Ambiguity and Endogenous Discounting

Antoine Bommier Asen Kochov François Le Grand*

December 13, 2018

Abstract

Existing work has shown that models of ambiguity aversion as well as expected-utility models of endogenous discounting share an important prediction about intertemporal behavior, namely, that agents are averse to positive autocorrelation in their consumption profile. This paper disentangles the intertemporal predictions of ambiguity aversion from those of endogenous discounting by identifying a form of autocorrelation that is disliked by ambiguity averse agents only. The analysis is supplemented by two representation theorems. The first delivers a novel axiomatization of endogenous discounting without restricting beliefs to be expected utility. The second restricts those beliefs to be of the maxmin form by leveraging our analysis of correlation and ambiguity aversion.

Keywords: Intertemporal Choice, Ambiguity, Correlation Aversion, Endogenous Discounting

*This paper supersedes a working paper by Kochov entitled "Stationary Cardinal Utility," which first circulated in February, 2015. We have benefited from the helpful of comments of John Duggan, Larry Epstein, Massimo Marinacci, and Tomasz Strzalecki. Bommier and Le Grand gratefully acknowledge the financial support of the Swiss Re Foundation and the ETH Zurich Foundation. **Bommier**: ETH Zurich, abommier@ethz.ch; **Kochov**: University of Rochester; asen.kochov@rochester.edu, **Le Grand**: emlyon business school and ETH Zurich, legrand@em-lyon.com.

1 Introduction

A recent paper by Kochov [17] considered the implications of ambiguity aversion for intertemporal behavior. A key lesson is that an ambiguity averse agent would seek to take different, negatively correlated bets in different time periods. As Figure 1 illustrates, doing so reduces the overall uncertainty faced by the agent as it implies that a bad outcome in some period t would be compensated by a good outcome in period $t+1$. Combining such behavior, which Kochov [17] called **intertemporal hedging**, with a notion of stationarity, which we call path stationarity and preview momentarily, Kochov [17] axiomatized the following dynamic version of the maxmin model of Gilboa and Schmeidler [15]:

$$V(c_0, c_1, \ldots) = \min_{p \in \mathcal{P}} \mathbb{E}_p[u(c_0) + \beta u(c_1) + \beta^2 u(c_2) \ldots], \tag{1.1}$$

where, as usual, $\mathcal{P}$ is a set of beliefs over the spate space Ω, representing the agent's perception of ambiguity.

A limitation of Kochov's analysis is that it depends critically on a third assumption which is evident from (1.1), namely, that the ranking of nonstochastic consumption streams be time separable. Without this auxiliary assumption, the behavior in Figure 1 need no longer be indicative of ambiguity aversion. To see this, consider the Uzawa [31] and Epstein [7] model of endogenous discounting,[1] which relaxes time separability by allowing the rate of time preference to vary with the consumption path:

$$V(c_0, c_1, \ldots) = \mathbb{E}_p[u(c_0) + b(c_0)u(c_1) + b(c_0)b(c_1)u(c_2) \ldots]. \tag{1.2}$$

It is known from Epstein [7] that this model, which has found applications in the theory of optimal growth and the study of small open economies, exhibits intertemporal hedging if and only if b is a decreasing function of the consumption level c. The latter property, known as **increasing marginal impatience**, is assumed in virtually all applications of the model as it insures the stability and uniqueness of steady states.

The main goal of this paper is to disentangle the intertemporal implications of ambiguity aversion from those of endogenous discounting. We do so by focusing on a special kind of uncertainty that can arise in a dynamic setting. Imagine an agent expecting a tax refund. The agent knows the amount to be refunded and plans to

[1] Uzawa [31] studied endogenous discounting in a setting with no uncertainty. Later, Epstein [7] showed how to incorporate uncertainty. Backus et al. [2] and Epstein and Hynes [10] survey applications of the model.

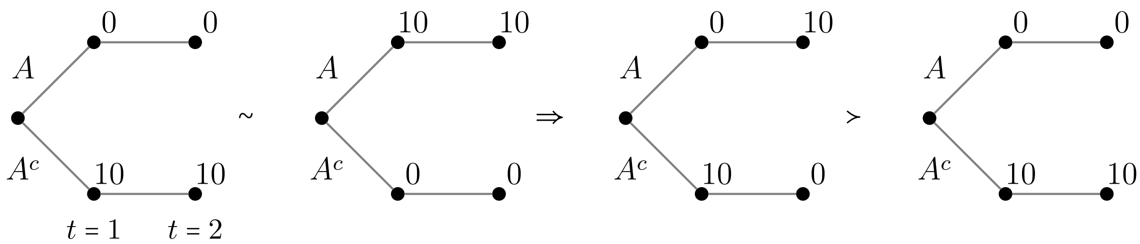

Figure 1: An Illustration of Intertemporal Hedging. The first ranking indicates that the agent views the event A and its complement as equally likely. Given that, the second ranking posits that the agent prefers consumption to be negatively rather than positively autocorrelated. This ranking could be strict under a model of ambiguity aversion, but not under the standard time-additive, expected-utility model. Intuitively, while negative correlation reduces overall uncertainty, it implies a consumption profile that is not smooth over time, which, in the context of the standard model, perfectly offsets the benefits from reducing uncertainty.

consume it as soon as the refund arrives. The uncertainty is *when* the tax refund will arrive. We show that seeking an intertemporal hedge against this special type of uncertainty, *about the timing of consumption within a given span of time*, is indicative of ambiguity aversion whether discounting is exogenous as in (1.1) or endogenous as in (1.2). In particular, such behavior cannot be rationalized by the expected utility model in (1.2).

The analysis is supplemented by two representation theorems. The first one concerns a general class of preferences which includes the models in (1.1) and (1.2) as special cases. The defining property of this class is an axiom which we call **path stationarity** and which is assumed in both Kochov [17] and Epstein [7]. The axiom extends Koopmans' classical notion of stationarity to a setting of uncertainty by positing the following implication. Consider an event A resolving in period t and note that A may affect contemporaneous consumption as well as consumption in a more distant period $t+k$. Path stationarity requires that the agent's attitudes toward uncertainty do not depend on the date on which consumption takes place and, in particular, on k.[2] This seemingly innocuous restriction on behavior turns out to be remarkably powerful. Its first implication is that the utility of a non-stochastic

[2]Epstein [7] and Kochov [17] refer to path stationarity simply as stationarity. As we explain in Section 5.1, we adopt a different name so as to distinguish path stationarity from another extension of Koopmans' axiom.

consumption stream is computed as in (1.2):

$$U(c_0, c_1, ...) = u(c_0) + b(c_0)u(c_1) + b(c_0)b(c_1)u(c_2) + ... \tag{1.3}$$

This implication sets the stage for our analysis of ambiguity aversion vis-á-vis endogenous discounting, which requires the existence of such a utility function. The second implication of path stationarity concerns the ranking of stochastic consumption streams. As the lifetime utility of any such stream could be random, the agent needs to assign an expectation $I(\xi) \in \mathbb{R}$ to each random variable $\xi : \Omega \to \mathbb{R}$. We show that the mapping $\xi \mapsto I(\xi)$, which we call the agent's **certainty equivalent**, must be translation invariant and positively homogeneous, which means that for all $\xi : \Omega \to \mathbb{R}, k \in \mathbb{R}, \alpha \in \mathbb{R}_+$,

$$I(\xi + k) = I(\xi) + k \qquad \text{and} \qquad I(\alpha\xi) = \alpha I(\xi). \tag{1.4}$$

Certainty equivalents of this form are known as **invariant biseparable** and have been studied extensively in the context of static choice under ambiguity. See Ghirardato et al. [14] and the references therein. In contrast, we characterize such certainty equivalents in terms of path stationarity, a property of intertemporal behavior.

Combining path stationarity with our notion of intertemporal hedging, our second result shows that the certainty equivalent I is also concave and, hence, takes the maxmin form

$$I(\xi) = \min_{p \in P} \mathbb{E}_p \xi. \tag{1.5}$$

Thus, we not only identify a more robust prediction of ambiguity aversion, we are also able to put this prediction to use and generalize Kochov's [17] characterization of the maxmin model.[3]

2 Choice Setting and Some Definitions

Time is discrete and varies over an infinite horizon: $t \in \{0, 1, 2, ...\} =: T$. Uncertainty is modeled by a filtered space $(\Omega, \{\mathcal{F}_t\}_t)$ where Ω is an arbitrary set of state of the world and $\{\mathcal{F}_t\}_t =: \mathcal{F}$ is a filtration, i.e., an increasing sequence of algebras such that $\mathcal{F}_0 = \{\Omega, \varnothing\}$. As usual, we interpret the algebra $\mathcal{F}_t$ to be the collection of all events

[3] An interesting aspect of Kochov's [17] analysis is that, unlike many other studies of ambiguity aversion, it does not require the existence of events with known, objectively given probabilities. This is true of the present analysis as well.

that resolve before or in period t. We assume that consumption outcomes lie in a compact, connected, and separable topological space X.[4] A stochastic consumption stream is then an X-valued and $\mathcal{F}$-adapted process, that is, a sequence $h = (h_0, h_1, ...)$ such that $h_t : \Omega \to X$ is $\mathcal{F}_t$-measurable for every t. As is standard in decision theory, we also refer to a process h as an **act**. An act $h = (h_0, h_1, ...)$ is **deterministic** if each function $h_t : \Omega \to X$ is constant, that is, if outcomes do not depend on the state of the world. We use $d, d' \in \mathcal{H}$ to denote such acts and, abusing notation, identify them with elements of X^∞. Given an algebra $\mathcal{F}' \subset \cup_t \mathcal{F}_t$, an act h is $\mathcal{F}'$**-adapted** if h_t is $\mathcal{F}'$-measurable for every t. An act h is **finite** if there is a finite algebra $\mathcal{F}' \subset \cup_t \mathcal{F}_t$ such that h is $\mathcal{F}'$-adapted. To avoid technical complications, we take **the choice domain** to be the space $\mathcal{H}$ of all finite acts.

We also let B^0 be the space of all simple and $\cup_t \mathcal{F}_t$-measurable functions $\xi : \Omega \to \mathbb{R}$. We refer to such functions as **random variables** and endow B^0 with the sup norm. Given a set $C \subset \mathbb{R}$, we use B^0_C to denote the set of all C-valued functions in B^0. Abusing notation, we 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 function $I : B^0 \to \mathbb{R}$ 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 $\alpha \in \mathbb{R}$, 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**. Endowing B^0 with the usual pointwise order, a function $I : B^0 \to \mathbb{R}$ is **increasing** if for all $\xi, \xi' \in B^0$, $\xi \geq \xi'$ implies that $I(\xi) \geq I(\xi')$. A normalized, increasing and norm-continuous function $I : B^0 \to \mathbb{R}$ is called a **certainty equivalent**. In later sections, we use certainty equivalents to model the individual's "beliefs" and think of $I(\xi)$ as "the expected value" assigned by the agent to the random variable $\xi \in B^0$. We use $\Delta(\Omega)$ to denote the space of all finitely additive probability measures p on the measurable space $(\Omega, \cup_t \mathcal{F}_t)$ and endow $\Delta(\Omega)$ with the weak* topology, i.e., the coarsest topology such that for every $\xi \in B^0$, the linear function $p \mapsto \mathbb{E}_p \xi$ from $\Delta(\Omega)$ into $\mathbb{R}$ is continuous.

Finally, a **preference relation** $\succeq$ on a set Y is a complete and transitive binary relation such that $y \succ y'$ for some $y, y' \in Y$. If Y is a topological space, then $\succeq$ is **continuous** if the upper and lower contour sets, $\{y' \in Y : y' \succeq y\}$ and $\{y' \in Y : y \succeq y'\}$, are closed for every $y \in Y$.

[4] Compactness is not essential. See Kochov [17, p.240] for details. Connectedness is needed for our representation theorems but not for the formulation of the axioms.

3 Path Stationary Preferences

This section introduces the class of path stationary preferences mentioned in the introduction, derives a representation for this class, and discusses the uniqueness of the representation. The analysis serves as a backdrop for our study of intertemporal hedging in Section 4, but may also be of independent interest as it generalizes several existing results concerning path stationary preferences and endogenous discounting.

3.1 Axioms

Let $\succeq$ be a preference relation $\succeq$ on the space $\mathcal{H}$ of finite acts. The first restriction we impose on $\succeq$ is a form of continuity familiar from Ghirardato and Marinacci [13]. It weakens the usual notion of topological continuity by allowing the agent's beliefs to be representable by finitely additive, as opposed to countably additive, probability measures, which is common in axiomatic work. To state the axiom, endow $\mathcal{H}$ with the product topology.

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

The next axiom posits that the tastes of the agent (how he evaluates deterministic consumption streams) are not influenced by the state of the world. Roughly, this means that if some deterministic act $d \in X^\infty$ is preferred to $d' \in X^\infty$ *unconditionally*, then d is also preferred to d' *conditional* on any event $A \in \cup_t \mathcal{F}_t$. As is well known from Kreps [22, p.108] and others, this requirement facilitates the measurement of the agent's beliefs vis-á-vis his tastes. To state the axiom formally, let $dA_t h$ be the act obtained from h by replacing all outcomes after period t and in states $\omega \in A$ with the respective outcomes of d. That is, given $h \in \mathcal{H}, d \in X^\infty, t \in T$, and $A \in \mathcal{F}_t$, $dA_t h$ is the act $g \in \mathcal{H}$ such that $g_k(\omega) = d_k$ for all $\omega \in A$ and $k \geq t$, and $g_k(\omega) = h_k(\omega)$ otherwise.

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 $A \in \mathcal{F}_t$ such that if $h_k = d_k = d'_k$ for all $k \leq t-1$ and $d \succ d'$, then $dA_t h \succ d' A_t h$ and $dA^c_t h \succ d' A^c_t h$.

The second part of State Independence posits the existence of an event $A \in \cup_t \mathcal{F}_t$ such that both A and A^c preserve strict as well as weak rankings. The existence of such events, called **essential** in the literature, is a mild technical requirement. In the context of the expected utility model in (1.2), an event A is essential if and only if $p(A) \in (0,1)$, that is, if neither A nor A^c are assigned probability zero. In the context of the maxmin model in (1.1), an event A is essential if and only if $p(A) \in (0,1)$ for every $p \in \mathcal{P}$.

The next axiom, Path Stationarity, is the main restriction we impose in this section. Consider an event A resolving in period t and note that A may affect contemporaneous consumption as well as consumption in a more distant period $t+k$. The axiom requires that the agent's attitude toward uncertainty do not depend on the delay k. To state it formally, let (x,h) be 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 $h \in \mathcal{H}$. That is, (x,h) is obtained from h by postponing the consumption date of each outcome by one period and inserting x in period $t = 0$.

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

Path Stationarity extends Koopmans' [19] classical notion of stationarity from the space X^∞ of deterministic acts to a setting that incorporates uncertainty. We should highlight, however, that Path Stationarity is not the only such extension. In particular, note that we work with a domain in which all acts are measurable with respect to a fixed filtration $\mathcal{F}$. This means that when we transform an act h into (x,h), we are postponing the dates on which the outcomes of h are consumed, but not the dates on which the relevant uncertainty resolves. If one were to postpone both the timing of consumption and the timing of resolution of uncertainty, one would obtain a different extension of Koopmans' axiom. Figure 3 in Section 5.1 illustrates the difference between the two extensions, both of which have appeared in the literature.

To simplify the exposition, from now on we say that a preference relation $\succeq$ on $\mathcal{H}$ is **path stationary** if it satisfies FC, SI, and PS. Similarly, a preference relation $\succeq$ on X^∞ is **stationary** if it is continuous in the product topology on X^∞ and stationary in the sense of Koopmans [19]. It would also be helpful to highlight two additional axioms satisfied by the class of path stationary preferences. The first is **History Independence**, which is implied by Path Stationarity and requires that for all $x, y \in X$ and $h, g \in \mathcal{H}$,

$$(x,h) \succeq (x,g) \quad \text{if and only if} \quad (y,h) \succeq (y,g).$$

The other axiom, **Monotonicity**, is implied by the conjunction of Path Stationarity and State Independence. Letting $h(\omega) := (h_0(\omega)h_1(\omega),\ldots) \in X^\infty$ be the consumption stream delivered by an act $h \in \mathcal{H}$ in state ω, the axiom requires that for all acts $h, g \in \mathcal{H}$,

$$h \geq g \quad \text{whenever} \quad h(\omega) \geq g(\omega) \text{ for every } \omega \in \Omega.$$

Thus, h is preferred to g whenever h gives a better consumption stream in every state of the world. A noteworthy implication of Monotonicity is that

$$h \sim g \quad \text{whenever} \quad h(\omega) \sim g(\omega) \text{ for every } \omega \in \Omega. \tag{3.1}$$

As formalized by equation (3.2) below and Lemma 6 in the appendix, (3.1) permits the construction of a utility representation in which the agent evaluates stochastic consumption streams by first computing lifetime utility state by state and then computing the expectation of lifetime utility, using some certainty equivalent I. Many models of intertemporal choice, including the ones in (1.1) and (1.2), are constructed in this manner. To understand what (3.1) entails on a behavioral level, suppose an act h delivers consumption that is uncertain only in a single period t, whereas g delivers consumption that is uncertain in more than one period. According to (3.1), such differences in the intertemporal distribution of uncertainty do not matter as long as h and g deliver equally desirable consumption streams in every state $\omega \in \Omega$. We refer the reader to Bommier et al. [3, p.1438] where this implication of Monotonicity is discussed in much greater detail and where we contrast Monotonicity with another axiom commonly used to construct intertemporal utility functions, namely, Recursivity.[5]

3.2 A Representation Theorem

Let (U, I) be a pair consisting of a continuous, nonconstant function $U : X^\infty \to \mathbb{R}$ and a certainty equivalent $I : B^0 \to \mathbb{R}$. For every act $h \in \mathcal{H}$, let $U \circ h \in B^0$ be the random variable $\omega \mapsto U(h(\omega))$ representing the **lifetime utility** induced by the act h and let

$$V(h) := I(U \circ h). \tag{3.2}$$

[5]The reader may notice that, abstracting from the existence of an essential event, Monotonicity is strictly stronger than State Independence and wonder why did not impose Monotonicity from the start and skip State Independence. Our goal is to highlight that, conditional on assuming State Independence, which as we noted earlier facilitates the measurement of beliefs, Monotonicity cannot be relaxed without relaxing Path Stationarity as well. In particular, the reader who finds the conclusions of Theorem 1 in Section 3.2 too strong should know that they cannot be escaped by relaxing Monotonicity while keeping Path Stationarity.

The pair (U, I) **represents** a preference relation $\succeq$ on $\mathcal{H}$ if the function $V : \mathcal{H} \to \mathbb{R}$ is well-defined and represents $\succeq$. If the function $U : X^\infty \to \mathbb{R}$ takes the form

$$U(x_0, x_1, ...) = u(x_0) + b(x_0)u(x_1) + b(x_0)b(x_1)u(x_2) + ...$$

for some continuous functions $u : X \to \mathbb{R}$ and $b : X \to (0, 1)$, we say that U an **Uzawa-Epstein utility function** and write (u, b) to denote it. Similarly, if a preference relation $\succeq$ on X^∞ admits an Uzawa-Epstein utility function, we say that $\succeq$ is an **Uzawa-Epstein preference relation**. Finally, a certainty equivalent $I : B^0 \to \mathbb{R}$ is **regular** if there is some event $A \in \cup_t \mathcal{F}_t, A \notin \{\Omega, \varnothing\}$, such that I is strictly increasing when restricted to the space of $\{A, A^c\}$-measurable functions $\xi \in B^0$. Regularity is needed to account for the existence of an essential event A, which was posited by State Independence.

Theorem 1 *A preference relation $\succeq$ on $\mathcal{H}$ is path stationary if and only if it has a representation (U, I) such that $U : X^\infty \to \mathbb{R}$ is an Uzawa-Epstein utility function (u, b) and the certainty equivalent I is regular, translation-invariant and $b(x)$-homogeneous for every $x \in X$. Furthermore, if $b(x) \neq b(y)$ for some $x, y \in X$, then I is positively homogeneous.*

Theorem 1 generalizes two existing results on path stationary preferences. Kochov [17] deduces the same restrictions on the certainty equivalent I, but under the additional assumption of **Future Independence**, which posits that for all $d, d' \in X^\infty$ and $x, y, x', y' \in X$,

$$(x, y, d) \succeq (x', y', d) \quad \text{if and only if} \quad (x, y, d') \succeq (x', y', d'). \tag{3.3}$$

In words, the choice of consumption in the first two periods is independent from the common continuation stream d. It is known from Koopmans [20] that given a stationary preference relation on X^∞, this axiom implies a standard time-additive utility function $U(x_0, x_1, ...) = \sum_t \beta^t u(x_t)$. Hence, the axiom rules out any intertemporal complementarities in the agent's tastes and forces the rate of time preference to be exogenous. On the other hand, Epstein [7] delivers an Uzawa-Epstein utility function $U : X^\infty \to \mathbb{R}$, but under the additional assumption of expected utility, which in the present context means having a representation (U, I) such that $I(\xi) = \mathbb{E}_p \xi$ for some belief $p \in \Delta(\Omega)$.

We note that our ability to deduce restrictions on the utility function U and the certainty equivalent I *simultaneously*, without making any assumptions other than Path Stationarity, requires techniques that are quite different from those in Kochov [17] and Epstein [7]. Section 5.2 gives an outline of the proof. There, we also compare Theorem 4 with our main result in Bommier et al. [3], which employed similar techniques.

3.3 Uniqueness of the Representation

This section studies the uniqueness of the representation derived in Theorem 1. We begin by considering a preference relation $\succeq$ on X^∞ with two Uzawa-Epstein utility functions.

Theorem 2 *Suppose a preference relation $\succeq$ on X^∞ has two Uzawa-Epstein utility functions:*

$$U(x_0, x_1, x_2, ...) = u(x_0) + b(x_0)u(x_1) + b(x_0)b(x_1)u(x_2) + ...$$

$$\hat{U}(x_0, x_1, x_2, ...) = \hat{u}(x_0) + \hat{b}(x_0)\hat{u}(x_1) + \hat{b}(x_0)\hat{b}(x_1)\hat{u}(x_2) + ...$$

Then, $b = \hat{b}$ and $U = \alpha\hat{U} + \gamma$ for some $\alpha \in \mathbb{R}_{++}, \gamma \in \mathbb{R}$.

Theorem 2 generalizes a well-known result by Koopmans [20] which delivers the same conclusions but in the special case of exogenous discounting, that is, when the function $b : X \to (0, 1)$ is constant.[6] Our theorem also generalizes a result by Epstein [7] which delivers the same conclusions but in a richer setting in which the preference relation $\succeq$ on X^∞ is the restriction of some expected-utility preference $\succeq'$ on the space of lotteries over X^∞. In such a setting, the cardinal uniqueness of $U : X^\infty \to \mathbb{R}$ follows directly from the von-Neuman-Morgenstern theorem. Using this, Epstein [7] goes on to prove that the function $b : X \to (0, 1)$ is unique. In contrast, we deduce both conclusions without the use of lotteries and the concomitant assumption of expected utility.[7]

Introducing uncertainty once again, our next result builds on Theorem 2 to show that the certainty equivalent I derived in Theorem 1 is unique whenever discounting is endogenous. To state the result, let (u, b, I) denote the representation obtained in Theorem 1.

Corollary 3 *Suppose a path stationary preference relation $\succeq$ on $\mathcal{H}$ has two representations (u, b, I) and $(\hat{u}, \hat{b}, \hat{I})$ as in Theorem 1. From Theorem 2, we know that $b = \hat{b}$. If the function $b : X \to (0, 1)$ is nonconstant, then $I = \hat{I}$.*

Theorem 4 in the next section delivers another instance in which the certainty equivalent I is unique. There, instead of making assumptions about discounting, we impose a notion of intertemporal hedging that restricts the certainty equivalent I to be concave.

[6] In this special case, the cardinal uniqueness of $U : X^\infty \to \mathbb{R}$ translates into the cardinal uniqueness of $u : X \to \mathbb{R}$.

[7] To illustrate the uniqueness of $b : X \to (0, 1)$, assume that $X \subset \mathbb{R}$ and that the function U is suitably differentiable. Then, given a constant stream $(x, x, ...) \in X^\infty$, Epstein [7] shows that $b(x)^{-1}$ is equal to the marginal rate of substitution between any two periods t and $t + 1$.

4 Intertemporal Hedging

This section tackles our main goal of disentangling the intertemporal implications of ambiguity aversion from those of endogenous discounting. We begin by recalling Kochov's [17] notion of intertemporal hedging and its connection to ambiguity aversion.

Intertemporal Hedging Against Uncertainty in Levels [IH-L]: For all $t \in T, d \in X^\infty$, and $h, g \in \mathcal{H}$,

$$(d_{-(t,t+1)}, h_t, h_t) \sim (d_{-(t,t+1)}, g_t, g_t) \quad \text{implies that} \quad (d_{-(t,t+1)}, h_t, g_t) \succeq (d_{-(t,t+1)}, h_t, h_t).$$

Figure 1 in the introduction illustrated the axiom by abstracting from the deterministic consumption in periods $k \neq t, t+1$.[8] As noted there, the choice between $(d_{-(t,t+1)}, h_t, g_t)$ and $(d_{-(t,t+1)}, h_t, h_t)$ can be understood as a choice between negative and positive autocorrelation. The benefit from negative autocorrelation is that it reduces the overall uncertainty faced by the agent; the benefit from positive autocorrelation is that it smoothes consumption over time. By choosing the act $(d_{-(t,t+1)}, h_t, g_t)$, the agent reveals that reducing uncertainty is of greater concern to him. Assuming a representation (U, I) with a time-additive utility function $U(x_0, x_1, ...) = \sum_t \beta^t u(x_t)$,[9] Kochov [17] showed that this concern translates into ambiguity aversion. Indeed, computing the lifetime utilities induced by the three acts in IH-L, we see that

$$U \circ (d_{-(t,t+1)}, h_t, g_t) = \frac{1}{1+\beta} \underbrace{U \circ (d_{-(t,t+1)}, h_t, h_t)}_{\xi} + \frac{\beta}{1+\beta} \underbrace{U \circ (d_{-(t,t+1)}, g_t, g_t)}_{\xi'}. \tag{4.1}$$

Consequently, IH-L can be rewritten as the implication:

$$I(\xi) = I(\xi') \quad \Rightarrow \quad I\Big(\frac{1}{1+\beta}\xi + \frac{\beta}{1+\beta}\xi'\Big) \geq I(\xi). \tag{4.2}$$

The above is nothing else but the requirement that the certainty equivalent I be quasiconcave. One may therefore view IH-L as an analogue of Schmeidler's [30]

[8] That this is w.l.o.g. follows from the assumption of a time-additive utility function $U : X^\infty \to \mathbb{R}$ made in Kochov [17].

[9] Recall from Section 3.2 that such a utility function can be deduced by imposing Future Independence in addition to Path Stationarity.

definition of ambiguity aversion, which delivers the same restriction on I but in a different domain.[10]

Unfortunately, equation (4.1) and, hence, the connection between IH-L and ambiguity aversion break down when U is not time-additive. Moreover, some path stationary preferences may exhibit a strict desire for intertemporal hedging, at least in some circumstances, even when the certainty equivalent I is of the expected utility form.

Example 1 (IH-L with Uzawa-Epstein Utility) *Assuming $X \subset \mathbb{R}$, consider the example in Figure 1 but replace 10 and 0 with two arbitrary monetary outcomes $x > y$.*[11] *The lifetime utility of a deterministic consumption stream $(x_t, x_{t+1}) \in X^2$ is given by $U(x_t, x_{t+1}) = u(x_t) + b(x_t)u(x_{t+1})$, while the certainty equivalent I is of the expected utility form, that is, $I(\xi) = \mathbb{E}_p(\xi)$. Under these assumptions, the first ranking in Figure 1 implies that $p(A) = p(A^c) = 0.5$. In turn, the second ranking holds if and only if $(b(x) - b(y))(u(y) - u(x)) \geq 0$. If $u : X \to \mathbb{R}$ is a strictly increasing function, we see that a strict preference for intertemporal hedging arises if and only if $b : X \to (0,1)$ is a strictly decreasing function. The latter assumption, known as **increasing marginal impatience**, is common in applied work. See Lucas and Stokey [23], Epstein [8], and Backus et al. [2]. What matters presently, however, is that a desire to hedge intertemporally arises not because of a stronger concern about uncertainty, but because of an intertemporal complementarity in the ranking of deterministic consumption streams, i.e., the assumption of increasing marginal impatience, which may be viewed as lessening the desire to smooth consumption across time.*

Example 1 shows that to pin down intertemporal implications of ambiguity aversion, we must find a way to bypass the taste complementarities associated with endogenous discounting. Interestingly, this goal takes us through a notion of impatience due to Koopmans [19].

Impatience: For all $t > 0, a, b \in X^t$, and $d \in X^\infty$,

$$(a, a, a, ...) \geq (b, b, b, ...) \quad \text{if and only if} \quad (a, b, d) \geq (b, a, d).$$

[10] A more direct analogy between IH-L and Schmeidler's [30] definition of ambiguity aversion is made in Kochov [17, p.242]. Note as well that there are definitions of ambiguity aversion which are neither implied nor imply quasiconcavity of the certainty equivalent. See Epstein [9] and Ghirardato et al. [14] for details.

[11] The formal, infinite-horizon version of IH-L reduces to the two-period example above by choosing d to be a constant stream $(z, z, ...)$ such that $u(z) = 0$.

The axiom posits that given two finite strings of outcomes, the agent wants to consume the better one first. Koopmans [19] conjectured that all stationary preferences on X^∞ have this property. While this conjecture was proved wrong by Koopmans et al. [21], Epstein [7] observed that Impatience holds for all Uzawa-Epstein preferences on X^∞. From Impatience, we deduce that for all $t > 0, a, b \in X^t$, and $d, d' \in X^\infty$,

$$(a, b, d) \geq (b, a, d) \quad \text{if and only if} \quad (a, b, d') \geq (b, a, d'). \tag{4.3}$$

To see why this implication is important to us, recall from Section 3.1 that path stationary preferences are history independent. It follows from the discussion of Future Independence in Section 3.2 that any intertemporal complementarity in the agent's tastes must take the form of a *future dependence*, as is the case when (3.3) is violated. But (4.3) shows that when the agent is asked about the order in which he wants to consume a and b, there are no such future dependencies. Namely, his choice is independent of the continuation stream d. This suggests the following notion of intertemporal hedging. Once again, the agent would seek to take different "bets" in different time periods. However, the uncertainty he would try to hedge would no longer concern the *level* of consumption in a given time period t, as was the case with IH-L, where the "bet" g_t taken in $t+1$ can be viewed as a hedge against the "bet" h_t taken in period t. Instead, the agent would try to hedge uncertainty about the *order* in which a given set of outcomes is consumed within a given set of periods. Since in these situations there are no intertemporal complementarities that stem from the agent's tastes, the behavior must be driven by the agent's attitudes toward uncertainty.

Figure 2 gives an example in which time spans a period of two years, with each year consisting of a fall and a spring semester. A payment of *\$10* is received in each year but the exact timing, fall or spring semester, depends on the realization of an event A. On the left, the timing of delivery in year 1 is perfectly negatively correlated with the timing of delivery in year 2. On the right, the correlation is positive. Assuming the events A and A^c are equally likely, Figure 2 indicates a preference for negative over positive correlation. As with IH-L, the benefit of negative correlation is that it reduces overall uncertainty, by insuring an early payment in at least one of the two years.

To formalize these ideas, fix some $t > 1$ 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 $\pi a = (x_{\pi(0)}, x_{\pi(1)}, ..., x_{\pi(t-1)}) \in X^t$ be the corresponding permutation of a. Say that $h \in \mathcal{H}$ is a **repeating permutation act (rp-act)** if there is some $t \in T, a \in X^t$, and permutations $\pi_\omega : \{0, 1, ..., t-1\} \to \{0, 1, ..., t-1\}$, one for each $\omega \in \Omega$, such that

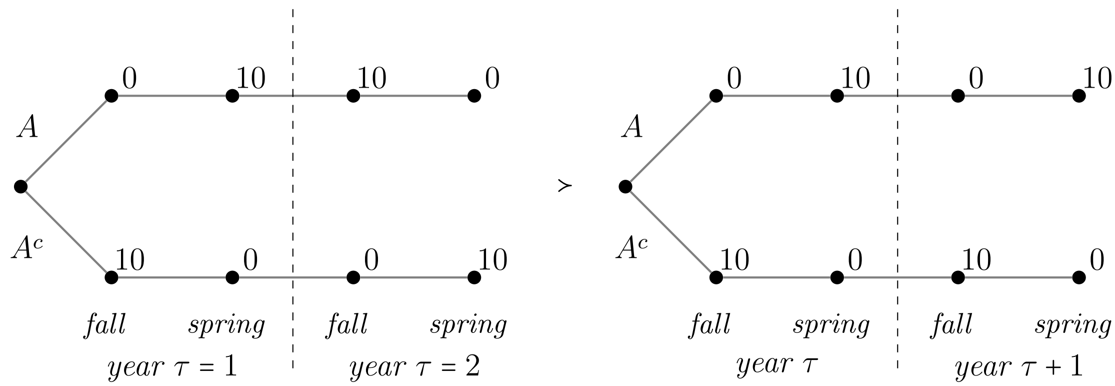

Figure 2: Intertemporal hedging when the relevant uncertainty concerns the timing of an outcome within a block of periods. The negatively correlated act on the left reduces overall uncertainty by guaranteeing an early payment in at least period. The positively correlated act on the right "smooths" consumption over time by delivering the same pattern of yearly consumption in each year. As in Figure 1, the ranking indicates that reducing uncertainty is more important.

$h(\omega) = (\pi_\omega a, \pi_\omega a, ...)$ for every $\omega \in \Omega$. The definition has three features: i) time is partitioned into blocks of equal length, ii) within each block, the only thing uncertain is the order in which the elements of a are consumed, iii) this uncertainty is perfectly positively correlated across blocks. As in Figure 2, an agent confronted with an rp-act may choose to reverse the positive correlation. Let $h, g \in \mathcal{H}$ be two rp-acts such that $h(\omega) = (\pi_\omega a, \pi_\omega a, ...)$ and $g(\omega) = (\tilde{\pi}_\omega a, \tilde{\pi}_\omega a, ...)$ for every $\omega \in \Omega$. If $h \sim g$, intertemporal hedging means that the agent would prefer the act $m \in \mathcal{H}$ such that $m(\omega) = (\pi_\omega a, \tilde{\pi}_\omega a, \tilde{\pi}_\omega a, ...)$ for every $\omega \in \Omega$. Such an act m appears on the left hand side of Figure 2. Below, we state a stronger axiom which allows one of the acts h, g to be arbitrary. This facilitates the proof of Theorem 4 and brings about a stronger prediction that we can associate with ambiguity aversion.

Intertemporal Hedging against Uncertainty in Timing [IH-T]: For every $g \in \mathcal{H}, t > 1, a \in X^t$, and rp-act $h \in \mathcal{H}$ such that $h(\omega) = (\pi_\omega a, \pi_\omega a, ...)$ for every ω, let $m \in \mathcal{H}$ be the act such that $m(\omega) = (\pi_\omega a, g(w))$ for every ω. If $h \succeq g$, then $m \succeq g$.

The next theorem formalizes the connection between IH-T and ambiguity aversion by showing that, within the class of path stationary preferences, IH-T delivers a certainty equivalent I of the maxmin form.

Theorem 4 *A path stationary preference relation* $\succeq$ *on* $\mathcal{H}$ *satisfies IH-T if and only if it has a representation* (u, b, I) *such that*

$$I(\xi) = \min_{p \in P} \mathbb{E}_p \xi$$

for some weak**-closed and convex set* P *of probability measures on* $(\Omega, \cup_t \mathcal{F}_t)$. *Moreover, the set* P *is unique.*

To sketch a proof of Theorem 4, let (u, b, I) be the representation obtained in Theorem 1 and observe that the utility of a consumption stream $(x, y, d') \in X^\infty$ can be written as

$$U(x, y, d') = (1 - b(x)b(y))U(x, y, x, y, ...) + b(x)b(y)U(d'). \tag{4.4}$$

Thus, $U(x, y, d')$ is a convex combination of the utility of the initial block (x, y), repeated ad infinitum, and the utility of the continuation stream d'. Similarly, the utility of a stream $(y, x, d'') \in X^\infty$ is equal to

$$U(y, x, d'') = (1 - b(x)b(y))U(y, x, y, x, ...) + b(x)b(y)U(d''). \tag{4.5}$$

Comparing (4.4) and (4.5), notice that the weights in the two convex combinations are exactly the same. With this in mind, let $h, g, m \in \mathcal{H}$ be acts as in the statement of IH-T and note that the various consumption paths $m(\omega) \in X^\infty$ induced by the act $m \in \mathcal{H}$ take a form similar to (x, y, d') and (y, x, d''), though the initial block could be of any length $t \geq 2$. The utility of each path $m(\omega)$ can thus be written as a suitable convex combination, in a manner that parallels (4.4) and (4.5), and, in addition, the weights in all these combinations are independent of ω. Letting $a = (x_0, x_1, ..., x_{t-1}) \in X^t$ be the list of elements used to define h and letting $b(a) := \Pi_{k=0}^{t-1} b(x_k)$, it follows that

$$U \circ m = (1 - b(a))[U \circ h] + b(a)[U \circ g]. \tag{4.6}$$

Thus, the lifetime utility $U \circ m$ induced by m is a convex combination of $U \circ h$ and $U \circ g$. This is the precise sense in which we previously said that the act m "reduces overall uncertainty." Letting $\xi := U \circ h, \xi' := U \circ g, \gamma := b(a)$, we also see that IH-T can be written as the implication

$$I(\xi) \geq I(\xi') \quad \Rightarrow \quad I\Big((1 - \gamma)\xi + \gamma\xi'\Big) \geq I(\xi'). \tag{4.7}$$

Thus, once again, intetemporal hedging is seen to be equivalent to the quasiconcavity of I and may be viewed as an analogue of Schmeidler's [30] definition of ambiguity aversion.

One caveat is that the quasiconcavity asserted in (4.7) is limited in two ways. First, the random variable ξ is not arbitrary, but restricted to be the lifetime utility $U \circ h$ of some rp-act $h \in \mathcal{H}$. Second, one cannot pick *any* weight $\gamma \in (0,1)$ to form a convex combination of ξ and ξ'. Instead, the weight γ is tied to ξ via the identity $\gamma = b(a)$, where $a = (x_0, ..., x_{t-1})$ is the list of outcomes used to construct h. The main challenge in the proof of Theorem 4 is to overcome these limitations and show that IH-T is in fact equivalent to the full-blown quasiconcavity of I. We do so by proving two structural results about rp-acts. The first is that they are dense in the space of all acts. The second and more challenging result is that the space of lifetime utilities $U \circ h$ attainable by such acts (or a suitable normalization thereof) contains an open set.[12]

Finally, it is worth highlighting that the derivation of (4.6) and (4.7) does not require the full force of Path Stationarity: we need U to be of the Uzawa-Epstein form, but do not need the positive homogeneity or scale invariance of I. Thus, if U is of the Uzawa-Epstein form, a quasiconcave I would always imply IH-T, so that, in particular, a strict desire to hedge intertemporally can be interpreted as a sign of ambiguity aversion. This is important since some well-known models of ambiguity aversion feature quasiconcave certainty equivalents I that are not positively homogeneous or translation-invariant. Among others, this is true of the variational model of Maccheroni et al. [27] and the model of Cerreia-Vioglio et al. [4]. On the other hand, we do require the full force of Path Stationarity to deduce that IH-T implies the full-blown quasiconcavity of I.

5 More On Path Stationarity

In this final section, we offer some additional remarks on path stationarity, the proof of Theorem 1 and its connection to our results in Bommier et al. [3].

5.1 Two Notions of Stationarity

Figure 3 depicts two distinct ways in which one can extend Koopmans' notion of stationarity from deterministic to stochastic environments. The top part depicts Path Stationarity, while the bottom part depicts another extension for which we reserve the name Stationarity. In both parts, the relevant uncertainty is the outcome

[12] An interesting implication of these results is that virtually any trade-off that can be formulated in the space of lifetime utilities can be construed as arising from uncertainty about the timing of consumption as captured by rp-acts.

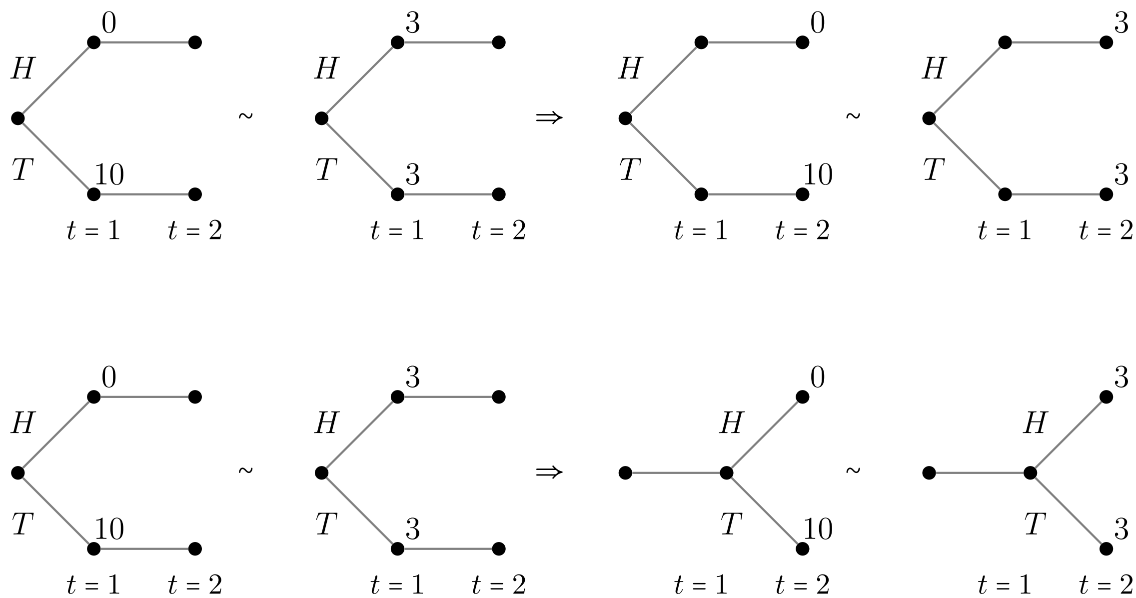

Figure 3: The top part of the figure depicts Path Stationarity. The bottom part depicts Stationarity. In both parts, the relevant uncertainty is the outcome of a coin toss. For the sake of simplicity, it is assumed that all consumption takes place in a single period.

of a coin toss. Consistent with the explanation in Section 3.1, Path Stationarity considers a situation in which the coin is always flipped in period $t = 1$, but one changes the date on which consumption takes place. The axiom requires that those changes do not affect the individual's attitudes toward uncertainty. The second extension, Stationarity, considers a situation in which one simultaneously changes the date on which the coin is flipped and the date on which consumption takes place. In particular, this is done so that the time distance between the two dates is kept unchanged.

The two extensions concern choice situations that are qualitatively different. Stationarity is of interest when the underlying uncertainty repeats itself. In such a case, an individual may confront the same uncertainty and the same decision problem at different moments in time, with Stationarity implying that the individual would make the same decision. In particular, Stationarity implies that choices are history-independent and time-consistent, an implication that is at the heart of *stationary dynamic programming*. On the other hand, Path Stationarity is of interest when the impact of an event can be delayed over time. Consider an agent who has to choose between two jobs whose career trajectories will be affected by some event A. The agent has to commit to a job prior to the realization of the event, but can opt to start immediately or delay the start date until after a one-month vacation. Path Stationarity says that the agent would take the same job whether he goes on vacation or not.

Interestingly, both extensions have appeared in the literature under the single name Stationarity. Thus, Epstein [7] and Kochov [17] call Stationarity what we call Path Stationarity. It is more typical, however, to reserve the name Stationarity for the other notion in Figure 3, as is done in Epstein and Zin [12] and Chew and Epstein [5], among others. To avoid confusion, we propose different names for the two extensions.

5.2 Theorem 1: A Proof Sketch

Theorem 1 in this paper uses techniques from Lundberg's [25] work on functional equations. In this section, we highlight the connection to Lundberg's work as well as compare Theorem 1 with our results in Bommier et al. [3], where we used similar techniques. To begin, recall from Koopmans [19] that a preference relation $\succeq$ on X^∞ is stationary if and only if every continuous utility function $U : X^\infty \to \mathbb{R}$ of $\succeq$ satisfies the recursion

$$U(x_0, x_1, x_2, ...) = \phi(x_0, U(x_1, x_2, ...)) \quad \forall (x_0, x_1, ...) \in X^\infty. \tag{5.1}$$

The function $\phi : X \times U(X^\infty) \to U(X^\infty)$ above, called **a time aggregator**, is strictly increasing in the second argument and continuous. With this in mind, if a preference relation $\succeq$ on $\mathcal{H}$ has representation (U, I) such that U can be written recursively as in (5.1), it is helpful to make ϕ explicit and write (U, ϕ, I) instead of (U, I). Given a $U(X^\infty)$-valued random variable $\xi \subset B^0$ and an outcome $x \in X$, let $\phi(x, \xi) \in B^0$ be the random variable $\omega \mapsto \phi(x, \xi(\omega))$. Take an act $(x, h) \in \mathcal{H}$, as in the statement of Path Stationarity, and consider the equalities:

$$V(x, h) = I(\phi(x, U \circ h)) = \phi(x, I(U \circ h)) \quad \forall x \in X, h \in \mathcal{H}. \tag{5.2}$$

The first equality combines the definition of V with (5.1). The interesting equality is the second one. It says that there are two ways to compute the utility of an act (x, h). The expression $I(\phi(x, U \circ h))$ means that one first aggregates utility across time and then across states. Conversely, the expression $\phi(x, I(U \circ h))$ means that one first computes the expectation $I(U \circ h)$ of future utility and then aggregates across time by adding the utility of the initial outcome x. When these two computations agree, as in (5.2), we say that the certainty equivalent I and the time aggregator ϕ **permute**. The next lemma shows that this is true if and only if Path Stationarity holds.

Lemma 5 *A preference relation $\succeq$ on $\mathcal{H}$ is path stationary if and only if it has a representation (U, ϕ, I) such that the time aggregator ϕ and the certainty equivalent I permute. The latter is true for all representations (U, ϕ, I) of a path stationary preference relation $\succeq$.*

Lemma 5 is the bridge to Lundberg's work. Thus, when $|\Omega| = 2$, the second equation in (5.2), viewed as a functional equation to be solved for ϕ and I, becomes what Lundberg calls a *distributivity equation*. He shows that the solutions to this equation are *well behaved locally*, by which we mean that, after a suitable monotone transformation of utility, and after restricting the functions ϕ and I to a suitable open set, they have the properties sought after in Theorem 1. Lundberg's work leaves us with three problems to address. First, we have to verify a technical restriction on the functions ϕ and I required by Lundberg. Second, we have to extend Lundberg's local solution to obtain a representation for the entire preference relation $\succeq$ on $\mathcal{H}$. Third, we have to extend Lundberg's analysis to state spaces Ω of arbitrary cardinality. The last of these problems requires care but is not the most insightful part of our proof. More interesting is the way we deal with the other problems. Regarding the first, we are able to show that Lundberg's restriction is satisfied whenever the functions ϕ and I are part of a representation (U, ϕ, I) of a path stationarity preference relation on $\mathcal{H}$. Intuitively, the infinite horizon implies that the time aggregator ϕ encodes a form of discounting, which turns out to be sufficient for Lundberg's restriction. For details, the reader should check Lemma 7 in the appendix and how this lemma is used in Section A.1.4.

To get a sense of how we deal with the local nature of Lundberg's solutions, let $O \subset \mathbb{R}$ be an open set such that the desired representation obtains whenever the acts $h, h' \in \mathcal{H}$ are such that $U \circ h, U \circ h'$ are O-valued. Fix some $x^* \in X$ such that $U(x^*, x^*, ...) \in O$ and take two *arbitrary* acts $h, h' \in \mathcal{H}$. From Path Stationarity, deduce that

$$h \succeq h' \Leftrightarrow (x^*, h) \succeq (x^*, h') \Leftrightarrow (x^*, x^*, h) \succeq (x^*, x^*, h') \Leftrightarrow \dots \tag{5.3}$$

and note that, as we increase the number of initial periods in which x^* is consumed, the acts in (5.3) converge to $(x^*, x^*, x^*, ...)$. Hence, their utilities are eventually contained in the set O in which the functions (U, ϕ, I) are well behaved. But the equivalences in (5.3) show that the restrictions of these functions to the set O represent the entire preference relation.

Lundberg's results were also used in our companion paper, Bommier et al. [3], where we assumed Monotonicity, Recursivity, and Stationarity (in the stochastic sense of Section 5.1). Comparing the two papers, the first thing to notice is that the preferences they characterize overlap, but aren't nested. Most notably, in our companion paper we do not assume Path Stationarity, with the consequence that the certainty equivalent I need not be positively homogeneous. As a result, the preferences in Bommier et al. [3] include the multiplier preferences of Hansen et al. [16] and the more general variational preferences of Maccheroni et al. [28], while these preferences are presently excluded. On the other hand, in this paper we do not assume recursivity which, in the context of Theorem 4, means that the set $\mathcal{P}$ of beliefs need not have the "rectangular structure" identified by Epstein and Schneider [11]. Given our focus on ambiguity aversion, this may be desirable since, as Epstein and Schneider [11] highlight, recursivity comes at the cost of ruling out some natural forms of ambiguity-averse behavior as well as some common specifications of the set $\mathcal{P}$ of beliefs.

Mathematically, the results in Bommier et al. [3] are more complicated in that, there, we had to solve a system of distributivity equations $I_t(\phi(x,\xi)) = \phi(x, I_{t+1}(\xi))$, $t \in T$, rather than a single equation with a single certainty equivalent I, as in (5.2). On the other hand, the lack of Path Stationarity meant that we had deal with the local nature of Lundberg's solutions in a different way. We did so by making the auxiliary assumption that X is an interval ($X \subset \mathbb{R}$) and that preferences are increasing in the implied pointwise order on X^∞. This assumption, which we called **Deterministic Monotonicity**, implies that Lundberg's local solutions are in fact global. The assumption also led to a more direct, but less general, proof of Lundberg's technical restriction.[13]

A Appendix

Given functions $f : X' \to Y'$ and $g : Y' \to Z'$, we use $g \circ f$ and gf interchangeably to denote the composition of f and g. Given an interval $C \subset \mathbb{R}$, B^0_C denotes the sets of all C-valued functions $\xi \in B^0$. Given a finite algebra $\mathcal{F}' \subset \cup_t \mathcal{F}_t$, $B^0(\mathcal{F}')$ is the set of all $\mathcal{F}'$-measurable functions $\xi \in B^0$. The set $B^0_C(\mathcal{F}') \subset B^0(\mathcal{F}')$ is similarly defined. A

[13] There is still a significant overlap between the two proofs. Given the complexity of both proofs however and the noted differences, we chose to provide another self-contained proof rather than ask the reader to piece together two different papers. Note, in particular, that our present proof, including the expository discussion of iteration groups in Section A.1.3, highlights the local nature of Lundberg's solutions, which wasn't needed in the earlier paper.

function $I : B^0_C \to \mathbb{R}$ is **finite continuous** if for every finite algebra $\mathcal{F}' \subset \cup_t \mathcal{F}_t$, the restriction of I to $B^0_C(\mathcal{F}')$ is norm-continuous.

A.1 Proof of Theorem 1

Necessity of the axioms is obvious. We prove sufficiency.

A.1.1 Preliminaries

Lemma 6 *State Independence and Path Stationarity imply Monotonicity.*

Proof. Let $h, h' \in \mathcal{H}$ be such that $h(\omega) \succeq h'(\omega)$ for every ω. Because h, h' are finite, 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 PS, $(a, h(\omega)) \succeq (a, h'(\omega))$ for every $\omega \in \Omega$. Moreover, $h \succeq h'$ if and only if $(a, h) \succeq (a, h')$, so it suffices to show the latter. Think of $(a, h), (a, h')$ as functions from Ω into X^∞. Since h, h' are finite, there is a coarsest partition $\{A_1, A_2, ..., A_n\}$ of Ω with respect to which both these functions are measurable.[14] Replace the infinite stream of (a, h') on A_1 by the respective infinite stream of (a, h). By SI, 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) and apply SI again. After n such steps, we see that $(a, h) \succeq (a, h')$. ■

Lemma 7 *Consider a path stationary preference relation $\succeq$ on $\mathcal{H}$. Then,*

1. *There are $x, y \in X, d \in X^\infty$ such that $(x, d) \succ (y, d)$.*
2. *For every $x \in X, h \in \mathcal{H}$, $(x, x, ...) \succeq h$ if and only if $(x, h) \succeq h$. Similarly, $h \succeq (x, x, ...)$ if and only if $h \succeq (x, h)$.*
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 $(z, d') \sim (x_n, z, d)$.*

[14] Measurability with respect to a partition means measurability with respect to the algebra generated by the partition.

Proof. Property (1) is proved in Kochov [17, Lemma 5]. Properties (2) and (3) can be proved as in Kochov [18, Lemmas 3.3-3.4]. Turn to (4). By way of contradiction, suppose $(z, z, d^*) \succeq (z^*, z, d^*)$. Since $d^* \succeq (z, d^*)$ and $\succeq$ is path stationary, we obtain

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

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

$$(z, z, d^*) \succeq (z^*, z, d^*) \succeq (z^*, z^*, z, d^*). \tag{A.1}$$

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

$$(z^*, z^*, z, d^*) \succeq (z^*, z^*, z^*, z, d^*). \tag{A.2}$$

Combining (A.1) and (A.2), we get

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

Repeating the argument gives $(z, z, d^*) \succeq ((z^*)^n, z, d^*)$ for every $n \in \mathbb{N}$. By FC , $(z, z, d^*) \succeq d^*$, contradicting property (2). Finally, turn to property (5). From property (2), we know that $(z^*, z, z, ...) \succ (z, z, ...)$. Because X is connected, there is a sequence $(x_n)_n$ in X converging to z such that $(x_n, z, z, ...) \succ (z, z, ...)$. By property (2) again, $(z, d^*) \succ (z, z, d^*)$. 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 construction, it is also the case that

$$(x_n, z, d^*) \succeq (x_n, z, z, z, ...) \succ (z, z, z, ...) \quad \forall n \in \mathbb{N}.$$

Combining the last two observations gives $(z, d^*) \succ (x_n, z, d^*) \succ (z, z, z, ...)$ for all $n \geq N$. By PS, it is also the case that

$$(x_n, z, d^*) \succeq (x_n, z, d) \succeq (x_n, z, z, ...) \succ (z, z, ...) \quad \forall d \in X^\infty, n \in \mathbb{N}.$$

Summing up, we have

$$(z, d^*) \succ (x_n, z, d) \succ (z, z, z, ...) \quad \forall n \geq N, d \in X^\infty.$$

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

A.1.2 Proof of Lemma 5

From now on, we adopt a more permissive notion of a **representation** (U, ϕ, I) according to which the certainty equivalent I is defined on $B^0_{U(X^\infty)}$ rather than the entire space B^0, and I is only finite continuous rather than norm-continuous. Using this notion, this section proves an analogue of Lemma 5. The proof of the lemma as stated in the main text would follow once we complete the proof of Theorem 1. See, e.g., Lemmas 16 and 23.

It is obvious that the existence of a representation such that I and ϕ permute implies Path Stationarity. To prove the opposite direction, assume that $\succeq$ is path stationary. Because X^∞ is connected and separable in the product topology, we know from Debreu [6] that there is a continuous function $U : X^\infty \to \mathbb{R}$ representing the restriction of $\succeq$ to X^∞. For every $x \in X$ and $k \in U(X^\infty)$, choose some $d_k \in X^\infty$ such that $U(d_k) = k$ and define $\phi(x, k) := U(x, d_k)$. Since U is continuous, ϕ is continuous in its first argument. Because the restriction of $\succeq$ to X^∞ is stationary, ϕ is strictly increasing in the second argument. Since U is continuous, the set $\phi(x, U(X^\infty))$ is connected for every $x \in X$. Conclude that ϕ is continuous in the second argument and, ultimately, that ϕ is a time aggregator for U. Turn to the construction of a certainty equivalent $I : B^0_{U(X^\infty)} \to \mathbb{R}$. Since X^∞ is connected, a standard argument shows that for every act $h \in \mathcal{H}$, there is an act $d_h \in X^\infty$ such that $h \sim d_h$. Extend U from X^∞ to $\mathcal{H}$ by letting $V(h) := U(d_h)$. Recall that $U \circ h$ denotes the function $\omega \mapsto U(h(\omega))$ and 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 letting $I(U \circ h) = V(h)$. By Lemma 6, $\succeq$ satisfies Monotonicity. It follows that I is well-defined and increasing. Because $\succeq$ satisfies FC, I is finite-continuous. By definition, $I(k) = k$ for all $k \in U(X^\infty)$, that is, I is normalized. Altogether, (U, ϕ, I) is a representation of $\succeq$.

The final step is to show that for all representations (U, ϕ, I) of $\succeq$, 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 Path Stationarity, $(x, d) \sim (x, h)$. Since the certainty equivalent I is normalized,

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

A.1.3 Iteration Groups

We need to introduce some mathematical concepts from Lundberg [25]. Let $C \subset \mathbb{R}$ be a nonempty, open interval and λ an extended real number in $\mathbb{R}_{++} \cup \{+\infty\}$. Let $\{g^\alpha : \alpha \in (-\lambda, \lambda)\}$ be a family of functions such that each function g^α is defined on an interval $C^\alpha \subset C$ and $g^\alpha(C^\alpha) \subset C$. Suppose each function g^α is continuous and strictly increasing, and its graph disconnects the Cartesian product C^2.[15] Suppose further that the graph of $g^\alpha \circ g^\beta$ is a subset of the graph of $g^{\alpha+\beta}$, with the latter holding 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** C. When such a group $\{g^\alpha : \alpha \in (-\lambda, \lambda)\}$ is given, we assume that the group is **maximal**, which means that there is no other iteration group $\{\tilde{g}^\alpha : \alpha \in (-\tilde{\lambda}, \tilde{\lambda})\}$ over C such that $\lambda < \tilde{\lambda}$ and $g^\alpha = \tilde{g}^\alpha$ for all $\alpha \in (-\lambda, \lambda)$. When no confusion arises, we may also suppress the interval C and the bound λ, and speak simply of an iteration group $\{g^\alpha\}$.

Given an iteration group $\{g^\alpha\}$ over C, g^0 is the identity function on C, which we denote as j. More generally, if α is an integer, then g^α is the α-iterate of the function g^1, with the domain of the α-iterate restricted so that the range does not exceed the set C. One can thus view an iteration group $\{g^\alpha\}$ as a way of defining the α-iterate of a function g^1 for all α, integer or not. Note as well that an iteration group $\{g^\alpha : \alpha \in (-\lambda, \lambda)\}$ over C induces an iteration group over any nonempty, open interval $D \subset C$. In particular, let $\hat{\lambda}$ be the supremum of all $\alpha \in (-\lambda, \lambda)$ such that the graph of $g^\alpha|_D$ intersects D^2. Then, $\{g^\alpha|_D : \alpha \in (-\hat{\lambda}, \hat{\lambda})\}$ is an iteration group over D.

Example 2 *Let $C = \mathbb{R}$, $\lambda = +\infty$. Let $f(k) = a + bk, a, b \in \mathbb{R}$ be an affine function defined for all $k \in \mathbb{R}$. If $b \neq 1$, there is a unique iteration group $\{g^\alpha\}$ such that $g^1 = f$. In fact, we can compute each function g^α explicitly:*

$$g^\alpha(k) = a\frac{1 - b^\alpha}{1 - b} + b^\alpha k. \tag{A.3}$$

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 over C 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.

An iteration group $\{g^\alpha\}$ over C is **fixed-point-free** if none of the functions g^α, $\alpha \neq 0$, has a fixed point. As noted in Lundberg [25], we can then find a so-called **Abel function** $L : C \to \mathbb{R}$ for the iteration group, which means that $g^\alpha(k) = L^{-1}(L(k) + \alpha)$ for all k in the domain of g^α and all $\alpha \in (-\lambda, \lambda)$. For future reference, observe that if L is an Abel function for $\{g^\alpha\}$, then so is the function $L + c$ where c is an arbitrary real number. Also, 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

[15] The graph of g^α disconnects C^2 if for every $(x, y), (x', y') \in C^2$ such that $x, x' \in C^\alpha$ and $y > g^\alpha(x), y' < g^\alpha(x')$, a continuous curve in C^2 that connects (x, y) to (x', y') must intersect the graph of g^α.

$\tilde{g}^\alpha := g^{-\alpha}$ for every $\alpha \in (-\lambda, \lambda)$. Under this assumption, any Abel function L for $\{g^\alpha\}$ is strictly increasing.

Going back to our example, suppose D is an interval such that all functions g^α, $\alpha \neq 0$, are fixed-point-free when restricted to D. If, for instance, we take $D = (k^*, +\infty)$, then $L(k) := \log_b(k - k^*)$ is an Abel function for the group $\{g^\alpha|_D\}$ on D.

Given a topological space $\mathcal{Z}$ and a set $A \subset \mathcal{Z}$, we use A° to denote the topological interior of A. 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 [25], a correspondence $f^* : C \rightrightarrows \mathbb{R}$, where C is an interval in $\mathbb{R}$, is called a **cliff function** if the set $f^*(k)$ is connected for every $k \in C$. A cliff function f^* is **increasing** if $k \leq k'$ and $l \in f^*(k), l' \in f^*(k')$ imply $l \leq l'$ 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.

A.1.4 Constructing an Iteration Group

Returning to the proof of Theorem 1, take a path stationary preference relation $\succeq$ on $\mathcal{H}$ with a representation (U, ϕ, I). Let $A \in \cup_t \mathcal{F}_t$ be an essential event and w.l.o.g. assume that $A \in \mathcal{F}_1$. Let $\mathcal{H}(A)$ be the subset of acts $h \in \mathcal{H}$ that are $\{A, A^c\}$-adapted and let $U \circ \mathcal{H}(A) := \{U \circ h : h \in \mathcal{H}(A)\}$ be the set of random variables $U \circ h$ generated by such acts. Think of $U \circ \mathcal{H}(A)$ as a subset of $\mathbb{R}^2$ and of I as a function on that subset. Let $C := \{U(z, d) : d \in X^\infty\}$ and note that C is a closed interval in $\mathbb{R}$ with nonempty interior C°. By SI and FC, I is a strictly increasing and continuous function on C^2.

We want to apply Theorem 4.16 in Lundberg [25] to find an iteration group $\{g^\alpha : \alpha \in (-\lambda, \lambda)\}$ over C° such that

$$g^\alpha I(k, k') = I(g^\alpha(k), g^\alpha(k')), \quad \forall \alpha \in (-\lambda, \lambda), \forall k, k' \in C^\circ. \tag{A.4}$$

The sufficient condition given in Lundberg [25] is that there is a sequence of real-valued functions g^n whose domains $\mathrm{Dom}\, g_n$ are contained in C and such that i) an analogue of (A.4) holds for the functions g^n, ii) $g_n \neq j$, iii) $\mathrm{Dom}\, g_n \to_L C$, and iv) $g_n \to_L j$.[16]

To that end, let $(x_n)_n$ be a sequence in X satisfying property (5) in Lemma 7. For every $k \in C$ and $n \in \mathbb{N}$, let $f_n(k) := \phi(x_n, k)$. Also, let $f(k) := \phi(z, k), k \in C$. By the

[16] This is the technical restriction we mentioned in Section 5.2.

choice of $(x_n)_n$, $f_n(C) \subset C$ for every n and $f(C) \subset C$. The next lemma shows that the time aggregator ϕ is uniformly continuous in its first argument. The technical proof, which invokes several results from Lundberg [25], can be skipped without loss of continuity.

Lemma 8 $f_n \to_L f$.

Proof. Because $\geq$ is continuous, $(f_n)_n$ converges pointwise to f. To prove the stronger form of convergence, identify each function f_n with an increasing cliff function. By Lundberg [25, Lemma 1.1], there is a subsequence $(f_{n_m})_m$ and a cliff function f^* such that $f_{n_m} \to_L f^*$. We wish to show that $f^* = f$ so that, in particular, f^* is a proper function. By Lemma 5, we know that ϕ and I permute, which implies that

$$f_{n_m} I(k, k') = I(f_{n_m}(k), f_{n_m}(k')) \quad \forall m \in \mathbb{N}, \forall k, k' \in C. \tag{A.5}$$

It follows from Lundberg [25, Lemma 4.8] that

$$f^*(I(k, k')) = \{I(l, l') : l \in f^*(k), l' \in f^*(k')\} \quad \forall k, k' \in C, \tag{A.6}$$

and from Lundberg [25, Lemma 4.7] that f^* is a proper function. Finally, it follows from Lundberg [25, Lemma 1.2] that $(f_{n_m})_m$ converges to f^* uniformly on all compact subsets A of the interior of C. But then, $f^*(k) = f(k)$ for all $k \in C^\circ$ and, since f, f^* are both continuous functions, that $f = f^*$. Since the convergent subsequence $(f_{n_m})_m$ was arbitrary, we are done. ■

For every n, let $g_n := f^{-1} \circ f_n$. By construction, an analogue of (A.4) holds for the functions g_n. By the choice of x_n, $g_n \neq j$. From Lemma 8 and Lundberg [24, Thm 5.3], deduce that $g_n \to_L j$ and, from Lundberg [24, Lemma 3.9], that $\operatorname{Dom} g_n \to_L C$. Thus, we can apply Lundberg [25, Theorem 4.16] and deduce that the sequence $(g_n)_n$ generates the sought-after iteration group $\{g^\alpha : \alpha \in (-\lambda, \lambda)\}$ over C°, where by generating we mean that for every $\alpha \in (-\lambda, \lambda)$ there is a sequence $(p_n)_n$ of integers such that $g_n^{p_n} \to_L g^\alpha$

The next lemma shows that an analogue of (A.4) continues to hold when I is viewed as a function of any C°-valued random variable $\xi \in B^0$, not only the binary ones.

Lemma 9 $I(g^\alpha \circ \xi) = g^\alpha I(\xi)$ *for all* $\alpha \in (-\lambda, \lambda)$ *and* $\xi \in B^0_{C^\circ}$.

Proof. 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 $\xi \in B^0_{C^\circ}$. Fix $\xi \in B^0_{C^\circ}$ and $\alpha \in (-\lambda, \lambda)$, and let $\mathcal{F}'$ be a finite algebra on Ω such that ξ is $\mathcal{F}'$-measurable. Because the sequence $(g_n)_n$ generates the iteration group $\{g^\alpha\}$, for every α there is a sequence $(p_n)_n$ of integers such that g^α is the closed limit of the sequences $(g_n^{p_n})_n$. Moreover, $g^\alpha \circ \xi$ and each function $g_n^{p_n} \circ \xi$ are $\mathcal{F}'$-measurable. The desired equality follows since I is finite-continuous and each function $g_n^{p_n}$ is continuous. ■

A.1.5 Constructing an Abel Function

The next step is to find a nonempty open interval $O \subset C^0$ on which the iteration group $\{g^\alpha : \alpha \in (-\lambda, \lambda)\}$ is fixed point free. The key is Theorem 4.13 in Lundberg [25] which shows that if an equation like (A.4) holds, then each $g^\alpha, \alpha \neq 0$, has at most one fixed point.

Lemma 10 *There exists a nonempty, open interval $O \subset C^\circ$ such that none of the functions $g^\alpha|_O, \alpha \neq 0$, has a fixed point.*

Proof. Let $C^\alpha \subset C$ be the domain of the function g^α. By Lundberg [25, Theorem 4.13], each function $g^\alpha, \alpha \neq 0$, has at most one fixed point. If none of the functions $g^\alpha, \alpha \neq 0$, has a fixed point, we can let $O = C^\circ$. Suppose instead that for some $\alpha^* \neq 0$, the function g^{α^*} has a fixed point $k^{\alpha^*} \in C^\circ$. Because C^{α^*} is an interval, there is $\varepsilon > 0$ such that either $(k^{\alpha^*} - \varepsilon, k^{\alpha^*}) \subset C^\alpha$ or $(k^{\alpha^*}, k^{\alpha^*} + \varepsilon) \subset C^\alpha$. Suppose the latter is true. An analogous argument applies to the other case. It is enough to show that there is $\varepsilon' \in (0, \varepsilon)$ such that no function $g^\alpha, \alpha \neq 0$, has a fixed point when restricted to the interval $(k^{\alpha^*}, k^{\alpha^*} + \varepsilon')$. If not, we can find a sequence $(\alpha_n)_n, \alpha_n \neq 0$, such that each function g^{α_n} has a fixed point k^{α_n} and $k^{\alpha_n} \searrow k^{\alpha^*}$. But then, $k^{\alpha_n} \in C^{\alpha^*}$ for all n large enough. From the properties of an iteration group, see Section A.1.3, we know that if some $g^\alpha, \alpha \neq 0$, has a fixed point k^α, then k^α is a fixed point of all other functions $g^{\alpha'}$ that are defined at k^α. Conclude that g^{α^*} has countably many fixed points, k^{α^*} and k^{α_n} for all n large enough, contradicting the fact that g^{α^*} has a unique fixed point. ■

As observed in Section A.1.3, the iteration group $\{g^\alpha : \alpha \in (-\lambda, \lambda)\}$ over C° induces an iteration group $\{g^\alpha|_O : \alpha \in (-\hat{\lambda}, \hat{\lambda})\}$ over O. Since the latter group is fixed point free, it has an Abel function $L : O \to \mathbb{R}$. As was further explained in Section A.1.3, it is w.l.o.g. to assume that L is strictly increasing. We summarize these observations in the next lemma.

Lemma 11 *The iteration group $\{g^\alpha|_O : \alpha \in (-\hat{\lambda}, \hat{\lambda})\}$ has an Abel function $L : O \to \mathbb{R}$ which is strictly increasing.*

A.1.6 A Local Abel Function is Enough

We plan to use the Abel function L to monotonically transform the representation (U, ϕ, I) into a new, more tractable representation. The problem is that L is defined only on a subinterval O of the range $U(X^\infty)$ of all possible utility levels. To address this problem, the next lemma uses Path Stationarity to "scale down" the representation (U, ϕ, I) into another representation whose utility levels are contained in O. Moreover, this is done without changing the certainty equivalent I, thus preserving the connection between I, the iteration group $\{g^\alpha\}$, and the Abel function L, established in Lemmas 9 and 11. The lemma formalizes a key step of the proof discussed in Section 5.2.

Lemma 12 *$\succeq$ has a representation $(\hat{U}, \hat{\phi}, \hat{I})$ such that $\hat{U}(X^\infty) =: D \subset O$ and $\hat{I}$ is equal to the restriction of I to B^0_D.*

Proof. 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 7, that $\{U(x, x, ...) : x \in X\} = U(X^\infty)$. Thus, we can find $x_0 \in X$ such that $U(x_0, x_0, ...) \in O$. 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 $f_0^N(U(X^\infty)) \subset O$. In addition, since f_0 is strictly increasing, f_0^N is strictly increasing. Thus, the function $\hat{U} := f_0^N \circ U : X^\infty \to \mathbb{R}$ represents the restriction of $\succeq$ to X^∞ and its range is contained in the set O. Moreover, if we let $\hat{\phi}(x, s) := f_0^N \phi(x, f_0^{-N}(s))$ for every $x \in X, s \in \hat{U}(X^\infty)$, then $\hat{\phi}$ is a time aggregator for $\hat{U}$. Let $\hat{I}$ be the restriction of I to B^0_D, where $D := \hat{U}(X^\infty) \subset O \subset C$. It remains to show that $(\hat{U}, \hat{\phi}, \hat{I})$ is a representation for $\succeq$. Let $x_0^N \in X^N$ be the N-dimensional vector, each coordinate of which is equal to $x_0 \in X$. For all $h, h' \in \mathcal{H}$, Path Stationarity implies that

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

completing the proof. ■.

A.1.7 A Monotone Transformation of Utility

As noted in Section A.1.3, if L is an Abel function for some iteration group, then so is the function $L + c$ where c is an arbitrary real number. Hence, we can assume that $0 \in L(D)^\circ$. Note that $L(D)^\circ$ is nonempty since $D^\circ = [\hat{U}(X^\infty)]^\circ$ is nonempty and L is strictly increasing. Now, use L to construct a monotone transformation of the representation $(\hat{U}, \hat{\phi}, \hat{I})$:

$$\begin{aligned}
\tilde{U} &:= L \circ \hat{U}, \\
\tilde{\phi}(x,s) &:= L\hat{\phi}(x, L^{-1}(s)) \quad \forall x \in X, s \in L(D), \\
\tilde{I}(\xi) &:= LI(L^{-1} \circ \xi) \quad \forall \xi \in B^0_{L(D)}.
\end{aligned}$$

By construction, $(\tilde{U}, \tilde{\phi}, \tilde{I})$ is a representation of $\succeq$. The next lemma shows that the certainty equivalent $\tilde{I} : B^0_{L(D)} \to \mathbb{R}$ is translation-invariant.

Lemma 13 *$\tilde{I}(\xi + \alpha) = \tilde{I}(\xi) + \alpha$ for all $\xi \in B^0_{L(D)}, \alpha \in (-\hat{\lambda}, \hat{\lambda})$ such that $\xi + \alpha \in B^0_{L(D)}$.*

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

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

The first equality follows from the definitions of $\tilde{I}$ and ξ', and the fact that $\hat{I}$ is the restriction of I to B^0_D. The second equality follows since L is an Abel function for the group $\{g^\alpha|_O : \alpha \in (-\hat{\lambda}, \hat{\lambda})\}$ and $D \subset O$. The third equality follows because, by Lemma 9, I and g^α permute. The fourth equality uses again the fact that L is an Abel function. The final two equalities follow from the definitions of ξ' and $\tilde{I}$ respectively. ■

A.1.8 A Functional Equation

In this section, we use the translation invariance of $\tilde{I}$ to deduce a functional equation which is solved in Lundberg [26]. Once again, 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

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

Lemma 14 *The function ψ is continuous and strictly increasing; $\psi(0) = 0$; $\psi(k) < k$ for all $k > 0$, while $\psi(k) > k$ for all $k < 0$. Finally, the function $k \mapsto \psi(k) - k$ is strictly decreasing.*

Proof. ψ is continuous and strictly increasing since $\tilde{I}$ is continuous and strictly increasing. Since $\tilde{I}(k,k) = k$ for all $k \in [c,c']$, $\psi(0) = 0$. To see that $k \mapsto \psi(k) - k$ is a strictly decreasing function, pick $k \in [0, c'-c)$ and $\varepsilon > 0$ such that $k + \varepsilon \in [0, c'-c]$. Then,

$$\psi(k+\epsilon) - \psi(k) - \varepsilon = \tilde{I}(c, c+k+\varepsilon) - \tilde{I}(c+\varepsilon, c+k+\varepsilon) < 0.$$

Conclude that $\psi(k) - k$ is strictly decreasing on $[0, c'-c]$. A similar argument for $k \in [c-c', 0]$ shows that $\psi(k) - k$ is strictly decreasing on its entire domain. ■

Since, by Lemma 13, $\tilde{I}$ is translation-invariant,

$$\tilde{I}(s,t) = s + \psi(t-s) \quad \forall s,t \in [c,c']. \tag{A.7}$$

For every $x \in X$, write $\tilde{f}_x$ for the function $\tilde{\phi}(x,\cdot)$ on $[c,c']$. Using (A.7), the fact that $\tilde{I}$ and $\tilde{\phi}$ permute implies that

$$\tilde{f}_x(s + \psi(t-s)) = \tilde{f}_x(s) + \psi(\tilde{f}_x(t) - \tilde{f}_x(s)) \quad \forall x \in X, s,t \in [c,c']. \tag{A.8}$$

The above is a special case of the functional equation

$$g(s + \psi(t-s)) = g(s) + \psi(g(t) - g(s)) \quad \forall s,t \in [c,c'], \tag{A.9}$$

which is solved in Lundberg [26] when the function g is continuous and strictly increasing and ψ has the properties listed in Lemma 14. Thinking of ψ as a known function and of g as a solution to (A.9), we break the remainder of the proof in two cases, depending on whether there is a solution g that is affine on some subinterval of its domain.

A.1.9 The Affine Case

Suppose that (A.9) is satisfied for some g that is affine on a subinterval of $[c,c']$. It follows from Lundberg [26, Thm 10.1,10.3] that all solutions g of (A.9) are affine on $[c,c']$. In particular,

$$\tilde{f}_x(k) = u(x) + b(x)k \quad \forall x \in X, k \in [c,c']. \tag{A.10}$$

Because $(\tilde{U}, \tilde{\phi}, \tilde{I})$ is a representation for $\succeq$ and $\succeq$ is continuous on X^∞, it follows 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 7 implies that $b(x) < 1$ for all $x \in X$. Thus, $\tilde{U} : X^\infty \to \mathbb{R}$ is an Uzawa-Epstein utility.

Turn to the certainty equivalent $\tilde{I} : B^0_{L(D)} \to \mathbb{R}$. By Lemma 13, we know that $\tilde{I}$ is translation invariant. The next lemma shows that $\tilde{I}$ is $b(x)$-homogeneous for all $x \in X$. To state the lemma, note that if $\xi \in B^0_{L(D)}$, then $b(x)\xi \in B^0_{L(D)}$. This is because $0 \in L(D)^\circ$ and $b(x) \in (0,1)$.

Lemma 15 *$\tilde{I}(b(x)\xi) = b(x)\tilde{I}(\xi)$ for all $x \in X, \xi \in B^0_{L(D)}$.*

Proof. 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 that

$$\tilde{I}\big[u(x) + b(x)\xi\big] = u(x) + b(x)\tilde{I}(\xi) \quad \forall x \in X, \xi \in B^0_{L(D)}. \tag{A.11}$$

Letting $x = x_0$, we obtain $\tilde{I}(\beta\xi) = \beta\tilde{I}(\xi)$ for all $\xi \in B^0_{L(D)}$. In turn, $\tilde{I}(\beta^t\xi) = \beta^t\tilde{I}(\xi)$ for all $\xi \in B^0_{L(D)}, t \in \mathbb{N}$.

Next, fix $x \in X$ and $\xi \in B^0_{L(D)}$. Choose t large enough so that $\beta^t u(x) \in (-\hat{\lambda}, \hat{\lambda})$. We claim that

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

The first equality follows from (A.11); the second because $\tilde{I}$ is translation-invariant; the final equality follows because, as we showed earlier in this proof, $\tilde{I}$ is β-homogeneous. ∎

Lemma 16 *$\tilde{I} : B^0_{L(D)} \to \mathbb{R}$ can be extended to a certainty equivalent $\tilde{I}^e : B^0 \to \mathbb{R}$ which is translation-invariant and $b(x)$-homogeneous for all $x \in X$.*

Proof. For every $\xi \in B^0$, pick $t \in \mathbb{N}$ large enough so that $\beta^t\xi \in B^0_{L(D)}$ and let $\tilde{I}^e(\xi) := \beta^{-t}\tilde{I}(\beta^t\xi)$. One can verify, see Kochov [17, Lemma 11], that $\tilde{I}^e$ is well-defined and extends $\tilde{I}$. 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 $\beta^t\alpha \in (-\hat{\lambda}, \hat{\lambda})$. Then,

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

Similar arguments show that the extension $\tilde{I}^e$ is $b(x)$-homogeneous for every $x \in X$. By construction, $\tilde{I}^e$ is increasing and normalized. Finally, because $\tilde{I}^e$ is translation-invariant, it is norm-continuous. Thus, $\tilde{I}^e$ is a certainty equivalent in the sense of Section 2. ∎

The next lemma shows that if $b : X \to (0,1)$ is nonconstant, then $\tilde{I}^e$ is positively homogeneous.

Lemma 17 *If a function $J : B^0 \to \mathbb{R}$ is γ-homogeneous for all γ in some nonempty open interval $(a, b) \subset (0, 1)$, then J is positively homogeneous.*

Proof. It is clear that J 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 J is γ-homogeneous for all $\gamma \in (0, b^k)$. Next pick any $\gamma > 0$ and $\xi \in B^0$. Choose $\beta \in (0, b^k)$ and t large enough so that $\beta^t \gamma \in (0, b^k)$. Because $\beta^t \in (0, b^k)$, J is β^t-homogeneous. Hence, $J(\beta^t \gamma \xi) = \beta^t J(\gamma \xi)$. Because J is $\beta^t \gamma$-homogeneous, $J(\beta^t \gamma \xi) = \beta^t \gamma J(\xi)$. The last two equalities prove that J is γ-homogeneous. ■

A.1.10 The Non-Affine Case

Suppose now that (A.9) has no solution g that is affine on a subinterval of $[c, c']$. It follows from Lundberg [26, Thm. 11.1] that all solutions g and, in particular, all functions $\tilde{f}_x$ take the form

$$\tilde{f}_x(k) = \tilde{\phi}(x, k) = \frac{1}{p} \ln(u(x) + b(x) e^{pk}) \quad \forall x \in X, k \in [c, c'], \tag{A.12}$$

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 $H(s) := -e^{ps}$ and 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,

$$\phi^*(x, k) = u(x) + b(x)k, \quad \forall x \in X, k \in D^*. \tag{A.13}$$

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

Turning attention to the certainty equivalent $I^* : B^0_{D^*} \to \mathbb{R}$, notice that the open interval $(H(-\hat{\lambda}), H(\hat{\lambda}))$ contains 1.

Lemma 18 *$I^*(\gamma\xi) = \gamma I^*(\xi)$ for all $\gamma \in (H(-\hat{\lambda}), H(\hat{\lambda}))$ and $\xi \in B^0_{D^*}$ such that $\gamma\xi \in B^0_{D^*}$.*

Proof. Let ξ and γ be as in the statement of the lemma. 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 $\alpha \in (-\hat{\lambda}, \hat{\lambda})$. we claim that

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

The first equality follows from the definition of I^*, the second from the definition of ξ', the third from the translation invariance of $\tilde{I}$ (Lemma 13), the fourth equality from the fact that H is an exponential function, and the final equality from the definition of I^*. ■

Because γ is restricted to lie in the interval $(H(-\hat{\lambda}), H(\hat{\lambda}))$, one can think of Lemma 18 as establishing a type of "local" homogeneity. Similarly, the next lemma establishes a type of "local" translation invariance.

Lemma 19 *For every ξ in the interior of $B^0_{D^*}$, there is $k_\xi > 0$ such that $I^*(\xi + k) = I^*(\xi) + k$ for all k such that $|k| \leq k_\xi$.*

Proof. Because I^* and ϕ^* permute,

$$u(x) + b(x)I^*(\xi) = I^*(u(x) + b(x)\xi) \quad \forall x \in X, \xi \in B^0_{D^*}. \tag{A.14}$$

Fix some ξ in the interior of $B^0_{D^*}$. For every $x, x' \in X$, define

$$\xi' = \frac{u(x') - u(x)}{b(x)} + \frac{b(x')}{b(x)}\xi.$$

Note that if x' is sufficiently close to x, then $\xi' \in B^0_{D^*}$. Using (A.14), deduce that

$$I^*(\xi) = \frac{u(x) - u(x')}{b(x')} + \frac{b(x)}{b(x')}I^*\Big(\frac{u(x') - u(x)}{b(x)} + \frac{b(x')}{b(x)}\xi\Big).$$

Moreover, if x' is sufficiently close to x, we can apply Lemma 18 and deduce that

$$I^*(\xi) = \frac{u(x) - u(x')}{b(x')} + I^*\Big(\frac{u(x') - u(x)}{b(x')} + \xi\Big).$$

If $u : X \to \mathbb{R}$ is a nonconstant function, then the above equality completes the proof. Suppose then that $u : X \to \mathbb{R}$ is constant. Because $U^* : X^\infty \to \mathbb{R}$ cannot be constant, it follows that $b : X \to (0,1)$ cannot be constant either. Moreover, from (A.14) we can deduce that

$$u(x) + b(x)u(x) + b^2(x)I^*(\xi) = I^*\Big(u(x) + b(x)u(x) + b^2(x)\xi\Big) \quad \forall x \in X, \xi \in B^0_{D^*}.$$

Letting $v(x) := u(x) + b(x)u(x)$ and $c(x) := b^2(x)$, we see that

$$v(x) + c(x)I^*(\xi) = I^*(v(x) + c(x)\xi) \quad \forall x \in X, \xi \in B^0_{D^*}.$$

Since by construction v is not constant, the proof reduces to the case when u is not constant. ■

Lemmas 20 and 21 below show that the local properties established by Lemmas 19 and 18 integrate into global properties.

Lemma 20 *$I^*(\xi + k) = I^*(\xi) + k$ for all $\xi \in B^0_{D^*}$ and $k \in \mathbb{R}$ such that $\xi + k \in B^0_{D^*}$.*

Proof. Fix some ξ in the interior of $B^0_{D^*}$ and some $k > 0$ such that $\xi + k \in B^0_{D^*}$. Analogous arguments apply when $k < 0$. Let $k^* \geq 0$ be the largest k' such that $I^*(\xi + k'') = I^*(\xi) + k''$ for all $k'' \in [0, k']$. By Lemma 19, $k^* > 0$. If $k \leq k^*$, we are done. Suppose $k > k^*$ and, by way of contradiction, that $I^*(\xi + k) \neq I^*(\xi) + k$. Because $k > k^*$, $\xi' := \xi + k^*$ is in the interior of $B^0_{D^*}$. By Lemma 19, there is $k^{**} > 0$ such that $I^*(\xi' + k') = I^*(\xi') + k'$ for all $k' \in [0, k^{**}]$. But then, for all such k',

$$I^*(\xi + k^* + k') = I^*(\xi + k^*) + k' = I^*(\xi) + k^* + k',$$

contradicting the definition of k^*. ■

The proof of the next lemma is analogous and omitted.

Lemma 21 *$I^*(\gamma\xi) = \gamma I^*(\xi)$ for all $\gamma \geq 0$ and $\xi \in B^0_{D^*}$ such that $\gamma\xi \in B^0_{D^*}$.*

We need one more property of I^*.

Lemma 22 *$I^*(\alpha\xi + (1-\alpha)k) = \alpha I^*(\xi) + (1-\alpha)k$ for all $\alpha \in [0,1], \xi \in B^0_{D^*}, k \in D^*$.*

Proof. Suppose ξ is in the interior of $B^0_{D^*}$. For all $\alpha \in [0,1]$ sufficiently close to 1, $\alpha\xi \in B^0_{D^*}$. By the translation invariance of I^* (Lemma 20) and the positive homogeneity of I^* (Lemma 21),

$$I^*(\alpha\xi + (1-\alpha)k) = I^*(\alpha\xi) + (1-\alpha)k = \alpha I(\xi) + (1-\alpha)k.$$

Arguments analogous to those in Lemma 20 show that the desired property holds for all $\alpha \in [0,1]$. ■

Lemma 23 *$I^* : B^0_{D^*} \to \mathbb{R}$ can be uniquely extended to a translation-invariant and positively homogeneous certainty equivalent $I^{*e} : B^0 \to \mathbb{R}$.*

Proof. Fix some k in the interior of D^*. For every $\xi \in B^0$, there exists $\alpha \in (0,1)$ such that $\alpha\xi + (1-\alpha)k \in B^0_{D^*}$. Let

$$I^{*e}(\xi) := \frac{1}{\alpha} I^*\big(\alpha\xi + (1-\alpha)k\big) - \frac{1-\alpha}{\alpha}k.$$

To see that I^{*e} is well-defined, take $\alpha_1 > \alpha_2$ such that

$$\xi_1 := \alpha_1\xi + (1-\alpha_1)k \in B^0_{D^*} \quad \text{and} \quad \xi_2 := \alpha_2\xi + (1-\alpha_2)k \in B^0_{D^*}.$$

By construction, $\xi_2 = \frac{\alpha_2}{\alpha_1}\xi_1 + (1 - \frac{\alpha_2}{\alpha_1})k$. By Lemma 22, $I^*(\xi_2) = \frac{\alpha_2}{\alpha_1} I(\xi_1) + (1 - \frac{\alpha_2}{\alpha_1})k$, which is equivalent to

$$\frac{1}{\alpha_1} I^*(\xi_1) - \frac{1-\alpha_1}{\alpha_1}k = \frac{1}{\alpha_2} I^*(\xi_2) - \frac{1-\alpha_2}{\alpha_2}k.$$

Thus, I^{*e} is well-defined. Messy but simple calculations show that I^{*e} is translation-invariant. Appealing to Lemma 22, one can then show that I^{*e} is positively homogeneous. By construction, I^{*e} is increasing and normalized. Finally, I^{*e} is norm-continuous because it is increasing and translation-invariant. Thus, I^{*e} is a certainty equivalent in the sense of Section 2. ■

A.2 Proof of Theorem 2

By Lemma 7, there is $z \in X$ such that $(z, z, ...)$ attains the minimum of $U : X^\infty \to \mathbb{R}$. It is w.l.o.g. to assume that $u(z) = 0$. Let $f : U(X^\infty) \to \hat{U}(X^\infty)$ be the function such that $f(U(d)) := \hat{U}(d)$ for every $d \in X^\infty$. Because U and $\hat{U}$ represent the same preference relation on X^∞, f is well defined and strictly increasing. Because $\hat{U}$ is continuous and X^∞ is connected, the set $\hat{U}(X^\infty) = f(U(X^\infty)) \subset \mathbb{R}$ is connected. Hence, f is continuous. Next, we want to show that f is differentiable at each point in the interior of its domain. For every $x \in X$, $s \in U(X^\infty)^\circ$, and $k \in \mathbb{R}$ small enough, we have, by construction,

$$\frac{f\big[u(x) + b(x)(s+k)\big] - f\big[u(x) + b(x)s\big]}{b(x)k} = \frac{\hat{b}(x)}{b(x)} \frac{f(s+k) - f(s)}{k}. \tag{A.15}$$

Thus, if f is differentiable at $s \in U(X^\infty)^\circ$, then f is differentiable at every point in the set

$$A(s) := \{u(x) + b(x)s : x \in X\}.$$

Let E be the set of all points $s \in U(X^\infty)^\circ$ at which f is differentiable. Each set $A(s), s \in E$, is an interval because X is connected and $u, b : X \to \mathbb{R}$ are continuous functions. In addition, we know that for every $k \in U(X^\infty)$, there is $x \in X$ such that $U(x, x, ...) = k$. It follows that $s \in A(s)$ for every $s \in E$. Fix $s \in E$ and, for every $n \in \mathbb{N}$, let $s_n := b(z)^n s$. Since $u(z) = 0$, $s_n \in A(s)$ for every n. Thus, f is differentiable at every s_n. Also,

$$s_{n+1} \in A(s_{n+1}) \cap A(s_n) \quad \forall n. \tag{A.16}$$

Conclude that the set $\cup_n A(s_n)$ is connected and, since $s_n \to_n 0$, that $(0,s] \subset \cup_n A(s_n)$. Thus, if f is differentiable at some $s \in U(X^\infty)^\circ$, then f is differentiable at each point in $(0,s]$. Because f is increasing, the set of points at which f is not differentiable has outer measure zero. See, e.g., Royden [29, Thm 3, p.100]). Hence, for every $s' \in U(X^\infty)^\circ$, there is $s \in E$ such that $s > s' > 0$ and so, f is differentiable in the interior of its domain.

Next, deduce from (A.15) that

$$b(x) f'\big[u(x) + b(x)s\big] = \hat{b}(x) f'(s) \quad \forall x \in X, s \in U(X^\infty)^\circ. \tag{A.17}$$

We want to show that the derivative f' is never 0. By way of contradiction, suppose $f'(s) = 0$ for some $s \in U(X^\infty)^\circ$. Since $b(x), \hat{b}(x) > 0$ for all $x \in X$, it follows from (A.17) that

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

Conclude that $f'(s') = 0$ for all $s' \in A(s)$, contradicting the fact that f is strictly increasing.

The next step is to show that $b = \hat{b}$. Take some $x \in X$ such that $s := u(x)(1 - b(x))^{-1} \in U(X^\infty)^\circ$. Plugging x and s into (A.17) gives

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

Since $f'(s) \neq 0$, we are done.

The final step is to show that f is affine. Since $b = \hat{b}$, (A.17) becomes

$$f'\big[u(x) + b(x)s\big] = f'(s) \quad \forall x \in X, \forall s \in U(X^\infty)^\circ. \tag{A.18}$$

Fix some $s \in U(X^\infty)^\circ$ and once again let $s_n := b(z)^n s$. It follows from (A.18) that f' is constant on each interval $A(s_n)$. Since $A(s_n) \cap A(s_{n+1}) \neq \varnothing$ for every n and $(0,s] \subset \cup_n A(s_n)$, f' is constant on $(0,s]$. Since s was arbitrary, f is affine on $U(X^\infty)^\circ$ and, by continuity, on $U(X^\infty)$.

A.3 Proof of Theorem 4

Let $\mathcal{H}^{rp}$ be the set of all rp-acts $h \in \mathcal{H}$.

Lemma 24 *$U \circ \mathcal{H}^{rp}$ is dense in $U \circ \mathcal{H}$.*

Proof. Take some act $h \in \mathcal{H}$. Since h is finite, there is some $t \in T$ and a finite partition $\Pi := \{A_1, ..., A_n\} \subset \mathcal{F}_t$ such that h is Π-adapted. Fix some ω_i from each set A_i and write $(x_0^i, x_1^i, ...)$ for $h(\omega_i) \in X^\infty$. For every $k \in T$ and every i, let $a_i^k := (x_0^i, ..., x_k^i)$. Let $b_1^k := (a_1^k, a_2^k, ..., a_n^k)$, $b_2^k := (a_2^k, a_3^k, ..., a_1^k)$,..., and $b_n^k := (a_n^k, a_1^k, ..., a_{n-1}^k)$. Let h^k be the Π-adapted act such that $h^k(\omega) = (b_i^k, b_i^k, ...)$ for all $\omega \in A_i$ and i. Observe that $h^k \in \mathcal{H}^{rp}$ and $h^k \to_k h$. ■

Given $t \in T$, write a for a list $(x_0, x_1, ..., x_{t-1}) \in X^t$ of outcomes as well as for a list $(f_0, ..., f_{t-1})$ of functions from Ω into X. Then, every act $h \in \mathcal{H}^{rp}$ can be written as a sequence $(a, a, ...)$ for some list $a = (f_0, f_1, ..., f_{t-1})$ of functions such that for every $\omega, \omega' \in \Omega$, the lists

$$(f_0(\omega), f_1(\omega), ..., f_{t-1}(\omega)) \in X^t \quad \text{and} \quad (f_0(\omega'), f_1(\omega'), ..., f_{t-1}(\omega')) \in X^t$$

are permutations of one another. Given a function $b : X \to (0,1)$, it follows that

$$\Pi_{k=0}^{t-1} b(f_k(\omega)) = \Pi_{k=0}^{t-1} b(f_k(\omega')) \quad \forall \omega, \omega' \in \Omega.$$

Given the rp-act $h = (a, a, ...)$, we can thus let

$$b(a) := \Pi_{k=0}^{t-1} b(f_k(\omega)) \tag{A.19}$$

and be certain that $b(a)$ is a number in $(0,1)$ independent of $\omega \in \Omega$.

Next, suppose $\succeq$ is a path stationary preference relation on $\mathcal{H}$ with a representation (u, b, I) and let $h, g, m \in \mathcal{H}$ be as in the statement of IH-T. Since h is an rp-act, $h = (a, a, ...)$ for some list a of functions. Defining $b(a)$ as in (A.19), observe that

$$U \circ m = (1 - b(a))[U \circ h] + b(a)[U \circ g].$$

Moreover, IH-T becomes equivalent to the implication:

$$I[U \circ h] \geq I[U \circ g] \Rightarrow I\Big((1 - b(a))[U \circ h] + b(a)[U \circ g]\Big) \geq I[U \circ g]. \tag{A.20}$$

On its own, (A.20) is strictly weaker than quasiconcavity since $U \circ h$ is restricted to lie in the set $U \circ \mathcal{H}^{rp}$ and since the mixing weight $b(a)$ is a function of h. Our strategy is to show that the set $U \circ \mathcal{H}^{rp}$ is sufficiently rich for (A.20) to imply quasiconcavity. From Lemma 24, we already know that $U \circ \mathcal{H}^{rp}$ is dense in $U \circ \mathcal{H}$. The bulk of the remaining proof is to show that $U \circ \mathcal{H}^{rp}$ (or some transformation thereof) contains an open set O.

Lemma 25 *For every rp-act $(a, a, ...) \in \mathcal{H}^{rp}$ and $x \in X$, if $(a, a, ...) \sim (x, x, ...)$, then $(a, x, a, x, ...) \sim (x, x, ...)$.*

Proof. Let $h \coloneqq (a, x, a, x, \ldots)$. First, we are going to show that $h \succeq (x, x, \ldots)$. By PS, $(x, a, a, a, \ldots) \sim (x, x, \ldots) \sim (a, a, \ldots)$. By IH-T, $(a, x, a, a, a, \ldots) \succeq (a, a, \ldots)$. By PS again,

$$(x, a, x, a, a, a, \ldots) \succeq (x, a, a, \ldots) \sim (a, a, \ldots).$$

By IH-T again, $(a, x, a, x, a, a, a, \ldots) \succeq (a, a, \ldots)$. Iterating the argument and using the continuity of $\succeq$ shows that $h \succeq (x, x, \ldots)$. Next, suppose by way of contradiction that $h \succ (x, x, \ldots)$ so that $h \succ (x, h) \succ (x, x, \ldots)$. Note that $(x, h) = (x, a, x, a, \ldots)$ is an rp-act. By IH-T, we conclude that $(x, a, h) \succeq (x, h)$. By PS, $(a, h) \succeq h$ and, hence, $(a, h) \succ (x, x, \ldots)$. The latter implies that $(a, h) \succ (x, a, h)$. To summarize, we have

$$(a, h) \succ (x, a, h) \succeq (x, h) = (x, a, x, a, \ldots).$$

By IH-T again, $(x, a, a, h) \succeq (x, h)$. By PS, $(a, a, h) \succeq h$. Similarly, $(a, a, a, h) \succeq h$. Iterating the argument and using the fact that $\succeq$ is continuous, we deduce that $(a, a, \ldots) \succeq h$. Altogether, we have

$$(a, a, \ldots) \succeq h \succ (x, x, \ldots) \sim (a, a, \ldots),$$

a contradiction. ■

Next, let $\Pi \subset \cup_t \mathcal{F}_t$ be a finite partition of Ω. To complete the proof of Theorem 4, it is enough to show that I is positively homogeneous and quasiconcave when I is restricted to the space of Π-measurable functions in B^0. Based on PS, it is w.l.o.g. to assume that $\Pi \subset \mathcal{F}_1$. To simplify the exposition, we also assume that Π contains three sets so that $\Pi = \{A_1, A_2, A_3\}$. The arguments extend naturally to all finite partitions Π.

Assume that utility is normalized so that $0 \in u(X)^\circ$. Let $x_0 \in X$ be such that $u(x_0) = 0$ and let $\beta \coloneqq b(x_0)$. Fix some $(y_0, y_1, y_2) \in X^3$ and consider the Π-measurable functions $f_0, f_1, \ldots, f_{11}$ from Ω into X defined as

	A_1	A_2	A_3		A_1	A_2	A_3		A_1	A_2	A_3
f_3:	x_0	x_0	y_0	f_7:	y_1	x_0	x_0	f_{11}:	x_0	y_2	x_0
f_2:	x_0	y_0	x_0	f_6:	x_0	x_0	y_1	f_{10}:	y_2	x_0	x_0
f_1:	y_0	x_0	x_0	f_5:	x_0	y_1	x_0	f_9:	x_0	x_0	y_2
f_0:	x_0	x_0	x_0	f_4:	x_0	x_0	x_0	f_8:	x_0	x_0	x_0

Let

$$\begin{aligned}
&a := (f_0, f_1, ..., f_{11}) \\
&h := (a, a, ...) \\
&h' := (a, x_0, x_0, ...) \\
&\hat{b}(y_0, y_1, y_2) := b(y_0)b(y_1)b(y_2)\beta^9.
\end{aligned}$$

Thus defined, h is a rp-act and

$$U \circ h = \frac{1}{1 - \hat{b}(y_0, y_1, y_2)} U \circ h'.$$

By varying the choice of $(y_0, y_1, y_2) \in X^3$, we obtain different acts $h = (a, a, ...)$ and $h' = (a, x_0, x_0, ...)$. Let $\mathcal{H}_3^{rp}$ be the space of all acts $h \in \mathcal{H}^{rp}$ obtained in this manner and let Φ be the function

$$(y_0, y_1, y_2) \mapsto U \circ h'.$$

Since $U \circ h' : \Omega \to \mathbb{R}$ is Π-measurable, we can identify $U \circ h'$ with a vector in $\mathbb{R}^3$ and Φ with a function from X^3 into $\mathbb{R}^3$.

Lemma 26 *$\Phi(X^3)$ has nonempty interior in $\mathbb{R}^3$. In particular, $\mathbf{0} \in \Phi(X^3)^\circ$.*

Proof. By construction,

$$\beta^{-1}\Phi(y_0, y_1, y_2) = \begin{bmatrix} 1 & \beta^5 b(y_0) & \beta^7 b(y_0)b(y_1) \\ \beta & \beta^3 b(y_0) & \beta^8 b(y_0)b(y_1) \\ \beta^2 & \beta^4 b(y_0) & \beta^6 b(y_0)b(y_1) \end{bmatrix} \begin{bmatrix} u(y_0) \\ u(y_1) \\ u(y_2) \end{bmatrix}.$$

Letting $v_0 := u(y_0)$, $v_1 := b(y_0)u(y_1)$ and $v_2 := b(y_0)b(y_1)u(y_2)$, we can rewrite the above expression as

$$\beta^{-1}\Phi(y_0, y_1, y_2) = \begin{bmatrix} 1 & \beta^5 & \beta^7 \\ \beta & \beta^3 & \beta^8 \\ \beta^2 & \beta^4 & \beta^6 \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ v_2 \end{bmatrix}$$

Since the 3×3 matrix in the last expression has full rank, the linear mapping

$$L(\tilde{v}_0, \tilde{v}_1, \tilde{v}_2) := \begin{bmatrix} 1 & \beta^5 & \beta^7 \\ \beta & \beta^3 & \beta^8 \\ \beta^2 & \beta^4 & \beta^6 \end{bmatrix} \begin{bmatrix} \tilde{v}_0 \\ \tilde{v}_1 \\ \tilde{v}_2 \end{bmatrix}$$

from $\mathbb{R}^3$ into $\mathbb{R}^3$ maps open sets into open sets. Now, consider the set

$$C := \{(u(y_0), b(y_0)u(y_1), b(y_0)b(y_1)u(y_2)) : (y_0, y_1, y_2) \in X^3\}$$

Let $\underline{b} := \min_{x \in X} b(X) \in (0,1)$, which is well defined since $b : X \to (0,1)$ is continuous and X is compact. Let $B_0 := u(X), B_1 := \underline{b}\, u(X)$, and $B_2 := \underline{b}^2 u(X)$. Note that $B_0 \times B_1 \times B_2$ is a rectangle in $\mathbb{R}^3$ with nonempty interior. We wish to show that $B_0 \times B_1 \times B_2 \subset C$. Take any $(v_0, v_1, v_2) \in B_0 \times B_1 \times B_2$. Let $y_0 \in X$ be such that $u(y_0) = v_0$. Recall that utility is normalized so that $0 \in u(X)^\circ$. Thus, by construction,

$$v_1 \in B_1 = \underline{b}\, u(X) \subset b(y_0)u(X),$$

which means that we can find $y_1 \in X$ such that $v_1 = b(y_0)u(y_1)$. Similarly,

$$v_2 \in B_2 = \underline{b}^2 u(X) \subset b(y_0)b(y_1)u(X).$$

Hence, we can find $y_2 \in X$ such that $v_2 = b(y_0)b(y_1)u(y_2)$. Conclude that $B_0 \times B_1 \times B_2 \subset C$. Hence, $L(B_1 \times B_2 \times B_3) \subset \Phi(X^3)$. Since $B_0 \times B_1 \times B_2$ has nonempty interior and L maps open sets into open sets, $\Phi(X^3)^\circ \neq \varnothing$. By construction, $\mathbf{0} \in [B_1 \times B_2 \times B_3]^\circ$ and, hence, $\mathbf{0} \in \Phi(X^3)^\circ$. ■

Next, consider the mapping

$$\Gamma(y_0, y_1, y_2) := \frac{1}{1 - \hat{b}(y_0, y_1, y_2)} \Phi(y_0, y_1, y_2).$$

Note that $\Gamma(X^3) = U \circ \mathcal{H}_3^{rp}$ and recall that we identify $U \circ \mathcal{H}_3^{rp}$ with a subset of $\mathbb{R}^3$. Focus on the restriction of I to $\mathbb{R}^3$. The remainder of the proof is broken in two cases.

Case 1: Suppose $b : X \to (0,1)$ is constant so that $b(x) = \beta$ for every $x \in X$. Then, $\Gamma = (1 - \beta^{12})^{-1}\Phi$, implying that $\Gamma(X^3)$ has nonempty interior in $\mathbb{R}^3$ and $\mathbf{0} \in O := \Gamma(X^3)^\circ$.

Lemma 27 *I is positively homogeneous.*

Proof. If not, we can find $\xi \in O$ and $\alpha \in (0,1)$ such that $I(\xi) = 0$ but $I(\alpha\xi) \neq 0$. Since I is continuous, there is $\alpha' \in (\alpha, 1]$ such that $I(\alpha'\xi) = 0$ and $I(\alpha''\xi) \neq 0$ for all $\alpha'' \in [\alpha, \alpha')$. But since $\alpha'\xi \in O$, there is an act $h \in \mathcal{H}_3^{rp}$ such that $U \circ h := \alpha'\xi$. As in Lemma 25, write h as a sequence $(a, a, ...)$ where a is some finite list of Π-measurable functions $f : \Omega \to X$. Let $g^1 := (a, x_0, a, x_0, ...)$, $g^2 := (a, a, x_0, a, a, x_0, ...)$, and so on. By construction, the sequence $(g^n)_n$ converges pointwise to h. For every n, a direct calculation shows that $U \circ g^n = \alpha^n U \circ h$ for some $\alpha_n \in (0,1)$ and $\alpha^n \nearrow_n 1$. Hence, for n large enough, $\alpha'\alpha^n \in [\alpha, \alpha']$ so that $I(U \circ g^n) \neq 0$. But Lemma 25 implies that $I(U \circ g^n) = 0$ for every n, a contradiction. ■

Lemma 28 *I is quasiconcave.*

Proof. Suppose not. Since I is positively homogeneous, we can find $\xi, \xi' \in O$ and $\gamma \in (0,1)$ such that $I(\xi) = I(\xi') = 0$ and $I(\gamma\xi + (1-\gamma)\xi') < 0$. Since I is continuous, we can choose the ξ and ξ' so that the preceding inequality obtains for every $\gamma \in (0,1)$. Since $\xi, \xi' \in O = U \circ \mathcal{H}_3^{rp}$, there are rp-acts h and h' such that $U \circ h = \xi$ and $U \circ h' = \xi'$. IH-T implies that $I(\gamma\xi + (1-\gamma)\xi') \geq 0$ for some $\gamma \in (0,1)$, a contradiction. ■

Case 2: Suppose $b : X \to (0,1)$ is not constant. By Theorem 1, I is positively homogeneous. It remains to show that I is quasiconcave.[17] If not, there are $\xi, \xi' \in U \circ \mathcal{H}$ and $\gamma \in (0,1)$ such that $I(\xi) = 0, I(\xi') > 0$ and $I(\gamma\xi' + (1-\gamma)\xi) < 0$. Since I is continuous, we can assume that the inequality holds for all γ in some interval $(0, \underline{\gamma}) \subset (0,1)$. Since I is continuous and $U \circ \mathcal{H}^{rp}$ is dense in $U \circ \mathcal{H}$, we can also assume that $\xi' \in U \circ \mathcal{H}^{rp}$. By Lemma 26, the ray S extending from $\mathbf{0} \in \mathbb{R}^3$ and passing through ξ' intersects the open set $\Phi(X^3)^\circ \ni \mathbf{0}$. Thus, we can find a sequence $(y_0^n, y_1^n, y_2^n)_n$ such that $\Phi(y_0^n, y_1^n, y_2^n) \in S$ for every n and $\Phi(y_0^n, y_1^n, y_2^n) \to_n \mathbf{0} \in \mathbb{R}^3$. Let

$$\xi_n := \frac{1}{1 - \hat{b}(y_0^n, y_1^n, y_2^n)} \Phi(y_0^n, y_1^n, y_2^n)$$
$$\lambda_n := 1 - \hat{b}(y_0^n, y_1^n, y_2^n)$$

Since $b(X)$ is a compact subset of $(0,1)$, there is $\varepsilon > 0$ such that $\lambda_n \in [\varepsilon, 1-\varepsilon] \subset (0,1)$ for each n. Since $\Phi(y_0^n, y_1^n, y_2^n) \to_n \mathbf{0}$, it follows that $\xi_n \to_n \mathbf{0}$. By construction, each ξ_n lies on the ray through ξ'. Hence, $\xi_n = k_n \xi'$ for some $k_n > 0$. Also, $k_n \to_n 0$. Since I is positively homogeneous, we know that $I(\xi_n) = k_n I(\xi') > 0 = I(\xi)$. For each n, there is an act $h_n \in \mathcal{H}_3^{rp}$ such that $U \circ h_n = \xi_n$. By IH-T,

$$0 \leq I(\lambda_n \xi_n + (1-\lambda_n)\xi) = I(\lambda_n k_n \xi' + (1-\lambda_n)\xi) \quad \forall n. \tag{A.21}$$

Let $\alpha_n := \frac{\lambda_n k_n}{\lambda_n k_n + (1-\lambda_n)}$. Since I is positively homogeneous, (A.21) implies that

$$0 \leq I(\alpha_n \xi' + (1-\alpha_n)\xi) \quad \forall n. \tag{A.22}$$

Recall that $k_n \to_n 0$, while the sequence $(\lambda_n)_n$ is bounded away from 1. Thus, $\alpha_n \to_n 0$. But then (A.22) contradicts the fact that $I(\gamma\xi' + (1-\gamma)\xi) < 0$ for all γ in the open interval $(0, \underline{\gamma})$.

[17] If one could once again show that $\Gamma(X^3)$ has nonempty interior, then Lemma 28 would deliver the desired result. When $b : X \to (0,1)$ is nonconstant however, we have not been able to prove that the interior of $\Gamma(X^3)$ is nonempty. Hence, we provide a different argument for the quasiconcavity of I.

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, A. Kochov, and F. Le Grand. On monotone recursive preferences. *Econometrica*, 85(5):1433–1466, 2017.

[4] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and L. Montrucchio. Uncertainty averse preferences. *Journal of Economic Theory*, 146:1275–1330, 2011.

[5] S. H. Chew and L. G. Epstein. Recursive utility under uncertainty. In A. M. Khan and N. C. Yannelis, editors, *Equilibrium Theory in Infinite Dimensional Spaces*, volume 1 of *Studies in Economic Theory*, pages 352–369. Springer Berlin Heidelberg, 1991.

[6] G. Debreu. Representation of a preference ordering by a numerical utility function. In R. Thrall, C. Coombs, and R. Davis, editors, *Decision Processes*. John Wiley & Sons, 1954.

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

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

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

[10] 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.

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

[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. Ambiguity made precise: A comparative foundation. *Journal of Economic Theory*, 102(2):251–289, 2002.

[14] P. Ghirardato, F. Maccheroni, and M. Marinacci. Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, 118(2):133–173, 2004.

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

[16] L. P. Hansen, T. J. Sargent, G. Turmuhambetova, and N. Williams. Robust control and model misspecification. *Journal of Economic Theory*, 128(1):45–90, 2006.

[17] A. Kochov. Time and No Lotteries: An Axiomatization of Maxmin Expected Utility. *Econometrica*, 83(1):239–262, 2015.

[18] A. Kochov. A separation result for stationary preferences. *Journal of Mathematical Economics*, 70:123–126, 2017.

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

[20] T. Koopmans. Representation of preference orderings over time. In C. McGuire and R. Radner, editors, *Decision and organization*, pages 79–100. Amsterdam: North Holland, 1972.

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

[22] D. Kreps. *Notes on the theory of choice*. Westview Press, 1988.

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

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

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

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

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

[28] F. Maccheroni, M. Marinacci, and A. Rustichini. Dynamic variational preferences. *Journal of Economic Theory*, 128(1):4–44, 2006.

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

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

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