Stationary Cardinal Utility

[preliminary and incomplete]

Asen Kochov*

February 9, 2015

Abstract

The paper provides a representation theorem for the class of all stationary preferences in a stochastic environment. A notion of ambiguity aversion applicable to such preferences is also proposed. The analysis helps discriminate between dynamic models of ambiguity aversion and expected utility models with endogenous discounting, which, as has been observed in both applied and decision-theoretic work, share a number of predictions concerning intertemporal behavior.

KEYWORDS: Intertemporal Choice, Stationarity, Endogenous Discounting, Ambiguity

*I thank John Duggan for his help in the proof of Theorem 6. University of Rochester, Rochester, NY. asen.kochov@rochester.edu

1 Introduction

The standard model of intertemporal choice assumes that utility is additively separable across states of the world and over time. The tight structure of the model has proved advantageous in many settings. The preferences are tractable and lead to strong predictions. Yet, an increasing number of empirical findings have proved difficult to rationalize within the standard model and have prompted researchers to explore alternative specifications. Backus et al. [2] provide a comprehensive survey of the literature, discussing specifications that relax each of the separability assumptions characterizing the standard model. The first objective of this paper is to unify some of the different strands in this literature by providing a general representation theorem for the class of all stationary preferences in a stochastic environment. As the paper shows, one reason to focus on stationary preferences is that they are sufficiently well behaved to permit a degree of separation between the individual's 'beliefs' and 'tastes'. Such separation is helpful in many economic problems. It is necessary to isolate the effects of learning on behavior and leads to sharp comparative statics. The second goal of the paper is to clarify how several prominent specifications within the class of stationary preferences differ in their predictions. In particular, the paper proposes a definition of ambiguity aversion that isolates the effects of ambiguity on intertemporal allocations and shows how these effects differ from those of competing models.

The main themes and results in the paper can be illustrated by considering the utility specifications

$$V(c_0, c_1, ...) = \mathbb{E}[u(c_0) + b(c_0)u(c_1) + b(c_0)b(c_1)u(c_2)...] \tag{1.1}$$

$$V(c_0, c_1, ...) = \min_{p\in\mathcal{P}} \mathbb{E}_p[\textstyle\sum_t \beta^t u(c_t)] \tag{1.2}$$

where $(c_0, c_1, ...)$ is a stochastic consumption stream. The first model relaxes time additivity by introducing a specific form of intertemporal complementarity: the rate of time preference between two consecutive periods t and $t+1$ depends on consumption in period t. It is common to say that discounting is endogenous. In a deterministic setting, a special case of that model was introduced by Uzawa [35]. The more general model in (1.1), including the extension to a stochastic framework, is due to Epstein [5]. Their work has spawned many applications in international economics and the study of small open economies in particular. See Epstein [6] and Epstein and Hynes [8], among others. In the second model in (1.2), the intertemporal utility index is additive and discounting takes the familiar geometric form. Its characteristic feature is that there is ambiguity, that is, the individual cannot quantify the relevant uncertainty by a single prior belief. Instead, he contemplates a set $\mathcal{P}$ of possible probability distributions and evaluates each consumption profile according to a worst case scenario. As has been understood since the work of Ellsberg [4] and Gilboa and Schmeidler [17], the specification in (1.2) departs from the standard model by relaxing the assumption that utility is additively separable across states of the world. The reader is referred to Epstein and Schneider [10] for a survey of the growing literature on ambiguity and some of its recent applications to problems in finance.

Work in decision theory has revealed that the models in (1.1) and (1.2), despite the obvious differences, share a number of important predictions concerning intertemporal behavior. The first such prediction extends Koopmans' [22] notion of stationarity to stochastic settings. Such an extension was originally proposed in Epstein [5], where it is used to characterize the model in (1.1). More recently, Kochov [21] showed that stationarity is one of the key predictions of the maxmin model in (1.2) that distinguish it from other models of ambiguity aversion. In addition to what is already assumed in Koopmans [22], the extension requires that the individual's attitude toward uncertainty does not depend on the date on which consumption takes place. In particular, it does not matter whether a given event affects immediate or future consumption. Accounting for the passage of time, the uncertainty exacts identical premia.

The second prediction shared by the models in (1.1) and (1.2) concerns a form of behavior called intertemporal hedging in Kochov [21].[1] Its intuitive meaning is that an individual who is concerned about uncertainty would seek to take different, negatively correlated bets in different time periods. Kochov [21] showed that intertemporal hedging can be viewed as the dynamic manifestation of ambiguity aversion. A limitation of that paper is that its conclusions depend critically on the assumption that the utility index is additively separable across time. In particular, it is observed in Epstein [5] that the model in (1.1) generates similar behavior whenever the discount factor $\beta(c)$ is decreasing in the level of consumption. This restriction on the rate of time preference is especially common in applications where it is used to guarantee the uniqueness and stability of steady states. The question arises whether the effects of ambiguity on intertemporal behavior can be successfully

[1] Conceptually similar conditions were formulated by Richard [30] and Epstein and Tanny [11] where they are called multivariate risk aversion and correlation aversion respectively. These papers consider environments in which the uncertainty is objective as in the von Neumann Morgenstern expected utility framework. Kochov [21] adopts an environment in which the uncertainty is subjective as in Savage [32].

disentangled from those of endogenous discounting and what kind of data or choice situations may serve this purpose. Before we address this question, it should be noted that its broader outlines have not escaped the attention of applied economists. In their comments to Backus et al. [2], both Hansen and Werning observe that it is not uncommon for different alternatives to the standard model to generate similar or, in some cases, identical predictions especially when the focus is on intertemporal behavior. Werning, in particular, calls for more work aimed at discriminating among the many alternatives reviewed in Backus et al. [2]. In what follows, we explain how stationarity, as well as the new notion of intertemporal hedging, may be helpful in this regard.

The first goal of this paper is to characterize the class of all stationary preferences in a stochastic framework. This is done without imposing any a priori restrictions how the individual may evaluate uncertainty or how he discounts the future. The analysis provides a stochastic counterpart of Koopmans [22] and generalizes the results in Epstein [5] and Kochov [21]. Two practical lessons emerge. Both of them stem from the fact that stationarity proves to be remarkably powerful in stochastic settings. First, the axiom implies that discounting takes the form in (1.1). This implication stands in sharp contrast to the conclusions reached by Koopmans [22] in a deterministic environment. As we explain, in such settings stationarity is consistent with a much broader class of preferences, permitting a greater and more nuanced array of intertemporal complementarities. It is also notable that, despite the richness of the preferences characterized by Koopmans [22], there has been little work, axiomatic or empirical, attempting to discriminate among them. It follows from the analysis in this paper that embedding these preferences in a stochastic framework may be helpful. A closely related lesson stems from the fact that many of the early applications that followed in Koopmans' path, e.g., Lucas and Stokey [24], adopt specifications which, when combined with the expected utility or maxmin criterion, fail to be stationary in the stochastic sense of this paper. To the extent that one considers applications in which the latter restriction on behavior may be desirable, care is therefore needed in making the transition from deterministic to uncertain environments.

The second implication of stationarity is that uncertainty is evaluated by a criterion that is 'almost' maxmin. To clarify, it is helpful to recall that the maxmin criterion in (1.2) is often viewed as exhibiting both constant absolute and constant relative ambiguity attitude. These notions, which we formalize in the paper, are the obvious analogues of the attitudes an individual may hold toward objective risk. For further discussion of this analogy, see Klibanoff et al. [19] and Strzalecki [34]. Stationarity implies that the individual evaluates uncertainty using a criterion that exhibits both constant absolute and constant relative ambiguity attitude. What is once again notable is that only a few ambiguity criteria fall into the same category. In addition to the maxmin criterion used in (1.2), popular alternatives include the Choquet criterion of Schmeidler [33] and the α-Hurwicz criterion recently axiomatized by Gul and Pesendorfer [18]. Within this class, the maxmin criterion is also unique in that every specification thereof, with the exception of expected utility, exhibits strict aversion to ambiguity. Before we explain what 'ambiguity aversion' means in the intertemporal setting of this paper, it may be helpful to summarize the implications of stationarity. In particular, taking into account its implications for discounting and for the way the individual evaluates uncertainty, we reach the surprising conclusion that the utility specifications in (1.1) and (1.2) isolate the essential features of all stationary preferences. Roughly speaking, these specifications may be viewed as two polar departures from the standard model, nesting every other stationary preference inbetween.

The second objective of this paper is to identify intertemporal behavior that can discriminate between the preference specifications in (1.1) and (1.2). In particular, we seek to modify the notion of intertemporal hedging used in Kochov [21] and isolate the demand for hedging that is driven purely by ambiguity aversion. As a preliminary step in this direction, we prove that the class of all stationary preferences is sufficiently well behaved to permit a degree of separation between the individual's 'tastes' and his 'beliefs'. Such a step is necessary to insure that the problem how to define ambiguity aversion, and how to discriminate between the models in (1.1) and (1.2), is feasible in theory. The specific solution we then propose is based on a straight-forward distinction between two forms of uncertainty that can arise in dynamic settings. To illustrate the main idea, it is helpful to focus on the special case whereby the consumption space consists of a single, infinitely divisible good. By way of contrast, consider first a static environment in which consumption takes place at a single point in time. It is then evident that uncertainty can only affect the level of consumption, that is, *how much* is consumed in each state of the world. In a dynamic environment, one can imagine a different form of uncertainty, namely, one concerning *when* the individual obtains a given, fixed level of consumption - think of an individual expecting a tax refund but not knowing when it will arrive. The paper shows that it is the latter type of uncertainty that brings out a sharp, qualitative difference between the models in (1.1) and (1.2). In particular, if the individual enters an intertemporal hedge to reduce or eliminate such uncertainty, it is necessarily because he perceives the environment to be ambiguous. The model in (1.1) cannot explain such behavior, no matter how the individual discounts the future. Formally, the paper proves that a maxmin criterion, like the one in (1.2), is implied once the new notion of intertemporal hedging is combined with the assumption that behavior is stationary.

2 Domain

Time is discrete and varies over an infinite horizon: $t \in \{0, 1, 2, ...\} =: T$. The available information 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 $\mathcal{F}_0 = \{\Omega, \varnothing\}$. Let X be a compact, connected, and separable topological space, interpreted as the set of consumption outcomes. Let h be 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 $t \in T$. Think of h as a stochastic consumption stream. Following Savage [32], we will also refer to a process h as an **act**. If $\mathcal{F}'$ is a collection of events such that $\mathcal{F}' \subset \cup_t \mathcal{F}_t$, then an act h is $\mathcal{F}'$**-adapted** if h_t is $\mathcal{F}'$-measurable for every $t \in T$. An act h is **finite** if there is a finite algebra $\mathcal{A} \subset \cup_t \mathcal{F}_t$ such that h is $\mathcal{A}$-adapted. To avoid technical complications, we take the choice domain to be the space $\mathcal{H}$ of all finite acts. 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 the outcomes of h do not depend on the state of the world. We use d, d' to denote such acts. As is common in the literature, such acts are identified with elements of X^∞.

We now take the opportunity to introduce some notation and a few mathematical concepts that will be used in the rest of the paper. Let B^0 be the space of all simple, real valued, $\cup_t \mathcal{F}_t$-measurable functions on Ω. Given a set $C \subset \mathbb{R}$, let B^0_C denote the set of all C-valued functions in B^0. To highlight the fact that Ω is a state space capturing the relevant uncertainty, we often refer to the functions $\xi \in B^0$ as random variables. We abuse notation and use k to denote both a real number and the function in B^0 that is identically equal to $k \in \mathbb{R}$. With this in mind, a functional $I : B^0 \to \mathbb{R}$ is **translation invariant** if $I(\xi + k) = I(\xi) + k$ for all $\xi \in B^0, k \in \mathbb{R}$. It is **normalized** if $I(k) = k$ for all $k \in \mathbb{R}$. Given $\alpha \in \mathbb{R}$, the functional I is α**-homogeneous** if $I(\alpha\xi) = \alpha I(\xi)$ for all $\xi \in B^0$. If I is α-homogeneous for all $\alpha \in \mathbb{R}_{++}$, then I is **positively homogeneous**. If we endow B^0 with the usual pointwise order, a functional $I : B^0 \to \mathbb{R}$ is **increasing** if for all $\xi, \xi' \in B^0$, $I(\xi) \geq I(\xi')$ whenever $\xi \geq \xi'$. A normalized and increasing functional $I : B^0 \to \mathbb{R}$ is called a **certainty equivalent**. In later sections, certainty equivalents will be used to specify the individual's 'beliefs'.

Thus, $I(\xi)$ will be interpreted as the expected value assigned by the individual to the random variable $\xi \in B^0$. Going back to the introduction, we should also observe that positive homogeneity and translation invariance are often interpreted as capturing constant relative and constant absolute ambiguity attitude. Finally, let Δ be the space of all finitely additive probability measures p on the measurable space $(\Omega, \cup_t \mathcal{F}_t)$. The space Δ is endowed with the weak* topology, that is, the coarsest topology such that for every $\xi \in B^0$, the linear functional $p \mapsto \mathbb{E}_p \xi$, mapping Δ into the reals, is continuous.

3 Axioms

Given a binary relation $\succeq$ on a set Y, the relations $\succ$ and $\sim$ are defined as usual. A **preference relation** $\succeq$ on a set Y is a complete and transitive binary relation such that $y \succ y'$ for some $y, y' \in Y$. The primitive of this paper is a preference relation on the space $\mathcal{H}$ of finite acts, representing behavior prior to the resolution of any uncertainty. The next axiom is a form of continuity familiar from Ghirardato and Marinacci [14].

Continuity (C): For every finite algebra $\mathcal{F}' \subset \cup_t \mathcal{F}_t$, the restriction of $\succeq$ to all $\mathcal{F}'$-adapted acts in $\mathcal{H}$ is continuous.

The next axiom requires that the tastes of the individual are state independent, that is, there are no taste shocks. The requirement goes back to the work of Savage [32]. Here, it is formulated for a dynamic choice setting. As is well understood, the importance of state independence stems from the fact that it is necessary for the separation of beliefs and tastes, a problem that is central to the analysis in this paper. Given $h \in \mathcal{H}, d \in X^\infty, t \in T$, and an event $A \in \mathcal{F}_t$, let dA^th denote the act $g \in \mathcal{H}$ such that $g_k(\omega) = d_k$ for all $k \geq t$ and all $\omega \in A$, and $g_k(\omega') = h_k(\omega')$ otherwise. Thus, dA^th is the act obtained from h by replacing its outcomes in the event A and after period t with the respective outcomes of d. Suppose now that the acts $h \in \mathcal{H}$ and $d, d' \in X^\infty$ are such that $h_k = d_k = d'_k$ for all $k \leq t-1$. The axiom requires that $dA^th \succeq d'A^th$ whenever $d \succeq d'$. Thus, evaluating d, d' conditional of the event A does not reverse their unconditional ranking. Note that this requirement does not preclude the possibility that $dA^th \sim d'A^th$ while $d \succ d'$. Such rankings can arise if the individual believes the event A to be impossible. In the maxmin model in (1.2), such rankings can also arise if the event is deemed possible but is assigned probability

0 under the worst case scenario. As other authors have recognized, in the study of uncertainty it is helpful to assume that there is at least one event A for which such possibilities do not apply, that is, for which the likelihood of both A and A^c is bounded away from zero. Formally, say that an event $A \in \mathcal{F}_t$, $t \in T$, is **essential** if $dA^t h \succ d'A^t h$ and $d[A^c]^t h \succ d'[A^c]^t h$ whenever $d \succ d'$. As part of the statement of state independence, we embed the requirement that there is at least one essential event.

State Independence (SI): For all $t \in T, A \in \mathcal{F}_t$, and acts $h \in \mathcal{H}, d, d' \in X^\infty$ such that $h_k = d_k = d'_k$ for all $k \leq t-1$, if $d \succeq d'$, then $dA^t h \succeq d'A^t h$. In addition, there is some $t \in T$ and some event $A \in \cup_t \mathcal{F}_t$ such that A is essential.

The next axiom is central to the analysis of this paper. It extends Koopmans' [22] notion of stationarity from deterministic to stochastic environments. Such an extension was first proposed in Epstein [5] where it was used to characterize the model in (1.1). More recently, Kochov [21] used the axiom to characterize the maxmin model in (1.2). In terms of choice under uncertainty, the primary implication of the axiom is that attitudes toward uncertainty do not depend on the date on which consumption takes place. To illustrate, fix some period $t \in T$ and an event $A \in \mathcal{F}_t$. Observe that the event A may affect consumption in period t as well as consumption at a more distant future date. Stationarity requires that, no matter how distant the ramifications, the individual exhibits the same degree of uncertainty aversion. To state the axiom formally, let (x,h) denote the act $g \in \mathcal{H}$ such that $g_0 = x$ and $g_t = h_{t-1}$ for all $t > 0$, where $x \in X$ and $h \in \mathcal{H}$. Thus, the act (x,h) postpones the consumption date of each outcome of h by one period. The axiom says that this has no effect on preferences.[2]

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

To simplify the exposition, from now on we say that a preference relation $\succeq$ on $\mathcal{H}$ is **stationary** if it satisfies the three axioms above. Similarly, we speak of a stationary preference on X^∞ to mean a continuous preference relation on X^∞ such that $d \succeq d'$ if and only if $(x,d) \succeq (x,d')$, for all $x \in X$ and $d, d' \in X^\infty$. Here, by continuous we mean simply that all the upper and lower contour sets are closed in the product topology on X^∞.

[2]For a more elaborate discussion of the axiom, see Kochov [21].

For future reference, we introduce two additional axioms that can be deduced as implications of Stationarity and State Independence. First, a preference relation $\succeq$ on $\mathcal{H}$ is said to satisfy **History Independence** if for all $x, y \in X, h, g \in \mathcal{H}$, the act (x, h) is preferred to (x, g) whenever (y, h) is preferred to (y, g). Thus, prior consumption does not affect how the individual evaluates the future. This axiom follows directly from Stationarity. Next, given an act $h \in \mathcal{H}$ and state $\omega \in \Omega$, let $h(\omega) := (h_0(\omega)h_1(\omega), ...) \in X^\infty$ be the consumption stream delivered by h in state ω. Say that $\succeq$ satisfies **Monotonicity** if an act $h \in \mathcal{H}$ is preferred to an act $g \in \mathcal{H}$ whenever $h(\omega)$ is preferred to $g(\omega)$ for all $\omega \in \Omega$. As one can verify, Monotonicity is stronger than State Independence, modulo the existence of an essential event. Lemma 7 in the appendix shows, however, that the axioms become equivalent under Stationarity. The lemma is of interest beyond the role it plays in the proof of Theorem 2. In particular, it is known that only a few specifications of the recursive preferences studied in Epstein and Zin [12] satisfy Monotonicity, whereas all such preferences can be formulated so as to be state independent. See Bommier and Grand [3] for details. It follows from Lemma 7 that the specifications in Epstein and Zin [12] that are not monotone are not stationary either and, therefore, fall outside the scope of this paper.

4 Representation

This section states two results concerning stationary preferences in a stochastic setting. The first reveals a feature that is shared by all utility representations of a single, stationary preference relation. As we explain, the result is helpful in verifying whether a given parametric specification induces a stationarity preference. The second result derives one specific utility representation. As we argue in Section 5, the latter is important in that it achieves a separation of the individual's tastes and beliefs.

It is helpful to begin by considering the implications of Stationarity for the ranking of deterministic consumption streams. In particular, suppose $U : X^\infty \to \mathbb{R}$ is a utility function for some preference on $\succeq$ on X^∞. Arguments familiar from Koopmans [22] show that $\succeq$ is stationary if and only if U can be written in the following recursive form[3]

$$U(x_0, x_1, x_2, \ldots) = \phi(x_0, U(x_1, x_2, \ldots)), \quad \forall (x_0, x_1, \ldots) \in X^\infty, \tag{4.1}$$

where $\phi : X \times U(X^\infty) \to U(X^\infty)$ is continuous in each argument and strictly increasing in the second. The function ϕ is commonly referred to as a **time aggregator** for the utility $U : X^\infty \to \mathbb{R}$. The name is motivated by the fact that ϕ computes overall, ex ante utility as an average of current consumption and the overall, continuation utility.

Turning attention to choice under uncertainty, suppose that a utility function $U : X^\infty \to \mathbb{R}$ representing the ranking of deterministic acts is given. In later sections, we refer to U as capturing the tastes of the individual. Given U, a stochastic act $h \in \mathcal{H}$ can be converted into a random variable in B^0 by assigning each state $\omega \in \Omega$ the utility of the consumption stream $h(\omega) \in X^\infty$. This random variable is denoted by $U \circ h \in B^0$. The choice of notation is inspired by the fact that one can think of an act $h \in \mathcal{H}$ as a function from Ω into X^∞. Similarly, $U \circ \mathcal{H}$ denotes the subset of all random variables $U \circ h \in B^0$ as the act h varies in $\mathcal{H}$. Given $U : X^\infty \to \mathbb{R}$ and a certainty equivalent $I : B^0 \to \mathbb{R}$, we can now define a utility function V on $\mathcal{H}$ by letting $V(h) := I(U \circ h)$ for every $h \in \mathcal{H}$. A tuple (U, I) is said to **represent** a preference relation $\succeq$ on $\mathcal{H}$ if the function V, thus defined, represents $\succeq$. Whenever the function $U : X^\infty \to \mathbb{R}$ can be written recursively as in (4.1), it is also useful to make the time aggregator ϕ explicit and write (U, ϕ, I) in place of (U, I). It should now be mentioned that a representation (U, ϕ, I) exists under very general restrictions on behavior. The proof of Lemma 1 in the appendix provides the details. The important point is that we have not yet accounted for the implications of Stationarity that concern the ranking of stochastic acts. To do so, we need some notation. Given a representation (U, ϕ, I) and a random variable $\xi \in U \circ \mathcal{H} \subset B^0$, let $\phi(x, \xi)$ denote the function $\omega \mapsto \phi(x, \xi(\omega))$. Observe that $\phi(x, \xi)$ is a random variable, that is, $\phi(x, \xi) \in B^0$. Now, take an act $(x, h) \in \mathcal{H}$, as such arise in the statement of Stationarity, and consider the equalities:

$$V(x, h) = I(\phi(x, U \circ h)) = \phi(x, I(U \circ h)), \quad \forall x \in X, h \in \mathcal{H}. \tag{4.2}$$

The first equality repeats the definition of V. The interesting equality is the second one. Recall from the discussion of Stationarity that postponing the outcomes of h by one period, while inserting x in period $t = 0$, does not affect how the individual feels about the uncertainty inherent in an act h. This invariance is captured in (4.2) by

[3] Koopmans [22] imposes the additional assumption that consumption in the first period is separable from the future. It is clear however that many of his results generalize.

the fact that the certainty equivalent I and the time aggregator ϕ **permute**. That is, there are two equivalent ways to compute the utility of an act (x, h). As a preliminary observation, note that when we consider the act (x, h), the random variable $U \circ h$ represents how the continuation utility of the act varies with the state of the world. Informally, the expression $I(\phi(x, U \circ h))$ thus means that we first aggregate utility across time, and then across states. Conversely, the expression, $\phi(x, I(U \circ h))$ means that we first compute the expectation of future utility, and then aggregate across time. The next lemma shows that the equivalence in (4.2) exhausts the implications of Stationarity.

Lemma 1 *Suppose (U, ϕ, I) is a representation for a preference relation $\geq$ on $\mathcal{H}$. Then, $\geq$ is stationary if and only if the time aggregator ϕ and the certainty equivalent I permute.*

As we explain at the end of this section, Lemma 1 is an important stepping stone in the proof of our main theorem. The lemma is of independent interest as well. To elaborate, observe that in Koopmans [22], and in axiomatic work more generally, the utility function $U : X^\infty \to \mathbb{R}$ is constructed first on the basis of the observed preference relation. A time aggregator ϕ for U is then defined via (4.1). In applications, preferences are typically deduced from a given representation. It is then convenient to reverse the way in which equation (4.1) is utilized. Namely, it becomes convenient to first specify a function $\phi : X \times \mathbb{R} \to \mathbb{R}$ and a certainty equivalent I. A utility function $U : X^\infty \to \mathbb{R}$ is then defined as the solution of the recursive equation in (4.1).[4] The convenience of this approach, which was pioneered by Lucas and Stokey [24], stems from the fact that a function ϕ on $X \times \mathbb{R}$ is a much simpler mathematical object than a function U on X^∞. In particular, ϕ is easier to specify. The price one has to pay is that many properties of behavior that arise in axiomatic work become difficult to verify. For example, in Section 6 we discuss an axiom from Koopmans et al. [23], which we call Impatience and which expresses a preference to advance the consumption of desirable outcomes. Despite being a conceptually simple restriction on behavior, the exact class of time aggregators ϕ that are consistent with the axiom is yet undetermined. The source of the problem is that for many ϕ there is no closed

[4] For some $\phi : X \times \mathbb{R} \to \mathbb{R}$, there may be no function $U : X^\infty \to \mathbb{R}$ that solves the recursive equation in (4.2). Another problem is that the equation may admit multiple solutions, in which case preferences on X^∞ are not well defined. We abstract from such problems since they do not concern the formal results in this paper. It should also be noted that such problems do not arise for the specific time aggregators characterized in Theorem 2. For the latter, there is in fact a closed form solution for the unique function U on X^∞ that solves the recursion in (4.1). It is given by the expression in (1.1).

form solution for the utility function U that solves (4.1). As a result, preferences on X^∞ are only implicitly defined and their properties difficult to assess. One lesson from Lemma 1 is that such a problem does not arise when one considers Stationarity. Given a function, $\phi : X \times \mathbb{R} \to \mathbb{R}$ and a certainty equivalent $I : B^0 \to \mathbb{R}$, one can confirm whether $I(\phi(x,\xi)) = \phi(x, I(\xi))$ for all $x \in X, \xi \in B^0$ without having to solve the recursion in (4.1). In particular, if I and ϕ permute and if we define U via (4.1), the preference relation on $\mathcal{H}$ induced by (U, I) is guaranteed to be stationary. The ease with which Stationarity can be verified is especially relevant given the fact that the literature on endogenous discounting and the literature on non-expected utility preferences have so far proceeded in isolation of one another. In particular, there is little knowledge which of the existing specifications of ϕ and I permute in the sense of Lemma 1.

To develop further intuition about Lemma 1, it is helpful to recall the preference specifications from the introduction and verify that the respective certainty equivalents and time aggregators permute. These observations would also provide perspective into the main theorem of this section. Consider first the model with endogenous discounting in (1.1). Its time aggregator ϕ and its certainty equivalent I take the form

$$\phi(x,k) = u(x) + b(x)k \quad \forall x \in X, k \in U(X^\infty), \quad \text{and} \quad I(\xi) = \mathbb{E}_p\xi \quad \forall \xi \in B^0, \tag{4.3}$$

where $u : X \to \mathbb{R}$ and $b : X \to (0,1)$ are continuous functions and $p \in \Delta$ is a probability measure over the states of the world. Observe that ϕ is linear in its second argument, that is, in the continuation utility $k \in U(X^\infty)$. Taking into account the familiar linearity properties of the expectation operator, it follows immediately that I and ϕ permute. A less obvious observation is that ϕ has to be linear in its second argument for it to permute with the expectation operator. This allows us to deduce Epstein's [5] characterization of the model in (4.3) from Lemma 1: It is enough to recall that Epstein maintains Stationarity and the expected utility hypothesis. When a function $U : X^\infty \to \mathbb{R}$ admits a time aggergator ϕ as in (4.3), we call U an **Uzawa-Epstein utility** and ϕ its **Uzawa-Epstein aggregator**. Such an aggregator will also be written as (u,b) where u and b are the functions that define ϕ as in (4.3). As part of the definition of an Uzawa-Epstein aggregator (u,b), it is also convenient to require that the function $x \mapsto u(x)(1 - b(x))^{-1}$, from X into the reals, is nonconstant. This requirement is necessary and sufficient for the function U on X^∞, defined via (4.1), to be nonconstant.

Next, consider the maxmin model in (4.1). The time aggregator and the certainty equivalent take the form

$$\phi(x,k) = u(x) + \beta k \quad \forall x \in X, k \in U(X^\infty), \quad \text{and} \quad I(\xi) = \min_{p\in\mathcal{P}} \mathbb{E}_p \xi \quad \forall \xi \in B^0, \qquad (4.4)$$

where $\beta \in (0,1)$ and $\mathcal{P} \subset \Delta$ is a weak*-closed, convex set of probability measures. When a functional $I : B^0 \to \mathbb{R}$ takes the form in (4.4) for some set $\mathcal{P}$, we refer to I as a **maxmin certainty equivalent**. Once again, a direct calculation confirms that ϕ and I permute in the sense of (4.2). In anticipation of Theorem 2, however, we should stress that the calculation exploits two well known properties of maxmin certainty equivalents, namely, each such functional I is positively homogeneous and translation invariant.

The next theorem is the main result of this paper. It shows that any stationary preference on $\mathcal{H}$ has a representation (U,ϕ,I) where ϕ is an Uzawa-Epstein aggregator and I shares the two features of maxmin certainty equivalents just noted. To state the theorem, say that a certainty equivalent I is **regular** if there is some event $A \in \cup_t \mathcal{F}_t$ such that I is strictly increasing when it is restricted to the subspace of functions $\xi \in B^0$ that are $\{A, A^c\}$-measurable. Regular certainty equivalents arise whenever there is an essential event A, which is one of the requirements of State Independence.

Theorem 2 *A preference relation $\succeq$ on $\mathcal{H}$ satisfies Continuity, State Independence, and Stationarity if and only if it has a representation (U,ϕ,I) such that ϕ is an Uzawa-Epstein time aggregator for some continuous functions $u : X \to \mathbb{R}$ and $b : X \to (0,1)$, and the certainty equivalent I is regular, translation invariant and $b(x)$-homogeneous for every $x \in X$. Furthermore, when b is not a constant function, I is positively homogeneous.*

It was observed in the introduction that Stationarity implies a certainty equivalent that is 'almost' maxmin. We can now make this claim precise. Specifically, it is known from Gilboa and Schmeidler [17] that a certainty equivalent I takes the maxmin form if and only if it is translation invariant, positively homogeneous and quasiconcave. Theorem 2 shows that Stationarity delivers the first two of these properties. In the decision-theoretic literature, the last property, quasiconcavity, is often associated with ambiguity aversion. For example, the behavioral definitions of ambiguity aversion introduced in Gilboa and Schmeidler [17] and Ghirardato et al. [15] are each necessary and sufficient for the certainty equivalent to be quasiconcave. In Section 6, we will propose a different behavioral definition of ambiguity aversion that is based on intertemporal choice and is applicable to all stationary preferences. We will then show that any stationary, ambiguity averse preference induces a certainty equivalent that I is also quasiconcave so that, indeed, I is a maxmin certainty equivalent.

Extending Koopmans' analysis and his notion of stationarity to stochastic settings proves to be powerful in another sense. In particular, the majority of stationary preferences on X^∞, that is, in a setting with no uncertainty, do not admit an Uzawa-Epstein utility. A discussion of this point is deferred until Section 6, when we discuss Impatience.

5 Uniqueness

This section investigates the uniqueness of the representation derived in Theorem 2. A second goal is to show that Stationarity imposes enough structure on preferences to permit the separation of beliefs and tastes. We begin the discussion with a result concerning the uniqueness of Uzawa-Epstein representations in settings without uncertainty.

Theorem 3 *Suppose a preference relation $\succeq$ on X^∞ has two Uzawa-Epstein utility functions $U, \hat{U} : X^\infty \to \mathbb{R}$ with time aggregators (u, b) and $(\hat{u}, \hat{b})$ respectively. Then, $b = \hat{b}$ and $U = \alpha \hat{U} + \gamma$ for some $\alpha \in \mathbb{R}_{++}, \gamma \in \mathbb{R}$.*

Fix a stationary preference relation $\succeq$ on X^∞ and let $\mathcal{U}$ be the space of all continuous functions $U : X^\infty \to \mathbb{R}$ that represent $\succeq$. Also, let $\mathcal{U}^* \subset \mathcal{U}$ be the subset of functions that admit an Uzawa-Epstein aggregator. To understand the rest of this section, it is important to observe that $\mathcal{U}^*$ is necessarily a strict subset of $\mathcal{U}$. In particular, suppose $\mathcal{U}^*$ is nonempty. Take some $U \in \mathcal{U}^*$ and a continuous, strictly increasing, non-affine function $g : \mathbb{R} \to \mathbb{R}$. If we let $\tilde{U} := g \circ U$, it is clear that $\tilde{U} \in \mathcal{U}$. Yet, Theorem 3 implies that $\tilde{U} \notin \mathcal{U}^*$. The question arises why one should single out Uzawa-Epstein representations as we did in Theorems 2 and 3. To further stress the importance of the question, note that each function $U' \in \mathcal{U}$ can be written recursively, as in (4.1), for some time aggregator ϕ. Thus, the familiar advantages of recursive representations do not allow us to discriminate in favor of Uzawa-Epstein utilities. We give two reasons.

First, the uniqueness result in Theorem 3 suggests that $b(x)$ has a well defined behavioral meaning. In fact, an analogy with the standard, time separable model suggests that $b(x)$ may be viewed as a local measure of impatience. Following Epstein [5] and Koopmans [22], the analogy can be made precise if we assume that $X \subset \mathbb{R}$

and that the function $U : X^\infty \to \mathbb{R}$ is differentiable. In particular, letting $\rho(x) := (1 - b(x))b(x)^{-1}$ for every $x \in X$, Epstein [5, p.137] observes that $\rho(x) + 1 = b(x)^{-1}$ is the marginal rate of substitution between consumption in periods 0 and 1 along the constant path $(x, x, ...)$. When consumption is in a neighborhood of $(x, x, ...)$, we can thus think of $\rho(x)$ as the rate of time preference and of $b(x)$ as the discount factor. One advantage of Uzawa-Epstein representations is now obvious: they incorporate the discount factor b explicitly into the functional form. The additive structure of the time aggregator and the close resemblance to the standard model is also beneficial analytically.

As we consider choice under uncertainty, a second reason emerges why a focus on Uzawa-Epstein representations is justified. It may be helpful to illustrate the essential points by first focusing on the model with endogenous discounting due to Epstein [5]. Thus, consider a preference relation $\succeq$ on $\mathcal{H}$ that admits a representation as in (1.1). Note that the certainty equivalent $I(\xi) = \mathbb{E}_p\xi$ is uniquely defined by the subjective belief $p \in \Delta$. Conversely, we can recover the belief $p \in \Delta$ from knowledge of I. There are of course other representations (U', I') for the same preference relation. If $g : \mathbb{R} \to \mathbb{R}$ is a continuous, strictly increasing function, it is enough to let $U' := g \circ U$ and $I'(\xi) := g(I(g^{-1} \circ \xi))$ for all $\xi \in B^0$. What makes the resulting representation (U', I') less satisfactory is that the tight connection between the certainty equivalent I and the subjective belief $p \in \Delta$ is not preserved. Specifically, knowledge of I' may be insufficient to identify the belief p. In the study of expected utility preferences, the standard response is to restrict attention to representations (U', I) where U' can vary but the certainty equivalent $I(\xi) = \mathbb{E}_p\xi$, $\xi \in B^0$, is kept fixed as an explicit representation of the individual's beliefs. What we have to stress is that U' cannot vary freely once I is fixed in this manner. In particular, we cannot change the 'curvature' of the utility index U'. If, for example, we were to make U' 'more concave' by applying a concave transformation $g : \mathbb{R} \to \mathbb{R}$, we would make the individual more risk averse, changing the entire preference relation. In the rest of this section, we consider arbitrary stationary preferences on $\mathcal{H}$ and try to identify a class of representations that shares the same features as expected utility: the utility index U' is free to vary up to positive affine transformations and the same certainty equivalent I works for all U' within the class. We show that such a class exists for any stationary preference. Moreover, it is unique, and it coincides with the representation derived in Theorem 2.

Let $\succeq$ be a stationary preference relation on $\mathcal{H}$. Since $\succeq$ induces a stationary preference on X^∞, we can define the sets $\mathcal{U}$ and $\mathcal{U}^*$ as above. If $\mathcal{U}' \subset \mathcal{U}$, say that $\mathcal{U}'$ **achieves a separation of tastes and beliefs** if for all representations (U, I) and (U', I')

of $\succeq$ such that $U, U' \in \mathcal{U}'$, we have $I = I'$. Our task is to identify sets $\mathcal{U}'$ that posses this property. Based on the discussion of expected utility preferences, we are led to consider sets $\mathcal{U}'$ such that any two functions $U, U' \in \mathcal{U}'$ are positive affine transformations of one another. Formally, say that a set $\mathcal{U}' \subset \mathcal{U}$ is **cardinal** if for any two functions $U, U' \in \mathcal{U}'$, there are $\alpha \in \mathbb{R}_{++}$ and $\gamma \in \mathbb{R}$ such that $U = \alpha U' + \gamma$. Once we restrict attention to cardinal sets, the obvious goal is to search for sets $\mathcal{U}'$ that are rich. Given the importance of recursive representations such as the one in (4.1), it is natural to require that the set $\mathcal{U}'$ inherits the recursive structure of the utility functions it contains. Formally, given a function $U \in \mathcal{U}$ and any $x \in X$, define the function U_x on X^∞ by letting $U_x(d) := U(x, d)$ for all $d \in X^\infty$. Say that the set $\mathcal{U}' \subset \mathcal{U}$ is **recursive** if for every $x \in X$ and $U \in \mathcal{U}'$, we have $U_x \in \mathcal{U}'$. To summarize the discussion so far, we have argued that to achieve a separation of beliefs and tastes, it is reasonable to restrict attention to cardinal, recursive sets $\mathcal{U}'$. Observe now that the set $\mathcal{U}^*$ is recursive by construction. By Theorem 3, $\mathcal{U}^*$ is also cardinal. Hence, $\mathcal{U}^*$ is one potential candidate for the separation of beliefs of tastes. The question remains: Are there other recursive, cardinal sets $\mathcal{U}$? Remarkably, the answer is no. Even though the arguments are formulated differently, this is proved in Epstein [5]. Specifically, it follows from the proof of Theorem 1 in Epstein [5] that if $\mathcal{U}'$ is a recursive, cardinal set, then $\mathcal{U}' \subset \mathcal{U}^*$. We collect these observations in the next corollary.

Corollary 4 *Let $\mathcal{U}' \subset \mathcal{U}$ be a nonempty, recursive, cardinal set of utility functions for some stationary preference relation on X^∞. Then, $\mathcal{U}' \subset \mathcal{U}^*$. In particular, $\succeq$ admits an Uzawa-Epstein utility, that is, $\mathcal{U}^* \neq \varnothing$. The set $\mathcal{U}^*$ is the largest cardinal, recursive set.*

So far we have argued that to attain a separation of beliefs and tastes, it is reasonable to restrict attention to representations (U, I) such that U has the Uzawa-Epstein form. The next result, which is an immediate corollary of Theorems 2 and 3, confirms that these representations do in fact attain the desired separation of beliefs and tastes.

Corollary 5 *Suppose a stationary preference relation $\succeq$ on $\mathcal{H}$ has two representations (U, ϕ, I) and $(\hat{U}, \hat{\phi}, \hat{I})$ such that $U, \hat{U} \in \mathcal{U}^*$ and $I, \hat{I}$ have the properties outlined in Theorem 2. Let $b : X \to (0, 1)$ be the associated discount factor which, by Theorem 3, does not depend the chosen representations. If b is a nonconstant function, then $I = \hat{I}$.*

The corollary does not guarantee the uniqueness of the certainty equivalent I in the special case when b is a constant function, that is, when the rate of time preference is fixed. Somewhat curiously, the standard model of time preference proves to be a strait jacket in the elicitation of beliefs. This gap in the conclusions of Corollary 5 will be closed once we impose further structure on preferences. See Section 6 for details.

It should be acknowledged that the analysis of this section is incomplete in that we do not give an explicit behavioral sense in which I represents the 'beliefs' of the individual. Instead, we followed an indirect route and sought representations that preserve certain 'nice' features of the expected utility representation, e.g., the cardinality of U, even though the underlying class of preferences is much more general. See Ghirardato et al. [16] for another paper pursuing the same route and for further discussion of its merits and limitations. It should be noted here that an explicit behavioral meaning of I can be given if we require that preferences are biseparable in the sense of Ghirardato and Marinacci [13]. Their work can also serve as a formal justification for restricting attention to cardinal sets $\mathcal{U}'$. Rather than pursuing these arguments at the present level of generality, the reader is once again referred to the results in Section 6. There, after imposing a behavioral notion of ambiguity aversion, we show that the certainty equivalent I, as singled out by Theorem 2 and Corollary 5, takes the familiar maxmin form. Thus, we have another example, which is more general than the expected utility preferences discussed earlier, in which the certainty equivalent I can be linked to specific behavior. The vast literature on maxmin preferences can also be used to substantiate the link further and clarify the extent in which I captures 'beliefs'. For example, one can carry out an analogue of the comparative static analysis in Ghirardato and Marinacci [14], or analyze the effects of learning on behavior and the chosen utility representation, as is done in Epstein and Schneider [9].[5]

To conclude this section, we should note that Theorem 3 is stronger than the uniqueness results obtained in Epstein [5]. The latter rely on the stochastic framework and the assumption that the certainty equivalent I has the expected utility form. In particular, the expected utility hypothesis implies directly that $U : X^\infty \to \mathbb{R}$ is cardi-

[5]The separation of 'beliefs' and 'tastes' delivered by Corollary 5 is also partial in the following sense. Consider a setting in which the environment is perceived to be ambiguous. The certainty equivalent I, as identified in Theorem 2 and Corollary 5, would then encode both the individual's perception of how ambiguous the environment is and his attitude toward the perceived ambiguity. While such attitudes may be more appropriately regarded as part of the individual's 'tastes', it has been recognized that a clearcut distinction between the individual's perception of ambiguity and his attitude toward it is often impossible. The response has been to treat both aspects of an individual's preference as part of his 'beliefs'. For recent work attempting to disentangle these aspects, see Klibanoff et al. [20]. Their results require more structure on the environment than is assumed here.

nally unique. Making use of this observation, Epstein [5] proves that $b = \hat{b}$. Theorem 3 is stronger in that it delivers the same conclusions on the basis of deterministic choice alone. The proof is also different. Thus, we first establish the uniqueness of the discount factor b. We then use the latter to prove that $U : X^\infty \to \mathbb{R}$ is cardinally unique.

6 Ambiguity Aversion

Intertemporal hedging expresses the idea that an individual who is concerned about uncertainty would seek to take different, negatively correlated bets in different time periods. To illustrate, consider an environment in which there are only two future periods, t and $t+1$, and all the uncertainty is whether an event A obtains in period t. Assuming $X = \mathbb{R}_+$, suppose that the individual is endowed with an asset that pays $x^* > 0$ in period t if the event A occurs, and $x < x^*$ otherwise. Prior to the resolution of uncertainty, the individual is then given the choice between two other assets, both of which deliver consumption in period $t+1$. The first asset, which we denote by P for 'positively correlated', pays x^* if the event A obtains and x otherwise. The second asset, which we denote by N for 'negatively correlated', pays x^* if A^c obtains and x otherwise. Suppose that the individual has no information about the likelihood of the event A. In particular, he has no reason to believe A is more likely than A^c. Finally, suppose that the individual has no other sources of consumption and cannot save. Kochov [21] defined intertemporal hedging as a preference for the asset N. The idea is that, by taking a different bet in period $t+1$, the individual can offset the uncertainty about his consumption in period t and, hence, smooth ex ante utility across states of the world. Such 'time diversification' has a downside, however, in that the individual is left with a consumption profile that fluctuates over time: if consumption is high in one period, it will be low in the other. As is explained in Kochov [21], when we employ the standard model of dynamic choice, the pros and cons of intertemporal hedging offset each other exactly. This is an implication of the fact the model treats variability across states and time symmetrically: the curvature of the period utility function $u : X \to \mathbb{R}$ determines both risk aversion and the elasticity of intertemporal substitution. Kochov [21] showed that many models of ambiguity aversion generate a strict preference for intertemporal hedging. The intuition is that ambiguity, being a more serious concern than risk, tips the scales in favor of smoothing consumption across states rather time. A limitation of that paper is that it assumes away intertemporal complementarities that may stem from the individual's tastes, that is, from the way he evaluates deterministic consumption streams. This assumption is not innocuous. It is shown in Epstein [5] that an expected utility model with endogenous discounting, such as the one in (1.1), generates similar behavior if the individual's degree of patience decreases with the level of consumption, that is, if $b(x) < b(y)$ whenever $U(x, x, ...) > U(y, y, ...)$, $x, y \in X$. The goal of this section is to modify the notion of intertemporal hedging given in Kochov [21] so as to insure that the demand for hedging is driven solely by ambiguity aversion. To that end, it is helpful to first consider an axiom from Koopmans [22], which expresses a preference to advance the consumption of desirable outcomes.

Impatience: For all $t \in T\backslash\{0\}, a, b \in X^t, d \in X^\infty$, the stream $(a, a, ...)$ is preferred to $(b, b, ...)$ if and only if (a, b, d) is preferred to (b, a, d).

Before we explain how this axiom is related to the study of ambiguity, we should point out that one of the main objectives of Koopmans [22] was to investigate whether stationary preferences on X^∞ satisfy Impatience. The project was continued by Koopmans et al. [23], who proved that, while a weaker version of the axiom is indeed implied, Impatience fails generically.[6] Epstein [5] revisited the problem in a stochastic environment. After deriving the model in (1.1), he observed that any preference on X^∞ that admits an Uzawa-Epstein utility satisfies Impatience. His observation and the discussion in Koopmans et al. [23] are also relevant here. They reveal a key prediction of the Uzawa-Epstein specification that is not shared by the larger class of all stationary preferences on X^∞. The fact that Impatience is implied once we adopt a stochastic analogue of stationarity is also of conceptual interest. As Epstein [5] remarked, his results provide an intriguing demonstration of the often held view that uncertainty about the future contributes to impatience. By relaxing the expected utility hypothesis, Theorem 2 provides a much stronger demonstration of the same view.

Still focusing on the ranking of deterministic consumption streams, recall now that Stationarity implies History Independence. That is, within the class of preferences we consider, current consumption does not affect how the individual ranks future outcomes. An intertemporal complementarity may therefore arise only when the anticipation of future outcomes affects the individual's decisions about prior consumption.

[6]The weaker property says that for every $t \in T$ and $a, b \in X^t$, if $(a, a, ...)$ is preferred to $(b, b, ...)$, then (a, b, d) is preferred to (b, a, d) for all $d \in X^\infty$ that are no better than $(a, a, ...)$ and no worse than $(b, b, ...)$.

The key observation that underlies the analysis in this section is that the scope of such 'future dependencies', while nontrivial, is restricted by Impatience. Specifically, the axiom implies that for all $t \in T, a, b \in X^t$ and $d, d' \in X^\infty$, if $(a, b, d) \succeq (b, a, d)$, then $(a, b, d') \succeq (b, a, d')$. Thus, when the individual is asked to choose the order in which a given set of outcomes is consumed, the intertemporal complementarities that may otherwise affect how he evaluates deterministic consumption streams do not play a role.

Taking the implications of Impatience into account, the way to modify intertemporal hedging becomes clear. Once again, the individual would seek to take different 'bets' in different 'time periods'. However, the uncertainty he is trying to hedge now no longer concerns the level of consumption in a give time period t. Instead, it concerns the order in which a given set of outcomes is consumed. It is in those choice situations that the taste complementarities, which confounded the elicitation of ambiguity attitudes, are no longer present. To proceed formally, fix some $t \in T$ and a finite stream $a := (x_0, x_1, ..., x_{t-1}) \in X^t$ of outcomes. For every permutation $\pi : \{0, 1, ..., t-1\} \to \{0, 1, ..., t-1\}$, let πa denote the stream outcomes obtained from a by permuting the order of its elements according to π. That is, $\pi a = (x_{\pi(0)}, x_{\pi(1)}, ..., x_{\pi(t)})$. Say that $h \in \mathcal{H}$ is a **permutation act** if there is some $t \in T$ and $a \in X^t$ such that for every $\omega \in \Omega$, $h(\omega) = (\pi_\omega a, \pi_\omega a, ...)$ for some permutation $\pi_\omega : \{0, 1, ..., t-1\} \to \{0, 1, ..., t-1\}$. Observe that any permutation act has a repeating structure: the order of outcomes in the first block of t periods is exactly the same as the order of outcomes in the second block of t periods, and so on. The repeating structure means that uncertainty is compounded: If an event A obtains in which the outcomes in the first t periods are ordered unfavorably, then the same unfavorable order will prevail in all subsequent periods. Confronted with such an act, an individual may choose to break the correlation by reversing the order in at least one block of t periods. In particular, take some $t \in T$ and $a \in X^t$. Let $h, g \in \mathcal{H}$ be two permutation acts that differ in the way the elements of a are ordered. Thus, for every $\omega \in \Omega$, $h(\omega) = (\pi_\omega a, \pi_\omega a, ...)$ and $g(\omega) = (\tilde{\pi}_\omega a, \tilde{\pi}_\omega a, ...)$ but the permutations $\pi_\omega, \tilde{\pi}_\omega$ are potentially different. If $h \sim g$, intertemporal hedging means that the individual would prefer the act $m \in \mathcal{H}$ where $m(\omega) = (\pi_\omega a, \tilde{\pi}_\omega a, \tilde{\pi}_\omega a, ...)$, $\omega \in \Omega$. Below, we state a stronger axiom whereby the individual may use any act $g \in \mathcal{H}$ to hedge the uncertainty inherent in a permutation act h. The stronger axiom is motivated by a desire to obtain the strongest distinction between the maxmin model in (1.2) and the expected-utility model with endogenous discounting in (1.1). Note that in the statement of the axiom we think of an act $h \in \mathcal{H}$ as a function, $\omega \mapsto h(\omega)$, from Ω into X^∞.

Intertemporal Hedging: Let $\{A_1, A_2, ..., A_n\} \subset \cup_t \mathcal{F}_t$ be a partition of Ω. If

$$h := \begin{cases} (\pi_1 a, \pi_1 a, ...) & \text{if } \omega \in A_1 \\ \dots & \\ (\pi_n a, \pi_n a, ...) & \text{if } \omega \in A_n \end{cases} \quad \succeq \quad g := \begin{cases} d_1 & \text{if } \omega \in A_1 \\ \dots & \\ d_n & \text{if } \omega \in A_n, \end{cases},$$

then

$$m := \begin{cases} (\pi_1 a, d_1) & \text{if } \omega \in A_1 \\ \dots & \\ (\pi_n a, d_n) & \text{if } \omega \in A_n \end{cases} \quad \succeq \quad g := \begin{cases} d_1 & \text{if } \omega \in A_1 \\ \dots & \\ d_n & \text{if } \omega \in A_n \end{cases},$$

where $a \in X^t$ for some $t > 1$, $\pi_i : \{0, 1, ..., t-1\} \to \{0, 1, ..., t-1\}$ is a permuation for every $i = 1, 2, ..., n$, and $h, g \in \mathcal{H}$.

The next result confirms that, within the class of stationary preferences, a strict desire to hedge intertemporally can be interpreted as evidence of ambiguity aversion.[7] The result is actually stronger and shows that, once the axiom is added to the hypothesis in Theorem 2, we obtain a maxmin certainty equivalent, that is, we can find a weak*-closed, convex set $\mathcal{P}$ of probability measures on $(\Omega, \cup_t \mathcal{F}_t)$ such that $I(\xi) = \min_{p \in \mathcal{P}} \mathbb{E}_p \xi$ for every $\xi \in B^0$. It should also be noted that the certainty equivalent I and, hence, the set $\mathcal{P}$ of prior beliefs are uniquely identified. In particular, the result closes the gap left by Corollary 5, which did not cover the case when the rate of time preference is fixed.

Theorem 6 *A stationary preference $\succeq$ on $\mathcal{H}$ satisfies Intertemporal Hedging if and only if it has a representation (U, ϕ, I) where ϕ is an Uzawa-Epstein time aggregator and I is a maxmin certainty equivalent. Moreover, if $(\hat{U}, \hat{\phi}, \hat{I})$ is any other such representation, then $I = \hat{I}$.*

The analysis in this section is based on the distinction between two forms of uncertainty: one affecting the order in which a given set of outcomes is consumed and

[7]Throughout the paper, we follow Ghirardato and Marinacci [14] and interpret any departure from expected utility as evidence of ambiguity. Epstein [5] observes that some specifications of the maxmin model are probabilistically sophisticated in the sense of Machina and Schmeidler [29] and develops a definition of ambiguity aversion that excludes those preferences. To interpret intertemporal hedging as evidence of ambiguity aversion in the sense of Epstein [5], we would have to exclude those preferences as well. See Maccheroni et al. [28] for an elegant characterization of such preferences.

one affecting the level of consumption in an isolated period. While this distinction is conceptually straight-forward, it should be acknowledged that it may be hard to find real-world applications in which the relevant uncertainty can be so finely categorized. It is more likely that uncertainty would affect the timing and the level of consumption simultaneously. The behavior we identify with ambiguity aversion can thus be criticized as narrow in scope. This is the price one has to pay for relaxing time additivity. If time additivity is maintained, it is enough for the individual to pursue an intertemporal hedge against any form of uncertainty. Such behavior would count as evidence of ambiguity aversion. This observation follows by combining the conclusions of Theorem 6 with the earlier results in Kochov [21]. In defence of the present analysis, it should be observed that providing a choice-theoretic definition of ambiguity aversion in general choice settings, with few auxiliary restrictions on behavior, is difficult. For example, the static definitions in Gilboa and Schmeidler [17] and Epstein [7] take advantage of the fact that the domain of choice includes a rich collection of unambiguous events, that is, events whose likelihood is known. Ambiguity aversion is then defined as the preference to bet on events with known rather than unknown probability. These definitions are not applicable in the present framework since we do not require the existence of events with known probability. On the other hand, the static definition in Ghirardato et al. [15] is applicable. In this respect, we should first point out that the present analysis is complementary as it concerns qualitatively different behavior, specific to the dynamic setting. The definitions differ in other dimensions as well, revealing alternative compromises one has to make to isolate the effects of ambiguity aversion. In particular, to formulate the definition in Ghirardato et al. [15], it is necessary to first elicit the certainty equivalents for an appropriately chosen family of bets.[8] Once this is done, a wide range of choice comparisons can be used to test if the individual is ambiguity averse. In contrast, the notion of ambiguity aversion we propose is based on a narrow set of choice comparisons, but does not require the prior elicitation of certainty equivalents. To emphasize this point, it should be noted that intertemporal hedging can be formulated even when the space X of outcomes is discrete and certainty equivalents may fail to exist.

[8]In a static environment, a **bet** is an act $f : \Omega \to X$ that depends on the realization of a single event, that is, there are outcomes $x, y \in X$ and an event $A \subset \Omega$ such that $f(\omega) = x$ for all $\omega \in A$ and $f(\omega) = y$ otherwise. By a **certainty equivalent** for the bet f, we mean an outcome $z \in X$ such that $z \sim f$.

A Appendix

Given functions, $f : X' \to Y'$ and $g : Y' \to Z'$, we use $g(f(x'))$ and $gf(x')$ interchangeably to denote the value of the function $g \circ f : X' \to Z'$ at a point $x' \in X'$.

A.1 Proof of Theorem 2

A.1.1 Preliminaries

Lemma 7 *State Independence and Stationarity imply Monotonicity.*

Proof. Let $h, h' \in \mathcal{H}$ be such that $h(\omega) \geq h'(\omega)$ for every ω. Because h, h' are simple, there is some $t \in T$ such that h_k, h'_k are $\mathcal{F}_t$-measurable for every k. Fix some $a = (x_0, ..., x_{t-1}) \in X^t$ and consider the acts $(a, h), (a, h')$. By construction, $(a, h)(\omega) = (a, h(\omega))$. By Stationarity, $(a, h(\omega)) \geq (a, h'(\omega))$ for every $\omega \in \Omega$. Moreover, $h \geq h'$ if and only if $(a, h) \geq (a, h')$, so it suffices to show the latter. Think of $(a, h), (a, h')$ as functions from Ω into X^∞. Both functions have finite range. Let $\{A_1, A_2, ..., A_n\}$ be the coarsest partition of Ω such that the functions are measurable. Replace the infinite stream of (a, h') on A_1 by the respective infinite stream of (a, h). By State Independence, the new act is preferred to (a, h'). Take the new act and replace its infinite stream on A_2 by the respective infinite stream of (a, h). Apply State Independence again. After n such steps, we obtain (a, h). By transitivity, (a, h) is preferred to the initial act (a, h'). ∎

The following lemma lists a number of additional properties of $\geq$ that will be used in subsequent results.

Lemma 8 *Suppose $\geq$ satisfies C, SI and S. Then,*

1. *There are $x, y \in X, d \in X^\infty$ such that $(x, d) \succ (y, d)$.*
2. *For every $x \in X, h \in \mathcal{H}$, $(x, x, ...) \geq h$ if and only if $(x, h) \geq h$. Similarly, $h \geq (x, x, ...)$ if and only if $h \geq (x, h)$.*
3. *The best and worst sequences in X^∞ are constant. Denote them by $(z^*, z^*, ...)$ and $(z, z, ...)$.*
4. *Writing d^* for $(z^*, z^*, ...)$, we have $(z^*, z, d^*) \succ (z, z, d^*)$.*
5. *There exists a sequence $(x_n)_n$ in X, converging to z such that $(x_n, z, z, ...) \succ$*

$(z, z, z, ...)$. *Moreover, for every* $n \in \mathbb{N}, d \in X^\infty$*, there is* $d' \in X^\infty$ *such that* $(z, d') \sim (x_n, z, d)$.

Proof. Property (1) is proved in Kochov [21, Lemma 5]. To prove (2), suppose $(x, x, ...) \succeq h$ but $h \succ (x, h)$. Then, by S, $h \succ (x, h) \succ (x, x, h)$. Repeating the argument, we obtain $h \succ (x^n, h)$ for every $n \in \mathbb{N}$, where x^n denotes the vector in X^n all of whose components are equal to x. By Continuity, $h \succ (x, x, ...)$, a contradiction. Analogous arguments establish the converse implication and the second equivalence postulated in (2). To prove (3), let $(z^*, z^*, ...)$ be the best among all constant sequences in X^∞. By Continuity, it is enough to show that $(z^*, z^*, ...) \succeq (x_1, x_2, ..., x_n, z^*, z^*, ...)$ for all $x_k \in X, k = 1, ..., n, n \in \mathbb{N}$. From (2), we know that $(z^*, z^*, ...) \succeq (x_k, z^*, z^*, ...)$ for every k. Making repeated use of this observation and Stationarity, we obtain

$$
\begin{aligned}
&(z^*, z^*, ...) \succeq (x_n, z^*, z^*, ...) \quad \Rightarrow \\
&(x_{n-1}, z^*, z^*, ...) \succeq (x_{n-1}, x_n, z^*, z^*, ...) \quad \Rightarrow \\
&(z^*, z^*, ...) \succeq (x_{n-1}, z^*, z^*, ...) \succeq (x_{n-1}, x_n, z^*, z^*, ...) \quad \Rightarrow \\
&... \\
&(z^*, z^*, ...) \succeq (x_1, x_2, ..., x_n, z^*, z^*, ...)
\end{aligned}
$$

Turn to (4). By way of contradiction, suppose $(z, z, d^*) \succeq (z^*, z, d^*)$. Together with the fact that $d^* \succeq (z, d^*)$ and preference is stationary, we obtain

$$(z, d^*) \succeq (z, z, d^*) \succeq (z^*, z, d^*)$$

By S, $(z^*, z, d^*) \succeq (z^*, z^*, z, d^*)$. By the contradiction hypothesis,

$$(z, z, d^*) \succeq (z^*, z, d^*) \succeq (z^*, z^*, z, d^*)$$

Repeating the argument, we obtain $(z, z, d^*) \succeq ((z^*)^n, z, d^*)$ for every $n \in \mathbb{N}$. By Continuity , $(z, z, d^*) \succeq d^*$, contradicting property (2). Finally, turn to property (5). From property (2) in that lemma, $(z^*, z, z, ...) \succ (z, z, ...)$. Because X is connected, there is a sequence $(x_n)_n$ in X converging to z such that $(x_n, z, z, ...) \succ (z, z, ...)$. Letting $d^* = (z^*, z^*, ...)$ once again, another implication of property (2) gives $(z, d^*) \succ (z, z, d^*)$. Since (x_n, z, d^*) converges to (z, z, d^*), as $n \to \infty$, we have $(z, d^*) \succ (x_n, z, d^*)$ for all n larger than some $N \in \mathbb{N}$. By choice of x_n, it is also the case that

$$(x_n, z, d^*) \succeq (x_n, z, z, z, ...) \succ (z, z, z, ...).$$

Combining the last two observations gives $(z, d^*) \succ (x_n, z, d^*) \succ (z, z, z, ...)$ for all $n \geq N$. But $(x_n, z, d^*) \succeq (x_n, z, d) \succ (z, z, ...)$ for all $d \in X^\infty, n \in \mathbb{N}$. Then for every $n \geq N$, we have

$$(z, d^*) \succ (x_n, z, d^*) \succeq (x_n, z, d) \succ (z, z, z, ...)$$

By Continuity and the connectedness of X^∞, there is $d' \in X^\infty$ such that $(z, d') \sim (x_n, z, d)$. ■

Proof of Lemma 1: Let $U : X^\infty \to \mathbb{R}$ represent the restriction of $\succeq$ to X^∞. Note that U is continuous in the product topology on X^∞. For every $x \in X, k \in U(X^\infty)$, define $\phi(x, k) := U(x, d_k)$ where $d_k \in X^\infty$ is such that $U(d_k) = k$. Stationarity implies that ϕ is strictly increasing in its second component. Since U is continuous, the set $\phi(x, U(X^\infty))$ is connected for every $x \in X$. Conclude that that ϕ is continuous in the second component. Continuity in the first component follows immediately from the continuity of U. Now focus on stochastic acts. Since X^∞ is connected, a standard argument shows that, for every act $h \in \mathcal{H}$, we can find an act $d_h \in X^\infty$ such that $h \sim d_h$. Extend U from X^∞ to $\mathcal{H}$ by letting $\tilde{U}(h) := U(d_h)$. Let $U \circ \mathcal{H} := \{U \circ h : h \in \mathcal{H}\} \subset B^\circ$. Define $I : U \circ \mathcal{H} \to \mathbb{R}$ by $I(U \circ h) = \tilde{U}(h)$. It follows from Lemma 7 that I is well defined and monotone. By definition, $I(k) = k$ for all $k \in U(X^\infty)$. we have thus shown that $\succeq$ admits a representation (U, ϕ, I). It remains to show that for all such representations, I and ϕ permute. Fix some $x \in X, h \in \mathcal{H}$ and note that $U \circ (x, h) = \phi(x, U \circ h)$. Choose $d \in X^\infty$ such that $d \sim h$. By Stationarity $(x, d) \sim (x, h)$. Since I is normalized, we have

$$I(\phi(x, U \circ h)) = U(x, d) = \phi(x, U(d)) = \phi(x, I(U \circ h)),$$

completing the proof of the lemma. ■

Take a strictly increasing function $g : \mathbb{R} \to \mathbb{R}$ and a representation (U, ϕ, I). Clearly, $g \circ U$ is a representation for the restriction of $\succeq$ to X^∞. We omit the obvious proof.

Lemma 9 *Suppose a preference relation $\succeq$ on $\mathcal{H}$ has a representation (U, ϕ, I). Given a strictly increasing function $g : U(X^\infty) \to \mathbb{R}$, let $V := g \circ U$ and*

$$\hat{I}(\xi) := gI(g^{-1} \circ \xi) \quad \textit{and} \quad \hat{\phi}(x, k) := g\phi(x, g^{-1}(k)), \quad \forall \xi \in V \circ \mathcal{H},\, k \in V(X^\infty).$$

Then, $\hat{\phi}$ is a time aggregator for V and $(V, \hat{\phi}, \hat{I})$ is a representation for $\succeq$ such that $\hat{\phi}$ and $\hat{I}$ permute.

A.1.2 Constructing an Iteration Group

We need to introduce some mathematical machinery developed in Lundberg [26]. Let C be an open interval in $\mathbb{R}$ and let $\{g^\alpha\}$ be a family of functions each from a subinterval of C into C, where the index α ranges over an interval $(-\lambda, \lambda)$ for some $\lambda \in \mathbb{R} \cup \{+\infty\}$. Suppose each function g^α is continuous and strictly increasing, and its graph disconnects C^2. Suppose further that the graph of $g^\alpha \circ g^\beta$ is contained in the graph of $g^{\alpha+\beta}$ for all $\alpha, \beta \in (-\lambda, \lambda)$ such that $\alpha + \beta \in (-\lambda, \lambda)$. We call such a family of functions an **iteration group over the interval** C.[9] Letting $j : C \to C$ denote the identify function, observe that $g^0 = j$. More generally, when α is an integer, then g^α is the α iterate of g^1. To streamline the notation, from now on, whenever a function g and an integer $m \in \mathbb{Z}$ are given, we use g^m to denote the m^{th} iterate of g. The following example will play a role in the rest of the proof and may clarify the notion of an iteration group. Let $C = \mathbb{R}$, $\lambda = +\infty$, and suppose that all functions g^α are defined on $\mathbb{R}$. Suppose further that g^1 is an affine function so that $g^1(k) = a + bk$ for some $a, b \in \mathbb{R}$. To be concrete, assume that $b \neq 1$. Then, we can compute all functions g^α explicitly:

$$g^\alpha(k) = a\frac{1 - b^\alpha}{1 - b} + b^\alpha k, \quad \forall k, \alpha \in \mathbb{R}. \tag{A.1}$$

Note that in this example all functions g^α, $\alpha \neq 0$, share the same fixed point $k^* = a(1-b)^{-1}$. This observation holds more generally. Thus, if $\{g^\alpha\}$ is an iteration group and $g^\alpha(k) = k$ for some $\alpha \neq 0$ and $k \in C$, then $g^\beta(k) = k$ whenever k is in the domain of g^β. Say that an iteration group $\{g^\alpha\}$ is **fixed point free** if none of the functions g^α, $\alpha \neq 0$, has a fixed point. As observed in Lundberg [26], we can then find a function $L : C \to \mathbb{R}$ such that $g^\alpha(k) = L^{-1}(L(k) + \alpha)$ for all k in the domain of g^α and all $\alpha \in (-\lambda, \lambda)$. The function L is called the **Abel function** for the group $\{g^\alpha\}$. For future reference, observe that if L is an Abel function for a group $\{g^\alpha\}$, then so is the function $L + c$ where c is an arbitrary real number. Also, note that when $\{g^\alpha\}$ is fixed point free, it it w.l.o.g. to assume that $g^\alpha(k) > k$ for all k in the domain of g^α and all $\alpha > 0$. Else, we can relabel the group by taking $\tilde{g}^\alpha := g^{-\alpha}$ for every $\alpha \in (-\lambda, \lambda)$. Under this assumption, any Abel function L for the iteration group is strictly increasing. In the example in (A.1), suppose D is an interval such that all functions g^α, $\alpha \neq 0$, are fixed point free when restricted to D. For instance we can take $D = (k^*, +\infty)$. Then, $L(k) := \log_b(k - k^*)$, $k \in D$, is an Abel function for the group $\{g^\alpha|_D\}$.

[9] When C is a proper subset of $\mathbb{R}$ and $\lambda < +\infty$, Lundberg [26] refers to the iteration group as truncated. We have no reason to distinguish between the cases and omit the 'quantifier' truncated.

If $\mathcal{Z}$ is a topological space and A is a set in $\mathcal{Z}$, we use A° to denote its topological interior. For a sequence $(A_n)_n$ of sets in $\mathcal{Z}$, we denote by $\mathrm{Ls}A_n \subset \mathcal{Z}$ and $\mathrm{Li}A_n \subset \mathcal{Z}$ the topological lim sup and lim inf of the sequence. See Aliprantis and Border [1, p.109] for precise definitions of these concepts. We write $A_n \to_L A$ if $A = \mathrm{Li}A_n = \mathrm{Ls}A_n$. The set A is called the **closed limit** of $(A_n)_n$. Following Lundberg [26], a correspondence $f^* : C \rightrightarrows \mathbb{R}$, where C is an interval in $\mathbb{R}$, is called a **cliff funciton** if the set $f^*(k)$ is connected for every $k \in C$. A cliff function f^* is **increasing** if $k \le k'$ and $l \in f^*(k), l' \in f^*(k')$ imply $l \le l'$ for all $k, k' \in C$. Observe that any increasing function $f : C \to \mathbb{R}$ is an increasing cliff function. If we identify every cliff function f^* with its graph in $\mathbb{R} \times \mathbb{R}$, we can also speak of the closed limit of a sequence $(f_n^*)_n$ of cliff functions.

Every iteration group $\{g^\alpha\}$ that arises in this paper can be constructed from a sequence of functions $(g_n)_n$ as follows. See Lundberg [26, Lemma 2.5] for details.

1. There is a sequence $(q_n)_n$ in $\mathbb{Z}$ such that $g_n^{q_n} \to_L g^1$.
2. For every $k \in \mathbb{Z}_+$ such that $m := 2^{-k} \in (-\lambda, \lambda)$, there is a sequence $(p_n)_n$ in $\mathbb{Z}$ such that $g_n^{p_n} \to_L g^m$.
3. For an arbitrary real number $\alpha \in (-\lambda, \lambda)$, the function g^α can be obtained as follows. Observe that we can write α as

$$\alpha = a_0 + \sum_{i=1}^{\infty} \frac{a_i}{2^i}$$

where a_0 is an integer and a_i is either 0 or 1 for all $i > 0$. Now, let $h_0 := g^{a_0}, h_i := g^{2-i}$ for $i > 0$ and $a_i = 1$. For all $n \in \mathbb{N}$, let $\hat{h}_n := h_0 \circ \ldots \circ h_n$. Then, $\hat{h}_n \to_L g^\alpha$.

Return to the proof of Theorem 2. Choose a sequence $(x_n)_n$ in X satisfying property (5) in Lemma 8. Let $C := \{U(z,d) : d \in X^\infty\}$. Note that C is a closed interval in $\mathbb{R}$ with nonempty interior. For every $k \in C$ and $n \in \mathbb{N}$, let $f_n(k) := \phi(x_n, k)$. Also, let $f(k) := \phi(z,k), k \in C$. Given (5), we know that $f_n(C) \subset C$ for every n, and $f(C) \subset C$. Now, let $A \in \cup_t \mathcal{F}_t$ be an essential event. It is w.l.o.g. to assume that $A \in \mathcal{F}_1$. Letting $\mathcal{H}(A)$ be the subset of acts $h \in \mathcal{H}$ that are $\{A, A^c\}$-adapted, identify $U \circ \mathcal{H}(A) := \{U \circ h : h \in \mathcal{H}(A)\}$ with a subset in $\mathbb{R}^2$. Observe that $C^2 \subset U \circ \mathcal{H}(A)$. Thus, I induces a function on C^2. Abusing notation, we denote this function by I as well and write $I(k,k')$ for its value at $(k,k') \in C^2$. Note that, by State Independence and Continuity respectively, I is strictly increasing and continuous on C^2.

Lemma 10 $f_n \to_L f$.

Proof. Continuity of the preference relation implies that $(f_n)_n$ converges pointwise to f. By Lundberg [26, Lemma 1.1], we can choose a subsequence of $(f_n)_n$ that has a closed limit f^*. It is enough to show that for any such subsequence, $f^* = f$. Thus, fix a subsequence with a closed limit f^*. Observe that, in principle, f^* may be a correspondence rather than a proper function. The first step is to show that this is not the case. The fact that ϕ and I permute can now be written as follows

$$f_n I(k, k') = I(f_n(k), f_n(k')) \quad \forall n \in \mathbb{N}, \forall k, k' \in C. \tag{A.2}$$

By Lundberg [26, Lemma 4.8], we also have

$$f^*(I(k, k')) = \{I(l, l') : l \in f^*(k), l' \in f^*(k')\} \quad \forall k, k' \in C.$$

It follows from Lundberg [26, Lemma 4.8] that f^* is a proper function and, from Lundberg [26, Lemma 1.2], that $(f_n)_n$ converges to f^* uniformly on all compact sets A, $A \subset C^\circ$. But then, $f^*(k) = f(k)$ for all $k \in C^\circ$. Since f, f^* are both continuous, conclude that $f = f^*$. ■

For every n, let $g_n := f^{-1} \circ f_n$. A direct calculation shows that the equality in (A.2) is preserved if we replaced the functions f_n with g_n. It follows from Lemma 10 and Lundberg [25, Thm 5.3] that $g_n \to_L j$. Letting $\operatorname{Dom} g_n$ denote the domain of g_n, it is also the case that $\operatorname{Dom} g_n \to_L C$. See Lundberg [25, Lemma 3.9]. Deduce from Lundberg [26, 4.16] that the sequence $(g_n)_n$ generates an iteration group $\{g^\alpha\}$ **over** C°, where α ranges over an interval $\Lambda := (-\lambda, \lambda)$ for some $\lambda \in \mathbb{R} \cup \{+\infty\}$. Furthermore,

$$g^\alpha I(k, k') = I(g^\alpha(k), g^\alpha(k')), \quad \forall \alpha \in \Lambda, \forall k, k' \in C^\circ. \tag{A.3}$$

Because all functions $g^\alpha, \alpha \neq 0$, share the same set of fixed points, we can find a closed interval $D \subset C^\circ$, with nonempty interior, such that all functions g^α, $\alpha \neq 0$, are fixed point free when they are restricted to D. From now on focus on these restrictions and write g^α for $g^\alpha|_D$. Let $L : D \to \mathbb{R}$ be the Abel function for the resulting iteration group. As was explained previously, it is w.l.o.g. to assume that L is strictly increasing.

Since $U : X^\infty \to \mathbb{R}$ is continuous and X is connected, we know that $\{U(x, x, ...) : x \in X\}$ is connected. It follows from property (3) in Lemma 8, that $\{U(x, x, ...) : x \in X\} = U(X^\infty)$. Thus, we can find $x_0 \in X$ such that $U(x_0, x_0, ...) \in D^\circ$. Let $f_0(k) := \phi(x_0, k)$ for all $k \in U(X^\infty)$. Because X is compact, there exists $N \in \mathbb{N}$ such that $f_0^N(U(X^\infty)) \subset D^\circ$. Let $\hat{U} = f_0^N \circ U$ and, for every $x \in X, s \in \hat{U}(X^\infty)$, let $\hat{\phi}(x, s) := f_0^N \phi(x, f_0^{-N}(s))$. Observe that $\hat{\phi}$ is a time aggregator for $\hat{U}$. Now let

$x_0^N \in X^N$ be the string of outcomes each element of which is equal to x_0. Stationarity implies

$$h \geq h' \Leftrightarrow (x_0^N, h) \geq (x_0^N, h') \Leftrightarrow I(f_0^N \circ U \circ h) \geq I(f_0^N \circ U \circ h') \Leftrightarrow I(\hat{U} \circ h) \geq I(\hat{U} \circ h').$$

for all acts $h, h' \in \mathcal{H}$. Conclude that $(\hat{U}, \hat{\phi}, I)$ is a representation for the preference relation $\geq$. As would become clear from the rest of the proof, this is the key implication of Continuity and Stationarity. It allows to deduce a representation in which the utility function takes values in D. We can the use the Abel function L to obtain another representation $(\tilde{U}, \tilde{\phi}, \tilde{I})$ such that $\tilde{I}$ is translation invariant. First, we need to show that an appropriate generalization of (A.3) obtains even when I is no longer restricted to C^2.

Lemma 11 *$I(g^\alpha \circ \xi) = g^\alpha I(\xi)$ for all $\alpha \in \Lambda$ and $\xi \in B_D^0$.*

Proof. Recall the definition of the functions g_n given after Lemma 10. Because I and ϕ permute, we know that $I(g_n^m \circ \xi) = g_n^m I(\xi)$ for all $n \in \mathbb{N}, m \in \mathbb{Z}$, and $\xi \in B_D^0$. Now take some $\alpha = 2^{-m}$ where $m \in \mathbb{Z}_+$ and $\alpha \in \Lambda$. Recall from the discussion of iteration groups that the function g^α can be obtained as the limit of a sequence $(g_n^{p_n})_n$, where $(p_n)_n$ is a suitable sequence of integers. The continuity of I and the functions $g_n^{p_n}$ insure that the desired equality holds for all such α. An analogous use of continuity shows that the equality holds for all $\alpha \in \Lambda$. ■

Recall now that if L is an Abel function for a group $\{g^\alpha\}$, then so is the function $L + c$ where c is an arbitrary real number. Hence, it is w.l.o.g. to choose L so that $0 \in L(D)^\circ$. Let $\tilde{U} := L \circ \hat{U}, \tilde{\phi}(x, s) := L\hat{\phi}(x, L^{-1}(s))$ for all $x \in X$, $s \in L(D)$, and $\tilde{I}(\xi) := LI(L^{-1} \circ \xi)$ for all $\xi \in B_{L(D)}^0$.

Lemma 12 *$\tilde{I}(\xi + \alpha) = \tilde{I}(\xi) + \alpha$ for all $\xi \in B_{L(D)}^0, \alpha \in \Lambda$ such that $\xi + \alpha \in B_{L(D)}^0$.*

Proof. Take ξ and α as in the statement of the lemma. Let $\xi' \in B_D^0$ be such that $L \circ \xi' = \xi$. Then,

$$\begin{aligned}\tilde{I}(\xi + \alpha) &= LI\big[L^{-1} \circ (L \circ \xi' + \alpha)\big] = LI[g^\alpha \circ \xi'] = Lg^\alpha I(\xi') = L(I(\xi')) + \alpha \\ &= LI[L^{-1} \circ \xi] + \alpha = \tilde{I}(\xi) + \alpha,\end{aligned}$$

completing the proof. ■

A.1.3 Time Aggregator

In this section, it is once again convenient to think of I as a function on $C^2 \subset \mathbb{R}^2$. Analogously, $\tilde{I}$ becomes a function on $L(D)^2$. Write $[c, c']$ for the closed interval $L(D)$ and define

$$\psi(k) := \begin{cases} \tilde{I}(c, c+k) - c & \text{if } k \in [0, c'-c] \\ \tilde{I}(c', c'+k) - c' & \text{if } k \in [c-c', 0] \end{cases}$$

By construction, $\tilde{I}(k,k) = k$ for all $k \in [c, c']$. Deduce that $\psi(0) = 0$. Since $\tilde{I}$ is continuous, we may conclude that ψ is a continuous function. Also, ψ is strictly increasing. Since $\tilde{I}$ is translation invariant, we have

$$\tilde{I}(s,t) = s + \psi(t-s), \quad \forall s,t \in [c,c'].$$

Also, $\psi(k) < k$ whenever $k > 0$ and $\psi(k) > k$ whenever $k < 0$. Finally, we want to show that $\psi(k) - k$ is a strictly decreasing function. Pick $k \in [0, c'-c)$ and $\varepsilon > 0$ such that $k + \varepsilon \in [0, c'-c]$. Then $\psi(k+\epsilon) - \psi(k) = I(c, c+k+\varepsilon) - I(c+\varepsilon, c+k+\varepsilon) < 0$. Conclude that $\psi(k) - k$ is a strictly decreasing on $[0, c'-c]$. A similar argument for $k \in [0, c'-c)$ confirms that $\psi(k) - k$ is strictly decreasing on its entire domain.

Now, for every $x \in X$ write $\tilde{f}_x$ for the function $\tilde{\phi}(x, \cdot)$ on $[c, c']$. The fact that $\tilde{I}$ and $\tilde{\phi}$ permute can then be written as follows:

$$\tilde{f}_x(s + \psi(t-s)) = \tilde{f}_x(s) + \psi(\tilde{f}_x(t) - \tilde{f}_x(s)), \quad \forall x \in X, s,t \in [c,c']. \tag{A.4}$$

The above is a special case of the functional equation

$$g(s + \psi(t-s)) = g(s) + \psi(g(t) - g(s)), \quad \forall s,t \in [c,c'], \tag{A.5}$$

where we think of ψ as known and of g as an unknown function. Suppose first that (A.5) is satisfied for some g that is affine on a subinterval in $[c, c']$. It follows from Lundberg [27, Thm 10.1,10.3] that all solutions g are affine on $[c, c']$. In particular,

$$\tilde{f}_x(k) = u(x) + b(x)k, \quad \forall x \in X, k \in [c,c']. \tag{A.6}$$

Because $(\tilde{U}, \tilde{\phi}, \tilde{I})$ is a representation for $\succeq$ and the preference is continuous, we know that $u, b : X \to \mathbb{R}$ are continuous functions. Since, each function $\tilde{f}_x$ is strictly increasing, we know that $b(x) > 0$. In addition, property (2) in Lemma 8 implies that $b(x) < 1$ for all $x \in X$.

Suppose now that (A.5) has no solution that is affine on a subinterval of $[c, c']$. It follows from Lundberg [27, Thm. 11.1] that all solutions take the form

$$\tilde{f}_x(k) = \tilde{\phi}(x, k) = \frac{1}{p} \ln(u(x) + b(x)e^{pk}), \quad \forall x \in X, k \in [c, c'], \tag{A.7}$$

where $u, b : X \to \mathbb{R}$, $p \in \mathbb{R}$ and $p \neq 0$. Assume that $p > 0$ and let $H(s) := e^{ps}$ for every $s \in \mathbb{R}$. If $p < 0$, we can let $H(s) := -e^{ps}$; the subsequent analysis would carry through in an analogous manner. Let $D^* := [e^{pc}, e^{pc'}]$ and define $U^* := H \circ \tilde{U}$, $\phi^*(x, k) := H\tilde{\phi}(x, H^{-1}(k))$, and $I^*(\xi) := H\tilde{I}(H^{-1} \circ \xi)$ for all $x \in X, k \in D^*, \xi \in B^0_{D^*}$. Then,

$$\phi^*(x, k) = u(x) + b(x)k, \quad \forall x \in X, k \in D^*. \tag{A.8}$$

By construction, (U^*, ϕ^*, I^*) is a representation for $\succeq$. Once again, it follows that the functions $u, b : X \to \mathbb{R}$ are continuous and $b(x) \in (0, 1)$ for every $x \in X$.

A.1.4 Certainty Equivalent

As in the preceding section, we break the proof in two cases depending on whether (A.5) admits an affine solution. Consider the affine case first. Since $0 \in L(D)^\circ$, we can find $x_0 \in X$ such that $u(x_0) = 0$. Let $\beta := b(x_0) \in (0, 1)$. The fact that $\tilde{I}$ and $\tilde{\phi}$ permute implies

$$\tilde{I}\big[u(x) + b(x)\xi\big] = u(x) + b(x)\tilde{I}(\xi), \quad \forall x \in X, \xi \in B^0_{L(D)} \tag{A.9}$$

Plugging x_0 above, we obtain $\tilde{I}(\beta\xi) = \beta\tilde{I}(\xi)$ for all $\xi \in B^0_{L(D)}$. Furthermore, $\tilde{I}(\beta^t\xi) = \beta^t\tilde{I}(\xi)$ for all $\xi \in B^0_{L(D)}, t \in \mathbb{N}$. Observe that $\beta^t\xi \in B^0_L(D)$ since $0 \in L(D)^\circ$. We will now show that $\tilde{I}\big[b(x)\xi\big] = b(x)\tilde{I}(\xi)$ for all $x \in X, \xi \in B^0_{L(D)}$. Choose t large enough so that $\beta^t u(x) \in \Lambda$ and $\beta^t b(x)\xi \in B^0_{L(D)}$. Making use of Lemma 12 and (A.9), we obtain

$$\begin{aligned}\beta^t u(x) + \beta^t b(x)\tilde{I}(\xi) &= \tilde{I}\big[\beta^t u(x) + \beta^t b(x)\xi\big] = \beta^t u(x) + \tilde{I}\big[\beta^t b(x)\xi\big] \\ &= \beta^t u(x) + \beta^t \tilde{I}[b(x)\xi].\end{aligned}$$

The next step is to extend the functional $\tilde{I}$ from $B^0_{L(D)}$ to B^0. Take $\xi \in B^0$ and pick $t \in \mathbb{N}$ large enough so that $\beta^t\xi \in B^0_{L(D)}$. Then, let $\tilde{I}^e(\xi) := \beta^{-t}\tilde{I}(\beta^t\xi)$. One can verify that $\tilde{I}^e$ is well defined and extends $\tilde{I}$. See Kochov [21]. The next step is to show that

$\tilde{I}^e$ is translation invariant. Take any $\xi \in B^0$ and any $\alpha \in \mathbb{R}$. Choose t large enough so that $\beta^t\xi, \beta^t(\xi + \alpha) \in B^0_L(D)$ and $\beta^t\alpha \in \Lambda$. Then,

$$\tilde{I}^e(\xi + \alpha) = \beta^{-t}\tilde{I}(\beta^t\xi + \beta^t\alpha) = \beta^{-t}(\tilde{I}(\beta^t\xi) + \beta^t\alpha) = \tilde{I}^e(\xi) + \alpha.$$

It follows from similar arguments that the extension $\tilde{I}^e$ is $b(x)$-homogeneous for every $x \in X$.

Suppose now that (A.4) has no affine solution. Recall the representation (U^*, ϕ^*, I^*) obtained in the preceding section and the set $\Lambda = (-\lambda, \lambda)$ as employed in Lemma 12. Recall that $H(s) = e^{ps}$ for every $s \in \mathbb{R}$ and let $O := H(\Lambda)$. Observe that O is an open interval such that $1 \in O$.

Lemma 13 *$I^*(\gamma\xi) = \gamma I^*(\xi)$ for all $\gamma \in O$ and $\xi \in B^0_{D^*}$ such that $\gamma\xi \in B^0_{D^*}$.*

Proof. Take γ and ξ as in the statement of the lemma. Observe that $H(s+t) = H(s)H(t)$ for all $s, t \in \mathbb{R}$. Let $\xi' := H^{-1} \circ \xi$ and $\alpha := H^{-1}(\gamma)$. Observe that $H^{-1} \circ (\gamma\xi) = \xi' + \alpha$. Also, $\gamma' + \alpha \in B^0_{L(D)}$ and $\alpha \in \Lambda$. From the construction of I^* and Lemma 12 deduce that

$$I^*(\gamma\xi) = H\tilde{I}\big[H^{-1} \circ (\gamma\xi)\big] = H\big[\tilde{I}(\xi' + \alpha)\big] = H\big[\tilde{I}(\xi') + \alpha\big] = H\big[\tilde{I}(\xi')\big]H(\alpha) = I^*(\xi)\gamma,$$

completing the proof of the lemma. ■

From Lemma 1, we know that I^* and ϕ^* permute. Let $C^* := \{V^*(z, d) : d \in X^\infty\}$. As in Section A.1.2, we can then construct an iteration group $\{g^\alpha\}$ over C^* where α varies over an interval $\Lambda^* := (-\lambda^*, \lambda^*)$ for some $\lambda^* \in \mathbb{N}$.[10] Since the functions g_n employed in that construction are now affine, the iteration group $\{g^\alpha\}$ can take one of two forms:

$$\text{either} \quad g^\alpha(k) = a\frac{1 - b^\alpha}{1 - b} + b^\alpha k \quad \forall\alpha, \quad \text{or} \quad g^\alpha(k) = \alpha + k \quad \forall\alpha,$$

where $a \in \mathbb{R}, b \in \mathbb{R}_{++}$. It is enough to consider the first possibility; in the other, the arguments are analogous and, in fact, simpler. From Lemma 11, we know that

$$I^*(a\frac{1 - b^\alpha}{1 - b} + b^\alpha\xi) = a\frac{1 - b^\alpha}{1 - b} + b^\alpha I^*(\xi) \tag{A.10}$$

for all $\alpha \in \Lambda^*$ and $\xi \in B^0_{C^*}$. Pick a proper interval $K \subset C^*$ and $\varepsilon > 0$ small enough such that $b^\alpha \in O$ for all $\alpha \in (-\varepsilon, \varepsilon)$, and $b^{-\alpha}\xi \in B^0_{C^*}$ for all $\alpha \in (-\varepsilon, \varepsilon)$ and $\xi \in B^0_K$. Let $O' := \{a\frac{1-b^\alpha}{1-b} : \alpha \in (-\varepsilon, \varepsilon)\}$. Observe that O' is an open interval such that $0 \in O'$.

[10] Note the abuse of notation: the functions g^α that we now construct are different from those in Section A.1.2.

Lemma 14 *$I^*(\xi+k)=I^*(\xi)+k$ for all $\xi\in B^0_K$ and $k\in O'$ such that $\xi+k\in B^0_K$.*

Proof. Pick ξ and k as in the statement of the lemma. By construction, we know that $k=a\frac{1-b^\alpha}{1-b}$ for some $\alpha\in(-\varepsilon,\varepsilon)$. Let $\xi':=b^{-\alpha}\xi$. By the choice of K and ε, we know that $\xi'\in B^0_{C^*}$ and $b^\alpha\in O$. Making use of (A.10) and Lemma 13, we obtain

$$I^*(\xi+k)=I^*(b^\alpha\xi'+a\frac{1-b^\alpha}{1-b})=b^\alpha I^*(\xi')+k=I^*(b^\alpha\xi')+k=I^*(\xi)+k,$$

completing the proof of the lemma. ■

Lemma 15 *There exists $\underline{\gamma}\in(0,1)$ and a certainty equivalent $I^e:B^0\to\mathbb{R}$ that is translation invariant, γ-homogeneous for every $\gamma\in(\underline{\gamma},1)$, and I^e extends $I^*:B^0_K\to\mathbb{R}$.*

Proof. Take $x\in X$ such that $U^*(x,x,...)=u(x)(1-b(x))^{-1}\in K^\circ$. Let $f(k):=u(x)+b(x)k$ for all $k\in\mathbb{R}$. Recall that $b(x)\in(0,1)$. Choose $\underline{\gamma}\in O\cap(0,1)$ such that $(1-\underline{\gamma})u(x)(1-b(x))^{-1}\in O'$ and $\underline{\gamma}u(x)(1-b(x))^{-1}\in K^\circ$. Based on the first inclusion, observe that

$$a(m,\gamma):=(1-\gamma)u(x)\frac{1-b^m(x)}{1-b(x)}\in O',\quad\forall m\in\mathbb{N},\gamma\in[\underline{\gamma},1].\tag{A.11}$$

Also, note that $[\underline{\gamma},1]\subset O$. Take a compact set K' such that $K'\supset K$. We are going to find an extension I^e of I^* from B^0_K to $B^0_{K'}$ such that (i) $I^e(\gamma\xi)=\gamma I^e(\xi)$ for all $\xi\in B^0_{K'}$ and $\gamma\in[\underline{\gamma},1]$ such that $\gamma\xi\in B^0_{K'}$, and (ii) $I^e(\xi+k)=I^e(\xi)+k$ for all $\xi\in B^0_{K'}$ and $k\in\mathbb{R}$ such that $\xi+k\in B^0_{K'}$. Letting $l(K')$ be the length of the interval K', note that $|k|\le l(K')$ for all $k\in\mathbb{R}$ such that $\xi+k\in B^0_{K'}$ for some $\xi\in B^0_{K'}$. Recall now that $b(x)\in(0,1)$. Hence, we can choose $n_1\in\mathbb{N}$ large enough so that $\gamma(f^n\circ\xi)\in B^0_K$ for all $\xi\in B^0_{K'}$ and all $\gamma\in[\underline{\gamma},1]$ such that $\gamma\xi\in B^0_{K'}$. Also, choose n_2 such that $b(x)^{n_2}l(K')\in O'$. Letting $n:=\max\{n_1,n_2\}$, define

$$I^e(\xi):=f^{-n}I^*(f^n\circ\xi),\quad\forall\xi\in B^0_{K'}.$$

Observe that $I^*(f^n\circ\xi)=f^n(I^*(\xi))$ for every $\xi\in B^0_K$. Thus, I^e is an extension of I^*. Next, take any $\gamma\in[\underline{\gamma},1]$ and $\xi\in B^0_{K'}$ such that $\gamma\xi\in B^0_{K'}$. Recalling (A.11), observe that

$$I^e(\gamma\xi)=f^{-n}I^*\big[f^n\circ(\gamma\xi)\big]=f^{-n}I^*\big[\gamma(f^n\circ\xi)+a(n,\gamma)\big]$$

By the choice of n, we know that $f^n\circ\xi,\gamma(f^n\circ\xi)\in B^0_K$. Since, also, $a(n,\gamma)\in O'$, we have

$$I^*\big[\gamma(f^n\circ\xi)+a(n,\gamma)\big]=I^*\big[\gamma(f^n\circ\xi)\big]+a(n,\gamma)=\gamma I^*\big[f^n\circ\xi\big]+a(n,\gamma).$$

Then, a direct calculation shows that $I^e(\gamma\xi) = \gamma I^e(\xi)$. To show translation invariance, pick any $\xi \in B^0_{K'}$ and $k \in \mathbb{R}$ such that $\xi + k \in B^0_{K'}$. Then,

$$I^e(\xi + k) = f^{-n} I^*\big[f^n \circ (\xi + k)\big] = f^{-n} I^*\big[(f^n \circ \xi) + b^n(x)k\big]$$

By the choice of n, we know that $b^n(x)k \in O'$. Hence,

$$I^*\big[f^n \circ \xi + b^n(x)k\big] = I^*[f^n \circ \xi] + b^n(x)k.$$

Once again, a direct calculation shows that $I^e(\xi + k) = I^e(\xi) + k$. Finally, take a sequence $(K_m)_m$ of compact sets such that $K \subset K_m \subset K_{m+1}$ for all $m \in \mathbb{N}$ and $\cup_m K_m = \mathbb{R}$. Define I_1 to be the extension of I^* to $B^0_{K_1}$ and, inductively, I_m to be the extension of I_{m-1} from $B^0_{K_{m-1}}$ to $B^0_{K_m}$, $m > 1$. Define I^e to be the projective limit of the sequence $(I_m)_m$. By construction, I^e is translation invariant, γ-homogeneous for all $\gamma \in [\underline{\gamma}, 1]$, monotone, and normalized. ■

Letting $f : \mathbb{R} \to \mathbb{R}$ be as in the proof of the lemma, pick $m \in \mathbb{N}$ large enough so that $f^m(U^*(X^\infty)) \subset K$. Let $U^{**} = f^m \circ U^*$ and $\phi^{**}(x,k) = f^m \phi^*(x, f^{-m}(k))$ for all $k \in U^{**}(X^\infty)$. As was shown previously, (U^{**}, ϕ^{**}, I^e) is a representation for $\succeq$. Summarizing the two cases, we have thus proved that $\succeq$ admits a representation (U, ϕ, I) such that U has an Uzawa-Epstein time-aggregator (u, b) and $I : B^0 \to \mathbb{R}$ is a regular certainty equivalent that is translation invariant and $b(x)$-homogeneous for every $x \in X$. The next lemma shows that I is positively homogeneous whenever the function b is not constant. The lemma completes the proof of Theorem 2.

Lemma 16 *If $I : B^0 \to \mathbb{R}$ is γ-homogeneous for all $\gamma \in (a, b) \subset (0, 1)$, where $a < b$, then I is positively homogeneous.*

Proof. It is clear that I is γ-homogeneous for all $\gamma \in (a^t, b^t)$ and all $t \in T$. Observe that $\log_b a > 1$ and pick k such that $1 + \frac{1}{k} < \log_b a$. Then, $b^{t+1} > a^t$ for all $t \geq k$. Conclude that $(0, b^k) \subset \cup_t (a^t, b^t)$ and, hence, that I is γ-homogeneous for all $\gamma \in (0, b^k)$. Suppose now that $I(\gamma a) \neq \gamma I(a)$ for some $\gamma > 0$ and $\xi \in B^0$. Choose $\beta \in (0, b^k)$ and t large enough so that $\beta^t \gamma \in (0, b^k)$. Then, $\beta^t I(\gamma a) \neq \beta^t \gamma I(a) = I(\beta^t \gamma a)$, contradicting the fact that I is $\beta^t\gamma$-homogeneous. ■

A.2 Proof of Theorem 3

Let $f : U(X) \to \hat{U}(X)$ be the continuous, strictly increasing function such that $f(U(d)) = \hat{U}(d)$ for all $d \in X^\infty$. Let $x_0 \in X$ be such that $(x_0, x_0, \ldots)$ attains the minimum of U. Renormalizing if necessary, we can assume that $u(x_0) = 0$. Let E be the set of all points in the interior of $U(X)$ at which f is differentiable. Because f is monotone, E^c has outer measure zero, see e.g. Royden [31, Thm 3, p.100]. For every $x \in X$, $a \in U(X)^\circ$, and $k \in \mathbb{R}$ small enough, we have

$$\frac{f\big[u(x) + b(x)(a+k)\big] - f\big[u(x) + b(x)a\big]}{b(x)k} = \frac{\hat{b}(x)}{b(x)} \frac{f(a+k) - f(a)}{k} \tag{A.12}$$

It follows that if f is differentiable at a, then f is differentiable at $u(x) + b(x)a$ for every x. Moreover, for every $a \in E$,

$$b(x) f'\big[u(x) + b(x)a\big] = \hat{b}(x) f'(a), \quad \forall x \in X, \forall a \in E. \tag{A.13}$$

For every $a \in U(X)$, let

$$A(a) \coloneqq \{u(x) + b(x)a : x \in X\}.$$

The set $A(a)$ is an interval because X is connected and $u, b : X \to \mathbb{R}$ are continuous. Note that $a \in A(a)$ for every $a \in U(X)$. We are going to show that f is differentiable at every $a > 0$. Because E^c has zero measure, for every a' in the interior of $U(X)$, there is $a \in E$ such that $a > a' > 0$. Thus it suffices to show that f is differentiable on $(0, a]$ for every $a \in E, a > 0$. Fix such an a and let $a^n \coloneqq b(x_0)^n a, n \in \{0, 1, 2, ...\}$. For every n, we have

$$a^{n+1} \in A(a^{n+1}) \cap A(a^n). \tag{A.14}$$

Conclude that $\cup_n A(a^n)$ is connected. Since a^n converges to zero, it follows that $(0, a] \subset \cup_n A(a^n)$. Since f is differentiable on every $A(a_n)$, the argument is complete.

For the final step, choose $x \in X$ such that $u(x) > 0$. Plugging that x and $a = 0$ in (A.12), one sees that f has a right derivative at zero. Denoting this derivative by $f'(0)$, note that $f'(0) = f'(u(x))$.

Suppose that $f'(a) = 0$ for some $a \in U(X)$. Since $b(x), \hat{b}(x) > 0$ for all x, it follows from (A.13) that

$$f'\big[u(x) + b(x)a\big] = 0, \quad \forall x \in X.$$

Conclude that $f'(a') = 0$ for all a' in the nondegenerate interval $A(a)$. This contradicts the fact that f is strictly increasing. Conclude that $f'(a) \neq 0$ for all a. For $x \in X$, let $a \coloneqq u(x)(1 - b(x))^{-1}$. Plugging these x, a into (A.13) gives:

$$b(x) f'(a) = \hat{b}(x) f'(a)$$

Since $f'(a) \neq 0$, it must be the case that $b(x) = \hat{b}(x)$ for every $x \in X$. But then (A.13) becomes

$$f'\big[u(x) + b(x)a\big] = f'(a), \quad \forall x \in X, a \in U(X). \tag{A.15}$$

Fix any $a > 0$ amd n, once again define $a^n := b(x_0)^n y$. It follows from (A.15) that f' is constant on each $A(a^n)$. As argued above, $A(a^n) \cap A(a^{n+1}) \neq \varnothing$ for every n and $(0, a] \subset \cup_n A(a^n)$. Thus f' is constant on $(0, a]$ for every a in the interior of $U(X)$. Conclude that f' is constant on the interior and, since f is continuous, that f is affine.

A.3 Proof of Convexity

To be added. Please contact me for details.

References

[1] C. Aliprantis and K. Border. *Infinite Dimensional Analysis*. Springer-Verlag, 2 edition, 1999.

[2] D. K. Backus, B. R. Routledge, and S. E. Zin. Exotic preferences for macroeconomists. In *NBER Macroeconomics Annual*, volume 19, pages 319–390. MIT Press, 2004.

[3] A. Bommier and F. L. Grand. A robust approach to risk aversion. working paper, 2012.

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

[5] L. G. Epstein. Stationary cardinal utility and optimal growth under uncertainty. *Journal of Economic Theory*, 31(1):133–152, 1983.

[6] L. G. Epstein. A simple dynamic general equilibrium model. *Journal of Economic theory*, 41(1):68–95, 1987.

[7] L. G. Epstein. A definition of uncertainty aversion. *The Review of Economic Studies*, 66(3):579–608, 1999.

[8] L. G. Epstein and J. A. Hynes. The rate of time preference and dynamic economic analysis. *The Journal of Political Economy*, 91(4):611–635, 1983.

[9] L. G. Epstein and M. Schneider. Recursive multiple-priors. *Journal of Economic Theory*, 113:1–31, 2003.

[10] L. G. Epstein and M. Schneider. Ambiguity and asset markets. *Annual Review of Financial Economics*, 2:315–346, 2010.

[11] L. G. Epstein and S. Tanny. Increasing generalized correlation: a definition and some economic consequences. *Canadian Journal of Economics*, 13:16–34, 1980.

[12] L. G. Epstein and S. Zin. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica*, 57 (4):937–969, 1989.

[13] P. Ghirardato and M. Marinacci. Risk, ambiguity, and the separation of utility and beliefs. *Mathematics of operations research*, 26(4):864–890, 2001.

[14] P. Ghirardato and M. Marinacci. Ambiguity made precise: A comparative foundation. *Journal of Economic Theory*, 102(2):251–289, 2002.

[15] P. Ghirardato, F. Maccheroni, M. Marinacci, and M. Siniscalchi. A subjective spin on roulette wheels. *Econometrica*, 71:1897–1908, 2003.

[16] P. Ghirardato, F. Maccheroni, and M. Marinacci. Certainty independence and the separation of utility and beliefs. *Journal of Economic Theory*, 120(1):129–136, 2005.

[17] I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18:141–153, 1989.

[18] F. Gul and W. Pesendorfer. Hurwicz expected utility and multiple sources. Mimeo, 2014.

[19] P. Klibanoff, M. Marinacci, and S. Mukerji. A smooth model of decision making under ambiguity. *Econometrica*, 73:1849–1892, 2005.

[20] P. Klibanoff, S. Mukerji, and K. Seo. Relevance and symmetry. 2014. forhcoming in Econometrica.

[21] A. Kochov. Time and no lotteries: An axiomatization of maxmin expected utility. forthcoming in Econometrica, 2014.

[22] T. Koopmans. Stationary ordinal utility and impatience. *Econometrica*, 28(2): 287–309, 1960.

[23] T. C. Koopmans, P. A. Diamond, and R. E. Williamson. Stationary utility and time perspective. *Econometrica*, pages 82–100, 1964.

[24] R. Lucas and N. Stokey. Optimal growth with many consumers. *Journal of Economic Theory*, 32(1):139–171, 1984.

[25] A. Lundberg. Cliff functions-a tool tor treating sequences of composed functions. *Aequationes Mathematicae*, 18:35–53, 1978.

[26] A. Lundberg. Generalized distributivity for real, continuous functions. i: Structure theorems and surjective solutions. *Aequationes Mathematicae*, 24(1):74–96, 1982.

[27] A. Lundberg. Generalized distributivity for real, continuous functions. ii: Local solutions in the continuous case. *Aequationes Mathematicae*, 28(1):236–251, 1985.

[28] F. Maccheroni, M. Marinacci, and A. Rustichini. Ambiguity aversion, robustness and the variational representation of preferences. *Econometrica*, 74:1447–1498, 2006.

[29] M. Machina and D. Schmeidler. A more robust definition of subjective probability. *Econometrica*, 60(4):745–780, 1992.

[30] S. Richard. Multivariate risk aversion, utility independence and separable utility functions. *Management Science*, 22:12–21, 1975.

[31] H. Royden. *Real Analysis*. Macmillan, New York, 3rd edition, 1988.

[32] L. Savage. *The Foundations of Statistics*. Wiley, New York, NY, 1954.

[33] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989.

[34] T. Strzalecki. Temporal resolution of uncertainty and recursive models of ambiguity aversion. *Econometrica*, 81(3):1039–1074, 2013.

[35] H. Uzawa. Time preference, the consumption function, and optimum asset holdings. In *Value, Capital and Growth*. Aldine, Chicago, 1968.