Maccheroni, Marinacci, and Rustichini (2006a).

[1]I thank Larry Epstein, Hari Govindan, Wolfgang Pesendorfer, William Thomson and three anonymous referees for their helpful comments.

 DOI: 10.3982/ECTA10886

This paper makes another contribution that may be of conceptual as well as practical interest. The original axiomatizations of the maxmin and variational models are cast in the choice framework of Anscombe and Aumann (1963), henceforth AA. In this domain, the objects of choice are mappings, $f:\Omega \to \Delta(X)$, from states of the world $\omega \in \Omega$ into lotteries over the space X of possible consumption outcomes. Two sources of uncertainty are therefore present. States of the world represent subjective uncertainty about which probabilities are unknown and potentially ambiguous. The lotteries represent a situation of objective risk in which probabilities are given. The inclusion of both sources has played an important role in the axiomatic study of ambiguity. Foremost, it has enabled decision theorists to draw sharp behavioral distinctions between risk aversion and ambiguity aversion. In fact, the latter is often associated with a preference to bet on the objective rather than the subjective source. Ellsberg (1961) provided early evidence of such behavior and his thought experiments have now been widely replicated. The analysis in this paper is qualitatively different in that it does not require the presence of objective risk. Instead, the paper shows how the temporal structure of the choice domain can be used to study ambiguity-averse behavior.

2. DOMAIN

Time is discrete and varies over an infinite horizon $t \in \{0, 1, 2, \dots\} =: T$. The information structure is described by a filtered space $(\Omega, \{\mathcal{F}_t\}_t)$, where Ω is an arbitrary set of states of the world and $\{\mathcal{F}_t\}_t =: \mathcal{F}$ is an increasing sequence of algebras such that $\mathcal{F}_0 = \{\Omega, \emptyset\}$. If each algebra $\mathcal{F}_t$ is finite, one can think of the information structure as an infinite event tree whose nodes correspond to time–state pairs $(t, \omega) \in T \times \Omega$. For every pair (t, ω), $\mathcal{F}_t(\omega)$ denotes the intersection of all sets in $\mathcal{F}_t$ containing the state $\omega \in \Omega$.

Let X be a connected, separable, and first-countable topological space. It is interpreted as the set of consumption outcomes. An *act* h is an X-valued, $\mathcal{F}$-adapted process, that is, a sequence $(h_t)_{t\in T}$ such that $h_t:\Omega \to X$ is $\mathcal{F}_t$-measurable for every $t \in T$. Such processes are the multiperiod counterparts of Savage acts. The requirement that X is connected bears emphasis since it rules out the class of discrete-choice dynamic problems that arise, for example, in the study of occupational decisions. On the other hand, every convex subset of a Euclidean space is connected. Thus, the standard problem of allocating a divisible consumption good over time fits naturally into the framework.

To avoid technical complications, preferences are defined on a restricted space of acts. An act h is *finite* if there is a finitely generated algebra $\mathcal{A} \subset \bigcup_t \mathcal{F}_t$ such that h_k is $\mathcal{A}$-measurable for every $k \in T$. An act h is *bounded* if there is a compact set $K \subset X$ such that $\bigcup_t h_t(\Omega) \subset K$. The choice domain is the space of all finite and bounded acts. This space is denoted by $\mathcal{H}$. A preference relation $\succeq$ on $\mathcal{H}$ can be uniquely extended to nonfinite acts by adopting the methods in Epstein and Wang (1996), and extended to certain unbounded acts by adopting the methods in Bleichrodt, Rohde, and Wakker (2008). An act $h \in \mathcal{H}$ is *deterministic* if, for every $t \in T$, $h_t : \Omega \to X$ is a constant function, that is, if its outcomes do not depend on the state of the world. Let $\mathcal{D}$ be the subset of all deterministic acts in $\mathcal{H}$. Its generic elements are denoted by $d, d' \in \mathcal{D}$. As is common in the literature, the space of deterministic acts $\mathcal{D}$ is identified with a subset of X^∞.

3. AXIOMS

Throughout the paper, a preference relation $\succeq$ is a complete and transitive binary relation that is *nondegenerate*, that is, whose strict component $\succ$ is nonempty. To simplify the exposition, this assumption is not made explicit from now on. The focus of this section is a preference relation on the space of bounded and finite acts. It represents behavior prior to the resolution of any uncertainty. The case of repeated choice over time is studied in Section 6. The first axiom we impose is a weak form of continuity. It is based on ideas of Kreps (1988, p. 68), who explained how standard continuity assumptions have to be modified to allow for an unbounded utility index over outcomes.

CONTINUITY (C): *For all compact subsets K of X and all acts $h \in \mathcal{H}$, the sets $\{d \in K^\infty : d \succeq h\}$ and $\{d \in K^\infty : h \succeq d\}$ are closed in the product topology on K^∞.*

The next axiom requires that an act is preferred to another if it gives a better outcome stream in every state of the world. An implication of the axiom is that preferences are state-independent, that is, there are no "taste shocks." To state the assumption formally, let $h(\omega) \in X^\infty$ be the infinite stream of outcomes delivered by an act h in state $\omega \in \Omega$.

MONOTONICITY (M): *For all $h, g \in \mathcal{H}$, if $h(\omega) \succeq g(\omega)$ for every $\omega \in \Omega$, then $h \succeq g$.*

The next two axioms are the core of the paper. For all $t \in T, d \in \mathcal{D}$ and $h, g \in \mathcal{H}$, let $(d_{-(t,t+1)}, h_t, g_t)$ denote the act that gives the same outcomes as d in all periods $k \notin \{t, t+1\}$ and the possibly uncertain prospects h_t and g_t in periods t and $t+1$, respectively.

INTERTEMPORAL HEDGING (IH): *For all $t \in T$, $h, g \in \mathcal{H}$, and $d \in \mathcal{D}$,*

$$\begin{aligned}&(d_{-(t,t+1)}, h_t, h_t) \sim (d_{-(t,t+1)}, g_t, g_t)\\ &\quad \Rightarrow \quad (d_{-(t,t+1)}, h_t, g_t) \succeq (d_{-(t,t+1)}, h_t, h_t).\end{aligned}$$

The choice between $(d_{-(t,t+1)}, h_t, g_t)$ and $(d_{-(t,t+1)}, h_t, h_t)$ captures the following trade-off. The latter act yields identical outcomes in periods t and $t+1$. Hence, it smooths consumption across the two time periods. If there is uncertainty, a potential downside is that the outcomes in the two periods become perfectly correlated. An individual may prefer to take different bets in different time periods so as to hedge or reduce uncertainty about his overall consumption profile. The act $(d_{-(t,t+1)}, h_t, g_t)$ affords such "time diversification." Choosing this act reveals that smoothing consumption across states of the world is more beneficial to the individual than smoothing consumption across time.

One can draw a parallel between Intertemporal Hedging and Ambiguity Aversion, one of the key axioms in Gilboa and Schmeidler (1989).[2] Recall from the Introduction that the objects of choice in their domain are one-shot acts of the form $f:\Omega \to \Delta(X)$. An advantage of that domain is that one can immediately define the mixture of two acts f and f'. Thus, for any $\alpha \in [0, 1]$, let $\alpha f \oplus (1-\alpha)f'$ be the act obtained by mixing the respective lotteries state by state.[3] If two acts f and f' are indifferent, Gilboa and Schmeidler (1989) require that the mixture should be preferred since it can smooth out, or hedge, ambiguity. To illustrate, consider one of Ellsberg's thought experiments. Suppose f is a monetary bet on some event $A \subset \Omega$ whereby the individual wins x^* if A happens and loses x, $x < x^*$, otherwise. Let f' be the analogous bet on the complement of A. If the individual has no information about the likelihood of A, the symmetry of the situation makes it plausible to assume that f and f' are indifferent. At the same time, Ellsberg (1961) argued that many people would prefer the fifty–fifty lottery between x^* and x. The reason is that, unlike the bets, the odds of winning x^* in the lottery are known and unambiguous. But since, by construction, the mixed act $\frac{1}{2}f \oplus \frac{1}{2}f'$ delivers that same lottery in every of the state world, it should be preferred to f or f' as well. In the dynamic environment of this paper, there is no objective risk and so the probabilistic mixtures used in Gilboa and Schmeidler (1989) are not available. Instead, as our axiom suggests, the individual can smooth consumption across states of the world by taking different bets in different time periods, which can be viewed as forming a "time mixture" of the bets. The analogy raises the question whether Intertemporal Hedging is indicative of ambiguity aversion. The answer is not immediate since our axiom does not entail a comparison between bets with unknown and known odds, which is how ambiguity can be directly differentiated from risk. We return to this question at the end of the section after we state the rest of the axioms.

The next axiom extends Koopmans' (1972) notion of stationarity from deterministic to stochastic environments. To illustrate the implications specific to choice under uncertainty, consider an event that takes place in period $t = 1$.

[2] Gilboa and Schmeidler (1989) call their axiom Uncertainty Aversion. However, they reserve the term "uncertainty" for situations in which probabilities are unknown. In some of the subsequent literature, whose terminology we follow, it has been common to use the term "uncertainty" to encompass situations of both risk and ambiguity, and, therefore, to rename the axiom Ambiguity Aversion.

[3] Formally, let $\Delta(X)$ denote the space of lotteries with finite support. Given $p, p' \in \Delta(X)$, their mixture $\alpha p \oplus (1-\alpha)p'$ is the lottery in $\Delta(X)$ that gives $x \in X$ with probability $\alpha p(x) + (1-\alpha)p'(x)$.

In principle, its realization may affect consumption in the same time period or at a more distant future date. Stationarity requires that no matter how distant the ramifications, the individual exhibits the same degree of uncertainty aversion. More formally, let $f : \Omega \to X$ be an uncertain, one-shot prospect whose outcome depends on what happens in period $t = 1$. Mathematically, f is an $\mathcal{F}_1$-measurable function. Given outcomes $y, z \in X$, suppose the acts $(y, f, y, y, \dots)$ and $(y, z, y, y, \dots)$ are indifferent. Interpret $z \in X$ as the certainty equivalent of f when both are consumed in period $t = 1$. Suppose now that the consumption date of f and z is postponed by one period, which delivers the acts $(y, y, f, y, \dots)$ and $(y, y, z, y, \dots)$. In considering the former act, note that the relevant uncertainty is still composed of the events that take place in period $t = 1$, but they now affect consumption in period $t = 2$. Stationarity requires that the new acts are indifferent as well or that z remains the certainty equivalent of f. Put differently, the individual's attitude toward uncertainty does not depend on the date on which consumption takes place. In the formulation of the axiom below, as in Koopmans (1972), we postpone entire acts rather than the outcomes in a single period. It is noteworthy, however, that the simple comparisons just described are sufficient to distinguish maxmin from the variational model of Maccheroni, Marinacci, and Rustichini (2006a). The point is made in Section 5. To state the axiom, let (x, h) be the act obtained from $h \in \mathcal{H}$ by postponing the consumption date of each outcome by one period and adding $x \in X$ in the initial period, that is, let (x, h) be the act h' such that $h'_0 = x$ and $h'_t = h_{t-1}$ for all $t > 0$.[4]

STATIONARITY (S): *For all* $h, g \in \mathcal{H}$ *and* $x \in X$, $h \succeq g$ *if and only if* $(x, h) \succeq (x, g)$.

In the formulation of Stationarity, we have also embedded the familiar requirement that the ranking of consumption streams is history-independent. Formally, say that $\succeq$ satisfies *History Independence* (HI) if (x, h) is preferred to (x, g) whenever (y, h) is preferred to (y, g), for all $x, y \in X$ and $h, g \in \mathcal{H}$. In Section 5, this assumption will be used to characterize the more general, variational model introduced by Maccheroni, Marinacci, and Rustichini (2006a).

Stationarity can be viewed as the dynamic counterpart of Certainty Independence, which is the second key axiom in the original axiomatization of the maxmin model in Gilboa and Schmeidler (1989). Recall from the preceding discussion the structure of the AA domain used in that paper and the role of probabilistic mixing in smoothing out ambiguity. Certainty Independence

[4]To avoid confusion, it should be emphasized that the term "stationarity" is often used in an alternative sense. Suppose that uncertainty is perceived to be independently and identically distributed over time so that the individual faces the same stochastic environment after any node in the event tree. Behavior is then said to be stationary if the ex ante and all ex post preferences are identical. This requirement is logically independent of ours. Thus, the recursive variational model in Strzalecki (2013) is stationary in this alternative sense but not in ours.

requires that mixing f with a constant act, one that yields the same lottery $p \in \Delta(X)$ in every state, provides no such benefit since it cannot offset, or hedge, the variability in f. In particular, mixing with a constant act does not reverse the ranking of any two acts f and f'. Formally, $f \succeq f'$ if and only if $\alpha f \oplus (1-\alpha)p \succeq \alpha f' \oplus (1-\alpha)p$. Stationarity entails a similar requirement, but once again the "mixing" of acts is implemented through time. Specifically, the axiom implies that incorporating the certain outcome $x \in X$ in the initial period, while postponing the consumption date of $h, g \in \mathcal{H}$, does not reverse their prior ranking and, hence, does not change how the individual feels about the uncertainty inherent in those acts.

Our last axiom is restrictive but common in modeling intertemporal choice. Consider two deterministic acts that yield identical outcomes from the third period onward. The axiom requires that their ranking does not depend on the common continuation.

TIME SEPARABILITY (TS): *For all* $x, x', y, y' \in X$ *and* $d, d' \in \mathcal{D}$, $(x, y, d) \succeq (x', y', d)$ *if and only if* $(x, y, d') \succeq (x', y', d')$.

Even though the axiom imposes only a limited form of separability, it is well known from Koopmans (1972) that, together with Stationarity and Continuity, it implies that the intertemporal utility index is additive. Thus, the ranking of consumption in any given subset of time periods is independent of consumption levels in all other time periods.

From the perspective of this paper, the assumption of time separability is an important benchmark. It keeps the focus on uncertainty and permits a direct comparison with the standard model of intertemporal choice. The latter is obtained as a special case of the maxmin model in (1.1) by taking the set $\mathcal{P}$ of beliefs to be a singleton. In effect, this means that the individual treats the uncertainty like risk and does not perceive any ambiguity. In terms of behavior, the standard model satisfies every axiom in this section. Its distinctive feature is that it does not permit a *strict* preference for intertemporal hedging. Anticipating Theorem 1, it follows that, within the class of preferences in (1.1), a desire to hedge intertemporally is possible if and only if the set $\mathcal{P}$ is not a singleton. This leads to the interpretation of such behavior as indicative of ambiguity aversion.[5] The intuition is that ambiguity, being a more serious concern than risk, is what tips the scales in favor of smoothing consumption

[5]Some specifications of the maxmin model are probabilistically sophisticated in the sense of Machina and Schmeidler (1992) even when the set $\mathcal{P}$ is a nonsingleton. Epstein (1999) develops a definition of ambiguity aversion according to which such specifications are ambiguity-neutral. To interpret Intertemporal Hedging as evidence of ambiguity aversion in the sense of Epstein (1999), we would first have to exclude all such specifications. See Maccheroni, Marinacci, and Rustichini (2006a) for an elegant characterization of that subclass. The discussion above is consistent with the analysis in Ghirardato and Marinacci (2002) according to which all maxmin preferences, other than expected utility, represent ambiguity-averse behavior.

across states rather than time. An important caveat we have to stress is that this interpretation may not be justified within a broader class of preferences. For example, Richard (1975) shows that a strict preference for intertemporal hedging can arise in an otherwise standard expected-utility model if one adopts a suitably nonadditive intertemporal index. It is an open problem if one can identify intertemporal choices that are indicative of ambiguity aversion when fewer additional assumptions on behavior have been imposed.

4. MAIN RESULT

Let Δ be the space of finitely additive probability measures on $\bigcup_t \mathcal{F}_t$. Endow Δ with the weak* topology generated by the space B° of all simple, real-valued, bounded, and $\bigcup_t \mathcal{F}_t$-measurable functions on Ω. The next theorem is the main result of this paper.

THEOREM 1: *A preference relation $\succeq$ on $\mathcal{H}$ satisfies the axioms C, M, IH, S, and TS if and only if there exist a discount factor $\beta \in (0, 1)$, a continuous function $u : X \to \mathbb{R}$, and a nonempty weak*-closed convex subset $\mathcal{P}$ of Δ such that $\succeq$ has a maxmin utility representation*

$$V(h_0, h_1, \ldots) = \min_{p \in \mathcal{P}} \mathbb{E}_p \left[\sum_t \beta^t (u \circ h_t) \right]. \tag{4.1}$$

Moreover, $\mathcal{P}$ and β are unique and $u : X \to \mathbb{R}$ is unique up to positive affine transformations.

A sketch of the proof of Theorem 1 may develop further intuition about the role of the temporal domain in characterizing the maxmin model. Based on arguments familiar from Koopmans (1972), the first step is to obtain a discounted utility representation $U(x_0, x_1, \ldots) = \sum_t \beta^t u(x_t)$ on the space of deterministic acts. A standard continuity argument then implies that we can extend the utility function to the space of all acts. In particular, for each act $h \in \mathcal{H}$, let $V(h) := U(d_h)$, where $d_h \in \mathcal{D}$ is such that $h \sim d_h$. Each dynamic act $h \in \mathcal{H}$ is then mapped into a one-shot utility (util) act, $U \circ h : \Omega \to \mathbb{R}$, by replacing every outcome stream $h(\omega)$ with its discounted utility $U(h(\omega))$. Monotonicity implies that any two acts $h, h' \in \mathcal{H}$ are indifferent whenever they induce identical one-shot util acts, that is, whenever $U \circ h = U \circ h'$. As a result, one can define a mapping I from the space of util acts, $\mathcal{U} \subset B^\circ$, into the reals by letting $I(U \circ h) := V(h)$. By construction, I is monotone and has the property that $\alpha = I(\alpha)$ for every $\alpha \in \mathbb{R}$, where we have abused notation and used α to denote a real number as well as the function in B° that is identically equal to α. A mapping I with these properties is called a *certainty equivalent functional*. It specifies an "expected value" of each util act.

It is well known from Lemma 3.5 in Gilboa and Schmeidler (1989) that a certainty equivalent functional $I:B^\circ \to \mathbb{R}$ has the maxmin form, that is, $I(a) = \min_{p\in\mathcal{P}} \mathbb{E}_p a$ for all $a \in B^\circ$ if and only if it is translation-invariant, positively homogeneous, and quasiconcave.[6] Our goal is to deduce these properties from the hypothesis of Theorem 1. Since Intertemporal Hedging implies a preference to smooth outcomes across states of the world, it is not surprising that it delivers the quasiconcavity of I. Next note that because U is additively separable, a change in initial consumption changes the overall discounted utility in each state of the world by the same absolute amount. By History Independence, such changes do not affect how the individual feels about future uncertainty. The expected value of any act must, therefore, increase or decrease by the same amount by which the utility of initial consumption changes. This is precisely the mathematical property of translation invariance. Strengthening History Independence to Stationarity further implies that I satisfies a limited form of homogeneity. The logic is once again clear. If consumption is postponed by one period, then the utility in each state of the world is scaled down by the rate of time preference $\beta \in (0, 1)$. Since Stationarity requires that the individual becomes neither more nor less uncertainty-averse, the functional I must be invariant to such changes in scale. Mathematically, we have $I(\beta a) = \beta I(a)$ for all $a \in \mathcal{U}$. This form of homogeneity is limited since the scaling factor β is restricted to equal the rate of time preference. For similar reasons, in fact, Intertemporal Hedging implies only a limited form of quasiconcavity, namely, that $I(\frac{1}{1+\beta}a + \frac{\beta}{1+\beta}b) \geq I(a)$ whenever $I(a) = I(b)$, $a, b \in \mathcal{U}$. Before we can apply Lemma 3.5 in Gilboa and Schmeidler (1989), several difficulties therefore remain. The functional I has to be extended from the subset $\mathcal{U}$ of util acts to the entire space B°. Also, one has to insure that the extension satisfies the full-fledged properties of homogeneity and quasiconcavity. The remainder of the proof shows that there is already enough homogeneity and continuity to overcome these difficulties.

A final aspect of the proof of Theorem 1 that bears emphasis is the role of the infinite time horizon. As we now argue, a finite horizon cannot be adopted without some loss of generality. Thus, suppose there is a terminal period N so that each act h becomes a finite sequence $(h_0, h_1, \ldots, h_N)$. In considering the formulation of Stationarity, one encounters the obvious difficulty that outcomes in the last period cannot be postponed into the future. A partial solution is to limit attention to acts h, g whose ranking does not depend on what happens in the last period. In particular, one can formulate Stationarity by re-

[6] A mapping $I:B^\circ \to \mathbb{R}$ is translation-invariant if $I(a + \alpha) = I(a) + \alpha$ for all $a \in B^\circ$, $\alpha \in \mathbb{R}$. It is positively homogeneous if $I(\alpha a) = \alpha I(a)$ for all $a \in B^\circ$, $\alpha \in \mathbb{R}_+$. Respectively, these properties of the certainty equivalent functional are sometimes viewed as capturing constant absolute and constant relative ambiguity attitude, due to the obvious analogies with the theory of choice under risk.

quiring that for all acts h, g and all outcomes $x, y \in X$,

$$(h_0, h_1, \dots, h_{N-1}, y) \succeq (g_0, g_1, \dots, g_{N-1}, y)$$
$$\Leftrightarrow \quad (x, h_0, h_1, \dots, h_{N-1}) \succeq (x, g_0, g_1, \dots, g_{N-1}).$$

The solution is partial since the axiom does not restrict how the individual evaluates uncertain prospects resolving in the last period. In the extreme case in which all the uncertainty resolves in period N, that is, when $\mathcal{F}_0 = \mathcal{F}_{N-1}$, the axiom has no "bite" beyond its implications on the ranking of deterministic acts. Analogous limitations arise when one considers Intertemporal Hedging. Thus, if a finite-horizon setting is adopted, our axioms cannot deliver a representation on the entire space of acts. Instead, a representation is implied for the subset of acts h for which h_N is $\mathcal{F}_{N-1}$-measurable.

To conclude this section, consider the related work of Strzalecki (2013), which shares our focus on intertemporal behavior. Strzalecki (2013) studies a time-homogeneous environment in which uncertainty is perceived to be independently and identically distributed across time. An important aspect of this setting is that one can find two consumption profiles whose only difference concerns when the relevant uncertainty resolves. Suppose in particular that the individual is indifferent between early or late resolution of uncertainty and refer to this assumption as KPI, after Kreps and Porteus (1978), who first studied such behavior. Restricting attention to a parametric class of recursive ambiguity-averse preferences modeled after Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2011), Strzalecki (2013) proves that maxmin is the only model that satisfies KPI. A potentially interesting implication of his result and Theorem 1 is that Stationarity and KPI, both of which are used to pin down the maxmin model, are related. In fact, within the environment studied in Strzalecki (2013), they are equivalent under recursivity. This is surprising since the assumptions are conceptually distinct. One axiom imposes invariance with respect to the date on which consumption takes place; the other imposes invariance with respect to the date on which uncertainty resolves.[7] A limitation of working with KPI is that the assumption cannot be formulated in settings in which uncertainty is not time-homogeneous or, as Strzalecki (2013, p. 1049) explains, when learning may take place. By focusing on Stationarity, we are able to obtain a representation in arbitrary information environments with or without learning. Another point of difference stems from the fact that Strzalecki (2013) does not discuss intertemporal hedging. By using the utility model in Cerreia-Vioglio et al. (2011), he relies implicitly on the static notion of ambiguity aversion in Gilboa and Schmeidler (1989). The latter, however,

[7]Thus, Stationarity is binding even when all the uncertainty resolves in a single period, whereas KPI is vacuous in such a setting. Conversely, KPI is binding even when all the consumption takes place at a single point in time, as long as uncertainty resolves gradually. In this case, Stationarity is vacuous.

is not applicable in a setting of purely subjective uncertainty. The final difference is that Theorem 1 does not require recursivity of the preference relation. While recursivity has a strong normative appeal, it is sometimes desirable to relax it, since, as Epstein and Schneider (2003) show, it restricts the scope of ambiguity-averse behavior one can model.

5. THE VARIATIONAL MODEL

This section considers the variational model of Maccheroni, Marinacci, and Rustichini (2006a). The representation takes the form

$$V(h) = \min_{p\in\Delta}\left\{\mathbb{E}_p\left[\sum_t \beta^t(u\circ h_t)\right] + c(p)\right\}, \tag{5.1}$$

where $c:\Delta\to[0,+\infty]$ is a convex, lower semicontinuous function whose minimum value is 0. The maxmin model with a set $\mathcal{P}$ of priors is obtained as a special case by letting $c(p)=0$ for all $p\in\mathcal{P}$ and $c(p)=+\infty$ otherwise. Other notable special cases, including the multiplier model of Hansen and Sargent (2001), are discussed in Maccheroni, Marinacci, and Rustichini (2006a).

To characterize the variational model, the main difference is that Stationarity has to be replaced with History Independence. The other changes are less essential and concern only the ranking of deterministic acts. To insure that discounting takes the geometric form in (5.1), we need a stationarity assumption restricted to the ranking of deterministic acts. Call this assumption *Deterministic Stationarity* (DS). One final requirement, which is not needed for the characterization of the maxmin model, is that the utility function over outcomes is unbounded. In terms of behavior, this means that the ranking between any two acts $d, d'\in\mathcal{D}$ can be reversed by changing only a single outcome arbitrarily far into the future. Formally, say that $\succeq$ satisfies *Unboundedness* (UB) if, for every $d, d'\in\mathcal{D}$, $t\in T$, there exist outcomes $x, y\in X$ such that $(d_{-t}, x)\succ d'\succ(d_{-t}, y)$. The next result summarizes the discussion.[8]

THEOREM 2: *A preference relation $\succeq$ on $\mathcal{H}$ satisfies the axioms C, M, IH, HI, TS, DS, and UB if and only if it has a utility representation as in* (5.1) *with the function $u:X\to\mathbb{R}$ unbounded. Moreover, the function $c:\Delta\to[0,+\infty]$ and the discount factor β are unique, and $u:X\to\mathbb{R}$ is unique up to positive affine transformations.*

[8]History Independence can be viewed as the dynamic counterpart of Weak Certainty Independence, which is the axiom introduced by Maccheroni, Marinacci, and Rustichini (2006a) to characterize the variational model in an AA domain. As the proof of Theorem 2 and the analysis in Maccheroni, Marinacci, and Rustichini (2006a) reveal, both axioms are used to insure that the implied certainty equivalent functional $I: B^\circ\to\mathbb{R}$ is translation-invariant.

To deduce further testable implications of the variational model and to sharpen the comparison with the subclass of maxmin preferences, the next proposition investigates how the variational model may violate Stationarity. Recall from the discussion of the axiom that the certainty equivalent of a one-shot prospect is required to remain the same if the consumption date of both is postponed by one period, while the date at which uncertainty resolves is kept fixed. In this sense, the axiom requires that the individual's attitude toward uncertainty does not depend on the date on which consumption takes place. In the case of the variational model, as we now show, the certainty equivalent of a prospect increases monotonically (in terms of its utility) when consumption is postponed. Therefore, the individual becomes less uncertainty-averse, the more removed consumption is. It will be helpful to introduce some auxiliary notation. Suppose $\succeq$ satisfies the axioms in Theorem 2. For every $t \in T$, let $\mathcal{H}_t$ be the space of simple, $\mathcal{F}_t$-measurable functions from Ω into X. Fix an arbitrary $d \in \mathcal{D}$ and, for all $k, t \in T$, $f, f' \in \mathcal{H}_t$, write $f \succeq_{t,k} f'$ when $(d_{-(t+k)}, f) \succeq (d_{-(t+k)}, f')$. The preference relation $\succeq_{t,k}$ represents the ranking of one-shot prospects $f, f' \in \mathcal{H}_t$ whose uncertainty resolves on or before period t but whose outcomes are consumed in period $t + k$. Note that Stationarity requires that these rankings are independent of k.

PROPOSITION 3: *For every* $k, t \in T$, $f \in \mathcal{H}_t$, *and* $x \in X$, *if* $f \succeq_{t,k} x$, *then* $f \succeq_{t,k+1} x$. *Moreover, if* $\succeq$ *is not a maxmin preference, one can find* $k, t \in T$, $f \in \mathcal{H}_t$, *and* $x \in X$ *such that* $f \sim_{t,k} x$ *and* $f \succ_{t,k+1} x$.

The fact that certainty equivalents increase monotonically leads to the question what happens in the limit, as consumption is postponed indefinitely. To give a formal answer, it is necessary to introduce a notion of convergence for preference relations. Here, we give a heuristic account of the results and defer the exact technical details to Section A.2 in the Appendix. Suppose first that the preference relation $\succeq$ has a certainty equivalent functional $I : B^\circ \to \mathbb{R}$, defined as in the previous section, that is differentiable at the origin of B°. This property is not satisfied by the maxmin model, but holds generically within the class of variational preferences. See Maccheroni, Marinacci, and Rustichini (2006a) for further discussion of this point. In that case, we prove that for every $t \in T$, the sequence $(\succeq_{t,k})_k$ converges to an expected-utility preference on $\mathcal{H}_t$. Thus, when events in period t affect consumption in the more distant future, the induced uncertainty is eventually treated like risk. If the functional I is not differentiable at the origin, the sequence $(\succeq_{t,k})_k$ of one-shot rankings converges to a maxmin preference on the space $\mathcal{H}_t$.

6. REPEATED CHOICE

This section extends the analysis to the study of repeated choice over time. By building on the axiomatization in Theorem 1, the objective is to obtain the recursive maxmin representation of Epstein and Schneider (2003) in a setting of purely subjective uncertainty, that is, without the use of lotteries. As we discuss in the Appendix, we also find that two of the behavioral assumptions in Epstein and Schneider (2003) are redundant.

As is common in the study of repeated choice, in this section it is assumed that every algebra $\mathcal{F}_t$, $t \in T$, is finitely generated. The behavioral primitive is an $\mathcal{F}$-adapted process of preference relations $\{\succeq_{t,\omega}\}$, where $\succeq_{t,\omega}$ represents the ranking of acts $h \in \mathcal{H}$ conditional on all the information available at the node (t, ω). The first axiom we impose requires that the conditional ranking at a point in time does not depend on past outcomes or outcomes in those states of the world that can no longer be realized.

CONDITIONAL PREFERENCE (CP): *For all $t \in T$ and $\omega \in \Omega$, and all acts $h, g \in \mathcal{H}$, if $h_k(\omega') = g_k(\omega')$ for all $k \geq t$ and $\omega' \in \mathcal{F}_t(\omega)$, then $h \sim_{t,\omega} g$.*

The next axiom, Dynamic Consistency, links together preferences at different points in time. It insures that prior "plans" remain optimal when they are reevaluated in the future.

DYNAMIC CONSISTENCY (DC): *For all $t \in T$, and $\omega \in \Omega$, and acts $h, g \in \mathcal{H}$ that yield identical outcomes up to and including period t, if $h \succeq_{t+1,\omega'} g$ for all $\omega' \in \mathcal{F}_t(\omega)$, then $h \succeq_{t,\omega} g$. Moreover, the latter ranking is strict if, in addition, one of the former rankings is strict.*

We can now introduce the recursive maxmin representation. It has two familiar components: a discount factor $\beta \in (0, 1)$ and a utility index $u : X \to \mathbb{R}$. In the standard expected-utility model, the recursive formulation is completed by specifying for each node in the event tree, a one-step-ahead probability measure. This transition probability represents beliefs about the next period conditional on all currently available information. By comparison, in the recursive maxmin representation, there is, for every node in the tree, a set of one-step-ahead probability measures. This set reflects ambiguity about events in the next period. Formally, let $\{P^{t,\omega}\}$ be an $\mathcal{F}$-adapted process such that for every $t \in T$ and $\omega \in \Omega$, $P^{t,\omega}$ is a weak*-closed, convex subset of probability measures on $(\Omega, \mathcal{F}_{t+1})$. In addition, each $p \in P^{t,\omega}$, $t \in T$, and $\omega \in \Omega$ is such that $p(\mathcal{F}_t(\omega)) = 1$ and $p(A) > 0$ for every nonempty $\mathcal{F}_{t+1}$-measurable subset A of $\mathcal{F}_t(\omega)$. The last two restrictions express the fact that the event $\mathcal{F}_t(\omega)$ is known at the node (t, ω) and that no successive node is null. Finally, say that the components $(u, \beta, \{P^{t,\omega}\})$ are a *recursive maxmin representation* for $\{\succeq_{t,\omega}\}$ if a utility function for each conditional preference can be defined by means of the recursive relation

$$V^t(h; \omega) = u\big(h_t(\omega)\big) + \beta \min_{p \in P^{t,\omega}} \mathbb{E}_p\big[V^{t+1}(h; \cdot)\big] \quad \forall t \in T, \omega \in \Omega, h \in \mathcal{H}, \tag{6.1}$$

where $V^t(h;\omega)$ denotes the conditional utility of an act $h \in \mathcal{H}$ at the node (t,ω), $t \in T$, $\omega \in \Omega$. The rest of the discussion recalls some familiar but essential properties of the recursive formulation, which allow us to relate it to the model and analysis in Section 3. The key is a result by Epstein and Schneider (2003) that delivers the following closed-form solution for the conditional utilities, $V^t(\cdot;\omega)$, satisfying the recursion in (6.1):

$$V^t(h;\omega) = \min_{p\in\mathcal{P}} \mathbb{E}_p\left[\sum_{k\geq t} \beta^{k-t}(u \circ h_k)\Big|\mathcal{F}_t(\omega)\right] \quad \forall t \in T, \omega \in \Omega, h \in \mathcal{H}. \tag{6.2}$$

Above, $\mathcal{P}$ is a suitably defined set of probability measures on $(\Omega, \bigcup_t \mathcal{F}_t)$ representing the ex ante beliefs of the individual over the entire state space. The most important lesson from (6.2) is that all conditional preferences satisfy the behavioral assumptions made in Section 3. If each such preference is looked at in isolation, the recursive model is, therefore, seen to be a special case of the one in (1.1). Another observation evident from (6.2) is that recursivity, or the requirement that behavior is dynamically consistent, delivers an explicit rule for updating the set $\mathcal{P}$ of prior beliefs. Simply apply Bayes' rule prior by prior. To completely characterize the relationship between the representations in (6.1) and (6.2), consider how the set $\mathcal{P}$ is constructed from the process of one-step-ahead beliefs $\{P^{t,\omega}\}$. Let $\{p^{t,\omega}\}$ be an $\mathcal{F}$-adapted process such that $p^{t,\omega} \in P^{t,\omega}$ for every $t \in T$ and $\omega \in \Omega$. Think of $\{p^{t,\omega}\}$ as an infinite conditional probability tree and let p be the unique measure on $(\Omega, \bigcup_t \mathcal{F}_t)$ induced by folding back the transition probabilities. Equivalently, let $p \in \Delta$ be such that $p(A|\mathcal{F}_t(\omega)) = p^{t,\omega}(A)$ for all $A \in \mathcal{F}_{t+1}$, $t \in T$, and $\omega \in \Omega$. The set $\mathcal{P}$ consists of all measures p on $(\Omega, \bigcup_t \mathcal{F}_t)$ that can be obtained by selecting different probability trees $\{p^{t,\omega}\}$ from $\{P^{t,\omega}\}$. A set $\mathcal{P}$ that can be constructed from a process $\{P^{t,\omega}\}$ in this manner is called *rectangular* in the literature.

The next theorem summarizes our characterization of the recursive maxmin representation. By changing the assumptions imposed on ex ante behavior as we did in Section 5, one can also obtain the recursive variational representation of Maccheroni, Marinacci, and Rustichini (2006b).

THEOREM 4: *A process $\{\succeq_{t,\omega}\}$ satisfies CP and DC, and $\succeq_0$ satisfies C, M, IH, S, and TS if and only if $\{\succeq_{t,\omega}\}$ has a recursive maxmin representation $(u, \beta, \{P^{t,\omega}\})$. Moreover, the discount factor β and the process $\{P^{t,\omega}\}$ are unique, and $u : X \to \mathbb{R}$ is unique up to positive affine transformations.*

APPENDIX

A.1. *Proof of Theorem 1*

Let $\mathcal{K}(X)$ be the space of all compact subsets of X. For every $x \in X$, let $\mathbf{x}$ denote the sequence $(x, x, \ldots) \in X^\infty$. With the exception of Continuity, necessity of the axioms is easily verified. To check Continuity, take some $K \in \mathcal{K}(X)$.

It suffices to show that the mapping $(x_0, x_1, \ldots) \mapsto \sum_t \beta^t u(x_t)$ is continuous in the product topology on K^∞. Because $\beta \in (0, 1)$ and the continuous function $u : X \to \mathbb{R}$ is bounded on K, the series $\sum_t \beta^t u(x_t)$ is uniformly convergent. Continuity on K^∞ follows from Rudin (1976, Theorems 7.10 and 7.11).

Turn to sufficiency. Say that $\succeq$ is *sensitive* if $(x, d) \succ (y, d)$ for some $x, y \in X$, $d \in \mathcal{D}$.

LEMMA 5: *Continuity, Monotonicity, and Stationarity imply Sensitivity.*

PROOF: By way of contradiction, suppose that $(x, d) \sim (y, d)$ for all $x, y \in X$ and $d \in \mathcal{D}$. By Stationarity, $(z, x, d) \sim (z', x, d) \sim (z', y, d)$ for all $z, z', x, y \in X$ and all $d \in \mathcal{D}$. Repeating the argument, conclude that any $d, d' \in \mathcal{D}$ that differ in at most finitely many periods are indifferent. Next, take some arbitrary $d = (x_0, x_1, \ldots)$ and $d' = (y_0, y_1, \ldots)$ in $\mathcal{D}$. Letting $d^t := (x_0, \ldots, x_{t-1}, y_t, y_{t+1})$ for every $t \in T$, the previous argument implies that $d^t \sim d$ for all $t \in T$. Because d and d' are bounded, there is some $K \in \mathcal{K}(X)$ such that $d, d', d^t \in K^\infty$ for every $t \in T$. By construction, $(d^t)_t$ converges to d' in K^∞. By Continuity, $d \sim d'$. Because $d, d' \in \mathcal{D}$ were arbitrary, Monotonicity implies that all $h, h' \in \mathcal{H}$ are indifferent, contradicting the assumption that $\succeq$ is nondegenerate. *Q.E.D.*

For every $d \in \mathcal{D}$, let $X^t(d)$ be the subset of all deterministic acts of the form $(x_0, x_1, \ldots, x_{t-1}, d)$, where $x_k \in X$ for all $k < t$. Observe that $X^t(d) \subset \mathcal{D}$ whenever $d \in \mathcal{D}$.

LEMMA 6: *For every $d \in \mathcal{D}$, $t \in T$, and $h \in \mathcal{H}$, the sets $\{d' \in X^t(d) : d' \succeq h\}$ and $\{d' \in X^t(d) : h \succeq d'\}$ are closed in the product topology on $X^t(d)$.*

PROOF: Fix $d \in \mathcal{D}$, $t \in T$, and $K \in \mathcal{K}(X)$ such that $d \in K^\infty$. Because $X^t(d)$ is homeomorphic to the finite Cartesian product X^t, it is first-countable. Thus, it suffices to show that $\{d' \in X^t(d) : d' \succeq h\}$ is sequentially closed. An analogous argument shows that all lower contour sets are closed. Take a sequence $((y_0^n, \ldots, y_{t-1}^n))_n$ in X^t converging to $(y_0, \ldots, y_{t-1})$ such that $(y_0^n, \ldots, y_{t-1}^n, d) \succeq h$ for every $n \in \mathbb{N}$. For every $k \in \{0, \ldots, t-1\}$, $y_k^n \to y_k$ and, hence, the set $C_k := \{y_k^n : n \in \mathbb{N}\} \cup \{y_k\}$ is compact in X. It follows that $C := \bigcup_{k=0}^{t-1} C_k$ is compact in X. By construction, $K \cup C$ is compact and $(y_0, \ldots, y_{t-1}, d), (y_0^n, \ldots, y_{t-1}^n, d) \in [K \cup C]^\infty$ for every n. By Continuity, $(y_0, \ldots, y_{t-1}, d) \succeq h$, as desired. *Q.E.D.*

LEMMA 7: *There exists a continuous function $u : X \to \mathbb{R}$ and a discount factor $\beta \in (0, 1)$ such that the restriction of $\succeq$ to $\mathcal{D}$ is represented by the utility function*

$$U(x_0, x_1, \ldots) = (1 - \beta) \sum_t \beta^t u(x_t). \tag{A.1}$$

Moreover, β is unique and $u : X \to \mathbb{R}$ is unique up to positive affine transformations.

PROOF: The proof follows the arguments in Koopmans (1972). There are two minor differences. First, Koopmans makes the additional structural assumption that the space X is metrizable. Second, he formulates continuity of the preference relation in terms of the implied uniform metric on X^∞. By working with continuity in the product topology, which is a stronger assumption, we are able to drop metrizability and two of Koopmans' axioms. In particular, Lemma 5 shows that Sensitivity, or Axiom P2 in Koopmans (1972), is implied. One can verify that the same is true for his axiom P5. Together with Lemma 6, we have thus confirmed that all conditions needed for Result F in Koopmans (1972, p. 88) are met. Note that metrizability of X is not used in the proof of this result. The latter is based on the arguments of Debreu (1960) and Gorman (1968) for which it suffices that X is separable and connected. The representation in (A.1) is then delivered by Proposition 3 in Koopmans (1972, p. 89). The proof of that proposition is, in fact, simplified once continuity in the product topology is assumed. *Q.E.D.*

Because X is connected, the range of $u: X \to \mathbb{R}$, $u(X)$, is an interval. Because $\succeq$ is sensitive, the interval has nonempty interior. Rescaling appropriately, it is without loss of generality (w.l.o.g.) to assume that $u(X)$ contains the interval $[-1, 1]$. Let $x^* \in X$ be an outcome such that $u(x^*) = 0$.

LEMMA 8: *For every $h \in \mathcal{H}$, there is some $d \in \mathcal{D}$ such that $h \sim d$.*

PROOF: Fix $h \in \mathcal{H}$ and let $K \in \mathcal{K}(X)$ be such that $\bigcup_t h_t(\Omega) \subset K$. Let $x, y \in X$ attain the maximum and, respectively, the minimum of $u: X \to \mathbb{R}$ over K. It follows from Monotonicity and Lemma 7 that $\mathbf{x} \succeq h \succeq \mathbf{y}$. If either $h \sim \mathbf{x}$ or $h \sim \mathbf{y}$, we are done. Suppose $\mathbf{x} \succ h \succ \mathbf{y}$. The sequence $d^1 := (x, \mathbf{y})$, $d^2 := (x, x, \mathbf{y}), \ldots$ in K^∞ converges pointwise to $\mathbf{x}$. By Continuity, there is some $t \in T$ such that $d^t \succ h \succ \mathbf{y}$. Both d^t and $\mathbf{y}$ belong to $X^t(\mathbf{y})$, which is connected. By Lemma 6, the preference relation has closed upper and lower contour sets in $X^t(\mathbf{y})$. A standard argument shows that there is some $d \in X^t(\mathbf{y})$ such that $d \sim h$. *Q.E.D.*

For every $h \in \mathcal{H}$, define $V(h) := U(d_h)$ for some $d_h \in \mathcal{D}$ such that $d_h \sim h$. The transitivity of $\succeq$ implies that $V: \mathcal{H} \to \mathbb{R}$ is well defined and represents the preference relation $\succeq$. For every act $h \in \mathcal{H}$, define the *util act* $U \circ h: \Omega \to \mathbb{R}$ as

$$[U \circ h](\omega) := (1 - \beta) \sum_t \beta^t u\big(h_t(\omega)\big). \tag{A.2}$$

Let $\mathcal{U} = \{U \circ h : h \in \mathcal{H}\}$ be the set of all util acts. Define the functional $I: \mathcal{U} \to \mathbb{R}$ as

$$I(U \circ h) := V(h). \tag{A.3}$$

In Section 4, I was referred to as the certainty equivalent functional of $\succeq$. Note that it is well defined by Monotonicity.

Let B° denote the set of all simple, real-valued, $\bigcup_t \mathcal{F}_t$-measurable functions on Ω. For every $t \in T$, B°_t is the subset of simple, $\mathcal{F}_t$-measurable functions whose range is in the interval $[-\beta^t, \beta^t]$. Endow B° with the sup-norm. For $a, b \in B^\circ$, write $a \ll b$ if $a(\omega) < b(\omega)$ for every $\omega \in \Omega$. Abusing notation, write α for the function in B° that is identically equal to $\alpha \in \mathbb{R}$.

LEMMA 9: *For every $t \geq 0$, $B^\circ_t \subset \mathcal{U}$. In particular, $\mathcal{U}$ is an absorbing subset of B°.*

PROOF: Let $a = \sum_{i=1}^k \alpha_i \mathbf{1}_{A_i}$ be the canonical representation of $a \in B^\circ_t$ and note that for every $i = 1, \dots, k$, $|\frac{\alpha_i}{\beta^t}| \leq 1$. Let x_i be such that $u(x_i) = \frac{\alpha_i}{\beta^t}$. Define $f(\omega) = x_i$ for every $\omega \in A_i$, $i \in \{1, \dots, k\}$. Define an act h by setting $h_\tau := x^*$ for all $\tau < t$ and $h_\tau = f$ for all $\tau \geq t$. By construction, $U \circ h = a$. To see that $\mathcal{U}$ is absorbing, fix any $a \in B^\circ$. Because a is simple, it must be $\mathcal{F}_t$-measurable for some t. Moreover, for k large enough, $\beta^k a(\omega) \in [-\beta^t, \beta^t]$ for every $\omega \in \Omega$. Conclude that $\beta^k a \in \mathcal{U}$ for k large enough. *Q.E.D.*

LEMMA 10: *For every $t, k \in T$ and $a \in B^\circ_t$, $I(\beta^k a) = \beta^k I(a)$.*

PROOF: Fix $t \in T$ and $a \in B^\circ_t$. It is enough to show that $I(\beta a) = \beta I(a)$. By Lemma 9, there is $h \in \mathcal{H}$ such that $U \circ h = a$. Let $d_h \in X^\infty$ be such that $d_h \sim h$. Consider the act (x^*, h) and note that $U \circ (x^*, h) = \beta a$. By Stationarity, $(x^*, h) \sim (x^*, d_h)$. Conclude that

$$
\begin{aligned}
I(\beta a) &= V\big(x^*, h\big) = U\big(x^*, d_h\big) = \beta U(d_h) \\
&= \beta V(h) = \beta I(U \circ h) = \beta I(a),
\end{aligned}
$$

as desired. *Q.E.D.*

LEMMA 11: *The functional $I : \mathcal{U} \to \mathbb{R}$ can be uniquely extended to a functional $\tilde{I} : B^\circ \to \mathbb{R}$ such that $\tilde{I}(\beta a) = \beta \tilde{I}(a)$ for every $a \in B^\circ$.*

PROOF: Fix $a \in B^\circ$ and let $k, t \in T$ be such that $\beta^k a \in B^\circ_t$. Define

$$
\tilde{I}(a) := \frac{1}{\beta^k} I\big(\beta^k a\big). \tag{A.4}
$$

To see that $\tilde{I}$ is well defined, let (k', t') be another pair such that $\beta^{k'} a \in B^\circ_{t'}$. If $k' \geq k$, then $\beta^{k'} a \in B^\circ_t$; if $k > k'$, then $\beta^k a \in B^\circ_{t'}$. It is, therefore, w.l.o.g. to assume that $t = t'$, $k' > k$, and $\beta^{k'} a, \beta^k a \in B^\circ_t$. Lemma 10 then implies that

$$
\frac{1}{\beta^{k'}} I\big(\beta^{k'} a\big) = \frac{1}{\beta^{k'}} I\big(\beta^{k'-k} \beta^k a\big) = \frac{1}{\beta^{k'}} \beta^{k'-k} I\big(\beta^k a\big) = \frac{1}{\beta^k} I\big(\beta^k a\big).
$$

Hence, $\tilde{I}$ is well defined. By construction, $\tilde{I}(\beta a) = \beta\tilde{I}(a)$ for every $a \in B^\circ$. Because $\mathcal{U}$ is absorbing, $\tilde{I}$ is the unique extension of I having this property. *Q.E.D.*

LEMMA 12: *For every $a \in B^\circ, \alpha \in \mathbb{R}$, $\tilde{I}(a+\alpha) = \tilde{I}(a) + \alpha$. That is, $\tilde{I}$ is translation-invariant.*

PROOF: Fix $a \in B^\circ$ and $\alpha \in \mathbb{R}$. By Lemma 11, it is enough to show that $\tilde{I}(\beta^k a + \beta^k \alpha) = \tilde{I}(\beta^k a) + \beta^k \alpha$ for some $k > 0$. Because a is simple, it is $\mathcal{F}_t$-measurable for some t. Choose k large enough so that $\beta^k a(\omega) \in [-\beta^{2t}, \beta^{2t}]$ for every ω and $\beta^k \alpha \in [\beta^t - 1, 1 - \beta^t]$. The first inclusion allows us to find an act $h \in \mathcal{H}$ such that $U \circ (x_0^*, \dots, x_{t-1}^*, h) = \beta^k a$. The second inclusion allows us to find $(x_0, \dots, x_{t-1}) \in X^t$ such that $(1-\beta)\sum_{\tau=0}^{t-1} u(x_\tau) = \beta^k \alpha$. Consider the act $(x_0, \dots, x_{t-1}, h)$ and notice that $U \circ (x_0, \dots, x_{t-1}, h) = \beta^k a + \beta^k \alpha$. If d_h is such that $d_h \sim h$, Stationarity implies that

$$(x_0, \dots, x_{t-1}, h) \sim (x_0, \dots, x_{t-1}, d_h),$$
$$\big(x_0^*, \dots, x_{t-1}^*, h\big) \sim \big(x_0^*, \dots, x_{t-1}^*, d_h\big).$$

Conclude that

$$\begin{aligned} I\big(\beta^k a + \beta^k \alpha\big) &= U(x_0, \dots, x_{t-1}, d_h) \\ &= U\big(x_0^*, \dots, x_{t-1}^*, d_h\big) + \beta^k \alpha \\ &= V\big(x_0^*, \dots, x_{t-1}^*, h\big) + \beta^k \alpha = I\big(\beta^k a\big) + \beta^k \alpha. \end{aligned}$$

Since $\tilde{I}$ is an extension of I, the above equalities complete the proof. *Q.E.D.*

LEMMA 13: *For every $a, b \in B^\circ$, $\tilde{I}(a) = \tilde{I}(b)$ implies $\tilde{I}(\frac{1}{1+\beta}a + \frac{\beta}{1+\beta}b) \geq \tilde{I}(a)$.*

PROOF: Take some $a, b \in B^\circ$ such that $\tilde{I}(a) = \tilde{I}(b)$. Let $t \in T$ be such that a and b are $\mathcal{F}_t$-measurable. Choose $k \in \mathbb{N}$ large enough so that

$$\frac{\beta^{k-t}}{1-\beta^2}a(\omega), \frac{\beta^{k-t}}{1-\beta^2}b(\omega) \in [-1, 1] \quad \forall \omega \in \Omega.$$

One can then find acts $h, g \in \mathcal{H}$ such that $u \circ h_t = \frac{\beta^{k-t}}{1-\beta^2}a$ and $u \circ g_t = \frac{\beta^{k-t}}{1-\beta^2}b$. By construction,

$$\begin{aligned} U \circ \big(\mathbf{x}^*{}_{-(t,t+1)}, h_t, h_t\big) &= \beta^t\big(1-\beta^2\big)(u \circ h_t) = \beta^k a, \\ U \circ \big(\mathbf{x}^*{}_{-(t,t+1)}, g_t, g_t\big) &= \beta^t\big(1-\beta^2\big)(u \circ g_t) = \beta^k b, \\ U \circ \big(\mathbf{x}^*{}_{-(t,t+1)}, h_t, g_t\big) &= \beta^t(1-\beta)(u \circ h_t + \beta u \circ g_t) \\ &= \beta^k\left(\frac{1}{1+\beta}a + \frac{\beta}{1+\beta}b\right). \end{aligned}$$

By Lemma 10, $\tilde{I}(a) = \tilde{I}(b)$ implies $\tilde{I}(\beta^k a) = \tilde{I}(\beta^k b)$. From the definition of $\tilde{I}$, $(\mathbf{x}^*_{-(t,t+1)}, h_t, h_t) \sim (\mathbf{x}^*_{-(t,t+1)}, g_t, g_t)$. By IH, $(\mathbf{x}^*_{-(t,t+1)}, h_t, g_t) \succeq (\mathbf{x}^*_{-(t,t+1)}, h_t, h_t)$. One more application of Lemma 10 gives $\tilde{I}(\frac{1}{1+\beta}a + \frac{\beta}{1+\beta}b) \geq \tilde{I}(a)$, as desired. *Q.E.D.*

LEMMA 14: *The functional* $\tilde{I}: B^\circ \to \mathbb{R}$ *is monotone, quasiconcave, and positively homogeneous.*

PROOF: That $\tilde{I}$ is monotone follows directly from Monotonicity and Lemma 10. To show quasiconcavity, first note that for any $a, b \in B^\circ$, we have $a \leq b + \|a - b\|$. By Lemma 12, $\tilde{I}(a) \leq \tilde{I}(b + \|a - b\|) = \tilde{I}(b) + \|a - b\|$. Reversing the roles of a and b gives $|\tilde{I}(a) - \tilde{I}(b)| \leq \|a - b\|$. Conclude that $\tilde{I}$ is sup-norm continuous. Next, suppose by way of contradiction that there are $a, b \in B^\circ$, $\gamma \in (0, 1)$, and $\alpha \in \mathbb{R}$ such that $\tilde{I}(a), \tilde{I}(b) \geq \alpha$ and $\tilde{I}(\gamma a + (1 - \gamma)b) < \alpha$. Because $\tilde{I}$ is norm-continuous, the following real numbers are well defined:

$$\lambda'' := \min\{\lambda \in [\gamma, 1] : \tilde{I}(\lambda a + (1 - \lambda)b) \geq \alpha\},$$
$$\lambda' := \max\{\lambda \in [0, \gamma] : \tilde{I}(\lambda a + (1 - \lambda)b) \leq \alpha\}.$$

By construction, $\lambda'' > \lambda'$ and $\tilde{I}(c) < \alpha$ for every $c \in B^\circ$ in the interior of the line segment connecting $\lambda' a + (1 - \lambda')b$ and $\lambda'' a + (1 - \lambda'')b$. Moreover, by the intermediate value theorem,

$$\tilde{I}(\lambda' a + (1 - \lambda')b) = \tilde{I}(\lambda'' a + (1 - \lambda'')b) = \alpha.$$

But then, by Lemma 13, there is some $c \in B^\circ$ in the linear segment connecting the points $\lambda' a + (1 - \lambda')b$ and $\lambda'' a + (1 - \lambda'')b$ such that $\tilde{I}(c) > \alpha$, establishing a contradiction. Turn to positive homogeneity. Take $\alpha > 0$ and $a \in B^\circ$. By Lemmas 10 and 12, respectively, it is w.l.o.g. to assume that $\alpha \in (0, 1)$ and $\tilde{I}(a) = 0$. Then the quasiconcavity of $\tilde{I}$ implies that $\tilde{I}(\alpha a) \geq 0$. By way of contradiction, suppose now that $k := \tilde{I}(\alpha a) > 0$ and let $b := \alpha a - k$. By Lemma 12, $I(b) = 0$. Choose $t \in T$ large enough so that $\beta^t < \alpha$. By construction, $b \ll \alpha a$ and, hence, $\frac{\beta^t}{\alpha} b \ll \beta^t a$. Because $a, b \in B^\circ$ are simple functions, one can choose $\varepsilon > 0$ so that $\frac{\beta^t}{\alpha} b \ll \beta^t a - \varepsilon$. Making use of the fact that $\tilde{I}$ is monotone and Lemma 12 in turn, we have

$$\tilde{I}\left(\frac{\beta^t}{\alpha} b\right) \leq \tilde{I}(\beta^t a - \varepsilon) = \tilde{I}(\beta^t a) - \varepsilon = -\varepsilon < 0.$$

But the quasiconcavity of $\tilde{I}$ implies that $\tilde{I}(\frac{\beta^t}{\alpha} a) \geq 0$, establishing a contradiction. *Q.E.D.*

Lemmas 12 and 14 show that $\tilde{I}$ satisfies all the properties necessary to apply Lemma 3.5 in Gilboa and Schmeidler (1989), which delivers the desired representation. The uniqueness of the set $\mathcal{P}$ of beliefs is established in Gilboa and Schmeidler (1989, p. 149).

A.2. *Proofs of Results From Section 5*

PROOF OF THEOREM 2: Necessity of the axioms is obvious. To prove sufficiency, adopt the same notation as in the proof of Theorem 1 and note that all results up to Lemma 9 continue to hold. Using UB, the first step is to show that the utility index $u: X \to \mathbb{R}$ is unbounded. In the statement of the axiom, let $d, d' \in \mathcal{D}$ be such that $d := (x^*, x^*, \dots)$ and $U(d') > 0$. Recall that x^* is such that $u(x^*) = 0$. It follows from UB that for every $t \in T$, there is $x_t \in X$ such that $\beta^t u(x_t) > U(d') > 0$. Because $\beta^t \to 0$ as $t \to +\infty$, the latter is possible only if $u(x_t) \to +\infty$. Thus, $u: X \to \mathbb{R}$ is unbounded above. An analogous argument shows that the function is unbounded below. Conclude that $\mathcal{U} = B^\circ$ and, hence, that the functional I, as defined in (A.3), takes B° for its domain. The next step is to establish that I is monotone, translation-invariant and quasiconcave. That I is monotone follows from Monotonicity. As we now show, translation invariance follows from HI. The difference from Lemma 12 is that we now make use of the fact that $u: X \to \mathbb{R}$ is unbounded, which obviates the need for the stronger stationarity axiom. Thus, take any $a \in B^\circ$ and $\alpha \in \mathbb{R}$. Because $a \in B^\circ$ is simple, it is $\mathcal{F}_t$-measurable for some $t \in T$. Because $u: X \to \mathbb{R}$ is unbounded, one can find $x \in X$ and $f: \Omega \to X$ such that $u(x) = (1-\beta)^{-1}\alpha$ and $u \circ f = (1-\beta)^{-1}\beta^{-(t+1)}a$. Let $h := (\mathbf{x}^*_{-t}, f)$. Note that $U \circ (x^*, h) = a$ and $U \circ (x, h) = a + \alpha$. Thus, $V(x^*, h) = I(a)$ and $V(x, h) = I(a + \alpha)$. By Continuity, there is $z \in X$ such that

$$\text{(A.5)} \qquad \begin{aligned} (x, h) &= \big(x, x_0^*, x_1^*, \dots, x_{t-1}^*, f, x_{t+1}^*, x_{t+2}^*, \dots\big) \\ &\sim \big(x, x_0^*, x_1^*, \dots, x_{t-1}^*, z, x_{t+1}^*, x_{t+2}^*, \dots\big). \end{aligned}$$

Hence, $I(a+\alpha) = (1-\beta)u(x) + (1-\beta)\beta^{t+1}u(z)$. Given (A.5), History Independence implies that

$$\begin{aligned} \big(x^*, h\big) &= \big(x^*, x_0^*, x_1^*, \dots, x_{t-1}^*, f, x_{t+1}^*, x_{t+2}^*, \dots\big) \\ &\sim \big(x^*, x_0^*, x_1^*, \dots, x_{t-1}^*, z, x_{t+1}^*, x_{t+2}^*, \dots\big). \end{aligned}$$

Hence, $I(a) = (1-\beta)\beta^{t+1}u(z)$. Combining all of the above arguments gives

$$I(a+\alpha) = (1-\beta)u(x) + (1-\beta)\beta^{t+1}u(z) = \alpha + I(a),$$

as desired. To prove that I is quasiconcave, one can mimic the proofs of Lemmas 13 and 14. To avoid the use of Stationarity, the only necessary change is in the proof of Lemma 13 in which one has to choose the acts $h, g \in \mathcal{H}$ so that $u \circ h_t = \beta^{-t}(1-\beta^2)^{-1}a$ and $u \circ g_t = \beta^{-t}(1-\beta^2)^{-1}b$. This is now possible because $u : X \to \mathbb{R}$ is unbounded. Having established that $I : B^\circ \to \mathbb{R}$ is monotone, translation-invariant, and quasiconcave, the proof of Theorem 2 is completed as in Maccheroni, Marinacci, and Rustichini (2006a). See, in particular, Lemmas 25 and 26 in their paper and how these lemmas are applied on p. 1480. *Q.E.D.*

PROOF OF PROPOSITION 3: For every $k, t \in T$, $\succeq_{t,k}$ has a utility representation

$$V_{t,k}(f) := \min_{p \in \Delta} \{ \mathbb{E}_p(u \circ f) + \beta^{-t-k} c(p) \} \quad \forall f \in \mathcal{H}_t. \tag{A.6}$$

Using this representation, the first assertion in Proposition 3 follows by direct verification. If $\succeq$ is not a maxmin preference, its certainty equivalent functional $I : B^\circ \to \mathbb{R}$ is not positively homogeneous. Hence, $I(\alpha a) \neq \alpha I(a)$ for some $\alpha > 0$ and $a \in B^\circ$. It is w.l.o.g. to assume that $\alpha < 1$. Else, take $a' := \alpha a$ and $\alpha' = \alpha^{-1}$. Because I is translation-invariant, one can also assume that $I(a) = 0$. Based on Theorem 2, we know that $I(b) = \min_{p \in \Delta}[\mathbb{E}_p b + c(p)]$ for all $b \in B^\circ$. Thus, I is the pointwise minimum of affine functionals on B°, which implies that I is concave. It follows that $I(\alpha a) > 0$ and, in fact, that $I(\gamma a) > 0$ for all $\gamma \in (0, 1)$. Pick $t \in T$ such that $a \in B^\circ$ is $\mathcal{F}_t$-measurable and any $k \in T$. Because $u : X \to \mathbb{R}$ is unbounded, there is $f \in \mathcal{H}_t$ such that $\beta^{t+k}(u \circ f) = a$. By construction, $f \sim_{t,k} x^*$ and $f \succ_{t,k+1} x^*$, as desired. *Q.E.D.*

Keeping the same notation and assumptions as in the preceding two proofs, we now address the problem raised at the end of Section 5. Fix some $t \in T$. The next result characterizes the limit of the sequence $(\succeq_{t,k})_k$ of preference relations on $\mathcal{H}_t$ as $k \to +\infty$. We use the standard notion of convergence for preference relations due to Kannai (1970). Endow the space $\mathcal{H}_t$ with the product topology and say that the sequence $(\succeq_k)_k$ converges to a preference relation $\succeq^*$ on $\mathcal{H}_t$ if, for all $f, f' \in \mathcal{H}_t$ and all sequences $(f_k)_k$ and $(f'_k)_k$ in $\mathcal{H}_t$ converging to f and f', respectively, if $f \succ^* f'$, then $f_k \succ_k f'_k$ for all k large enough. Now, let $\partial I(0)$ be the set of all $q \in \Delta$ such that $\mathbb{E}_q a \geq I(a)$ for all $a \in B^\circ$. Because $I : B^\circ \to \mathbb{R}$ is translation-invariant and monotone, it is norm-continuous. This was pointed out in Lemma 10. In the proof of Proposition 3, it was also observed that I is concave. It follows from Phelps (1993, Proposition 1.11) that $\partial I(0)$ is nonempty and weak*-closed. Define $V^*(f) := \min_{q \in \partial I(0)} \mathbb{E}_q(u \circ f)$ for all $f \in \bigcup_t \mathcal{H}_t$. Finally, observe that if I is differentiable at $0 \in B^\circ$, by which we mean that $\partial I(0)$ is a singleton, then the function V^* induces an expected-utility preference on the space $\bigcup_t \mathcal{H}_t$.

PROPOSITION 15: *Assume that $\mathcal{F}_t$ is finitely generated for every $t \in T$. Then, for every $t \in T$, the sequence $(\succeq_{t,k})_k$ converges to the maxmin preference relation on $\mathcal{H}_t$ represented by V^*.*

PROOF: Fix $t \in T$. It follows from Maccheroni, Marinacci, and Rustichini (2006a, Propositions 6 and 12) that as $k \to +\infty$, the functions $V_{t,k}$ in (A.6) converge pointwise to V^* on $\mathcal{H}_t$. In general, the pointwise convergence of utility functions is not strong enough to imply that the underlying preferences converge in a meaningful way. See Border and Segal (1994, p. 185). As we now show, however, convergence in the stronger sense of Kannai (1970) is implied. Take $f, f' \in \mathcal{H}_t$ such that $V^*(f) > V^*(f')$, and sequences (f_k) and (f'_k) in $\mathcal{H}_t$ that converge to f and f', respectively. Since the pointwise topology on $\mathcal{H}_t$ is first-countable, the set C consisting of the elements of both sequences and their limit points is compact in $\mathcal{H}_t$. Because the algebra $\mathcal{F}_t$ is finite, each function $V_{t,k}$ is continuous in the product topology on $\mathcal{H}_t$. It is also clear from (A.6) that the functions $V_{t,k}$ converge monotonically to V^* on $\mathcal{H}_t$. It follows from Dini's theorem that the functions $V_{t,k}$ converge uniformly to V^* on C. But then $V_{t,k}(f_k) > V_{t,k}(f'_k)$ for all k large enough, as desired. *Q.E.D.*

A.3. *Proof of Theorem 4*

The proof begins by deducing two properties of the process $\{\succeq_{t,\omega}\}$ that are assumed a priori in Epstein and Schneider (2003). First, Lemma 16 shows that under CP and DC, if the ex ante preference $\succeq_0$ satisfies the axioms in Section 3, then so does every conditional preference. In contrast, the axiom Multiple Prior in Epstein and Schneider (2003, p. 5) imposes a priori that all conditional preferences satisfy the axioms used in their derivation of the maxmin model. The long but elementary proof of the lemma is omitted.

LEMMA 16: *Suppose $\{\succeq_{t,\omega}\}$ satisfies CP and DC. If $\succeq_0$ satisfies C, M, IH, S, and TS, then so does every conditional preference $\succeq_{t,\omega}$, $t \in T$, $\omega \in \Omega$.*

Next, say that the process $\{\succeq_{t,\omega}\}$ satisfies *Conditional State Independence* if, for every $t \in T$, $\omega \in \Omega$,

$$(x_0, x_1, \dots, x_{t-1}, d) \succeq_0 (x_0, x_1, \dots, x_{t-1}, d')$$
$$\Leftrightarrow \quad (x_0, x_1, \dots, x_{t-1}, d) \succeq_{t,\omega} (x_0, x_1, \dots, x_{t-1}, d')$$

for all $x_k \in X$, $k < t$, and all $d, d' \in \mathcal{D}$. The axiom is common in the analysis of dynamic choice and requires that tastes do not change over time. It is imposed in Epstein and Schneider (2003) as part of a composite assumption called Risk Preference. Lemma 17 below shows that under a minor regularity condition, the axiom is implied whenever the ex ante ranking $\succeq_0$ satisfies Monotonicity and the process $\{\succeq_{t,\omega}\}$ satisfies CP and DC. Lemmas 18 and 19 then confirm that the regularity condition is met whenever the ranking of deterministic acts at any node (t, ω) is stationary and continuous. Conditional State Independence is thus implied by the hypothesis of Theorem 4 and is redundant in Epstein and Schneider (2003). To proceed, say that two preference relations

$\succeq$ and $\succeq'$ on $\mathcal{D}$ *disagree* if there are $d, d' \in \mathcal{D}$ such that $d \succ d'$ and $d' \succ' d$. The proof of the next lemma is straightforward and is therefore omitted.

LEMMA 17: *Suppose that for every $t \in T$, $\omega, \omega' \in \Omega$, the restrictions of $\succeq_{t,\omega}$ and $\succeq_{t,\omega'}$ to $\mathcal{D}$ disagree whenever they are not identical. If $\{\succeq_{t,\omega}\}$ satisfies CP and DC, and $\succeq_0$ satisfies M, then $\{\succeq_{t,\omega}\}$ satisfies Conditional State Independence.*

A pair $(d', d'') \in \mathcal{D}^2$ is a *gap* for a preference relation $\succeq$ on $\mathcal{D}$ if $d'' \succ d'$ and there is no $d \in \mathcal{D}$ such that $d'' \succ d \succ d'$. The preference relation $\succeq$ on $\mathcal{D}$ is *continuous* if, for every $K \in \mathcal{K}(X)$, the upper and lower contour sets of the restriction of $\succeq$ to K^∞ are closed in K^∞.

LEMMA 18: *A continuous preference relation $\succeq$ on $\mathcal{D}$ has no gaps.*

PROOF: Suppose (d', d'') is a gap. Let $d' =: (x_0, x_1, \dots)$ and $d'' =: (y_0, y_1, \dots)$, and define $d^t := (y_0, \dots, y_{t-1}, x_t, x_{t+1}, \dots)$ for every $t \in T$. Let $K \in \mathcal{K}(X)$ be such that $d'', d' \in K^\infty$. By construction, the sequence $(d^t)_t$ lies in K^∞ and converges to d''. Continuity implies that $d^t \succ d'$ for some $t \in T$. Because (d', d'') is a gap, $d^t \succeq d''$. Let $d := (x_t, x_{t+1}, \dots)$. By construction, $d^t, d' \in X^t(d)$. Conclude that the sets $\{\tilde{d} \in X^t(d) : \tilde{d} \succeq d''\}$ and $\{\tilde{d} \in X^t(d) : d' \succeq \tilde{d}\}$ are nonempty. An argument analogous to the proof of Lemma 6 shows that both sets are closed in $X^t(d)$. Because (d', d'') is a gap, they partition the connected space $X^t(d)$ into two nonempty closed sets, a contradiction. *Q.E.D.*

LEMMA 19: *Let $\succeq$ and $\succeq'$ be two continuous, stationary preference relations on $\mathcal{D}$. If $\succeq$ and $\succeq'$ are not identical, then one can find alternatives $d, d' \in \mathcal{D}$ such that $d \succ d'$ and $d' \succ' d$.*

PROOF: A sequence $d \in X^\infty$ is *repeating* if there is some $n \in \mathbb{N}$ and a vector $a \in X^n$ such that $d = (a, a, a, \dots)$. For example, $(x, x', x, x', \dots)$ is repeating with $a := (x, x') \in X^2$. Note that X^{rep} is a subset of $\mathcal{D}$ and, for every $K \in \mathcal{K}(X)$, the intersection $X^{\text{rep}} \cap K^\infty$ is dense in K^∞. Furthermore, because $\succeq$ and $\succeq'$ are continuous, the restrictions of $\succeq$ and $\succeq'$ to X^{rep} are nondegenerate and distinct. Therefore, we can find $d, d' \in X^{\text{rep}}$ such that $d \succ d'$ and $d' \succeq' d$. If $d' \succ' d$, the proof is complete, so suppose that $d' \sim' d$. Because the restriction of $\succeq'$ to X^{rep} is nondegenerate, there is some $d'' \in X^{\text{rep}}$ such that $d'' \nsim' d' \sim' d$. Suppose that $d'' \succ' d' \sim' d$; the opposite case follows from analogous arguments. If $d \succ d''$, the proof is complete, so suppose that $d'' \succeq d \succ d'$. Let $K \in \mathcal{K}(X)$ be such that $d, d', d'' \in K^\infty$. Because d' is repeating, $d' = (a, a, \dots)$ for some $a \in X^n$, $n \in \mathbb{N}$. Define $d^1 := (a, d'')$, $d^2 := (a, a, d'')$, and so on. By construction, the sequence $(d^n)_n$ lies in K^∞ and converges to $d' \in K^\infty$. Since $d \succ d'$, continuity on K^∞ implies that $d \succ d^n$ for all $n \in \mathbb{N}$ large enough. On the other hand, it follows from $d'' \succ' d'$ and the stationarity of $\succeq'$ that

$$d^n = (a, \dots, a, d'') \succ' (a, \dots, a, d') = d' \sim' d \tag{A.7}$$

for every $n \in \mathbb{N}$. Conclude that for n large enough, $d \succ d^n$ and $d^n \succ' d$, as desired. *Q.E.D.*

To summarize, we have shown that every conditional preference $\succeq_{t,\omega}$ has a maxmin representation as in (1.1). In addition, all preferences share the same instantaneous utility function $u: X \to \mathbb{R}$ and the same discount factor $\beta \in (0, 1)$. The proof of Theorem 4 is now completed by mimicking the steps in Epstein and Schneider (2003). The only part of their argument that needs to be modified is Lemma A2. The necessary changes to that lemma can be deduced from the proof of Theorem 1 in this appendix.

REFERENCES

ANSCOMBE, F. J., AND R. AUMANN (1963): "A Definition of Subjective Probability," *The Annals of Mathematical Statistics*, 34, 199–205. [239,240]

BLEICHRODT, H., K. I. ROHDE, AND P. P. WAKKER (2008): "Koopmans' Constant Discounting for Intertemporal Choice: A Simplification and a Generalization," *Journal of Mathematical Psychology*, 52 (6), 341–347. [240]

BORDER, K. C., AND U. SEGAL (1994): "Dynamic Consistency Implies Approximately Expected Utility Preferences," *Journal of Economic Theory*, 63 (2), 170–188. [259]

CERREIA-VIOGLIO, S., F. MACCHERONI, M. MARINACCI, AND L. MONTRUCCHIO (2011): "Uncertainty Averse Preferences," *Journal of Economic Theory*, 146, 1275–1330. [247]

DEBREU, G. (1960): "Topological Methods in Cardinal Utility Theory," in *Mathematical Methods in the Social Sciences*. Stanford: Stanford University Press. [253]

ELLSBERG, D. (1961): "Risk, Ambiguity and the Savage Axioms," *Quarterly Journal of Economics*, 75, 643–669. [240,242]

EPSTEIN, L. G. (1999): "A Definition of Uncertainty Aversion," *Review of Economic Studies*, 66 (3), 579–608. [244]

EPSTEIN, L. G., AND M. SCHNEIDER (2003): "Recursive Multiple-Priors," *Journal of Economic Theory*, 113, 1–31. [248,250,251,259,261]

EPSTEIN, L. G., AND T. WANG (1996): "Beliefs About Beliefs Without Probabilities," *Econometrica*, 64 (6), 1343–1373. [240]

GHIRARDATO, P., AND M. MARINACCI (2002): "Ambiguity Made Precise: A Comparative Foundation," *Journal of Economic Theory*, 102 (2), 251–289. [244]

GILBOA, I., AND D. SCHMEIDLER (1989): "Maxmin Expected Utility With Non-Unique Prior," *Journal of Mathematical Economics*, 18, 141–153. [239,242,243,246,247,257]

GORMAN, W. M. (1968): "The Structure of Utility Functions," *Review of Economic Studies*, 35, 367–390. [253]

HANSEN, L. P., AND T. J. SARGENT (2001): "Robust Control and Model Uncertainty," *American Economic Review*, 91 (2), 60–66. [248]

KANNAI, Y. (1970): "Continuity Properties of the Core of a Market," *Econometrica*, 38 (6), 791–815. [258,259]

KOOPMANS, T. C. (1972): "Representation of Preference Orderings Over Time," in *Decision and Organization*, ed. by C. B. McGuire and R. Radner. Amsterdam: North-Holland, 79–100. [239, 242-245,253]

KREPS, D. (1988): *Notes on the Theory of Choice*. Boulder, CO: Westview Press. [241]

KREPS, D., AND E. PORTEUS (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46 (1), 185–200. [247]

MACCHERONI, F., M. MARINACCI, AND A. RUSTICHINI (2006a): "Ambiguity Aversion, Robustness and the Variational Representation of Preferences," *Econometrica*, 74, 1447–1498. [239, 243,244,248,249,258,259]

——— (2006b): "Dynamic Variational Preferences," *Journal of Economic Theory*, 128 (1), 4–44. [251]

MACHINA, M., AND D. SCHMEIDLER (1992): "A More Robust Definition of Subjective Probability," *Econometrica*, 60 (4), 745–780. [244]

PHELPS, R. R. (1993): *Convex Functions, Monotone Operators and Differentiability*. Berlin: Springer. [258]

RICHARD, S. F. (1975): "Multivariate Risk Aversion, Utility Independence and Separable Utility Functions," *Management Science*, 22, 12–21. [245]

RUDIN, W. (1976): *Principles of Mathematical Analysis*. New York: McGraw-Hill. [252]

SAVAGE, L. (1954): *The Foundations of Statistics*. New York: Wiley. [239]

STRZALECKI, T. (2013): "Temporal Resolution of Uncertainty and Recursive Models of Ambiguity Aversion," *Econometrica*, 81 (3), 1039–1074. [243,247]

Dept. of Economics, University of Rochester, Rochester, NY 14627, U.S.A.; asen.kochov@rochester.edu.

Manuscript received June, 2012; final revision received July, 2014.