On Monotone Recursive Preferences*

Antoine Bommier Asen Kochov François Le Grand†

July 8, 2016

We explore the set of preferences defined over temporal lotteries in an infinite horizon setting. We provide utility representations for all preferences that are both recursive and monotone. Our results indicate that the class of monotone recursive preferences includes Uzawa and risk-sensitive preferences, but leaves aside several of the recursive models suggested by Epstein and Zin (1989). Our representation result is derived in great generality using Lundberg (1982, 1985)'s work on functional equations.

Keywords: recursive utility, monotonicity, stationarity, temporal lotteries, risk aversion.
JEL codes: D90, D81.

1 Introduction

Intertemporal decisions lie at the heart of many applied economic problems. It is well understood that the analyses of such problems and the related policy recommendations depend critically on the structure of the intertemporal utility functions, and therefore on the underlying decision theoretic assumptions. A popular assumption, first introduced by Koopmans (1960) in a deterministic setting, is stationarity. It implies that an agent can, at all dates, evaluate future prospects using the same history and time-independent preference relation and be time-consistent. In the presence of uncertainty, stationarity is most often complemented by the assumption of recursivity, allowing one to preserve time consistency and history independence. Recursivity is moreover extremely useful in applications, as it permits the use of

*This paper is a substantial revised version of Bommier and LeGrand (2014). We thank Jaroslav Borovička, Rose-Anne Dana, Sujoy Mukerji, Karl Schlag, Joel Sobel, Jean-Marc Tallon, Håkon Tretvoll, Katsutoshi Wakai, Stanley Zin, and seminar participants at the Aix-Marseille University, BI Norwegian Business School, EMLyon Business School, Paris-Dauphine University, University of Lyon, as well as conference participants at RUD 2012, FUR XV 2012, SAET 2013, EEA 2013, NASM 2014, ESEM 2014 and BI-SHoF 2015.

†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.
Bommier and Le Grand gratefully acknowledge support from the Swiss-Re Foundation and the ETH Zurich Foundation.

dynamic programming methods. The assumptions of stationarity and recursivity, although of a different nature, are so often coupled together that the single adjective "recursive" is typically used to describe their conjunction.

The so-called *recursive preferences* are the object of analysis in the current paper. More precisely, we study recursive preferences that satisfy another popular assumption, monotonicity, resulting in the class of *monotone recursive preferences*. As is best explained in Chew and Epstein (1990, p. 56), monotonicity (called "ordinal dominance" in their paper) roughly "states that if two random sequences, C and C', are such that in every state of the world, the deterministic consumption stream provided in C is weakly preferred to that provided in C', then C should be weakly preferred to C'." Thus, monotonicity requires that a decision maker would never choose an action if another available action is preferable in every state of the world.

The additively separable expected utility model with exponential discounting, by far the most widely used model of intertemporal choice, is a particular case of monotone recursive preferences. This model has however been criticized for its lack of flexibility and in particular for being unable to disentangle risk aversion from the degree of intertemporal substitution. The search for greater flexibility has led researchers to consider either non-recursive or non-monotone preferences. For example, Chew and Epstein (1990, p. 56) explain that "given the inflexibility of the [intertemporal expected utility function] we are forced to choose which of recursivity and ordinal dominance to weaken," with non-recursive preferences explored in Chew and Epstein (1990) and non-monotone preferences explored in Epstein and Zin (1989). The latter article has provided a widespread alternative to the standard model of intertemporal choice.

The current paper explores if and how flexibility can be obtained within the set of monotone recursive preferences. The core of the analysis is developed in the risk setting, where preferences are defined over temporal lotteries.[1] Our main finding is that recursive preferences are monotone if and only if they admit a recursive utility representation $U_t = W(c_t, I[U_{t+1}])$ with a time aggregator W and a certainty equivalent I belonging to one of the following two cases:[2]

- $W(c,x) = u(c) + \beta x$ and I is translation-invariant, or
- $W(c,x) = u(c) + b(c)x$ and I is translation- and scale-invariant.

[1] Section A of the Appendix extends the analysis to a setting of subjective uncertainty.

[2] From the perspective of period t, the continuation utility U_{t+1} may be random. A certainty equivalent I provides a general way of computing the "expected value" of U_{t+1}. This expected value is then combined with current consumption c_t via the time aggregator W in order to compute U_t. Formal definitions are given in Section 4.1.

We should emphasize that these restrictions are derived in great generality. In particular, we do not assume any form of separability of ordinal preferences except the kind implied by stationarity.[3]

Our results contribute to a growing literature that seeks to understand how alternative models of intertemporal choice differ in their predictions. Notably, the specifications we derive constrain both ordinal preferences and risk preferences. Regarding ordinal preferences, Koopmans (1960), whose analysis was restricted to a deterministic setting, showed that any monotone time aggregator W generates a stationary preference relation. Here, we obtain that only ordinal preferences that can be represented by affine time aggregators can be extended to monotone recursive preferences once risk is introduced. Moreover, the restrictions related to the certainty equivalent I drastically reduce the set of admissible risk preferences. A direct consequence of our results is that the specifications suggested by the general recursive approach of Epstein and Zin (1989) are monotone only in very specific cases.[4] In particular, the most widely used isoelastic Epstein-Zin specification is not monotone, unless it reduces to the standard additively separable model of intertemporal choice or the elasticity of intertemporal substitution is assumed to be equal to one.

As a corollary of our main result, we obtain novel characterizations of two models that have featured prominently in applied work. We do so by restricting attention to preferences à la Kreps and Porteus (1978) or alternatively by imposing additional but familiar restrictions on the individual's attitudes toward the timing of resolution of uncertainty. In each case, we find that preferences admit a representation fulfilling one of the following two recursive equations:

- $U_t = u(c_t) - \beta \frac{1}{k} \log(E[e^{-kU_{t+1}}])$, or
- $U_t = u(c_t) + b(c_t)E[U_{t+1}]$.

The first case corresponds to the risk-sensitive preferences of Hansen and Sargent (1995), while the second case corresponds to the class of Uzawa (1968) preferences, axiomatized in discrete time by Epstein (1983). With Uzawa preferences, the degree of risk aversion cannot be modified without affecting ordinal preferences. In contrast, it is known from Chew and Epstein (1991) that the parameter k that enters the recursion defining risk-sensitive preferences has a direct interpretation in terms of risk aversion: the greater the value of k, the greater the risk aversion. In particular, we reach the conclusion that risk-sensitive preferences are the only

[3]Ordinal preferences are preferences over deterministic consumption paths.

[4]Epstein and Zin (1989) consider utility representations fulfilling the recursion $U_t = (c^\rho + \beta(I[U_{t+1}])^\rho)^{\frac{1}{\rho}}$. To make the link with our paper more explicit, consider the renormalized utility function $V_t = \frac{U_t^\rho}{\rho}$, which fulfills the recursion $V_t = \frac{c^\rho}{\rho} + \frac{\beta}{\rho}(I[(\rho V_{t+1})^{\frac{1}{\rho}}])^\rho$. It follows from our results that these preferences are monotone if and only if the certainty equivalent $\mu \to \frac{1}{\rho}(I((\rho\mu)^{\frac{1}{\rho}}))^\rho$ is translation-invariant.

Kreps-Porteus preferences that admit a separation of risk and intertemporal attitudes, while being monotone.

The proof of our main result employs powerful techniques that may be of interest outside the scope of this paper. Namely, we show that combining stationarity, recursivity and monotonicity is only possible if the utility function satisfies a system of generalized distributivity equations. Such equations were studied by Aczél (1966) and solved in great generality by Lundberg (1982, 1985) with methods imported from group theory and the study of iteration groups in particular.[5] Our proof shows how to apply these methods to the study of recursive utility. In the process, we provide a brief introduction to these methods as well as several extensions.

The remainder of the paper is organized as follows. Section 2 introduces our choice setting and Section 3 our axioms. Section 4 presents our main representation result and its corollaries. Section 5 develops some intuition for our main result. In Section 6, we use a consumption-savings example to contrast the consequences of using monotone and non-monotone preferences. Section 7 concludes the paper and discusses several venues for future work.

2 Choice setting

Time is discrete and indexed by $t = 0, 1, \ldots$ For the sake of simplicity, we assume that per period consumption lies in a compact interval $C = [\underline{c}, \overline{c}] \subset \mathbb{R}$ where $0 < \underline{c} < \overline{c}$. The infinite Cartesian product C^∞ represents the space of deterministic consumption streams. To introduce uncertainty, we follow Epstein and Zin (1989) and construct a space of infinite temporal lotteries. We should note that our account of the construction is brief and at times heuristic; the reader is referred to Epstein and Zin (1989) and Chew and Epstein (1991) for the formal details. To proceed, we need a few mathematical preliminaries. The Cartesian product of topological spaces is endowed with the product topology. Given a topological space X, the Borel σ-algebra on X is denoted $\mathcal{B}(X)$. The space of Borel probability measures on X is denoted $M(X)$ and endowed with the topology of weak convergence. As is typical, we identify each $x \in X$ with the Dirac measure on x. When convenient, we can therefore view X as a subset of $M(X)$.

We define the space D of temporal lotteries in two steps. First, let $D_0 := C^\infty$ and for all $t \geq 1$, let $D_t := C \times M(D_{t-1})$. For each t, D_t is the set of temporal lotteries for which all

[5] The key results trace back to the relationship between solutions of the translation equation –i.e., bivariate functions that solve $f(f(x, y), z) = f(x, y + z)$– and iteration groups. See Aczél (1966, Chap. 6). Lundberg (1982, 1985, 2005) proved that these results can be used to solve the so-called distributivity equation $f(g(x, y), z) = f(g(x, z), g(y, z))$ as well as some more general versions thereof.

uncertainty resolves in or before period t. The second step is to define temporal lotteries for which the uncertainty may resolve only asymptotically. To that end, note that each temporal lottery in D_{t+1}, $t > 0$, can be projected into a temporal lottery in D_t by assuming that all the uncertainty that resolves in period $t + 1$ resolves in period t instead. We can then define the space D of all temporal lotteries as the projective limit of the sequence $(D_t)_t$. The space D serves as the choice domain in this paper. One can visualize its elements as potentially infinite probability trees, each branch of which is a consumption stream in C^∞. As a subset of $D_0 \times D_1 \times \ldots$, the set D inherits the appropriate relative topology.[6]

It is known from Epstein and Zin (1989) that the set D is homeomorphic to $C \times M(D)$. Subsequently, we write (c, m) for a generic temporal lottery in D. There is clear intuition for this homeomorphism: Each temporal lottery can be decomposed into a pair (c, m) where $c \in C$ represents initial consumption, which is certain, and $m \in M(D)$ represents uncertainty about the future, that is, about the temporal lottery to be faced next period. Since we identify D with a subset of $M(D)$, we can also write $(c_0, (c_1, m)) \in D$ for a temporal lottery that consists of two periods of deterministic consumption, c_0 and c_1, followed by the lottery $m \in M(D)$. More generally, for any consumption vector $c^t = (c_0, \ldots, c_{t-1}) \in C^t$ and $m \in M(D)$, the temporal lottery $(c_0, (c_1, (c_2, (\ldots, (c_{t-1}, m))) \ldots) \in D$ is one that consists of t periods of deterministic consumption followed by the lottery m. For simplicity, we shorten the last expression by writing $(c^t, m) \in D$.

Being a space of probability measures, $M(D)$ is a mixture space in the sense of Herstein and Milnor (1953). We write $\pi m \oplus (1 - \pi)m' \in M(D)$ for the mixture of $m, m' \in M(D)$ given $\pi \in [0, 1]$.[7] The mixture of n lotteries $(m_i)_{1 \le i \le n}$ with a probability vector $(\pi_i)_{1 \le i \le n}$ will be denoted $\bigoplus_{i=1}^n \pi_i m_i$.

3 Axioms

The behavioral primitive in this paper is a binary relation $\succeq$ on the space D of temporal lotteries. In this section, we introduce the main axioms we impose on this relation. The first two are standard.

Axiom 1 (Weak order) *The binary relation $\succeq$ is complete and transitive.*

[6]Each set D_t is homeomorphic to a subset of D, with the homeomorphism defined as follows: for any $m_t \in D_t$, let m_k, $k \le t$, be the projection of m_t onto D_k and identify m_t with the sequence $(m_0, \ldots, m_{t-1}, m_t, m_t, m_t, \ldots) \in D \subset \times_{t'} D_{t'}$. As is usual, we do not distinguish between the set D_t and its embedding in D.

[7]In particular, $\pi m \oplus (1 - \pi)m'$ is the probability measure in $M(D)$ such that $[\pi m \oplus (1 - \pi)m'](B) = \pi m(B) + (1 - \pi)m'(B)$ for every Borel subset B of D.

Axiom 2 (Continuity) *For all* $(c,m) \in D$*, the sets* $\{(c',m') \in D | (c',m') \succeq (c,m)\}$ *and* $\{(c',m') \in D | (c,m) \succeq (c',m')\}$ *are closed in* D*.*

The next axiom, Recursivity, ensures that ex-ante choices remain optimal when they are evaluated ex-post.

Axiom 3 (Recursivity) *For all* $n,t > 0$*, consumption vectors* $c^t \in C^t$*, temporal lotteries* $(c_i,m_i),(c_i',m_i') \in D$*,* $i = 1,2,\ldots,n$*, and* $(\pi_1,\ldots,\pi_n) \in (0,1)^n$ *such that* $\sum_i \pi_i = 1$*, if for every* $i = 1,\ldots,n$*:*

$$\big(c^t,(c_i,m_i)\big) \succeq \big(c^t,(c_i',m_i')\big),$$

then

$$\big(c^t,\bigoplus_{i=1}^{n} \pi_i(c_i,m_i)\big) \succeq \big(c^t,\bigoplus_{i=1}^{n} \pi_i(c_i',m_i')\big). \tag{1}$$

Moreover, the latter ranking is strict if, in addition, one of the former rankings is strict.

The next two axioms, History Independence and Stationarity, are complementary assumptions expressing Koopmans' (1960, p. 294) idea that "the passage of time does not have an effect on preferences."

Axiom 4 (History Independence) *For all* $c,c' \in C$ *and* $m,m' \in M(D)$*,* $(c,m) \succeq (c,m')$ *if and only if* $(c',m) \succeq (c',m')$*.*

Axiom 5 (Stationarity) *For all* $c_0 \in C$ *and* $(c,m),(c',m') \in D$*,*

$$(c_0,(c,m)) \succeq (c_0,(c',m')) \quad \textit{if and only if} \quad (c,m) \succeq (c',m').$$

Following Chew and Epstein (1991), we refer to preferences satisfying Axioms 1 through 5 as *recursive preferences*.[8]

The next axiom requires that in the absence of uncertainty higher consumption is always better. To state it, let $\geq$ denote the usual pointwise order on C^∞.

Axiom 6 (Monotonicity for deterministic prospects) *For all* $c^\infty, c'^\infty \in C^\infty$*, if* $c^\infty \geq c'^\infty$*, then* $c^\infty \succeq c'^\infty$*. Moreover, the latter ranking is strict whenever* $c^\infty \gneq c'^\infty$*.*

The next and final axiom is central to the analysis of this paper.

[8]Unlike us, Chew and Epstein (1991) adopt $M(D)$ as the domain of choice for their work on recursive preferences. The difference is immaterial since any recursive preference relation on D extends uniquely to a recursive preference relation on $M(D)$. We should note however that if we had chosen $M(D)$ as the domain of choice, then History Independence and Stationarity could have been combined into a single assumption stating that for all $m,m' \in M(D)$ and all $c \in C$, $m \succeq m'$ if and only if $(c,m) \succeq (c,m')$. It is for this reason that we often use the term "stationarity" to mean the conjunction of Axioms 4 and 5.

Axiom 7 (Monotonicity) *For all $n, t > 0$, consumption vectors $c^t, c'^t \in C^t$, consumption streams $c_i^\infty, c_i'^\infty \in C^\infty$, $i = 1, 2, \ldots, n$, and $(\pi_1, \ldots, \pi_n) \in [0,1]^n$ such that $\sum_i \pi_i = 1$, if for every $i = 1, \ldots, n$:*

$$\left(c^t, c_i^\infty\right) \succeq \left(c'^t, c_i'^\infty\right),$$

then

$$\left(c^t, \bigoplus_{i=1}^n \pi_i c_i^\infty\right) \succeq \left(c'^t, \bigoplus_{i=1}^n \pi_i c_i'^\infty\right). \tag{2}$$

Monotonicity corresponds to the notion of Ordinal Dominance in Chew and Epstein (1990). It is noteworthy that if the consumption levels during the first t periods are identical, that is, if $c^t = c'^t$, then the requirement in equation (2) is implied by Recursivity. Monotonicity extends the requirement to the case when $c^t \neq c'^t$. Another important observation is that, as in the statement of Recursivity, the consumption streams $\left(c^t, c_i^\infty\right)$ and $\left(c'^t, c_i'^\infty\right)$ are mixed at the *same* date t on both sides of equation (2). This explains why this notion of monotonicity allows for non-trivial attitudes toward the timing of resolution of uncertainty and, hence, for the separation of risk and intertemporal attitudes. See Sections 4.2 and 4.3 for a detailed discussion of this point.

Monotonicity is a consistency requirement between preferences over temporal lotteries and preferences over deterministic consumption streams. The preference relation over D induces a preference relation over C^∞. Axiom 7 stipulates that a temporal lottery be preferred whenever it provides a better consumption stream in every state of the world. Monotonicity is satisfied by the standard additively separable model of intertemporal choice and more generally by the recursive preferences of Epstein (1983), as a direct consequence of the von-Neumann Morgenstern independence axiom. Monotonicity is also assumed in Chew and Epstein (1990). In a setting of subjective uncertainty, the axiom is found in Epstein and Schneider (2003b), Maccheroni, Marinacci, and Rustichini (2006), and Kochov (2015). It is noteworthy that in those papers a stronger version of the axiom is actually used. See Axiom A.7 in Appendix A and the discussion therein. By comparison, the models introduced in the seminal papers of Selden (1978) and Kreps and Porteus (1978) are typically non-monotone, an aspect which is not discussed in these papers.[9]

[9] This deviation from Monotonicity is directly related to the way the independence axiom is formalized in Selden (1978) and Kreps and Porteus (1978). For example, Kreps and Porteus (1978) assume that $(c^t, m_1) \succeq (c^t, m_1')$ implies that $(c^t, \pi m_1 \oplus (1-\pi) m_2) \succeq (c^t, \pi m_1' \oplus (1-\pi) m_2)$ for any temporal lottery m_2, the ordering being strict when $(c^t, m_1) \succ (c^t, m_1')$ and $0 < \pi < 1$. Another possible assumption would have been to state that $(c^t, m_1) \succeq (c'^t, m_1')$ and $(c^t, m_2) \succeq (c'^t, m_2')$ implies that $(c^t, \pi m_1 \oplus (1-\pi) m_2) \succeq (c'^t, \pi m_1' \oplus (1-\pi) m_2')$, with moreover a strict ordering when $(c^t, m_1) \succ (c'^t, m_1')$ or $(c^t, m_2) \succ (c^t, m_2')$. In an atemporal setting, where we would simply mix lotteries with no initial vector of deterministic consumption, both assumptions would be equivalent. This is why the literature on atemporal lotteries did not have to worry about the choice between one formulation or the other. But when mixtures are not always possible, these definitions are no longer equivalent. The first one, chosen by Kreps and Porteus (1978), is weaker and does not imply Monotonicity,

We should stress that Monotonicity, as well as Recursivity, implies indifference to some forms of uncertainty, reflecting the underlying separability properties implied by these assumptions. Consider for example the case of two different consumption streams (c, c_1^∞) and (c, c_2^∞), such that $(c, c_1^\infty) \sim (c, c_2^\infty)$. Note that these consumption streams start with the same first period consumption c. If Monotonicity or Recursivity (or both) holds, we have $(c, \pi c_1^\infty \oplus (1-\pi) c_2^\infty) \sim (c, c_1^\infty)$ for every $\pi \in (0,1)$. Thus, there is indifference between a risky temporal lottery and a degenerate lottery, whatever the agent's degree of risk aversion. This may of course be seen as disputable: one may argue that an agent who strongly dislikes risk should strictly prefer the deterministic consumption stream (c, c_1^∞) to the temporal lottery $(c, \pi c_1^\infty \oplus (1-\pi) c_2^\infty)$, which allows for future, per period consumption levels to be uncertain. Aversion to such risk is however ruled out whenever Monotonicity or Recursivity is assumed. A similar feature also appears when Monotonicity is used in a static setting with a set of outcomes that is not totally ordered (meaning that indifference between different outcomes is possible). For example, consider applying the expected utility theory of von Neumann and Morgenstern to a setup where outcomes are multidimensional consumption bundles. Then, the degree of concavity of the utility index does not reflect the aversion to inequalities in possible outcomes, but the aversion to inequalities in the welfare levels associated with the possible outcomes. More generally, Monotonicity implies that substituting a possible outcome of a lottery with another outcome which is considered equally good leaves the evaluation of the lottery unaffected. But, as soon as the set of outcomes is not totally ordered, this requirement embeds a non-trivial separability property.[10] Our axiom makes no exception in this respect.

As with any separability assumption, one may wonder whether Monotonicity is appealing or excessively restrictive. Our aim is not to take a position on this point but to explore the flexibility that remains when Monotonicity is introduced. The current paper contributes to the literature by fully characterizing the class of monotone recursive preferences. We should also mention that there has been little discussion of the implications of Monotonicity within the type of specific intertemporal decision problems that arise in applications. This is in spite of the fact, which is made clear by our results, that Monotonicity is a key difference between some of the main utility specifications used in practice. For example, the problem of saving under uncertainty has been addressed with monotone specifications in Drèze and Modigliani (1972) and Kimball (1990), and with non-monotone preferences in Kimball and Weil (2009), but there is no discussion on the potential impact of monotonicity breakdowns. Section 6 provides insights into the role of Monotonicity in a standard consumption-savings

while the second one does impose Monotonicity.

[10]This is also the case when Monotonicity is formulated in a setting of subjective uncertainty, à la Savage, with a set of consequences that is not totally ordered. A particular example is the setting of Anscombe and Aumann (1963), in which the set of consequences is that of roulette lotteries.

problem. The discussion illustrates that the restrictions imposed by Monotonicity come with the advantage of providing unambiguous and intuitive conclusions about the role of risk aversion.

As formulated above, Monotonicity is restricted to temporal lotteries that resolve in a single period of time. Bommier and LeGrand (2014) show that a stronger notion of monotonicity can be formulated, extending the consistency requirement to lotteries that resolve sequentially over many periods. This stronger notion builds on the work of Segal (1990), who provides such an extension for lotteries that resolve in two periods of time. We decided not to pursue this direction here since the extension is quite involved and since our main results can be obtained using only the weaker axiom. We should stress however that every preference relation that satisfies Axioms 1 through 7 is monotone in the stronger sense of Bommier and LeGrand (2014).

In the remainder of the paper, we refer to preferences satisfying Axioms 1 through 7 as *monotone recursive preferences.*

4 Representation results

4.1 Monotone recursive preferences

Our main representation result uses the notion of a certainty equivalent. Formally, a *certainty equivalent* I is a mapping from $M(\mathbb{R}_+)$ into $\mathbb{R}_+$ which is continuous, increasing with respect to first-order stochastic dominance, and such that $I(x) = x$ for every $x \in \mathbb{R}_+$.[11] Informally, one can think of I as specifying an 'expected value' to each probability distribution over the reals.

Two additional properties of certainty equivalents play a major role in the subsequent analysis. For every $x \in \mathbb{R}_+$ and $\mu \in M(\mathbb{R}_+)$, let $\mu + x$ be the probability measure in $M(\mathbb{R}_+)$ such that $[\mu + x](B + x) = \mu(B)$ for every set $B \in \mathcal{B}(\mathbb{R}_+)$. Similarly, for every $\lambda \in \mathbb{R}_+$ and $\mu \in M(\mathbb{R}_+)$, let $\lambda\mu$ be the probability measure such that $[\lambda\mu](\lambda B) = \mu(B)$ for every $B \in \mathcal{B}(\mathbb{R}_+)$.[12] In words, $\mu + x$ is obtained from μ by adding x to each y in μ's support, while $\lambda\mu$ is obtained from μ by scaling each y in μ's support by λ. A certainty equivalent I is *translation-invariant* if $I(x + \mu) = x + I(\mu)$ for all $x \in \mathbb{R}_+$ and $\mu \in M(\mathbb{R}_+)$. It is *scale-invariant* if $I(\lambda\mu) = \lambda I(\mu)$ for all $\lambda \in \mathbb{R}_+$ and $\mu \in M(\mathbb{R}_+)$. Translation invariance has an obvious analogue in the notion of constant *absolute* risk aversion, while scale invariance is related to the notion of constant *relative* risk aversion. In what follows, however, certainty equivalents are applied to distributions of utility levels rather than consumption levels. Thus,

[11] Recall that we abuse notation and identify a degenerate probability distribution on x with x itself.

[12] As is standard, $B + x$ denotes the set $\{y + x : y \in B\}$, and λB denotes the set $\{\lambda y : y \in B\}$.

the two invariance properties have no direct implications in terms of risk attitudes with respect to consumption.

We proceed by recalling an important result from Chew and Epstein (1991). It delivers a representation for the class of all recursive preferences. Given a function $U : D \to [0,1]$ and a probability measure $m \in M(D)$, define the *image measure* $m \circ U^{-1} \in M([0,1])$ by letting:

$$[m \circ U^{-1}](B) := m\left(\{(c', m') \in D | U(c', m') \in B\}\right), \quad \forall B \in \mathcal{B}([0,1]). \tag{3}$$

The following result holds:

Lemma 1 (Chew and Epstein, 1991) *A binary relation $\succeq$ on D satisfies Axioms 1 through 5 if and only if it can be represented by a continuous utility function $U : D \to [0,1]$ such that for all $(c, m) \in D$,*

$$U(c, m) = W(c, I(m \circ U^{-1})), \tag{4}$$

where $I : M(\mathbb{R}_+) \to \mathbb{R}_+$ is a certainty equivalent and $W : C \times [0,1] \to [0,1]$ is a continuous function, strictly increasing in its second argument.

We call the representation in (4), which we denote as (U, W, I), a *recursive representation for $\succeq$*. The function W is called a *time aggregator* and I a *certainty equivalent.* Faced with a temporal lottery (c, m), the individual first evaluates the uncertain future by assigning the value $I(m \circ U^{-1})$ to the distribution $m \circ U^{-1}$ of continuation utilities; this value is then combined with current consumption c via W, so as to compute the overall utility of the lottery (c, m).

We are ready to state the main result of our paper. It delivers a representation for the class of monotone recursive preferences.

Proposition 1 *A binary relation $\succeq$ on D is a monotone recursive preference relation if and only if it admits a recursive representation (U, W, I) such that either:*

1. *$W(c, x) = u(c) + \beta x$ and I is translation-invariant, where $\beta \in (0,1)$ and $u : C \to [0,1]$ is a continuous, strictly increasing function such that $u(\underline{c}) = 0$ and $u(\overline{c}) = 1 - \beta$, or*

2. *$W(c, x) = u(c) + b(c)x$ and I is translation- and scale-invariant, where $u, b : C \to [0,1]$ are continuous functions such that $b(C) \subset (0,1)$, the functions u and $u + b$ are strictly increasing, and $u(\underline{c}) = 0$, $u(\overline{c}) = 1 - b(\overline{c})$.*

The formal proof of Proposition 1 is given in Appendix B. Some intuition is provided in Section 5. Here, we can mention one of the main lessons of our result. The work of Chew and Epstein (1991) facilitated the construction of recursive intertemporal utility functions by allowing one to select a certainty equivalent I from the large literature on *atemporal*, non-expected utility preferences and integrating it into an intertemporal utility function via the recursion in (4). However, the question arose as to which specifications of the certainty equivalent I, and of the time aggregator W should be used, and what their implications for *intertemporal* behavior would be. The literature on this matter is active and growing. In their comments to Backus, Routledge and Zin (2005), a paper that surveys the literature on recursive utility, both Hansen and Werning focus on this problem, with Werning in particular emphasizing the need for more work aimed at discriminating among alternative utility specifications. Proposition 1 shows that Monotonicity greatly restricts the admissible specifications and is thus a powerful criterion to consider.

We conclude this section with a few remarks about the uniqueness of the representations introduced in this section. First, note that utility has so far been normalized so that $U(D) = [0, 1]$, which provides a simple way to express Axiom 6 in terms of the representation.[13] Uniqueness results can however be stated even if we drop this normalization. Doing so leaves the definition of a recursive representation (U, W, I) essentially unchanged: the only difference is that W becomes a function from $C \times U(D)$ into $U(D)$. So long as the time aggregator W is restricted to be affine, the specific representations obtained in Proposition 1 have sharp uniqueness properties.[14] In the first case, these properties are familiar: the discount factor β is unique and the instantaneous utility function $u : C \to \mathbb{R}$ is unique up to positive affine transformations. While the certainty equivalent I is not unique, the extent of the non-uniqueness is well understood from the literature on ambiguity aversion, in which translation-invariant certainty equivalents have played a prominent role. We refer the reader to Maccheroni, Marinacci, and Rustichini (2006) and Cerreia-Vioglio, Maccheroni, Marinacci, and Rustichini (2015). Regarding the second case, we can show that the function $b : C \to (0, 1)$ is unique, the certainty equivalent I is unique, and the utility function $U : D \to \mathbb{R}$ is unique up to positive affine transformations. This result is proved in our companion paper, Bommier, Kochov, and LeGrand (2016).

[13]When utility is normalized so that $U(D) = [0, 1]$, Axiom 6 becomes equivalent to the restriction that the functions u and $u + b$ be strictly increasing. If we drop the normalization, there is no simple way of expressing Axiom 6 in terms of the representation. See Lemma 17 in the Appendix for details.

[14]If we do not constrain the time aggregator to be affine, then there are no straightforward uniqueness results. It is actually known that the general recursive representation obtained in Lemma 1 is not unique: for every (continuous) utility function $U' : D \to \mathbb{R}$, there exists a recursive representation (U', W', I'). This implies that the properties of the certainty equivalent I cannot be interpreted independently of the structure assumed for the time aggregator W.

4.2 Risk-sensitive preferences

This section considers three additional properties of recursive preferences and their representations: (i) preference for early resolution of uncertainty, (ii) betweenness of the certainty equivalent I, and (iii) the ability to disentangle risk aversion from the degree of intertemporal substitution. All three properties have received attention in the literature and are commonly assumed in practice. We show that within the class of monotone recursive preferences, only the risk-sensitive preferences of Hansen and Sargent (1995) possess all three. The result complements the work of Strzalecki (2011) who provides a different axiomatization of these preferences. First, we recall the definition of preference for early resolution of uncertainty from Kreps and Porteus (1978).

Definition 1 *A binary relation $\succeq$ on D exhibits a preference for early resolution of uncertainty if for all $n > 0, c_0, c_1 \in C, (m_i) \in (M(D))^n$, and $(\pi_i) \in [0,1]^n$ such that $\sum_{i=1}^n \pi_i = 1$ we have:*

$$A := (c_0, \bigoplus_{i=1}^{n} \pi_i (c_1, m_i)) \succeq (c_0, (c_1, \bigoplus_{i=1}^{n} \pi_i m_i)) =: B. \tag{5}$$

If the above ranking is one of indifference, then $\succeq$ exhibits indifference toward the timing of the resolution of uncertainty.

Lotteries A and B in equation (5) deliver identical and certain consumption levels in periods $t = 0, 1$. At the same time, there is uncertainty about the continuation lottery that will prevail in period $t = 2$. The two lotteries differ in the way this uncertainty resolves. If A is chosen, the uncertainty resolves gradually: in period $t = 1$ the agent learns $m_i \in M(D)$; in period $t = 2$ she learns the outcome of m_i. If B is chosen, the same uncertainty resolves at once, in period $t = 2$. The ranking in (5) indicates that early resolution of uncertainty is beneficial.

A certainty equivalent I satisfies *betweenness* if for all $\mu_1, \mu_2 \in M(\mathbb{R}_+)$ and $\pi \in [0,1]$, $I(\mu_1) = I(\mu_2)$ implies that $I(\pi\mu_1 \oplus (1-\pi)\mu_2) = I(\mu_1)$, that is, if I has linear, but not necessarily parallel, indifference curves.[15] Certainty equivalents of this form have proved useful in econometric work: they lead to first order conditions that are linear in probabilities and, hence, can be estimated by the method of moments. Details can be found in Epstein and Zin (2001) and Backus, Routledge and Zin (2005, p. 338). From Chew and Epstein (1991), it is also understood that such certainty equivalents have clear implications for the agent's attitudes toward the timing of the resolution of uncertainty. Consider equation (5) and

[15] The notion of betweenness was first introduced by Chew (1983, 1989) and Dekel (1986) in the context of atemporal choice under risk. Accordingly, certainty equivalents that satisfy betweenness are often called Chew-Dekel certainty equivalents.

suppose that $(c_1, m_i) \sim (c_1, m_j)$ for every $i, j = 1, 2, \ldots, n$, i.e., that all continuation lotteries m_i yield the same expected utility. The benefits of early resolution are then less apparent and it is more plausible that the agent will be indifferent between lotteries A and B in (5). If $\succeq$ is a recursive preference relation with representation (U, W, I), then such indifference obtains if and only if I satisfies betweenness.

Turning to the separation of risk and intertemporal attitudes, it is helpful to recall another result from Chew and Epstein (1991). If $\succeq_1$ and $\succeq_2$ are two recursive preferences, then $\succeq_1$ is more risk averse than $\succeq_2$ if and only if we can find representations (U_1, W_1, I_1) and (U_2, W_2, I_2) such that $U_1|_{C^\infty} = U_2|_{C^\infty}$ (same ordinal preferences), $W_1 = W_2$ (same time aggregator), and $I_1 \leq I_2$. The result shows that within the class of recursive preferences, a partial separation of risk and intertemporal attitudes is achieved: one can vary the degree of risk aversion by changing the certainty equivalent I without affecting the ranking of deterministic consumption streams. The class of Uzawa preferences can be obtained by considering an affine time aggregator and letting $I = E$, where E is the standard expectation operator.[16] This class, which includes the standard additively separable model of intertemporal choice, is notable for the fact that it affords no separation of risk and intertemporal attitudes: an Uzawa preference relation $\succeq$ on D is fully determined by its restriction to the space C^∞ of deterministic consumption streams.[17] In practice, it is common to work with recursive preferences that are at least as risk averse as some Uzawa preference relation. For example, this assumption ensures that a high level of risk aversion can coexist with a high elasticity of intertemporal substitution, which, as the empirical literature on asset returns has emphasized, is needed to reconcile the high equity premium and the low risk-free rate. So long as the time aggregator W is restricted to be affine (which by Proposition 1 is possible for the class of monotone recursive preferences), the assumption is naturally expressed by the inequality $I \leq E$, which we maintain in our next result.[18]

Proposition 2 *Consider a monotone recursive preference relation $\succeq$ with a representation (U, W, I) as in Proposition 1. Suppose I satisfies betweenness and $I \leq E$. Then, $\succeq$ exhibits a preference for early resolution of uncertainty if and only if the representation (U, W, I) is such that either*

[16] For any $\mu \in M(\mathbb{R}_+)$, we define $E[\mu] = \int_{\mathbb{R}+} x\mu(dx)$. More generally for any strictly increasing function ϕ, we define $\phi^{-1}E\,\phi$ as: for any $\mu \in M(\mathbb{R}_+)$, $\left(\phi^{-1}E\,\phi\right)[\mu] = \phi^{-1}\left(\int_{\mathbb{R}+} \phi(x)\mu(dx)\right)$.

[17] This interdependence is particularly sharp when $W(c, x) = u(c) + \beta x$ and u is of the CRRA form. As is well known, the coefficient of relative risk aversion is then equal to the inverse of the elasticity of intertemporal substitution.

[18] When discounting is exogenous, that is, when $W(c, x) = u(c) + \beta x$, one can replace $I \leq E$ with the assumption of intertemporal autocorrelation aversion. When discounting is endogenous, the behavioral implication of $I \leq E$ can be related to a more sophisticated notion of intertemporal hedging, that we will detail in our companion paper, Bommier et al. (2016).

1. Risk-sensitive case*: $W(c,x) = u(c) + \beta x$ and $I = \phi^{-1} E\,\phi$ where $\phi(x) = -\exp(-kx)$, $k \in \mathbb{R}_+$, $\beta \in (0,1)$ and $u : C \to [0,1]$ is a continuous, strictly increasing function such that $u(\underline{c}) = 0$ and $u(\overline{c}) = 1-\beta$, or*

2. Uzawa case*: $W(c,x) = u(c) + b(c)x$ and $I = E$, where $u, b : C \to [0,1]$ are continuous functions such that $b(C) \subset (0,1)$, the functions u and $u+b$ are strictly increasing, and $u(\underline{c}) = 0$, $u(\overline{c}) = 1 - b(\overline{c})$.*

It is important to observe that of the two cases obtained in Proposition 2, only the first affords a separation of risk and intertemporal attitudes. In the risk-sensitive case, one can vary the parameter k so as to change risk aversion, without affecting the elasticity of intertemporal substitution.

The proof of Proposition 2, which we provide in Appendix C, consists of two steps. First, we use a result from Grant, Kajii, and Polak (2000) to show that the certainty equivalent I must be of the *expected utility form*, that is, there is a function $\phi : \mathbb{R} \to \mathbb{R}$ such that $I = \phi^{-1} E\,\phi$.[19] The second step uses the properties of I which we derived in Proposition 1: Because I is translation-invariant, the function ϕ must be exponential. If I is also scale-invariant, then ϕ must be linear.

The latter step can be used to deliver another result of interest. Observe first that if a recursive preference relation $\succeq$ has a representation (U, W, I) where I is of the expected utility form, then the same is true for all other representations (U', W', I') of $\succeq$.[20] Like betweenness, having a certainty equivalent of the expected utility form is thus a restriction on behavior. Recursive preferences that fulfill this restriction were introduced by Kreps and Porteus (1978) and are often referred to as *Kreps-Porteus preferences*. It follows from the analysis so far that if a Kreps-Porteus preference relation $\succeq$ is monotone, then it has a representation (U, W, I) where W is affine, I is of the expected utility form, and I has the invariance properties derived in Proposition 1. As in the proof of Proposition 2, we deduce that Uzawa and the risk-sensitive preferences of Hansen and Sargent (1995) are the only Kreps-Porteus preferences that are monotone.

[19]It is worth noting that the results of Grant et al. (2000) are not directly applicable due to differences in the domain of choice: in their paper consumption takes place at a single, terminal point in time, whereas in this paper consumption takes place in every period. As we explain in Section 5, however, if Monotonicity holds, then the agent behaves as if she identifies each temporal lottery with a compound lottery over lifetime utility. Intuitively, we can then apply the results in Grant et al. (2000) to the latter domain. Although the formal proof of Proposition 2, which we provide in Appendix C, is not written this way, the reader will notice that the proof relies on the invariance properties of I derived in Proposition 1 and therefore on Monotonicity.

[20]Suppose $\succeq$ has a representation (U, I, W) where $I = \phi^{-1} E\,\phi$. Let $f : \mathbb{R} \to \mathbb{R}$ be a strictly increasing function. Then the utility function $U' := f \circ U$ is part of the recursive representation (U', W', I') where $W'(c,x) := fW(c, f^{-1}(x))$, $I' := \psi^{-1} E \psi$ and $\psi := \phi \circ f^{-1}$.

4.3 Time-dependent risk attitudes

In their seminal contribution, Kreps and Porteus (1978) briefly mention that it is possible to make a link between the agent's attitudes toward the timing of resolution of uncertainty and how risk aversion changes with time distance. In this section, we provide several results that elaborate on this point. Even though the results are rather immediate, or known from previous contributions, they serve several purposes. First, they clarify the behavioral implications of having a scale-invariant certainty equivalent I, which is the main difference between the two cases obtained in Proposition 1. Second, they facilitate the discussion in Section 5.2 in which we provide some intuition behind Proposition 1. Lastly, the results help clarify a link between our paper and the recent work of Strzalecki (2013), where a setting of subjective uncertainty is adopted.

We begin by introducing a definition similar to Definition 1, but where the lotteries $(m_i) \in (M(D))^n$ are constrained to be deterministic consumption streams.

Definition 2 *A binary relation $\succeq$ on D exhibits a degree of risk aversion that increases with time distance if for all $n > 0, c_0, c_1 \in C, (c_i^\infty) \in (C^\infty)^n$, and $(\pi_i) \in [0,1]^n$ such that $\sum_{i=1}^n \pi_i = 1$ we have:*

$$A := (c_0, \bigoplus_{i=1}^{n} \pi_i(c_1, c_i^\infty)) \succeq (c_0, (c_1, \bigoplus_{i=1}^{n} \pi_i c_i^\infty)) =: B. \tag{6}$$

If the above ranking is one of indifference, then $\succeq$ exhibits time-independent risk attitudes.

The temporal lotteries A and B in (6) present a simple tradeoff: either all uncertainty is resolved in period $t = 1$ or in period $t = 2$. One can thus view the ranking in (6) as indicative of an agent who exhibits greater risk aversion toward uncertainty resolving at the later date. Of course, the same ranking can also be viewed as a special instance of preference for early resolution of uncertainty. Definition 2 adopts different terminology so as to avoid any confusion with Definition 1. As the next result shows, the two definitions are in general not equivalent.

Proposition 3 *Consider a monotone recursive preference relation $\succeq$ with representation (U, W, I) as in Proposition 1. Then:*

- *$\succeq$ exhibits indifference toward the timing of resolution of uncertainty if and only if $I = E$;*
- *if I is scale-invariant, then $\succeq$ exhibits time-independent risk attitudes. The converse is true if I is concave in prizes;*

– *if $\succeq$ is a Kreps-Porteus preference relation, that is, if the certainty equivalent I is of the expected utility form, then a degree of risk aversion that increases with time distance is equivalent to a preference for early resolution of uncertainty.*

Proof. The first point is proved in Chew and Epstein (1991). The second is proved in Strzalecki (2013).[21] As for the Kreps-Porteus case, we know that the certainty equivalent is given by $I = \phi^{-1} E\,\phi$ where $\phi(x) = -\exp(-kx)$ for some $k \in \mathbb{R}$. One can easily check that a degree of risk aversion that increases with time distance is equivalent to having $k \geq 0$. From Kreps and Porteus (1978), this is known to imply a preference for early resolution of uncertainty. ■

The first part of Proposition 3 shows that a separation of risk and intertemporal attitudes is possible only if the temporal resolution of uncertainty matters.[22] The significance of this point, which we discuss further in Section 7, has been recently emphasized by Epstein, Farhi, and Strzalecki (2014). Another lesson from Proposition 3 is that it is possible for a monotone recursive preference relation $\succeq$ to exhibit time-independent risk attitudes without exhibiting indifference toward the timing of resolution of uncertainty. A simple example is obtained by adopting a certainty equivalent I based on the dual approach of Yaari (1987). The final part of Proposition 3 shows that this possibility disappears if we restrict I to be of the expected utility form.

We can use Proposition 3 to highlight an important lesson from Strzalecki (2013) regarding the domain of choice. The point is that in a setting of subjective uncertainty, such as the one we use in Section A, there is no appropriate analogue of Definition 1. Instead, attitudes toward the timing of uncertainty resolution are defined by rankings analogous to those in Definition 2, which means that indifference toward the timing of uncertainty resolution becomes an effectively weaker requirement. In particular, to model such indifference in a subjective setting, one is free to choose any translation- and scale-invariant certainty equivalent I. The parallel result in a setting of risk is provided by the second point of Proposition 3, which shows that such certainty equivalents imply time-independent risk attitudes. Unless we have $I = E$, however, they do not imply indifference toward the timing of uncertainty resolution in the strong sense of Definition 1.

[21] Strzalecki (2013) proves these results when discounting is exogenous, that is, when $W(c, x) = u(c) + \beta x$. Extending his results to the case of endogenous discounting is not difficult.

[22] The result in Chew and Epstein (1991) is in fact stronger: it shows that Uzawa preferences are the only recursive preferences that exhibit indifference toward the timing of resolution of uncertainty. In particular, recursive preferences that exhibit indifference to the timing of resolution of uncertainty are necessarily monotone.

5 Intuition behind Proposition 1

This section provides intuition behind Proposition 1, our main result. First, we show that any monotone recursive preference relation can be represented by two distinct recursions. The compatibility of the two recursions restricts the time aggregator W to be affine, after a suitable renormalization of utility. Second, we use the structure of W to deduce the restrictions on the certainty equivalent I obtained in Proposition 1. We should note that Monotonicity and Stationarity play the most important role in the subsequent discussion. The other axioms are assumed to hold, but we do not mention them explicitly.

5.1 Monotonicity and Stationarity: Two different recursive representations

Recursive representation related to Monotonicity. We explain here that Monotonicity implies a recursive representation whose form is different from the one deduced in Lemma 1. As there is no need to be fully rigorous in this section, we restrict the demonstration to the case of temporal lotteries that resolve in at most two periods of time and have finite support, that is, to temporal lotteries of the form:

$$(c_0, \bigoplus_i \pi_i(c_i, \bigoplus_j \pi_{ij} c_{ij}^{\infty})) \in D_2, \tag{7}$$

where $c_i \in C$, $c_{ij}^{\infty} \in C^{\infty}$, $\pi_i, \pi_{ij} \in (0,1)$, such that $\sum_i \pi_i = 1$ and, for all i, $\sum_j \pi_{ij} = 1$. This includes the case where all the uncertainty resolves in the first period (the index j being irrelevant) and the case where all the uncertainty resolves in the second period (the index i being irrelevant).

A recursive preference relation $\succeq$ on D induces a preference relation on C^{∞}. Let $V : C^{\infty} \to Im(V) \subset \mathbb{R}$ be a continuous function representing the latter. First focus on the case where all the uncertainty resolves in the first period, that is, on temporal lotteries of the form:

$$(c_0, \bigoplus_i \pi_i(c_i, c_i^{\infty})) \in D_1. \tag{8}$$

Each such temporal lottery can be associated with a lottery $\bigoplus_i \pi_i V(c_0, c_i, c_i^{\infty}) \in M^f(Im(V))$, where $M^f(Im(V))$ is the set of finite-support lotteries with outcomes in $Im(V)$. Moreover, because of Monotonicity, if two temporal lotteries as in (8) induce the same lottery in $M^f(Im(V))$, they have to be equally preferred. One can therefore use the preference relation $\succeq$ on D to define a preference relation $\succeq_1$ on $M^f(Im(V))$.[23] Let $I_1 : M^f(Im(V)) \to Im(V)$

[23]To be fully precise, for domain reasons, the preference relation $\succeq$ generates a preference relation $\succeq_1$ on a

be a continuous function representing $\succeq_1$ such that $I_1(x) = x$ for all $x \in Im(V)$. Monotonicity requires I_1 to be increasing with respect to first-order stochastic dominance. Thus, in our terminology, I_1 is a certainty equivalent.

Next, consider the case where all the uncertainty resolves in the second period of time. Every temporal lottery $(c_0, (c_i, \bigoplus_j \pi_{ij} c_{ij}^{\infty}))$ can be associated with a lottery $\bigoplus_j \pi_{ij} V(c_0, c_i, c_{ij}^{\infty}) \in M^f(Im(V))$. As above, the preference relation $\succeq$ over D generates a preference relation $\succeq_2$ over $M^f(Im(V))$, which can be represented by a certainty equivalent I_2. Remark that none of our assumptions constrains $\succeq_1$ and $\succeq_2$ to be identical. The certainty equivalents I_1 and I_2 may therefore be different.

Turning to the general case, consider a temporal lottery as in (7). For any i, there exists $c_i^{\infty} \in C^{\infty}$ such that:

$$(c_0, (c_i, \bigoplus_j \pi_{ij} c_{ij}^{\infty})) \sim (c_0, (c_i, c_i^{\infty})). \tag{9}$$

In terms of utility, equation (9) can be expressed as:

$$I_2(\bigoplus_j \pi_{ij} V(c_0, c_i, c_{ij}^{\infty})) = V(c_0, c_i, c_i^{\infty}). \tag{10}$$

Consider another temporal lottery $(c_0', \bigoplus_i \pi_i'(c_i', \bigoplus_j \pi_{ij}' c_{ij}'^{\infty}))$ as in (7) and construct $c_i'^{\infty}$ as in (9). The following equivalences hold:

$$\begin{aligned}
&(c_0, \bigoplus_i \pi_i (c_i, \bigoplus_j \pi_{ij} c_{ij}^{\infty})) \succeq (c_0', \bigoplus_i \pi_i' (c_i', \bigoplus_j \pi_{ij}' c_{ij}'^{\infty})) \\
\Leftrightarrow &(c_0, \bigoplus_i \pi_i (c_i, c_i^{\infty})) \succeq (c_0', \bigoplus_i \pi_i' (c_i', c_i'^{\infty})) \quad \text{(because of Recursivity)} \\
\Leftrightarrow &I_1(\bigoplus_i \pi_i V(c_0, c_i, c_i^{\infty})) \geq I_1(\bigoplus_i \pi_i' V(c_0, c_i', c_i'^{\infty})) \quad \text{(by definition of } I_1\text{)} \\
\Leftrightarrow &I_1(\bigoplus_i \pi_i I_2(\bigoplus_j \pi_{ij} V(c_0, c_i, c_{ij}^{\infty}))) \geq I_1(\bigoplus_i \pi_i' I_2(\bigoplus_j \pi_{ij}' V(c_0, c_i', c_{ij}'^{\infty}))) \quad \text{(see eq. (10))}
\end{aligned}$$

We thus observe that using the ex-post lifetime utility function V and the certainty equivalents I_1 and I_2 in a recursive way affords a utility representation for temporal lotteries that resolve in two periods of time.

The procedure can be extended to temporal lotteries that resolve in t periods of time, that is, to elements of D_t. We eventually find that, given a utility representation V for preferences on C^{∞}, there corresponds a utility representation for preferences on D_t, which is given by the

subset of $M^f(Im(V))$. This preference relation $\succeq_1$ can however be extended to the whole domain $M^f(Im(V))$. The formal proof of Proposition 1 addresses these technicalities.

endpoint V_0 of the recursion:

$$\begin{cases} V_\tau &= V(c_0, c_1, \ldots, c_t, \ldots) \text{ for } \tau = t, \\ V_\tau &= I_{\tau+1}([V_{\tau+1}]) \text{ for all } \tau < t. \end{cases} \tag{11}$$

In words, a temporal lottery can be associated with a compound lottery over lifetime utilities, which is then evaluated recursively as in Segal (1990), using a sequence I_1, $I_2, \ldots$ of certainty equivalents. Time-dependent risk attitudes are obtained whenever the I_τ depend on τ.

Here, it may be helpful to contrast the above recursion with what Kreps and Porteus (1978) call the standard "pay-off vector approach" for evaluating temporal lotteries. The latter approach consists in first computing a compound lottery over lifetime utility, and then evaluating this lottery using the reduction of compound lottery axiom. The recursion in (11) requires one to preserve the first step of the pay-off vector approach, but not the latter, since the compound lottery over lifetime utility is evaluated recursively without reducing it to a one-stage lottery.

Recursive representation related to Stationarity. While Monotonicity is related to the recursion in (11), Stationarity requires preferences to admit a recursive representation:

$$U_\tau = W(c_\tau, I[U_{\tau+1}]), \tag{12}$$

where the time aggregator W and the certainty equivalent I are independent of τ.[24] In other words, with Stationarity, a temporal lottery can be evaluated by first computing the expected value of the *continuation utilities* $U_{\tau+1}$ and then aggregating this value with current consumption c_τ, using a time and history invariant time aggregator W. In a sense, aggregation across states precedes aggregation across time, while the aggregation has to occur in the reverse order to get Monotonicity.

Reconciling the two recursions constrains ordinal preferences. Based on the preceding discussion, the recursions in (11) and (12) may appear quite antagonistic. They are however not incompatible: the standard model of intertemporal choice is a well-known example when both recursions hold. Proposition 1 characterizes all preferences that can be represented by each of the two recursions. The compatibility of the two recursions is also the key to the proof of Proposition 1. In Appendix B, we show that combining (11) and (12) leads to a system of generalized distributivity equations, which we solve using the work of

[24]This recursion is exactly the same as the recursion in (4). To simplify notation, however, we now write $U_{\tau+1}$ for the image measure $m_\tau \circ U^{-1}$. In essence, we are identifying a random variable with its distribution.

Lundberg (1982, 1985). After a suitable monotone transformation of utility, fulfilling these distributivity equations restricts the time aggregator W to be affine, that is, of the form $W(c, x) = u(c) + b(c)x$.[25] In Section 5.2 below, we use this fact to explain why Stationarity and Monotonicity restrict the certainty equivalents I_τ used in recursion (11). In turn, these restrictions translate directly into restrictions on the certainty equivalent I used in recursion (12) and in the formulation of our results.

5.2 Risk aversion, time discounting and attitudes toward the timing of resolution of uncertainty

Consider an agent comparing temporal lotteries that provide the same consumption profile $(c_0, \ldots, c_{N-1})$ during the first $N > 0$ periods of time, but may differ thereafter. On the one hand, with stationary preferences, the N initial periods of consumption do not matter and the ranking has to be independent of $(c_0, \ldots, c_{N-1})$. On the other hand, Monotonicity implies that future risks are evaluated in terms of their impact on lifetime utility, which includes the utility derived from the first N periods. To see how this tension can be resolved, first consider the case of a time aggregator $W(c, x) = u(c) + \beta x$. Then, the first N periods impact lifetime utility through an additive term, $\sum_{i=0}^{N-1} \beta^i u(c_i)$, and a factor β^N that multiplies continuation utility. Changing the consumption levels $c_0, \ldots, c_{N-1}$ impacts the additive term, shifting lifetime utility by a constant. For the ranking of temporal lotteries sharing the same history $(c_0, \ldots, c_{N-1})$ to be independent of that history, preferences must exhibit constant absolute risk aversion with respect to lifetime utility, that is, the certainty equivalents I_τ have to be translation-invariant.

Next, consider an increase in N, the number of initial periods. This has two effects. First, because of the discount factor β^N, the utility risk stemming from future consumption is scaled down. Second, the resolution of that risk is postponed. The two effects generate a break-down of Stationarity, unless they are both separately neutralized or they cancel each other out. The first possibility involves assuming constant relative risk aversion with respect to lifetime utility, which neutralizes the scaling effect, and a degree of risk aversion that is independent of time distance, which neutralizes the postponement of uncertainty resolution. Formally, the certainty equivalents I_τ have to be scale-invariant and independent of τ. The second possibility is to let the certainty equivalents I_τ depend on τ in a manner that offsets the scaling effect due to time discounting. To be more precise, this amounts to using certainty

[25]The existence of an affine time aggregator W imposes non-trivial restrictions on ordinal preferences. It is known that preferences that admit such a time aggregator comprise a much smaller class than the stationary preferences of Koopmans (1960): the former exhibit a strong form of impatience, which fails generically within the broader class. See Koopmans, Diamond, and Williamson (1964) and Epstein (1983) for details.

equivalents I_τ that are related via the equation:

$$I_\tau(\mu) = \beta^\tau I\left(\frac{1}{\beta^\tau}\mu\right), \text{ for all } \mu \in M(\mathbb{R}_+). \tag{13}$$

Notably, the I_τ are derived by combining a certainty equivalent I (independent of τ) with an amplification of the risk μ by the factor $\frac{1}{\beta^\tau}$. This amplification offsets the decrease of utility due to the discount factor β^τ.

The risk-sensitive preferences of Hansen and Sargent (1995) provide one example in which Stationarity is preserved via the "amplification mechanism" in equation (13). In this case, the equation (13) takes on a more concrete interpretation as well. First, recall that the class of risk-sensitive preferences is obtained by letting the certainty equivalent I in (12) take the form $I = \phi^{-1}E\phi$, where $\phi(x) = -\exp(-kx)$. Suppose $k > 0$, so that risk aversion is greater than what one can attain using a standard expected utility specification, that is, by letting $I = E$. Then, both the certainty equivalent I and the certainty equivalents I_τ are concave in prizes. It follows that $I_\tau \gneqq I_{\tau+1}$. But, given equation (11), this means that risk aversion increases with time distance. In particular, to preserve Stationarity and attain a separation of risk and intertemporal attitudes ($k > 0$), one needs strict preference for early resolution of uncertainty. Another lesson to keep in mind is that this interplay arises because one has to counteract the effects of discounting ($\beta < 1$). We return to this observation in Section 7.

When the time aggregator is of the Uzawa kind, $W(c, x) = u(c) + b(c)x$, that is, when the function b is non-constant, the contribution of the first N periods to lifetime utility is slightly more complex, involving both an additive term $\sum_{i=0}^{N-1} u(c_i^i)\Pi_{j=0}^{i-1}b(c_j)$ and a term $\Pi_{j=0}^{N-1}b(c_j)$ that multiplies continuation utility.[26] For a given N, it is possible to change the consumption levels $c_0, \ldots, c_{N-1}$, so as to impact the additive term, without changing the multiplicative one.[27] As in the previous case, we can thus deduce that for the ranking of temporal lotteries to be independent of the history $(c_0, \ldots, c_{N-1})$ the certainty equivalents I_τ have to be translation-invariant. More generally, changes in $(c_0, \ldots, c_{N-1})$ will affect the multiplicative term $\Pi_{j=0}^{N-1}b(c_j)$ as well. For this new effect to be neutralized, the certainty equivalents I_τ have to be scale-invariant. Finally, the postponement of the uncertainty associated with an increase in N is once again countered if the certainty equivalents I_τ are independent of τ. The other option, which involved using different I_τ with an amplification mechanism akin to the one of equation (13), has no analogue in the present case.

Altogether, we observe that translation invariance of the certainty equivalents I_τ is always

[26] If $W(c, x) = u(c) + b(c)x$, then $V(c_0, c_1, \ldots) = u(c_0) + b(c_0)u(c_1) + b(c_0)b(c_1)u(c_2) + \ldots$ The expressions for the additive and multiplicative term follow directly from this formula.

[27] For example, one may consider permutations of the consumption levels during the first N periods: e.g., using $(c_1, c_0, c_2, \ldots, c_{N-1})$ instead of $(c_0, c_1, c_2, \ldots, c_{N-1})$.

required. Moreover, either scale invariance or a specific form of time-dependent risk attitudes, as formalized in equation (13), are necessary for Stationarity to hold. Finally, knowing that the time aggregator W is affine, all restrictions on the certainty equivalents I_τ translate directly into restrictions on the certainty equivalent I used in recursion (12), which we employed in the formulation of our results. Indeed, when the certainty equivalents I_τ are scale-invariant, one can verify directly that $I_\tau = I$ for every τ. Thus, I inherits all the properties of the I_τ. When the certainty equivalents I_τ are only translation-invariant and discounting is exogenous, the relationship between the I_τ and I is given by equation (13). But then if the I_τ are translation-invariant, so is I.

6 Monotonicity in a two-period consumption-savings problem

This section illustrates the implications of Monotonicity in the context of a standard consumption-savings problem. One notable conclusion is that Monotonicity permits simple comparative statics regarding the role of risk aversion on the optimal level of saving. We consider a two-period economy. At date $t = 0$, the agent receives income y_0 which is certain and which she can allocate between consumption and savings. At date $t = 1$, one of two states, h or l, is realized. The states occur with probabilities $\pi_h \in (0, 1)$ and $\pi_l = 1 - \pi_h$, and determine both the level of income at date $t = 1$, equal to y_1^h or y_1^l, and the gross return on savings, equal to R_h or R_l.

Throughout this section, we assume that preferences over deterministic consumption paths are represented by the function $U(c_0, c_1) = (c_0^\rho + \beta c_1^\rho)^{\frac{1}{\rho}}$, where $1 > \rho \neq 0$ and $\beta > 0$. Risk preferences will either be unspecified, though assumed to be monotone (for Lemma 2) or preferences à la Epstein-Zin (for Lemma 3 and Figure 6), or risk-sensitive (for the monotone case in Figure 6).

In the presence of uncertainty, the agent has to choose a level of savings before observing the state of the world. We denote by c_0^* the optimal consumption at date $t = 0$ and by s^* the optimal level of savings. The budget constraints can be expressed as follows:

$$\begin{aligned} y_0 - s^* &= c_0^* \geq 0, \\ y_1^\kappa + R_\kappa s^* &= c_{1,\kappa}^* \geq 0 \text{ for } \kappa = h, l, \end{aligned}$$

where $c_{1,\kappa}^*$ denotes consumption at date $t = 1$ if state κ occurs. We use s_κ to denote the level of savings chosen if the agent had perfect foresight, that is, if she knew that state κ would occur for sure. We have:

$$s_\kappa = \frac{y_0 - y_1^\kappa (\beta R_\kappa)^{\frac{1}{\rho-1}}}{1 + R_\kappa (\beta R_\kappa)^{\frac{1}{\rho-1}}} \text{ for } \kappa = h, l.$$

Lemma 2 (Savings with monotone preferences) *Consider the savings problem described above. If preferences are monotone, then* $\min(s_h, s_l) \leq s^* \leq \max(s_h, s_l)$.

Proof. Assume that $s^* > \max(s_h, s_l)$, the case $s^* < \min(s_h, s_l)$ being completely symmetric. Since ordinal preferences are strictly convex, choosing $\hat{s} = \max(s_h, s_l)$ provides higher utility in both states of the world. This means that $(y_0 - \hat{s}, y_1^\kappa + R_\kappa \hat{s}) \succ (y_0 - s^*, y_1^\kappa + R_\kappa s^*)$ for both states $\kappa = h, l$, where $\succ$ denotes the strict preference relation. Then, Monotonicity implies:

$$(y_0 - \hat{s}, \pi_l(y_1^l + R_l \hat{s}) \oplus \pi_h(y_1^h + R_h \hat{s})) \succ (y_0 - s^*, \pi_l(y_1^l + R_l s^*) \oplus \pi_h(y_1^h + R_h s^*)),$$

which contradicts the optimality of s^*. ■

The above result reflects the fact that with monotone preferences, an agent would never choose a level of savings if another choice gives higher lifetime utility in both states of the world.

The result does not extend to non-monotone preferences. Indeed, assume that preferences can be represented by the function:

$$U^{EZ}(c_0, \tilde{c}_1) = \left(c_0^\rho + \beta \left(E[\tilde{c}_1^\alpha] \right)^{\frac{\rho}{\alpha}} \right)^{\frac{1}{\rho}}, \tag{14}$$

where E is the standard expectation operator and $\alpha \neq 0$ a parameter driving risk aversion, with larger α indicating lower risk aversion. These preferences were one of the recursive specifications introduced in Epstein and Zin (1989). Because they belong to the class of Kreps-Porteus preferences as well, we know from the analysis in Section 4.2 that they are not monotone whenever $\rho \neq \alpha$. Let s^{EZ} be the optimal level of savings for an agent with such preferences, that is, let

$$s^{EZ} = \arg \max_{s \in (-\min(\frac{y_1^l}{R_l}, \frac{y_1^h}{R_h}), y_0)} U^{EZ}(y_0 - s, \tilde{y}_1 + \tilde{R} s), \tag{15}$$

where $\tilde{y}_1$ and $\tilde{R}$ denote the state-contingent income and asset returns.

Lemma 3 (Savings with Epstein-Zin preferences) *Consider the savings problem described in equation (15). If* $\rho \neq \alpha$*, there exist values of* R_κ *and* y_1^κ*,* $\kappa = h, l$*, for which the agent chooses a level of savings* $s^{EZ} \notin [\min(s_h, s_l), \max(s_h, s_l)]$.

Proof. Assume $y_1^h \neq y_1^l$ and $R_\kappa = \frac{1}{\beta}(\frac{y_1^\kappa}{y_0})^{1-\rho}$ in state $\kappa = h, l$. In that case $s_h = s_l = 0$, so that $[\min(s_h, s_l), \max(s_h, s_l)] = \{0\}$. However, we have

$$\frac{d}{ds}\left(\log U^{EZ}(y_0 \ - s, \tilde{y}_1 + \tilde{R}s)\right)\Big|_{s=0} = \frac{y_0^{\rho-1}}{U^{EZ}(y_0 \ , \tilde{y}_1)}\left(\frac{E[\tilde{z}^{1-\frac{\rho}{\alpha}}]}{E[\tilde{z}]^{1-\frac{\rho}{\alpha}}} - 1\right), \tag{16}$$

where $\tilde{z} = \tilde{y}_1^\alpha$. Since $\rho \neq 0$ and $\rho \neq \alpha$, the function $x \mapsto x^{1-\frac{\rho}{\alpha}}$ is either strictly concave or strictly convex. Using Jensen inequality, the derivative (16) cannot be equal to zero. Thus $s^{EZ} \neq 0$ and therefore $s^{EZ} \notin [\min(s_h, s_l), \max(s_h, s_l)]$. Moreover, if the agent had perfect foresight, choosing $s = 0$ would be preferred to choosing s^{EZ}, no matter the state of the world. ■

Lemma 3 shows that an agent endowed with non-monotone preferences may choose a level of savings s^{EZ} even though a different choice leads to higher lifetime utility in both states of the world. To better understand the role of Monotonicity and the different conclusions of Lemmas 2 and 3, note that the proof of Lemma 3 builds on the particular case where the states h and l are such that, with perfect foresight, the saving decisions in both states would be identical, that is, $s_l = s_h$. The lifetime utilities in those states are however different. An agent, who lacks perfect foresight and has non-monotone preferences, may prefer to reduce the difference in lifetime utilities even if this reduces lifetime utility in both states. The saving decision s^{EZ} then responds to uncertainty and depends on the probabilities π_l and π_h. In contrast, Monotonicity implies that the willingness to reduce risk, no matter how strong, cannot lead to a choice that reduces lifetime utility in all states of the world. In the special case when $s_l = s_h$, this also means that the agent's saving decision is unaffected by the uncertainty.

Building on the two-period example in this section, we may also emphasize that, because of the restrictions it imposes, Monotonicity affords an intuitive understanding of the role of risk aversion and simple comparative statics. Indeed, choice under uncertainty can then be seen as making a trade-off between state-specific utilities. If preferences are monotone and convex, the agent's optimal choice has to maximize a (possibly endogenous) convex combination of ex-post lifetime utilities, just like a Pareto optimum has to maximize a convex combination of individual utilities. Risk aversion is then reflected in the weights that appear in this convex combination. In particular, stronger risk aversion requires that higher weights be assigned to the "bad states." Bommier, Chassagnon, and LeGrand, (2012) formalize this reasoning and show that, whenever Monotonicity is assumed, simple dominance arguments make it possible to derive general and intuitive conclusions about the role of risk aversion in many problems of interest.

A precautionary saving example. To illustrate the last point, consider a simpler version of the above consumption-savings problem whereby only income is random with $y_1^h > y_1^l$. Since the asset return is the same in both states, we have $c_{1,h} > c_{1,l}$ whatever the agent's saving decision. One can thus regard state h as the "good state" and state l as the "bad state." Saving choices are such that $s_h < s_l$. With monotone preferences, the optimal saving choice has to lie in the interval (s_h, s_l). Moreover, as is demonstrated in Bommier, Chassagnon, and LeGrand, (2012), an increase in risk aversion involves selecting a level of savings that is closer to s_l, the best response in the bad state. Intuitively, in the presence of income uncertainty, saving provides an imperfect insurance device which is more intensively used when the degree of risk aversion increases. Non-monotone preferences may deliver different results: (i) the agent may choose to save more than she would if any of the states, including the worst one, were to occur for sure, and (ii) the role of risk aversion may be non-monotonic. Figure 6 illustrates the contrast between the saving patterns obtained with monotone and non-monotone preferences.[28]

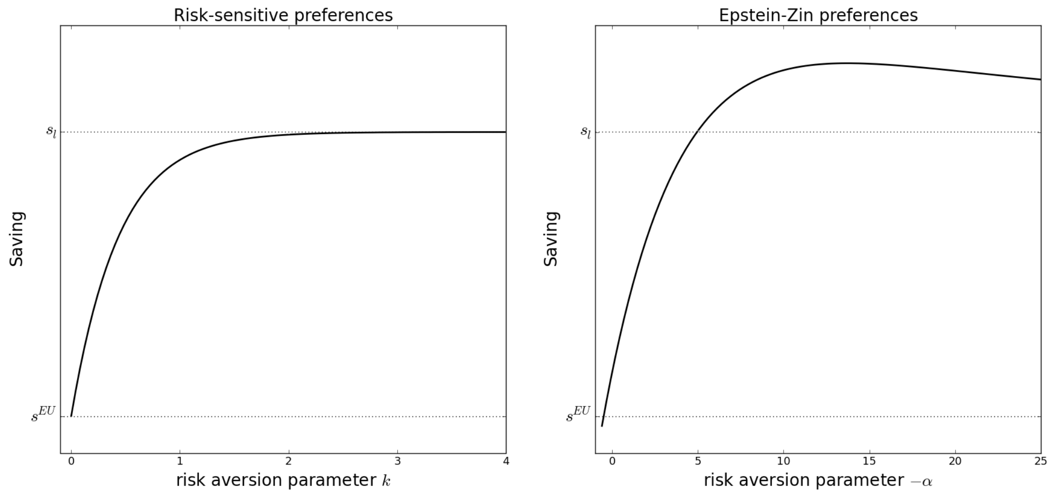

Figure 1: The relation between risk aversion and savings

[28]The graphs are built using risk-sensitive preferences, $U^{RS}(c_0, \tilde{c}_1) = c_0^\rho - \frac{\beta}{k} \log \left(E[e^{-k\tilde{c}_1^\rho}] \right)$, and Epstein-Zin preferences (equation 14). We plot the optimal savings as a function of the risk aversion parameter, k or $-\alpha$. We use the following parameters: $\rho = \frac{1}{2}$, implying an intertemporal elasticity of substitution equal to 2, $R_l = R_h = \beta = 1$, $y_0 = 100$, $y_1^l = 100$, $y_1^h = 125$ and $\pi_h = 1 - \pi_l = 5\%$. In other words, the agent has a probability of 5% of earning a bonus equal to a quarter of the base wage in the next period. The amount of savings s^{EU} reported on the graphs corresponds to what is obtained with the standard additive model (i.e. when $\alpha = \rho$ or when $k = 0$). The optimal amount of saving in the good state, s_h, lies further below and is not reported for reasons of scale.

7 Discussion

Our contribution may provide insight into the difficult question about which preference specification to use when modeling dynamic choice. Of course, there is no simple take-home message, as the answer may surely depend on the context. We may however stress some pros and cons that emerge from our analysis. As we explain below, the choice of the domain is not innocuous either.

First, the standard model of intertemporal choice, which is at the intersection of all models we mentioned, is well-behaved in all aspects and very tractable. Its main drawback, which is well-known, is its lack of flexibility. Risk aversion is indeed fully determined by the properties of ordinal preferences.

Maintaining recursivity and stationarity leaves a few options to get flexibility. One is to impose monotonicity, which leads to the preferences studied in this paper. If we restrict the certainty equivalent to be of the expected utility form, then we arrive at the class of risk-sensitive preferences. A particular feature of risk-sensitive preferences is that they are not homothetic, unless the elasticity of substitution is equal to one.[29] Homotheticity can be achieved by adopting a certainty equivalent I that is both translation- and scale-invariant, but is not of the expected utility form. Certainty equivalents based on the dual approach of Yaari (1987) fall into this category. By Proposition 2, however, attitudes toward the timing of uncertainty will be unstable: in some cases the agent will prefer early resolution of uncertainty, in others late resolution. This may be deemed unappealing if one believes that attitudes toward the timing of uncertainty are a fundamental trait of agent's preferences, like risk aversion.

Another option is to depart from monotonicity. The most popular specification from the work of Epstein and Zin (1989), see equation (14), is homothetic and has a certainty equivalent of the expected utility form. These preferences have proved to be very tractable and have been used in a large number of studies. As we explain in Section 6, however, there are applications in which abandoning monotonicity may result in comparative statics that are difficult to interpret.

One can also gain flexibility, while preserving recursivity and monotonicity, by weakening the notion of stationarity. One possibility, which maintains history independence, is to use the recursion in (11) along with certainty equivalents I_τ that are translation-invariant but not linked through the amplification mechanism of equation (13). The frameworks of Pye (1973) and van der Ploeg (1993) fall into this category. One can go further and allow for history

[29]Homotheticity has well-known advantages. Recent contributions, however, have emphasized that departing from homotheticity may help explain some empirical regularities such as the relationship between trade flows and income per capita (Fieler, 2011) or between wealth and stock holdings (Wachter and Yogo, 2010).

dependence, while preserving sufficient structure so as to maintain reasonable tractability. A solution is to consider a time-additive function $V : C^{\infty} \to \mathbb{R}$ together with certainty equivalents of the form $I_\tau = \phi_\tau^{-1} E \phi_\tau$. Though generating some history dependence, the history can be summarized by a single variable, the stock of accumulated welfare. Economic problems with such preferences can still be analyzed using standard dynamic programming techniques, with the introduction of only one additional state variable.[30]

The last point we want to make is that the choice of domain deserves careful consideration. Our findings about the relationship between risk aversion, time discounting, and the agent's attitudes toward the timing of resolution of uncertainty complete the initial results of Koopmans (1960, 1965) by showing that key features of preferences become intertwined when imposing stationarity in an infinite horizon setting. As is explained in Koopmans' papers, the infinite horizon setting constrains an agent with stationary preferences to exhibit a non-trivial rate of time preference, an aspect which several authors, e.g. Ramsey (1928), considered ethically indefensible. In comments that can be found in Koopmans (1965, pp. 298–300), Fisher argues that one should rather depart from the infinite horizon setting, which was initially introduced by Koopmans (1960, p. 287) so as "to avoid complications connected with the advancing age and finite life span of the individual consumer," than accept the existence of time preferences.[31] Following the same line of arguments, one might want to abandon the infinite horizon setting, or the stationarity assumption, so as to avoid the intertwinement of risk aversion, time preferences, and the agent's attitudes toward the timing of uncertainty. A possibility, while maintaining stationarity, is to replace the assumption of an infinite horizon by that of a possibly uncertain (but always finite) time horizon. This is done in Bommier (2013) who suggests a multiplicatively separable expected utility specification, which actually corresponds to the risk-sensitive preferences discussed in the current paper, with β being set to one. A simple separation between risk aversion and the elasticity of intertemporal substitution is obtained, without introducing a preference for early or late resolution of uncertainty. Interest in this approach is however restricted to cases where imposing strong asymptotic constraints, which can be justified by the inevitability of death as in Bommier (2013) or by the convergence to a maximal satisfaction level as in Ramsey (1928), may be considered appealing.

[30]Bommier (2008) uses such preferences to study life-cycle behavior, while assuming indifference toward the timing of the resolution of uncertainty (i.e., with functions ϕ_τ independent of τ) .

[31]According to Fisher, "The obvious conclusion from Koopmans' paper, therefore, seems to me to be that one ought to abandon the use of infinite horizons – not that one ought to abandon certain ethical notions."

References

Aczél, J. (1966): *Lectures on Functional Equations and Their Applications.* Academic Press, New York – London.

Aliprantis, C., and K. Border (1999): *Infinite Dimensional Analysis.* Springer-Verlag, 2nd edn.

Anscombe, F. J., and R. J. Aumann (1963): "A Definition of Subjective Probability," *The Annals of Mathematical Statistics*, 34(1), 199–205.

Backus, D. K., B. R. Routledge, and S. E. Zin (2005): "Exotic Preferences for Macroeconomists," *NBER Macroeconomics Annual 2004*, 19, 319–414.

Bommier, A. (2008): "Rational Impatience?," HAL-SHS hal-00441880, CNRS.

——— (2013): "Life Cycle Preferences Revisited," *Journal of European Economic Association*, 11(6), 1290–1319.

Bommier, A., A. Chassagnon, and F. LeGrand (2012): "Comparative Risk Aversion: A Formal Approach with Applications to Saving Behaviors," *Journal of Economic Theory*, 147(4), 1614–1641.

Bommier, A., A. Kochov, and F. LeGrand (2016): "Ambiguity or Correlation Aversion?," Discussion paper.

Bommier, A., and F. LeGrand (2014): "A Robust Approach to Risk Aversion," Discussion paper, ETH Zurich, http://e-collection.library.ethz.ch/eserv/eth:47441/eth-47441-01.pdf.

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and A. Rustichini (2015): "The Structure of Variational Preferences," *Journal of Mathematical Economics*, 57(C), 12–19.

Chew, S. H. (1983): "A Generalization of the Quasilinear Mean with Applications to the Measurement of Income Inequality and Decision Theory Resolving the Allais Paradox," *Econometrica*, 51(4), 1065–1092.

——— (1989): "Axiomatic Utility Theories with the Betweenness Property," *Annals of Operations Research*, 19(1), 273–298.

Chew, S. H., and L. G. Epstein (1990): "Nonexpected Utility Preferences in a Temporal Framework with an Application to Consumption-Savings Behaviour," *Journal of Economic Theory*, 50(1), 54–81.

——— (1991): "Recursive Utility Under Uncertainty," in *Equilibrium Theory in Infinite Dimensional Spaces*, ed. by A. M. Khan, and N. C. Yannelis, vol. 1 of *Studies in Economic Theory*, pp. 352–369. Springer Berlin Heidelberg.

Dekel, E. (1986): "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom," *Journal of Economic Theory*, 40(2), 304–318.

Drèze, J. H., and F. Modigliani (1972): "Consumption Decisions under Uncertainty," *Journal of Economic Theory*, 5(3), 308–335.

Epstein, L. G. (1983): "Stationary Cardinal Utility and Optimal Growth under Uncertainty," *Journal of Economic Theory*, 31(1), 133–152.

Epstein, L. G., E. Farhi, and T. Strzalecki (2014): "How Much Would You Pay to Resolve Long-Run Risk?," *American Economic Review*, 104(9), 2680–2697.

Epstein, L. G., and M. Schneider (2003a): "IID: Independently and Indistinguishably Distributed," *Journal of Economic Theory*, 113(1), 32–50.

Epstein, L. G., and M. Schneider (2003b): "Recursive Multiple-Priors," *Journal of Economic Theory*, 113(1), 1–31.

Epstein, L. G., and S. E. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57(4), 937–969.

——— (2001): "The Independence Axiom and Asset Returns," *Journal of Empirical Finance*, 8(5), 537–572.

Fieler, A. C. (2011): "Nonhomotheticity and Bilateral Trade: Evidence and a Quantitative Explanation," *Econometrica*, 79(4), 1069–1101.

Grant, S., A. Kajii, and B. Polak (2000): "Temporal Resolution of Uncertainty and Recursive Non-Expected Utility Models," *Econometrica*, 68(2), 425–434.

Hansen, L. P., and T. J. Sargent (1995): "Discounted Linear Exponential Quadratic Gaussian Control," *IEEE Transactions on Automatic Control*, 40(5), 968–971.

Hayashi, T. (2005): "Intertemporal Substitution, Risk Aversion and Ambiguity Aversion," *Economic Theory*, 25(4), 933–956.

Herstein, I. N., and J. Milnor (1953): "An Axiomatic Approach to Measurable Utility," *Econometrica*, 21(2), 291–297.

Johnsen, T. H., and J. B. Donaldson (1985): "The Structure of Intertemporal Preferences under Uncertainty and Time Consistent Plans," *Econometrica*, 53(6), 1451–1458.

Ju, N., and J. Miao (2012): "Ambiguity, Learning and Asset Returns," *Econometrica*, 80(2), 423–472.

Kimball, M. S. (1990): "Precautionary Savings in the Small and in the Large," *Econometrica*, 58(1), 53–73.

Kimball, M. S., and P. Weil (2009): "Precautionary Saving and Consumption Smoothing Across Time and Possibilities," *Journal of Money, Credit, and Banking*, 41(2-3), 245–284.

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

Koopmans, T. C. (1960): "Stationary Ordinal Utility and Impatience," *Econometrica*, 28(2), 287–309.

——— (1965): "On the Concept of Economic Growth," Cowles Foundation Paper 238, Reprinted from Academiae Scientiarum Scripta Varia 28(1), 1965.

Koopmans, T. C., P. A. Diamond, and R. E. Williamson (1964): "Stationary Utility and Time Perspective," *Econometrica*, 32(1/2), 82–100.

Kreps, D. M., and E. L. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46(1), 185–200.

Lundberg, A. (1982): "Generalized Distributivity for Real, Continuous Functions. I: Structure Theorems and Surjective Solutions," *Aequationes Mathematicae*, 24(1), 74–96.

——— (1985): "Generalized Distributivity for Real, Continuous Functions. II: Local Solutions in the Continuous Case," *Aequationes Mathematicae*, 28(1), 236–251.

——— (2005): "Variants of the Distributivity Equation Arising in Theories of Utility and Psychophysics," *Aequationes Mathematicae*, 69(1-2), 128–145.

Maccheroni, F., M. Marinacci, and A. Rustichini (2006): "Dynamic Variational Preferences," *Journal of Economic Theory*, 128(1), 4–44.

Pye, G. (1973): "Lifetime Portfolio Selection in Continuous Time for a Multiplicative Class of Utility Functions," *American Economic Review*, 63(5), 1013–1016.

Ramsey, F. P. (1928): "A Mathematical Theory of Saving," *Economic Journal*, 38(152), 543–559.

Segal, U. (1990): "Two-Stage Lotteries without the Reduction Axiom," *Econometrica*, 58(2), 349–377.

Selden, L. (1978): "A New Representation of Preferences over "Certain x Uncertain" Consumption Pairs: The "Ordinal Certainty Equivalent" Hypothesis," *Econometrica*, 46(5), 1045–1060.

Strzalecki, T. (2011): "Axiomatic Foundations of Multiplier Preferences," *Econometrica*, 79(1), 47–73.

——— (2013): "Temporal Resolution of Uncertainty and Recursive Models of Ambiguity Aversion," *Econometrica*, 81(3), 1039–1074.

Uzawa, H. (1968): "Time Preference, the Consumption Function, and Optimal Asset Holdings," in *Capital and Growth: Papers in Honor of Sir John Hicks*, ed. by J. N. Wolfe, pp. 65–84. Aldine, Chicago.

van der Ploeg, F. (1993): "A Closed-Form Solution for a Model of Precautionary Saving," *Review of Economic Studies*, 60(2), 385–395.

Wachter, J. A., and M. Yogo (2010): "Why Do Household Portfolio Shares Rise in Wealth?," *Review of Financial Studies*, 23(11), 3929–3965.

Yaari, M. E. (1987): "The Dual Theory of Choice under Risk," *Econometrica*, 55(1), 95–105.

Appendix

A Monotonicity in IID ambiguity models

In this section, we derive a result analogous to Proposition 1 in a stationary IID ambiguity setting, similar to the one of Strzalecki (2013).[32] By stationary IID ambiguity we mean (i): restricting the analysis to cases where the passing of time has no impact on the structure of the domain of choice; and (ii): introducing a set of assumptions implying that a decision maker who uses, at all dates, the same history independent preference relation is time consistent. Clearly, this is a restrictive approach, as it precludes the use of an arbitrary state space and rules out non trivial belief updating. This framework has however proved very insightful in several instances. The exploration of more general settings is left for further work.[33] For mathematical rigor, we provide an axiomatic derivation using assumptions that parallel Axioms 1 to 5 of the main body of the paper. This axiomatization implies a recursive utility representation. The main contribution involves then showing that, like in the risk setting, significant restrictions are further obtained when imposing monotonicity.

A.1 Setup

We consider a setup similar to that of Strzalecki (2013). Let S be a finite set representing the states of the world to be realized in each period. We assume that S has at least three elements and let $\Sigma := 2^S$ be the associated algebra of events. The full state space is $\Omega := S^\infty$, with a state $\omega \in \Omega$ specifying a complete history $(s_1, s_2, \ldots)$.[34] In each period $t > 0$, the individual knows the partial history $(s_1, \ldots, s_t)$. Such knowledge can be represented by a filtration $\mathcal{G} = (\mathcal{G}_t)_t$ on Ω where $\mathcal{G}_0 := \{\emptyset, \Omega\}$ and for every $t > 0$, $\mathcal{G}_t := \Sigma^t \times \{\emptyset, S\}^\infty$. To introduce the domain of choice, we again let $C = [\underline{c}, \overline{c}]$ be the set of all possible consumption levels. A *consumption plan*, or an *act*, is a C-valued, $\mathcal{G}$-adapted stochastic process, that is, a sequence $h = (h_0, h_1, \ldots)$ such that $h_t : \Omega \to C$ is $\mathcal{G}_t$-measurable for every t. The set of all consumption plans is denoted by $\mathcal{H}$ and endowed with the topology of pointwise convergence.

We consider a binary relation $\succeq$ on $\mathcal{H}$ and introduce a set of axioms similar to those of Section 3. The axioms Weak order, Continuity, and Monotonicity for Deterministic Prospects

[32]The notion of IID ambiguity was first introduced in Epstein and Schneider (2003a) in the case of max-min expected utility representation.

[33]An investigation of the role of Monotonicity in the subjective uncertainty of Ju and Miao (2012) can also be found in Bommier and LeGrand (2014).

[34]By setting $\Omega = S^\infty$ we constrain the state space to have a stationary structure (i.e., $\Omega = S \times \Omega$). If no such stationary structure were assumed, the passing of time would impact the structure of the preference domain. This means that (independently of time consistency issues) the same preferences could not be used at all dates. The domain of choice would simply change with time, which would require the use of different preference relations.

require no major modification. Below we state appropriate analogues for Axioms 3, 4, 5, and 7. Some notation is needed first. Given an act $h \in \mathcal{H}$ and state $\omega \in \Omega$, let $h(\omega) \in C^\infty$ be the deterministic consumption stream induced by h in state $\omega \in \Omega$, that is, $h(\omega) = (h_0, h_1(\omega), \ldots)$. Moreover for any act $h \in \mathcal{H}$ and any $s \in S$ we define the *conditional act* $h^s \in \mathcal{H}$ by

$$\forall \omega = (s_1, s_2, \ldots) \in \Omega : \; h^s(s_1, s_2, \ldots) = h(s, s_2, \ldots) = (h_0, h_1(s, s_2, \ldots), h_2(s, s_2, \ldots), \ldots). \tag{17}$$

The act h^s is therefore the act obtained from h when knowing that the first component of the state of the world is equal to $s \in S$. It is noteworthy that $h^s(s_1, s_2, \ldots)$ is independent of s_1.

We can construct the *continuation act* $h^{s,1} \in \mathcal{H}$ from the conditional act h^s by removing the first period consumption. Formally, for any act $h = (h_0, h_1, h_2, \ldots) \in \mathcal{H}$ and any $s \in S$, the continuation act $h^{s,1}$ is given by

$$\forall \omega = (s_1, s_2, \ldots) \in \Omega : \; h^{s,1}(s_1, s_2, \ldots) = (h_1(s, s_2, \ldots), h_2(s, s_2, \ldots), \ldots). \tag{18}$$

The continuation act $h^{s,1}$ can be viewed as the consumption plan implied by h starting at date 1 (ignoring date 0 consumption) and where the information revealed at the beginning of date 1 (i.e., s_1) is equal to s.

Last, for any consumption $c \in C$ and any act $h \in \mathcal{H}$, we define the *concatenated act* $(c, h) \in \mathcal{H}$ by

$$(c, h) : \; \omega = (s_1, s_2, \ldots) \in \Omega \mapsto (c, h)(\omega) = (c, h(s_2, \ldots)) \in C^\infty. \tag{19}$$

The notions of conditional, continuation and concatenated acts are related to each other. In particular, the conditional act is the concatenation of first period consumption and the continuation act. Formally:

$$h = (h_0, h_1, h_2, \ldots) \in \mathcal{H} \text{ and } s \in S \Rightarrow h^s = (h_0, h^{s,1}). \tag{20}$$

Moreover, any concatenated act (c, h) has continuation h. In mathematical terms, for any $c \in C$, $h \in \mathcal{H}$ and $s \in S$:

$$(c, h)^{s,1} = h. \tag{21}$$

We can now state the axioms that parallel those given in the risk setting in Section 3.

Axiom A.3 *For all acts $h = (h_0, h_1, h_2, \ldots)$ and $h' = (h'_0, h'_1, h'_2, \ldots)$ in $\mathcal{H}$ such that $h_0 = h'_0$:*

$$\left(\forall s \in S, \; h^s \succeq h'^s\right) \Rightarrow h \succeq h'.$$

If, in addition, one of the former rankings is strict, then the latter ranking is strict as well.

Axiom A.3 can be viewed as a restricted notion of monotonicity that replicates the one of Axiom 3. This axiom is a concise statement that embeds both a property of recursivity and of state independence, the latter being implicit in the risk setting.[35] To be precise, recursivity alone would involve stating that for any $h, h' \in \mathcal{H}$ and $\sigma \in S$ such that $h \succeq h'$, and $h^s = h'^s$ for all $s \neq \sigma$:

$$\big(g \in \mathcal{H},\ g' \in \mathcal{H},\ g^\sigma = h^\sigma,\ g'^\sigma = h'^\sigma \text{ and } g^s = g'^s \text{ for all } s \neq \sigma\big) \Rightarrow g \succeq g'.$$

Such a property of recursivity makes it possible to combine time consistency and consequentialism in dynamic frameworks (see Johnsen and Donaldson, 1985). State independence extends the requirement of having $g \succeq g'$ to cases where there exists a state of the world $\sigma' \in S$ (possibly different from σ) such that $g^{\sigma'} = h^\sigma$, $g'^{\sigma'} = h'^\sigma$ and $g^s = g'^s$ for all $s \neq \sigma'$. When plugged into a dynamic framework, the state independence property translates into a form of history independence, in the sense that preferences regarding the future have to be independent of which states realized in past periods. Many papers relax the state-independent assumption allowing for non-trivial updating of beliefs. A prominent example is Hayashi (2005), who provides axiomatic foundations for more general recursive preferences, in a more complex setting that combines both objective and subjective uncertainty. As already mentioned, we leave for further work the exploration of the consequences of assuming Monotonicity in such more general settings.

Axiom 4 rewrites as follows:

Axiom A.4 *For all acts $h = (h_0, h_1, h_2, \ldots)$ and $h' = (h_0, h'_1, h'_2 \ldots)$ in $\mathcal{H}$, and $h'_0 \in C$,*

$$(h_0, h_1, h_2, \ldots) \succeq (h_0, h'_1, h'_2, \ldots) \Leftrightarrow (h'_0, h_1, h_2, \ldots) \succeq (h'_0, h'_1, h'_2, \ldots).$$

Regarding stationarity, Axiom 5 becomes:

Axiom A.5 *For all $c \in C$ and $h, h' \in \mathcal{H}$,*

$$(c, h) \succeq (c, h') \Leftrightarrow h \succeq h'.$$

This assumption basically states that the comparison of two acts that assume the same deterministic consumption in period 0 and whose continuation acts are independent of the information revealed in the first period, can be done by comparing their respective continuation acts (with the same preference relationship $\succeq$).

[35] In the risk setting, state independence is readily imposed by the fact that preferences are defined over lotteries, and not over random variables.

To avoid confusion, we shall emphasize that our stationary assumption differs from that of Kochov (2015). In Kochov's paper, the information tree has an arbitrary exogenous structure, which does not allow him to define stationarity in the same way as we do. Kochov's stationarity is a property of preference invariance when changing the timing of consumption, while holding fixed the timing of resolution of uncertainty. In contradistinction, our stationarity assumption (Axiom A.5) is a property of preference invariance when changing both the timing of consumption and the timing of resolution of uncertainty. Indeed, for a given $c \in C$ and a given $h \in \mathcal{H}$, the concatenated act (c, h), as defined in equation (19) is obtained by adding one initial period consumption c and postponing the timing of resolution of uncertainty by one period. For example, if h only depends on information revealed in the first period, then (c, h) only depends on the information revealed in the second period. This simply aims at reflecting that today is the day after yesterday. With respect to the mathematical formalism, the difference between Kochov's approach and ours lies in the way the concatenation operation is defined.[36] This eventually leads to assumptions of different nature, unless the agent exhibits indifference to the timing of uncertainty resolution.

Central to our analysis is the assumption of Monotonicity:

Axiom A.7 (Monotonicity) *For any h and h' in $\mathcal{H}$:*

$$\big(h(\omega) \succeq h'(\omega) \text{ for all } \omega \in \Omega\big) \Rightarrow h \succeq h'. \tag{22}$$

The above monotonicity axiom can be found in Epstein and Schneider (2003b), Maccheroni, Marinacci, and Rustichini (2006), and Kochov (2015). It is important to note that this axiom is "stronger" than the one we used in the risk setting. An exact analogue of the risk axiom would restrict the acts h and g to depend on the uncertainty resolving in a single period only. As the analysis in the risk setting suggests, the representation result we state in Proposition 4 below would continue to hold even if we were to weaken Axiom A.7 accordingly. We adopt Axiom A.7 because the axiom is standard in the literature on subjective uncertainty and because we want to emphasize that the preferences we consider are in fact monotone in the strong sense of Axiom A.7.[37]

As in the risk setting, we say that a binary relation $\succeq$ on $\mathcal{H}$ is a *monotone recursive preference relation* if it satisfies Axioms 1, 2, A.3, A.4, A.5, 6, and A.7.

[36] In Kochov (2015) the concatenation (c, h) is defined by $(c, h)(s_1, s_2, \ldots) = (c, h(s_1, s_2, \ldots))$ which differs from the definition introduced in equation (19).

[37] As already mentioned, the axiom employed in the risk setting can be strengthened so as to obtain an analogue of Axiom A.7. The appropriate formulation is provided in Bommier and LeGrand (2014).

A.2 Representation result

Let $B_0(\Sigma)$ be the set of simple, Σ-measurable simple functions from S into $\mathbb{R}_+$. The next few definitions parallel those in Section 4.1. A *certainty equivalent* $I : B_0(\Sigma) \to \mathbb{R}_+$ is a continuous, strictly increasing function such that $I(x) = x$ for any $x \in \mathbb{R}_+$. A certainty equivalent is *translation-invariant* if for all $x \in \mathbb{R}_+$ and $f \in B_0(\Sigma)$, $I(x + f) = x + I(f)$. It is *scale-invariant* if for all $\lambda \in \mathbb{R}_+$ and $f \in B_0(\Sigma)$, $I(\lambda f) = \lambda I(f)$. Given a function $U : \mathcal{H} \to \mathbb{R}$ and an act $h \in \mathcal{H}$, we let $U \circ h^1$ denote the function $s \in S \mapsto U(h^{s,1})$. If U is a utility function, then $U \circ h^1$ is the state contingent profile of continuation utilities induced by the act h in period $t = 1$. Letting $W : C \times [0,1] \to [0,1]$ be a time aggregator as before, a *recursive representation* for $\succeq$ is a tuple (U, W, I) such that the function $U : \mathcal{H} \to \mathbb{R}$ represents $\succeq$ and satisfies the recursion:

$$U(h) = W(h_0, I(U \circ h^1)), \tag{23}$$
$$\text{where: } U \circ h^1 : \ s \in S \mapsto U(h^{s,1}).$$

It is relatively simple to show that Axioms 1, 2, A.3, A.4 and A.5 are necessary and sufficient conditions for preferences to have a recursive representation (Lemma 19 in Appendix D). Our contribution involves showing that further restrictions on the recursive representation appear when assuming preference monotonicity.

Proposition 4 *A binary relation $\succeq$ on $\mathcal{H}$ is a monotone recursive preference relation if and only if it admits a recursive representation (U, W, I) such that either:*

1. *I is translation-invariant and $W(c, x) = u(c) + \beta x$ satisfies the conditions listed in the first case of Proposition 1, or*

2. *I is translation- and scale- invariant and $W(c, x) = u(c) + b(c)x$ satisfies the conditions listed in the second case of Proposition 1.*

This proposition parallels Proposition 1 obtained in the risk setting. Its proof can be found in Section D of the Appendix.

B Proof of Proposition 1

Necessity of the axioms is obvious. Proving sufficiency is a long task, but a good account of the proof can be found by reading Section B.1 and the roadmap provided there. Before getting to that section, we introduce some notational conventions. $\mathbb{N}$ denotes the set of natural numbers including 0, while $\mathbb{N}_+$ denotes the set $\mathbb{N} \setminus \{0\}$. A vector in a Euclidean space $\mathbb{R}^m$ is denoted as a tuple $(x_1, \ldots, x_m)$ or by using a bold faced symbol $\mathbf{x}, \mathbf{y}$, etc. The composition of two functions f and g, when it is well defined, is denoted as fg or $f \circ g$. Given $n \in \mathbb{N}$ and a function $f : X \to X$, f^n denotes the n^{th}-iterate of the function f. Thus, for example, f^2 stands for the function ff. In what follows, we often work with an ambient space X and real valued functions f, f' that are defined on proper subsets of X. When we write $f(x)$, it is implicitly understood that x lies in the domain of f. Similarly, when we write $f > f'$, the expression is understood to hold for those $x \in X$ for which the functions f, f' are both defined.

B.1 Deriving a distributivity equation

Our first step is to show that under the axioms of Section 3 utility satisfies a distributivity equation of the form studied Lundberg (1982).

From Lemma 1, the preference relation $\succeq$ has a recursive representation (U, W, I). It is w.l.o.g. to assume that $U(D) = [0, 1]$.[38] Fix some $m \in \mathbb{N}_+, m > 2$. Let $\mathcal{W}_0 := [0, 1]^m$ and let

$$\mathcal{W}_1 := \{(W(c, x_1), \ldots, W(c, x_m)) : c \in C, (x_1, \ldots, x_m) \in \mathcal{W}_0\},$$
$$\mathcal{W}_2 := \{(W(c, x_1), \ldots, W(c, x_m)) : c \in C, (x_1, \ldots, x_m) \in \mathcal{W}_1\}.$$

Note that $\mathcal{W}_0 \supset \mathcal{W}_1 \supset \mathcal{W}_2$. Now fix a vector $(\pi_1, \ldots, \pi_m) \in (0, 1)^m$ such that $\sum_i \pi_i = 1$. For every vector $(x_1, \ldots, x_m) \in [0, 1]^m$, let $(\pi_1, x_1; \ldots; \pi_m, x_m)$ be the lottery in $M([0, 1])$ that gives x_k with probability π_k. Define a function $G_0 : \mathcal{W}_0 \to [0, 1]$ by letting

$$G_0(x_1, \ldots, x_m) := I((\pi_1, x_1; \ldots; \pi_m, x_m)) \quad \forall (x_1, \ldots, x_m) \in [0, 1]^m. \tag{24}$$

For $k \in \{1, 2\}$, define a function $G_k : \mathcal{W}_k \to [0, 1]$ inductively by letting

$$G_{k+1}(W(c, x_1), \ldots, W(c, x_m)) := W(c, G_k(x_1, \ldots, x_m)). \tag{25}$$

The functions $G_k, k \in \{1, 2\}$, are well defined by Monotonicity. For every $c \in C$, let F_c denote the function $x \mapsto W(c, x)$ from $[0, 1]$ into $[0, 1]$. Each function F_c is continuous and strictly increasing. With this notation, equation (25) becomes

$$G_{k+1}(F_c(x_1), \ldots, F_c(x_m)) = F_c G_k(x_1, \ldots, x_m)), \tag{26}$$

[38] Lemma 1 is an immediate consequence of Theorem 3.1 in Chew and Epstein (1991). One minor difference is that the latter paper studies preferences on $M(D)$ while we study preferences on D. Observe however that a recursive preference relation on D can be uniquely extended to $M(D)$: fix some $c \in C$ and for every $m, m' \in M(D)$ let $m \succeq^e m'$ if $(c, m) \succeq (c, m')$. By construction, $\succeq^e$ is a recursive preference relation in the sense of Chew and Epstein (1991). Theorem 3.1 in their paper delivers a representation for $\succeq^e$ which is also a recursive representation (U, W, I) for $\succeq$.

which holds for every $c \in C, k \in \{0,1\}$, and $(x_1,\ldots,x_m) \in \mathcal{W}_k$. From (25), we need to derive one more equation that plays a key role in the rest of the proof. To simplify our notation, let $\beta := W(\underline{c},1)$. If $F_{\overline{c}}(0) > \beta$, let c^* be such that $F_{c^*}(0) = \beta$. Alternatively, if $F_{\overline{c}}(0) \le \beta$, let $c^* := \overline{c}$. In each case, we have $F_c^{-1}[0,\beta] = [0, F_c^{-1}(\beta)] \ne \emptyset$ for every $c < c^*$. For every $c < c^*, k \in \{0,1\}$, and $(x_1,\ldots,x_m) \in [0,F_c^{-1}(\beta)]^m \cap \mathcal{W}_k$, we can apply $F_{\underline{c}}^{-1}$ to both sides of (26) to deduce that

$$F_{\underline{c}}^{-1} G_{k+1}(F_c(x_1),\ldots,F_c(x_m)) = F_{\underline{c}}^{-1} F_c G_k(x_1,\ldots,x_m). \tag{27}$$

Letting $c = \underline{c}$, conclude that

$$F_{\underline{c}}^{-1} G_{k+1}(F_{\underline{c}}(x_1),\ldots,F_{\underline{c}}(x_m)) = G_k(x_1,\ldots,x_m). \tag{28}$$

Combining (27) and (28) gives

$$G_k(F_{\underline{c}}^{-1}F_c(x_1),\ldots,F_{\underline{c}}^{-1}F_c(x_m)) = F_{\underline{c}}^{-1} F_c G_k(x_1,\ldots,x_m), \tag{29}$$

which holds for all $c < c^*, k \in \{0,1\}$, and $(x_1,\ldots,x_m) \in [0,F_c^{-1}(\beta)]^m \cap \mathcal{W}_k$. To simplify the exposition, let $f_c := F_{\underline{c}}^{-1}F_c$. With this change of notation, equation (29) becomes

$$G_k(f_c(x_1),\ldots,f_c(x_m)) = f_c G_k(x_1,\ldots,x_m). \tag{30}$$

The latter is a *distributivity equation* of the form studied in Lundberg (1982). When such an equation is specified, we say that f_c *solves the distributivity equation for* G_k. When G_k is clear from the context, we say simply that f_c solves the distributivity equation.

Roadmap for the remainder of the proof. The proof of Proposition 1 exploits the three distributivity equations obtained by setting $k = 0$, 1 and 2 in (30), and using (26) that relates the solutions of these distributivity equations. The proof is long (sections B.2 to B.9), so a brief roadmap may be helpful. In a first step (Sections B.2 and B.3), we follow Lundberg (1982) to show that one can build a collection of functions, formally an iteration group $\{f^\alpha\}_{\alpha\in(-1,1)}$, such that (30) continues to hold when f_c is replaced by f^α for any $\alpha \in (-1,1)$. One of the difficulties we have to address is that the domains $\mathcal{W}_1$ and $\mathcal{W}_2$ of G_1 and G_2 are not rectangles, which is a departure from Lundberg's (1982) framework. Once the iteration group $\{f^\alpha\}_{\alpha\in(-1,1)}$ is constructed, we can use the associated Abel function (Lemma 6) to proceed to a convenient utility renormalization (Section B.4), after which the renormalized certainty equivalents G_k become translation-invariant (Lemmas 8 and 10). We eventually obtain a pair of linear distributivity equations, (37) and (38) in Section B.6, which are solved in great generality in Lundberg (1985). The key here is equation (26) which constrains the solutions of these equations to be related in a specific way. This allows us to rule out a number of the potential solutions identified in Lundberg (1985). We are left with only two possible cases, which we explore further in Sections B.7 and B.8, and which end up providing the two cases listed in Proposition 1. The final part of the proof elicits the restrictions imposed by Axiom 6 and addresses some remaining technical issues.[39]

B.2 Constructing an iteration group

The proof requires some mathematical machinery from Lundberg (1982). First, given a proper interval $A \subset \mathbb{R}$, let $\mathcal{D}(A)$ be the set of all continuous, strictly increasing functions f whose domain and range are intervals contained in A and whose graphs disconnect A^2. Given $\lambda \in \mathbb{R} \cup \{+\infty\}$, a collection $\{f^\alpha : \alpha \in (-\lambda, \lambda)\} \subset \mathcal{D}(A)$ is *an iteration group on A* if $f^{\alpha+\alpha'} = f^\alpha f^{\alpha'}$ for all $\alpha, \alpha', \alpha + \alpha' \in (-\lambda, \lambda)$.[40] When no confusion arises, we suppress λ and the interval A and write $\{f^\alpha\}$ for an iteration group. A few remarks about the definition of an iteration group are in order. First, f^0 is necessarily the identity function on A. Moreover, if $1 \in (-\lambda, \lambda)$ and α is any other integer in $(-\lambda, \lambda)$, then f^α is the α-iterate of the function f^1. In fact, let $f := f^1$. We know how to define the α-iterate of f for any integer α. One can think of an iteration group as a way to define an α-iterate of the function f for any real number α, while ensuring (i) that the definition is consistent with the usual definition of an iterate for integer α, and (ii) that the different 'iterates', $f^\alpha, f^{\alpha'}$, and $f^{\alpha+\alpha'}$, do in fact 'iterate'. We should also point out that the index α has no meaning beyond encoding this second property. Formally, let $\gamma \neq 1$ be any real number and for every $\alpha \in (-\lambda, \lambda)$, define $g^{\alpha\gamma} := f^\alpha$. Then, $\{g^{\tilde{\alpha}} : \tilde{\alpha} \in (-\gamma\lambda, \gamma\lambda)\}$ is an iteration group on A and $\{g^{\tilde{\alpha}}\} = \{f^\alpha\}$. Thus, $\{g^{\tilde{\alpha}}\}$ is just a relabeling of $\{f^\alpha\}$. When we specify an iteration group $\{f^\alpha\}$, we assume that the group is *nontrivial*, that is, that $f^\alpha \neq f^0$ for at least one $\alpha \neq 0$. If the group is nontrivial, then $f^\alpha \neq f^0$ for all $\alpha \neq 0$. It should also be observed that $\lambda < +\infty$ whenever A is a bounded interval. For example, if $f^1(x) > x$ for all $x \in A$, then the graph of f^n lies outside of $A \times A$ for all n large enough, so that $f^n \notin \mathcal{D}(A)$. Finally, when we specify an iteration group $\{f^\alpha : \alpha \in (-\lambda, \lambda)\}$ on a bounded interval A, we assume that the group is *maximal*, that is, there is no other iteration group $\{g^\alpha : \alpha \in (-\lambda', \lambda')\} \subset \mathcal{D}(A)$ such that $\lambda' > \lambda$ and $g^\alpha = f^\alpha$ for all $\alpha \in (-\lambda, \lambda)$.

Let $(f_n)_n$ be a sequence of functions $f_n \in \mathcal{D}(A)$. A function $f \in \mathcal{D}(A)$ is the *closed limit of* $(f_n)_n$, which we denote as $f_n \to_L f$, if the graph of f is the closed limit of the graphs

[39] For some parts of the proof we have to work on a smaller domain, where consumption is restricted to be in $(\underline{c}, \overline{c})$ instead of $[\underline{c}, \overline{c}]$ but we show that the representation result extends by continuity.

[40] When A is a proper subset of $\mathbb{R}$, Lundberg (1982) calls the iteration group *truncated.* We have no occasion to distinguish between truncated and untruncated groups and use the term iteration group to denote both.

of the functions f_n.[41] If A is a closed interval and the graphs of f_n and f are closed, then $f_n \to_L f$ if and only if the graphs of f_n converge to the graph of f in the Hausdorff metric. We write $f_n \to_H f$ to denote the latter type of convergence. The sequence $(f_n)_n, f_n \in \mathcal{D}(A)$, *generates the iteration group* $\{f^\alpha\}$ on A if for every $\alpha \in (-\lambda, \lambda)$, there exists a sequence $(p_n)_n$ of integers such that $f_n^{p_n} \to_L f^\alpha$.

We come back to the proof of the theorem. Let j be the identity function on $[0, 1]$. Fix a sequence $(c_n)_n$ such that $c_n \in (\underline{c}, c^*)$ for every n and the sequence decreases monotonically to $\underline{c}$. Let $(f_{c_n})_n$ be the associated sequence of functions where $f_{c_n} = F_{\underline{c}}^{-1} F_{c_n}$ for every n. We note several properties of the sequence $(f_{c_n})_n$. First, $f_{c_n} > f_{c_{n+1}} > j$ for every n. Second, each function f_{c_n} has domain $\mathrm{Dom}_n := [0, F_{c_n}^{-1}(\beta)]$ and range $[f_{c_n}(0), 1]$. It follows that the graph of each function f_{c_n} disconnects $[0, 1]^2$ so that $f_{c_n} \in \mathcal{D}([0, 1])$. Another immediate implication is that $\mathrm{Dom}_n \to_H [0, 1]$. The latter implies that for every $x \in (0, 1)$, there is $k > 0$ such that $f_{c_n}(x)$ is defined for all $n \geq k$. The sequence $(f_{c_k}(x), f_{c_{k+1}}(x), \ldots)$ converges to x. The next lemma, whose proof is technical and can be skipped without loss of continuity, shows that the convergence is in fact uniform.

Lemma 4 (uniform convergence) $f_{c_n} \to_H j$.

Proof. Let Gr_n denote the graph of f_{c_n}. Let E' be a limit point of the sequence $(\mathrm{Gr}_n)_n$ in the Hausdorff metric. Let $E := \{(x, x) : x \in [0, 1]\}$, that is, E is the diagonal of the unit square $[0, 1]^2$. It is also the graph of the identity function j. For every $a \in (0, 1)$ and every n large enough, the functions f_{c_n} are defined on the interval $[0, a]$. Since the functions f_{c_n} converge monotonically to the identity function, we can apply Dini's theorem to conclude that the convergence is uniform when the functions are restricted to the interval $[0, a]$. But the uniform convergence of functions is equivalent to the Hausdorff convergence of their graphs. We conclude that the $E \cap ([0, a] \times [0, 1]) = E' \cap ([0, a] \times [0, 1])$. Since this is true for every $a < 1$, the intersections of E and E' with $[0, 1) \times [0, 1]$ coincide. Since the set E' is closed, we know that $(1, 1) \in E'$. Moreover, since $f_{c_n} > j$ for all n, the set E' 'lie above' E, that is, there is no pair $(1, x) \in [0, 1]^2$ such that $x < 1$ and $(1, x) \in E'$. We conclude that $E' = E$. Since the limit point E' of $(\mathrm{Gr}_n)_n$ was arbitrary, this concludes the proof. ■

The next two lemmas are key in terms of solving the distributivity equation.

Lemma 5 (constructing an iteration group) *There is an iteration group* $\{f^\alpha : \alpha \in (-\lambda, \lambda)\}$ *on* $(0, 1)$ *such that* $\lambda > 1$, $f^\alpha > j$ *for all* $\alpha > 0$, *and*

$$f^\alpha G_0(x_1, \ldots, x_m) = G_0(f^\alpha(x_1), \ldots, f^\alpha(x_m)) \tag{31}$$

[41]See Aliprantis and Border(1999, p. 109) for the definition of a closed limit.

for all $(x_1, \ldots, x_m) \in [0,1]^m$ *and* $\alpha \in (-\lambda, \lambda)$ *for which the equation is well defined.*

Proof. We know that $f_{c_n} \to_L j$, $f_{c_n} \neq j$ for every n, and $\mathrm{Dom}_n \to_L (0,1)$. Theorem 4.16 in Lundberg (1982) shows that the sequence $(f_{c_n})_n$ generates an iteration group $\{f^\alpha : \alpha \in (-\lambda, \lambda)\}$ on $(0,1)$ such that $\lambda > 1$ and (31) holds. Using the fact that $f_{c_n} > j$ for every n, one can furthermore show that $f^\alpha > j$ for all $\alpha > 0$. In particular, fix some $\alpha \in (0, \lambda)$. Since the iteration group is nontrivial, $f^\alpha \neq j$. Since $(f_{c_n})_n$ generates the group, there is a sequence $(p_n)_n$ of integers such that $f_{c_n}^{p_n} \to_L f^\alpha$. Since $f_{c_n} > j$ for every n, we can conclude that $f_{c_n}^{p_n} > j$ for every n and, hence, that $f^\alpha \geq j$. Lemmas 4.7 and 4.8 in Lundberg (1982) show that whenever a function f solves a distributivity equation such as the one in (31), $f \geq j$, and $f \neq j$, then $f > j$. ■

Lemma 6 (constructing an Abel function) *There is a continuous strictly increasing function* $L : (0,1) \to \mathbb{R}$ *such that* $f^\alpha(x) = L^{-1}(L(x) + \alpha)$ *for all* x *in the domain of* f^α *and all* $\alpha \in (-\lambda, \lambda)$.

Proof. We know that $f^\alpha > j$ for all $\alpha \in (0, \lambda)$. Since f^α is the inverse of $f^{-\alpha}$, the latter implies that $f^\alpha < j$ for all $\alpha \in (-\lambda, 0)$. In particular, none of the functions $f^\alpha, \alpha \neq 0$, has a fixed point. It follows that the iteration group has an *Abel function*, that is, a continuous function $L : (0,1) \to \mathbb{R}$ such that $f^\alpha(x) = L^{-1}(\alpha + L(x))$ for every $\alpha \in (-\lambda, \lambda)$ and every x in the domain of f^α. See Lundberg (1982, p. 79) for more details about Abel functions. Since $f^\alpha > j$ for all $\alpha > 0$, the function L is strictly increasing. Since each function f^α is continuous, the function L is continuous . ■

Recall that each function f_{c_n} is defined in a right neighborhood of 0. It follows that each function $f^\alpha, \alpha > 0$, is defined in a right neighborhood of zero. Similarly, each function $f^\alpha, \alpha < 1$, is defined in a left neighborhood of 1. For each $\alpha > 0$, let $f^\alpha(0) := \lim_{x \searrow 0} f^\alpha(x)$ and for each $\alpha < 0$, let $f^\alpha(1) := \lim_{x \nearrow 1} f^\alpha(x)$. Assume now that the iteration group $\{f^\alpha\}$ is such that $f^1(0) > 0$ and $f^{-1}(1) < 1$; Section B.9 shows how to modify the proof if either $f^1(0) = 0$ or $f^{-1}(1) = 1$. Under this assumption, we have $f^\alpha(0) > 0$ for all $\alpha > 0$ and $f^\alpha(1) < 1$ for all $\alpha < 0$. Another implication is that $L(0) := \lim_{x \searrow 0} L(x) < -\infty$ and $L(1) := \lim_{x \nearrow 1} L(x) < +\infty$. Using the latter, we now argue that the Abel function L can be chosen so that $L(0) = 0$ and $L(1) = 1$. First, observe that if L is an Abel function for the iteration group $\{f^\alpha\}$, then so is the function $L + l$ where $l \in \mathbb{R}$ is a constant. Thus, we can choose L so that $L(0) = 0$. To see that L can be chosen so that $L(1) = 1$, observe that $\lambda = \lim_{\alpha \nearrow \lambda} f^\alpha(0) = L(1)$. Relabeling the iteration group $\{f^\alpha : \alpha \in (-\lambda, \lambda)\}$ so that $\lambda = 1$ implies that $L(1) = 1$.

B.3 An extended distributivity equation for G_1

For every $c \in C$, let $A_c := [F_c(0), F_c(1)]$ and let A_c^m be the Cartesian product of m-copies of the set A_c. Observe that $\mathcal{W}_1 = \cup_c A_c^m$, i.e., the domain of G_1 is the union of product sets situated along the diagonal in $[0,1]^m$. Lemma 7 below shows that equation (31) continues to hold when the function G_0 is replaced with G_1. One important caveat is that the equation is only guaranteed to hold 'locally', that is, within each product set A_c^m rather than across the entire domain $\mathcal{W}_1$ of G_1. The proof of Lemma 7 clarifies why we can obtain only a local analogue of Lemma 5.

Lemma 7 *For every $c \in C$, $(x_1, \ldots, x_m) \in A_c^m$, and $\alpha \in (-1, 1)$ such that $(f^\alpha(x_1), \ldots, f^\alpha(x_m)) \in A_c^m$, we have $G_1(f^\alpha(x_1), \ldots, f^\alpha(x_m)) = f^\alpha G_1(x_1, \ldots, x_m)$.*

Proof. We need to establish a preliminary property of the distributivity equation in (30). Fix $c \in C$ and let f be a function such that $f > j$ and $fG_1(x_1, \ldots, x_m) = G_1(f(x_1), \ldots, f(x_m))$ for all $(x_1, \ldots, x_m) \in A_c^m$ such that $(f(x_1), \ldots, f(x_m)) \in A_c^m$. Let $p > 1$ be an integer and let $(x_1, \ldots, x_m) \in A_c^m$ be such that $(f^p(x_1), \ldots, f^p(x_m)) \in A_c^m$. We want to show that $f^p G_1(x_1, \ldots, x_m) = G_1(f^p(x_1), \ldots, f^p(x_m))$. Suppose first that $p = 2$. Then,

$$f^2 G_1(x_1, \ldots, x_m) = ffG_1(x_1, \ldots, x_m) = fG_1(f(x_1), \ldots, f(x_m)) \tag{32}$$

$$= G_1(f^2(x_1), \ldots, f^2(x_m)), \tag{33}$$

as desired. Next, fix an integer $p > 2$. For every integer p' such that $0 < p' < p$, we have

$$(x_1, \ldots, x_m) \leq (f^{p'}(x_1), \ldots, f^{p'}(x_m)) \leq (f^p(x_1), \ldots, f^p(x_m))$$

where $\leq$ is the pointwise order on $\mathbb{R}^m$. Since A_c^m is a product set and the vectors $(x_1, \ldots, x_m)$, $(f^p(x_1), \ldots, f^p(x_m))$ belong to A_c^m, it follows that the vector $(f^{p'}(x_1), \ldots, f^{p'}(x_m))$ belongs to A_c^m. But then a chain of equalities analogous to those in (32) and (33) shows that f^3 solves the distributivity equation. By induction, so do the functions $f^4, f^5, \ldots$, and f^p.

We can now complete the proof of the lemma. Take some $\alpha > 0$; symmetric arguments apply when $\alpha < 0$. Fix any $c \in C$. Take some $(x_1, \ldots, x_m)$ in the interior of A_c^m and $\alpha \in (0, \lambda)$ such that $(f^\alpha(x_1), \ldots, f^\alpha(x_m)) \in A_c^m$. We know that there is a sequence $(p_n)_n$ of integers such that $f_{c_n}^{p_n} \to_L f^\alpha$. For n large enough, we know that $(f_{c_n}^{p_n}(x_1), \ldots, f_{c_n}^{p_n}(x_m)) \in A_c^m$. Since (30) holds for each f_{c_n}, we know that $G_1(f_{c_n}^{p_n}(x_1), \ldots, f_{c_n}^{p_n}(x_m)) = f_{c_n}^{p_n} G_1(x_1, \ldots, x_m)$ for all n large enough. Since $f_{c_n}^{p_n} \to_L f^\alpha$ and G_1 is continuous, we have $G_1(f^\alpha(x_1), \ldots, f^\alpha(x_m)) = f^\alpha G_1(x_1, \ldots, x_m)$, as desired. ■

B.4 A monotone transformation of utility

Since $L : [0,1] \to [0,1]$ is strictly increasing, the function $\tilde{U} := LU : D \to [0,1]$ represents $\succeq$ on D. Moreover, the function $\tilde{U}$ is part of a recursive representation $(\tilde{U}, \tilde{W}, \tilde{I})$ where

$$\tilde{W}(c,x) := LW(c, L^{-1}(x)) \quad \forall x \in [0,1], c \in C,$$
$$\tilde{I}(\mu) := LI(\mu \circ L^{-1}) \quad \forall \mu \in M([0,1]).$$

For every $c \in C$, let $\tilde{F}_c := LF_cL^{-1}$. For $k \in \{0,1,2\}$, let $\tilde{G}_k(x_1, \ldots, x_m) := LG_k(L^{-1}(x_1), \ldots, L^{-1}(x_m))$. As before, define $\tilde{\mathcal{W}}_0 := [0,1]^m$ and inductively for $k \in \{1,2\}$,

$$\mathcal{W}_k := \{(\tilde{F}_c(x_1), \ldots, \tilde{F}_c(x_m)) : c \in C, (x_1, \ldots, x_m) \in \tilde{\mathcal{W}}_{k-1}\}. \tag{34}$$

Observe that, by definition, the function $\tilde{G}_k, k \in \{0,1,2\}$, has domain $\tilde{\mathcal{W}}_k$. Also, $\tilde{\mathcal{W}}_0 \supset \tilde{\mathcal{W}}_1 \supset \tilde{\mathcal{W}}_2$.

Lemma 8 (translation-invariance $\tilde{G}_0$) *For every $(x_1, \ldots, x_m) \in \tilde{\mathcal{W}}_0, \alpha \in (-1,1)$ such that $(\alpha + x_1, \ldots, \alpha + x_m) \in \tilde{\mathcal{W}}_0$, we have $\tilde{G}_0(\alpha + x_1, \ldots, \alpha + x_m) = \alpha + \tilde{G}_0(x_1, \ldots, x_m)$.*

Proof. Let $(x_1, \ldots, x_m)$ and α be as in the statement of the lemma. Let $y_i = L^{-1}(x_i)$ for $i = 1, \ldots, m$. Then,

$$\begin{aligned}
&\tilde{G}_0(\alpha + x_1, \ldots, \alpha + x_m) = \tilde{G}_0(\alpha + L(y_1), \ldots, \alpha + L(y_m)) = \\
&LG_0(L^{-1}(\alpha + L(y_1)), \ldots, L^{-1}(\alpha + L(y_m))) = LG_0(f^\alpha(y_1), \ldots, f^\alpha(y_m)) = \\
&Lf^\alpha(G_0(y_1, \ldots, y_m)) = L(G_0(y_1, \ldots, y_m)) + \alpha = L(G_1(L^{-1}(x_1), \ldots, L^{-1}(x_m))) + \alpha = \\
&\tilde{G}_0(x_1, \ldots, x_m) + \alpha.
\end{aligned}$$

■

For every $x_1, x_2 \in [0,1]$, define

$$\phi_0(x_2 - x_1) := \tilde{G}_0(x_1, x_2, x_2, \ldots, x_2) - x_1.$$

To see that ϕ_0 is well defined, take $x_1, x_2, y_1, y_2 \in [0,1]$ such that $x_2 - x_1 = y_2 - y_1$ and let $\mathbf{z} := (x_1, x_2, \ldots, x_2), \mathbf{z}' := (y_1, y_2, \ldots, y_2) \in \tilde{\mathcal{W}}_0$. Let $\alpha := y_1 - x_1 > 0$. By construction, $\mathbf{z}' = \mathbf{z} + \alpha$ and $\alpha \in (-1,1)$. But then $\tilde{G}_0(\mathbf{z} + \alpha) = \tilde{G}_0(\mathbf{z}) + \alpha$, which is equivalent to

$$\tilde{G}_0(x_1, x_2, x_2, \ldots, x_2) - x_1 = \tilde{G}_0(y_1, y_2, y_2, \ldots, y_2) - y_1,$$

showing that ϕ_0 is well defined. Finally observe that ϕ_0 has domain $[-1,1]$.

Lemma 9 *The function ϕ_0 is continuous and strictly increasing. The function $j - \phi_0$ is strictly decreasing.*

Proof. Only the second statement requires proof. Fix some x in $(-1, 1)$, that is, in the interior of ϕ_0's domain. Then, there is some $x' \in [0, 1]$ such that for all $\varepsilon > 0$ small enough, the vectors $\mathbf{z} := (x', x' + x, \ldots, x' + x), \mathbf{z} + \varepsilon$, and $\mathbf{z}' := \mathbf{z} + (0, \varepsilon, \varepsilon, \ldots, \varepsilon)$ belong to $\tilde{\mathcal{W}}_0$. Observe that $\mathbf{z}' < \mathbf{z} + \varepsilon$. It is enough to show that $\phi_0(x + \varepsilon) - \phi_0(x) - \varepsilon < 0$. From the definition of ϕ_0 and Lemma 8, we can deduce that $\phi_0(x+\varepsilon) - \phi_0(x) - \varepsilon = G_0(\mathbf{z}') - G_0(\mathbf{z}+\varepsilon)$, which is less than 0 since G_0 is strictly increasing. ■

B.5 An analogous construction for $\tilde{G}_1$ and $\tilde{G}_2$

For every $c \in C$, let $\tilde{A}_c := [\tilde{F}_c(0), \tilde{F}_c(1)]$ and let $\tilde{A}_c^m$ be the Cartesian product of m-copies of the set $\tilde{A}_c$. The proof of the next lemma parallels that of Lemma 8 and is omitted. As was the case with Lemmas 5 and 7, Lemma 10 is only a partial analogue of Lemma 8 in that its conclusion holds only within each separate rectangle $\tilde{A}_c^m$, rather than within the entire domain of $\tilde{G}_1$.

Lemma 10 (translation-invariance for $\tilde{G}_1$) *For every $c \in C$, $(x_1, \ldots, x_m) \in \tilde{A}_c^m$, and $\alpha \in (-1, 1)$ such that $(\alpha + x_1, \ldots, \alpha + x_m) \in \tilde{A}_c^m$, we have $\tilde{G}_1(\alpha + x_1, \ldots, \alpha + x_m) = \alpha + \tilde{G}_1(x_1, \ldots, x_m)$.*

Using the above lemma, we can define for every $c \in C$, $(x_1, x_2, \ldots, x_2) \in \tilde{A}_c^m$,

$$\phi_1^c(x_2 - x_1) := \tilde{G}_1(x_1, x_2, \ldots, x_2) - x_1.$$

An analogue of Lemma 9 shows that ϕ_1^c is a continuous, strictly increasing function and that $j - \phi_1^c$ is a strictly decreasing function. We omit the details. We should observe however that if $(x_1, x_2, \ldots, x_2) \in \tilde{A}_c^m$, then $(x_2, x_1, \ldots, x_1) \in \tilde{A}_c^m$. Thus, the domain of ϕ_1^c is an interval of the form $[-a_1^c, a_1^c]$. Since $\tilde{A}_c \subset [0, 1]$, we also know that $[-a_1^c, a_1^c] \subset [-1, 1]$.

We need an analogous construction for $\tilde{G}_2$ as well. Fix $c \in C$ and let

$$\tilde{A}_{cc}^m := \{(\tilde{F}_c(x_1), \ldots, \tilde{F}_c(x_m) : (x_1, \ldots, x_m) \in \tilde{A}_c^m\}.$$

An analogue of Lemma 10 shows that $\tilde{G}_2$ is translation invariant within the rectangle $\tilde{A}_{cc}^m$. Hence, we can define a function ϕ_2^c such that $\phi_2^c(x_2 - x_1) = \tilde{G}_2(x_1, x_2, \ldots, x_2) - x_1$ for all $(x_1, x_2, \ldots, x_2) \in \tilde{A}_{cc}^m$. Once again we omit the details. Finally, note that $\tilde{A}_{cc}^m \subset \tilde{A}_c^m$ and so ϕ_2^c is defined on an interval $[-a_2^c, a_2^c] \subset [-a_1^c, a_1^c]$.

B.6 Two linear distributivity equations

The functions $\tilde{G}_0, \tilde{G}_1, \tilde{G}_2$, and $\tilde{F}_c$ satisfy analogues of equation (26), that is,

$$\tilde{F}_c\tilde{G}_0(x_1, \ldots, x_m) = \tilde{G}_1(\tilde{F}_c(x_1), \ldots, \tilde{F}_c(x_m)), \tag{35}$$

$$\tilde{F}_c\tilde{G}_1(x_1, \ldots, x_m) = \tilde{G}_2(\tilde{F}_c(x_1), \ldots, \tilde{F}_c(x_m)), \tag{36}$$

where the first equation holds for all $c \in C$ and $(x_1, \ldots, x_m) \in \tilde{\mathcal{W}}_0$, while the second holds for all $c \in C$ and $(x_1, \ldots, x_m) \in \tilde{\mathcal{W}}_1$. From these equations and from the definitions of the functions $\phi_0, \phi_1^c, \phi_2^c$, we can deduce the two following linear distributivity equations

$$\tilde{F}_c(x_1 + \phi_0(x_2 - x_1)) = \tilde{F}_c(x_1) + \phi_1^c(\tilde{F}_c(x_2) - \tilde{F}_c(x_1)), \tag{37}$$

$$\tilde{F}_c(x_1 + \phi_1^c(x_2 - x_1)) = \tilde{F}_c(x_1) + \phi_2^c(\tilde{F}_c(x_2) - \tilde{F}_c(x_1)), \tag{38}$$

where the first equation holds for all $c \in C$ and $x_1, x_2 \in [0, 1]$, while the second holds for all c and x_1, x_2 such that $(x_1, x_2, x_2, \ldots, x_2) \in \tilde{A}_c^m$. Focus on the first equation:

$$\tilde{F}_c(x_1 + \phi_0(x_2 - x_1)) = \tilde{F}_c(x_1) + \phi_1^c(\tilde{F}_c(x_2) - \tilde{F}_c(x_1)), \tag{39}$$

which holds for all $c \in C$, $x_1, x_2 \in [0, 1]$. Think of (39) as a system of functional equations, one for each $c \in C$. Equations of this form are studied in Lundberg (1985). His results, Theorem 11.1 in particular, are applicable since the functions $\tilde{F}_c, \phi_0, \phi_1^c, \phi_2^c$ are strictly increasing and, by Lemma 9, the functions $j - \phi_0, j - \phi_1^c, j - \phi_2^c$ are strictly decreasing. Also, all functions are continuous. For any given $c \in C$, Theorem 11.1 in Lundberg (1985) shows that there are four cases for the functions $\tilde{F}_c, \phi_1^c$ that solve (39). As in Lundberg (1985), we enumerate those cases: a),b),c),d). In addition, we let $\Omega_{a)}$ be the set of all $c \in C$ such that the functions $\tilde{F}_c, \phi_1^c$ belong to case a). The sets $\Omega_{b)}, \Omega_{c)}, \Omega_{d)}$ are defined analogously. The next lemma shows that all but one of the sets 'Ω' is empty. Namely, the system of equations in (39) is solved by functions that belong to the same case.

Lemma 11 *$C = \Omega_k$ for some $k \in \{a), b), c), d)\}$.*

Proof. The four sets $\Omega_{a)}, \Omega_{b)}, \Omega_{c)}$, and $\Omega_{d)}$ form a partition of C. Since C is connected, it is enough to show each of these sets is open in C. For every $c \in C$, write (a_c, b_c) for the interval $(\tilde{F}_c(0), \tilde{F}_c(1))$. If $\Omega_{a)}$ is empty, it is necessarily open. So suppose $\Omega_{a)}$ is nonempty and fix some $c' \in \Omega_{a)}$. Since the functions $c \mapsto a_c$ and $c \mapsto b_c$ are continuous, we can find $\varepsilon > 0$ such that for all $c'' \in (c', c' + \varepsilon) \cap C$, we have $a_{c''} \in (a_{c'}, b_{c'})$ and for all $c'' \in (c' - \varepsilon, c') \cap C$, we have $b_{c''} \in (a_{c'}, b_{c'})$. In other words, for all c'' sufficiently close to c', the intervals $(a_{c'}, b_{c'})$ and $(a_{c''}, b_{c''})$ have a nonempty intersection. To show that $\Omega_{a)}$ is open in C, it is enough to show that the neighborhood $(c'-\varepsilon, c'+\varepsilon)\cap C$ of c' is a subset of $\Omega_{a)}$. First, take some $c'' \in (c', c'+\varepsilon)\cap C$. For every $x_1, x_2 \in (a_{c''}, b_{c'}) = (a_{c'}, b_{c'}) \cap (a_{c''}, b_{c''})$, we know from the definitions of $\phi_1^{c'}, \phi_1^{c''}$ that

$$\phi_1^{c'}(x_2 - x_1) = \tilde{G}_1(x_1, x_2, x_2, \ldots, x_2) \tag{40}$$

$$\phi_1^{c''}(x_2 - x_1) = \tilde{G}_1(x_1, x_2, x_2, \ldots, x_2). \tag{41}$$

Note that if $x_1, x_2 \in (a_{c''}, b_{c'})$, then $x_2 - x_1 \in (a_{c''} - b_{c'}, b_{c'} - a_{c''})$. From (40), conclude that $\phi_1^{c'}, \phi_1^{c''}$ coincide on the interval $(a_{c''} - b_{c'}, b_{c'} - a_{c''})$, which is a symmetric, nontrivial neighborhood of 0. From Theorem 11.1 in Lundberg, if $\phi_{c'}^1, \phi_{c''}^1$ belong to different cases, they cannot coincide on any nontrivial interval. We conclude that $c'' \in \Omega_{a)}$. Analogous arguments show that $(c'-\varepsilon, c')\cap C \subset \Omega_{a)}$ and, hence, that $\Omega_{a)}$ is open in C. Similarly, the sets $\Omega_{b)}, \Omega_{c)}, \Omega_{d)}$ are open in C, completing the proof of the lemma. ■

We now argue that either $C = \Omega_{a)}$ or $C = \Omega_{d)}$, that is, we can rule out cases b) and c). To do so, fix some $c \in C$ and consider the pair of equations:

$$\tilde{F}_c(x_1 + \phi_0(x_2 - x_1)) = \tilde{F}_c(x_1) + \phi_1^c(\tilde{F}_c(x_2) - \tilde{F}_c(x_1)), \tag{42}$$

$$\tilde{F}_c(x_1 + \phi_1^c(x_2 - x_1)) = \tilde{F}_c(x_1) + \phi_2^c(\tilde{F}_c(x_2) - \tilde{F}_c(x_1)). \tag{43}$$

The first important observation is that the two equations are linked by the function $\tilde{F}_c$ which appears in both. This implies that if the solutions to one of the equations belong to a given case, then so do the solutions to the other equation. Second observe that the function ϕ_1^c appears in both equations but in a 'different position' within each equation. This rules out cases b) and c) in Lundberg (1985) since in those cases functions that appear in 'different positions' cannot be the same. Suppose in particular that both equations have solutions belonging to case b). Applied to the first equation, Theorem 11.1 in (1985) tells us that ϕ_c^1 must be piecewise linear. Applied to the second equation, the theorem tells us that ϕ_c^1 is not piecewise linear, a contradiction. An analogous argument rules out case c).

We now focus on cases a) and b) which we refer to as the affine and non-affine case respectively.

B.7 The affine case

If all functions $\tilde{F}_c$ belong to case a), then $\tilde{F}_c(x) = u(c) + b(c)x$ for every $c \in C$ and $x \in [0,1]$. Moreover, the functions $u, b : C \to \mathbb{R}$ are continuous and $b(C) \subset (0,1)$. If b is a constant function, there is little left to prove since we already know that $\tilde{G}_0$ is translation invariant. See Section B.9 for the remaining details. Here, suppose that the function b is not constant.

We begin with a general lemma concerning the scale invariance of a real valued function G' defined on a convex set in a Euclidean space.

Lemma 12 *Let $\mathcal{W}'$ be a convex set in $\mathbb{R}^m$ containing the origin and G' be a continuous function from $\mathcal{W}'$ into $\mathbb{R}$. Suppose that for every $\mathbf{x} \in \mathcal{W}'$, there is $\varepsilon \in (0,1)$ such that $G'(\alpha\mathbf{x}) = \alpha G'(\mathbf{x})$ for all $\alpha \in (1-\varepsilon, 1]$. Then, $G'(\alpha\mathbf{x}) = \alpha G'(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{W}'$ and all $\alpha > 0$ such that $\alpha\mathbf{x} \in \mathcal{W}'$.*

Proof. Pick $\mathbf{x} \in \mathcal{W}'$. It is enough to show that $G'(\gamma\mathbf{x}) = \gamma G'(\mathbf{x})$ for all $\gamma \in (0,1]$. We proceed by way of contradiction. Let assume that there is $\gamma' \in (0,1)$ such that $G'(\gamma\mathbf{x}) = \gamma G'(\mathbf{x})$ for all $\gamma \in [\gamma', 1]$ and $G'(\gamma\mathbf{x}) \neq \gamma G'(\mathbf{x})$ for all γ in a left neighborhood of γ'. But we know that there is $\varepsilon_{\gamma'\mathbf{x}} > 0$ such that $G'(\alpha\gamma'\mathbf{x}) = \alpha G'(\gamma'\mathbf{x})$ for all $\alpha \in (1-\varepsilon_{\gamma'\mathbf{x}}, 1]$. Also, by the definition of γ', $\alpha G'(\gamma'\mathbf{x}) = \alpha\gamma' G'(\mathbf{x})$ and, hence, $G'(\alpha\gamma'\mathbf{x}) = \alpha\gamma' G'(\mathbf{x})$ for all $\alpha \in (1-\varepsilon_{\gamma'\mathbf{x}}, 1]$, contradicting the fact that $G'(\gamma\mathbf{x}) \neq \gamma G'(\mathbf{x})$ for all γ in some left neighborhood of γ'. ■

When a function $G' : \mathcal{W}' \to \mathbb{R}$ has the property deduced in Lemma 12, we say that G' is *scale invariant on $\mathcal{W}'$*.

Lemma 13 *If $b : C \to (0,1)$ is non-constant, then $\tilde{G}_0$ is scale invariant on $\tilde{\mathcal{W}}_0$.*

Proof. For every $c, c' \in C$ and $\mathbf{x}$ in the interior of $\tilde{\mathcal{W}}_0$, let

$$\mathbf{y} := \frac{u(c') - u(c)}{b(c)} + \frac{b(c')}{b(c)}\mathbf{x}. \tag{44}$$

Observe that if c' is sufficiently close to c, then $\mathbf{y}$ is close to $\mathbf{x}$ and hence $\mathbf{y} \in \tilde{\mathcal{W}}_0$. Similarly, we can insure that $\frac{b(c')}{b(c)}\mathbf{x} \in \tilde{\mathcal{W}}_0$. From now on, assume that c, c' are chosen so that both inclusions hold. From the definition of $\mathbf{y}$, conclude that $u(c) + b(c)\mathbf{y} = u(c') + b(c')\mathbf{x}$. Hence, $\tilde{G}_1(u(c) + b(c)\mathbf{y}) = \tilde{G}_1(u(c') + b(c')\mathbf{x})$. Since $\tilde{G}_0$ and $\tilde{G}_1$ satisfy an analogue of equation (26), conclude that $u(c) + b(c)\tilde{G}_0(\mathbf{y}) = u(c') + b(c')\tilde{G}_0(\mathbf{y})$. Substituting the expression for $\mathbf{y}$ from (44), we get

$$\tilde{G}_0(\mathbf{x}) = \frac{u(c) - u(c')}{b(c')} + \frac{b(c)}{b(c')}\tilde{G}_0\Big(\frac{u(c') - u(c)}{b(c)} + \frac{b(c')}{b(c)}\mathbf{x}\Big). \tag{45}$$

Since $\mathbf{y} \in \tilde{\mathcal{W}}_1$ and $\frac{b(c')}{b(c)}\mathbf{x} \in \tilde{\mathcal{W}}_1$, we can apply Lemma 8 and deduce that

$$\tilde{G}_0(\mathbf{x}) = \frac{b(c)}{b(c')}\tilde{G}_0\big(\frac{b(c')}{b(c)}\mathbf{x}\big).$$

Since b is non-constant, we can also choose c, c' so that $b(c) > b(c')$. Since b is continuous and C a connected set, we can also vary c, c' so that $\frac{b(c')}{b(c)}$ spans an open interval of the form $(1-\varepsilon, 1]$. It follows from Lemma 12 that $\tilde{G}_0$ is scale invariant on $\tilde{\mathcal{W}}_0$. ■

B.8 The non-affine case

If all functions $\tilde{F}_c$ belong to case d), then

$$\tilde{F}_c(x) = \frac{1}{a}\log(u(c) + b(c)e^{ax}) \quad \forall c \in C, x \in [0,1], \tag{46}$$

where $u, b : C \to \mathbb{R}$ are continuous functions, $a \in (0, +\infty)$, and $b(C) \subset (0,1)$. Let $H(x) := e^{ax}$ and observe that $H([0,1]) = [1, e^a]$ and $H^{-1}(y) = \frac{1}{a}\log y$. For every $c \in C$, let $\hat{F}_c := H\tilde{F}H^{-1}$. Each function $\hat{F}_c$ has domain $[1, e^a]$ and, by construction, $\hat{F}_c(x) = u(c) + b(c)x$ for every $c \in C$ and $x \in [1, e^a]$. Also, let $\hat{\mathcal{W}}_0 := [1, e^a]^m$ and

$$\hat{\mathcal{W}}_k := \{(\hat{F}_c(x_1), \ldots, \hat{F}_c(x_m)), c \in C, (x_1, \ldots, x_m) \in \hat{\mathcal{W}}_{k-1}\}$$

for $k \in \{1, 2\}$. For $k \in \{0, 1, 2\}$ and every $(x_1, \ldots, x_m) \in \hat{\mathcal{W}}_k$, let

$$\hat{G}_k(x_1, \ldots, x_m) := H\tilde{G}_k(H^{-1}(x_1), \ldots, H^{-1}(x_1)).$$

The next two lemmas focus on the function $\hat{G}_0$.

Lemma 14 *The function $\hat{G}_0 : \hat{\mathcal{W}}_0 \to \mathbb{R}$ is scale invariant on $\hat{\mathcal{W}}_0$.*

Proof. Fix $\mathbf{x} = (x_1, \ldots, x_m) \in \hat{\mathcal{W}}_0$ and $\alpha > 0$ such that $\alpha\mathbf{x} \in \hat{\mathcal{W}}_0$. Since $x_i, \alpha x_i \in [1, e^a]$ for every i, we know that $H^{-1}(x_i) \in [0,1]$ and $H^{-1}(\alpha x_i) = H^{-1}(x_i) + H^{-1}(\alpha) \in [0,1]$ for every $i = 1, \ldots, m$. Using Lemma 8 and the definition of $\hat{G}_0$, deduce that

$$\begin{aligned}
\hat{G}_0(\alpha\mathbf{x}) &= H\tilde{G}_0\big(H^{-1}(\alpha x_1), \ldots, H^{-1}(\alpha x_m)\big) \\
&= H\tilde{G}_0\big(H^{-1}(x_1) + H^{-1}(\alpha), \ldots, H^{-1}(x_m) + H^{-1}(\alpha)\big) \\
&= H\big[\tilde{G}_0(H^{-1}(x_1), \ldots, H^{-1}(x_m)) + H^{-1}(\alpha)\big] = \alpha H\tilde{G}_0(H^{-1}(x_1), \ldots, H^{-1}(x_m)) \\
&= \alpha\hat{G}_0(\mathbf{x}).
\end{aligned}$$

■

Lemma 15 *For every $\mathbf{x}$ in the interior of $\hat{\mathcal{W}}_0$, there is some $\delta_{\mathbf{x}} > 0$ such that $\hat{G}_0(\mathbf{x} + \delta) = \hat{G}_0(\mathbf{x}) + \delta$ for all $\delta \in [0, \delta_{\mathbf{x}}]$.*

Proof. Suppose first that u is a non-constant function. By construction, the functions $\hat{F}_c, \hat{G}_0$, and $\hat{G}_1$ satisfy an analogue of equation (35), that is,

$$u(c) + b(c)\hat{G}_0(\mathbf{x}) = \hat{G}_1\big(u(c) + b(c)\mathbf{x}\big) \tag{47}$$

for every $c \in C$ and $\mathbf{x} \in \hat{\mathcal{W}}_0$. For every $c, c' \in C$ and every $\mathbf{x}$ in the interior of $\hat{\mathcal{W}}_0$, let

$$\mathbf{y} := \frac{u(c') - u(c)}{b(c)} + \frac{b(c')}{b(c)}\mathbf{x}.$$

Using (47), deduce that

$$\hat{G}_0(\mathbf{x}) = \frac{u(c) - u(c')}{b(c')} + \frac{b(c)}{b(c')}\hat{G}_0\Big(\frac{u(c') - u(c)}{b(c)} + \frac{b(c')}{b(c)}\mathbf{x}\Big). \tag{48}$$

If c, c' are close to one another, then $\frac{b(c)}{b(c')}\mathbf{y}, \mathbf{y} \in \hat{\mathcal{W}}_0$. From Lemma 14, we can conclude that $\frac{b(c)}{b(c')}\hat{G}_0(\mathbf{y}) = \hat{G}_0(\frac{b(c)}{b(c')}\mathbf{y})$. Then, (48) becomes

$$\hat{G}_0(\mathbf{x}) = \frac{u(c) - u(c')}{b(c')} + \hat{G}_0\Big(\frac{u(c') - u(c)}{b(c')} + \mathbf{x}\Big). \tag{49}$$

Summarizing the arguments so far, we can insure that (49) holds for all $\mathbf{x}$ in the interior of $\hat{\mathcal{W}}_0$, all $c \in C$, and all c' in some neighborhood $O_{\mathbf{x},c}$ of c. Since $u : C \to \mathbb{R}$ is non-constant, we can choose c such that u is non-constant in some right neighborhood of c. But then (49) implies that $\hat{G}_0(\mathbf{x}) = -\delta + \hat{G}_0(\delta + \mathbf{x})$ for all $\mathbf{x}$ in the interior of $\hat{\mathcal{W}}_0$ and all $\delta > 0$ less than some $\delta_{\mathbf{x}} > 0$, as we wanted to prove.

Finally suppose that u is a constant function. The functions $\hat{F}_c, \hat{G}_0, \hat{G}_1$, and $\hat{G}_2$ satisfy equations analogous to equations (35) and (36). Deduce that

$$u(c) + b(c)u(c) + b(c)b(c)\hat{G}_0(\mathbf{x}) = \hat{G}_2\big(u(c) + b(c)u(c) + b(c)b(c)\mathbf{x}\big) \tag{50}$$

for every $c \in C$ and $\mathbf{x} \in \hat{\mathcal{W}}_0$. For every $c \in C$, let $v(c) := u(c)(1 + b(c))$ and $\gamma(c) := b(c)b(c)$. Observe that if $u : C \to \mathbb{X}$ is a constant function, then $b : C \to \mathbb{R}$ is necessarily non-constant. Otherwise, $\succeq$ fails to be strictly increasing in the pointwise order on C^{∞}. Conclude that v is necessarily a non-constant function. Then, (50) becomes

$$v(c) + \gamma(c)\hat{G}_0(\mathbf{x}) = \hat{G}_2\big(v(c) + \gamma(c)\mathbf{x}\big), \tag{51}$$

which holds for every $c \in C, \mathbf{x} \in \hat{\mathcal{W}}_0$. But this equation is an exact analogue of equation (47), with the function v non-constant. Hence, the proof can be completed in an identical manner.
■

The next lemma shows that the local property obtained in Lemma (15) 'integrates' into a global property. The proof is analogous to that of Lemma 12 and is omitted.

Lemma 16 *For every $\mathbf{x} \in \hat{\mathcal{W}}_0$ and every $\delta \in \mathbb{R}$ such that $\mathbf{x} + \delta \in \hat{\mathcal{W}}_0$, we have $\hat{G}_0(\mathbf{x} + \delta) = \hat{G}_0(\mathbf{x}) + \delta$.*

B.9 Concluding the proof of Proposition 1

The preceding arguments show that is always possible to renormalize the utility representation so as to have an affine time aggregator and a renormalized certainty equivalent ($\tilde{G}_0$ in the affine case and $\hat{G}_0$ in the non-affine case) which is translation invariant (Lemmas 8 and 16). Moreover this renormalized certainly equivalent has to be scale invariant in the case of endogenous discounting (Lemmas 13 and 14). Recall from (24) that G_0 was defined by fixing $m > 1$ and a probability vector $(\pi_1, \ldots, \pi_m)$ and projecting I onto $[0,1]^m$. Since m and $(\pi_1, \ldots, \pi_m)$ were arbitrary, we obtain that the recursive representation (U, W, I) of $\succeq$ can be renormalized so that:

- $W(c, x) = u(c) + \beta x$ and I is translation invariant on $M^f(\mathcal{U})$
- $W(c, x) = u(c) + b(c)x$ and I is translation- and scale-invariant on $M^f(\mathcal{U})$,

where $\mathcal{U} := U(D)$ and $M^f(\mathcal{U})$ is the set of simple lotteries with prizes drawn from the interval $\mathcal{U}$. In the first case, $u : C \to \mathbb{R}$ is continuous and $\beta \in (0,1)$. In the second, the functions $u, b : C \to \mathbb{R}$ are continuous and $b(C) \subset (0,1)$. Since $M^f(\mathcal{U})$ is dense in $M(\mathcal{U})$ and the certainty equivalent $I : M(\mathcal{U}) \to \mathcal{U}$ is continuous, we know that if I is translation-invariant on $M^f(\mathcal{U})$, then I is also translation-invariant on $M(\mathcal{U})$. An identical argument holds for scale invariance. It remains to show that in the second case the functions $u, u + b : C \to \mathbb{R}$ are strictly increasing whenever utility is normalized so that $\mathcal{U} = [0,1]$. To that end, say that a preference relation $\succeq$ on C^∞ has an *Uzawa representation* (u, b, U) if it is represented by the utility function

$$
\begin{aligned}
U(c_0, c_1, \ldots) &= u(c_0) + b(c_0)u(c_1) + b(c_0)b(c_1)u(c_2) + \ldots \\
&= u(c_0) + b(c_0)U(c_1, c_2, \ldots),
\end{aligned}
$$

where $u, b : C \to \mathbb{R}$ are continuous functions and $b(C) \subset (0,1)$.

Lemma 17 *Suppose a preference relation $\succeq$ on C^∞ has an Uzawa representation (u, b, U) normalized so that $U(C^\infty) = [0,1]$. The preference relation $\succeq$ is strictly increasing in the pointwise order on C^∞ if and only if the functions $u, u + b : C \to \mathbb{R}$ are strictly increasing.*

Proof. Suppose first that U is strictly increasing in the pointwise order on C^∞. Since U is normalized so that $U(C^\infty) = [0,1]$, we have $U(c, \underline{c}, \underline{c}, \ldots) = u(c)$ and $U(c, \bar{c}, \bar{c}, \ldots) = u(c) + b(c)$ for every $c \in C$. Thus, the functions $u, u + b : C \to \mathbb{R}$ are strictly increasing. Conversely, suppose that

$$u(c') > u(c)$$
$$u(c') + b(c') > u(c) + b(c)$$

for all $c, c' \in C$ such that $c' > c$. Fix $x \in [0,1]$. After multiplying the first inequality by $(1-x)$ and the second by x, and adding the resulting inequalities, we obtain

$$u(c') + b(c')x > u(c) + b(c)x. \tag{52}$$

We conclude that $U(c', c_1, c_2, \ldots) > U(c, c_1, c_2, \ldots)$ for all $c' > c$ and $(c_1, c_2, \ldots) \in C^\infty$. Since U is recursive and continuous, a standard backward-induction argument shows that U is strictly increasing on C^∞. ■

Lemma 17 completes the proof of Proposition 1 provided that the iteration group $\{f^\alpha\}$ obtained in Lemma 5 is such that $f^1(0) > 0$ and $f^{-1}(1) < 1$. This meant that the Abel function $L : (0,1) \to \mathbb{R}$ is bounded, which allowed us to extend L continuously from $(0,1)$ to $[0,1]$. If either $f^1(0) = 0$ or $f^{-1}(1) = 1$, then the Abel function $L : (0,1) \to \mathbb{R}$ is no longer bounded and cannot be continuously extended to $[0,1]$. To see how this affects the preceding proof, note that we started with a utility function $U : D \to [0,1]$ and then obtained the desired representations by looking at the functions LU or HLU, depending on whether we were in the affine or non-affine case. If L is unbounded on $(0,1)$ however, the functions LU or HLU are not well defined on the entire domain D: we have to exclude the best and worst temporal lotteries in D, namely, the deterministic consumption streams $(\overline{c}, \overline{c}, \ldots)$ and $(\underline{c}, \underline{c}, \ldots)$. In particular, let $D^\circ \subset D$ be the subset of all temporal lotteries whose consumption levels are drawn from the open interval $C^\circ := (\underline{c}, \overline{c})$. Following the preceding arguments, we can then obtain the desired representations on D°. It remains to show that these representations can be extended from D° to the entire domain D. The only nontrivial part in this argument is to show that an Uzawa representation on $(C^\circ)^\infty$ can be extended to an Uzawa representation on C^∞. The next lemma provides the details, thus completing the proof of Proposition 1.

Lemma 18 *If $\succeq$ is continuous on C^∞ and has an Uzawa representation on $(C^\circ)^\infty$, then $\succeq$ has an Uzawa representation on the entire domain C^∞.*

Proof. Let (u, b, U) be the Uzawa representation on $(C^\circ)^\infty$. In particular, note that u, b are functions on C° and U is a function on $(C^\circ)^\infty$. First we are going to show that $\lim_{c \nearrow \overline{c}} U(c, c, \ldots) < +\infty$. If $\lim_{c \nearrow \overline{c}} U(c, c, \ldots) = +\infty$, then we can find a sequence $(c_n)_n$ such that $c_n \in C^\circ$ and $\beta^n(c')U(c_n, c_n, \ldots) \geq 1$ for every n. Fix some $c', c'' \in C^\circ$ such that $U(c', c', \ldots) < U(c'', c'', \ldots) < U(c', c', \ldots) + \frac{1}{2}$ and consider the consumption streams

$d_1 := (c', c_1, c_1, \ldots), d_2 := (c', c', c_2, c_2, \ldots)$, and so on. Since the sequence $(d_n)_n$ converges pointwise to $(c', c', \ldots)$ and $\succeq$ is continuous in the product topology on C^∞, we know that $(c'', c'', \ldots) \succ d_n$ for all n large enough. But for every n,

$$U(d_n) = U(c', c', \ldots)(1 - b(c')^n) + b(c')^n U(c_n, c_n, \ldots) \geq U(c', c', \ldots)(1 - b(c')^n) + 1.$$

Hence, $U(d_n) > U(c', c', \ldots) + \frac{1}{2} > U(c'', c'', \ldots)$ for all n large enough, a contradiction.

Next we are going to show that $\lim_{c \nearrow \overline{c}} b(c) < 1$. The proof is once again by contradiction. Let $(c_n)_n$ be a sequence such that $c_n \nearrow \overline{c}, b(c_n) \nearrow 1$, and $c_n \in C^\circ$ for every n. Fix some $c, c' \in C^\circ$ such that $(\overline{c}, c, c, \ldots) \succ (c', c, c, \ldots) \succ (c, c, \ldots)$. Since $\succeq$ is continuous, we know that $(c_n, c, c, \ldots) \succ (c', c, c, \ldots)$ for all n large enough. Also,

$$U(c_n, c, c, \ldots) = (1 - b(c_n))U(c_n, c_n, \ldots) + b(c_n)U(c, c, \ldots) \quad \forall n.$$

Since $\lim_n U(c_n, c_n, \ldots) < \infty$ and $b(c_n) \nearrow 1$, it follows that $\lim_n U(c_n, c, c, \ldots) = U(c, c, \ldots)$. But then $U(c_n, c, c, \ldots) < U(c', c, c, \ldots)$ for all n large enough, contradicting the fact that U represents $\succeq$ on $(C^\circ)^\infty$. Analogous arguments show that $\lim_{c \searrow \underline{c}} U(c, c, \ldots) > -\infty$ and $\lim_{c \searrow \underline{c}} b(c) > 0$. Since $u(c) = (1 - b(c))^{-1} U(c, c, \ldots)$ for every $c \in C^0$ and the function U is bounded, we can conclude that $u : C^\circ \to \mathbb{R}$ is bounded. By taking limits, we can extend the functions $u, b : C^\circ \to \mathbb{R}$ from C° to C. Let $(u'.b', U')$ be the ensuing Uzawa representation on C^∞. By construction U' agrees with U on $(C^\circ)^\infty$ and hence represents $\succeq$ on $(C^\circ)^\infty$. Since U' is the continuous extension of U from $(C^\circ)^\infty$ to C^∞, the function U' represents $\succeq$ on C^∞ as well. ■

C Proof of Proposition 2

The first step is to show that $I : M([0,1]) \to [0,1]$ is *convex in probabilities*, i.e., that $\pi I(\mu_1) + (1-\pi)I(\mu_2) \geq I(\pi\mu_1 \oplus (1-\pi)\mu_2)$ for every $\pi \in [0,1]$, $\mu_1, \mu_2 \in M([0,1])$. The arguments are based on Lemma 1 in Grant, Kajii and Polak (2000), with several adjustments arising from the fact that we employ a different choice setting. Fix $\mu_1, \mu_2 \in M([0,1])$ and let $m_1, m_2 \in M(D)$ be such that $\mu_1 = m_1 \circ U^{-1}$ and $\mu_2 = m_2 \circ U^{-1}$. From preference for early resolution, we know that for all $c_0, c_1 \in C$ and $\pi \in [0,1]$

$$(c_0, \pi(c_1, m_1) \oplus (1-\pi)(c_1, m_2)) \succeq (c_0(c_1, \pi m_1 \oplus (1-\pi)m_2)). \tag{53}$$

Suppose $\succeq$ has a representation with a fixed discount factor $\beta \in (0,1)$ and a translation-invariant certainty equivalent I. Identical arguments apply when the discounting is endogenous. From (53), deduce that

$$u(c_0)+\beta I\Big(\pi[u(c_1)+\beta I(\mu_1)]\oplus(1-\pi)[u(c_1)+\beta I(\mu_2)]\Big) \geq u(c_0)+\beta u(c_1)+\beta^2 I(\pi\mu_1\oplus(1-\pi)\mu_2).$$

Using the fact that I is translation-invariant and canceling terms, we obtain:

$$I\big(\pi[\beta I(\mu_1)] \oplus (1-\pi)[\beta I(\mu_2)]\big) \geq \beta I(\pi\mu_1 \oplus (1-\pi)\mu_2).$$

Since $I \leq E$, we also know that

$$\pi\beta I(\mu_1) + (1-\pi)\beta I(\mu_2) \geq I\big(\pi[\beta I(\mu_1)] \oplus (1-\pi)[\beta I(\mu_2)]\big).$$

Combining the last two inequalities shows that I is convex in probabilities. Since I satisfies betweenness, Lemma 2 in Grant, Kajii and Polak (2000) shows that I is of the expected utility form, that is, there is a function $\phi : [0,1] \to [0,1]$ such that $I(\mu) = \phi^{-1}E_\mu\phi$. Since $I \leq E$, ϕ is concave. Since I is translation-invariant, ϕ is either exponential (of the CARA form) or linear. If I is also scale-invariant, then ϕ is linear.

D Proof of Proposition 4

We first establish a result similar to Lemma 1, after which the proof of Proposition 1 can be almost readily applied.

Lemma 19 *A binary relation $\succeq$ on $\mathcal{H}$ admits a recursive representation (U, W, I) (as defined in equation 23) if and only if it fulfills Axioms 1, 2, A.3, A.4 and A.5.*

Proof. Let us start with the necessity of the axioms. It is obvious that representation (23) implies that Axioms 1, 2 and A.4 hold. Remark that (21) and (23) imply that for any $c \in C$ and $h \in \mathcal{H}$, we have

$$U(c,h) = W(c, U(h)),$$

which proves that Axiom A.5 holds. Last, for Axiom A.3, let consider two acts $h = (h_0, h_1, \ldots)$ and $h' = (h_0, h'_1, \ldots)$ such that $U(h^s) \geq U(h'^s)$ for all s. Using (20) we get $U(h^s) = W(h_0, U(h^{s,1}))$ and $U(h'^s) = W(h_0, U(h'^{s,1}))$, so that the inequality $U(h^s) \geq U(h'^s)$ for all s provides $U(h^{s,1}) \geq U(h'^{s,1})$ for all s. Thus, we deduce that $U(h) \geq U(h')$, which implies that Axiom A.3 holds and concludes the necessity part.

We now demonstrate that the axioms are sufficient. Let us denote by $\mathcal{H}_e$ the set of consumption plans $h = (h_0, h_1, \ldots)$ whose consumption h_0 at date 0 is not constrained to be deterministic, that is the set $C-$valued and $\mathcal{G}-$adapted processes. States will be denoted

$(s_0, s_1, \ldots)$ to emphasize the difference with our setting, where h_0 was supposed to be constant (while here it may depend on s_0). Let $c_0 \in C$ be a given constant consumption level. We define a binary relation $\succeq_e$ on the set $\mathcal{H}_e$ as follows:

$$\forall (h, h') \in \mathcal{H}_e^2, \ h \succeq_e h' \Leftrightarrow (c_0, h) \succeq (c_0, h'),$$

where (c, h) is defined similarly to equation (19). Because of Axiom A.4, the binary relation $\succeq_e$ is independent of c_0 and defines a preference relation on $\mathcal{H}_e$. Moreover, for any $c \in C$ and $h \in \mathcal{H}$, Axioms A.4 and A.5 imply that $(c, h) \succeq_e (c, h') \Leftrightarrow h \succeq_e h'$. The preference relation $\succeq_e$ fulfills therefore a property similar to the one defined in Axiom A.5. By continuity of $\succeq$ and thus of $\succeq_e$, there exists a continuous utility representation, whose corresponding utility function is denoted U.

For any $c \in C$, the functions $h \in \mathcal{H}_e \mapsto U(h)$ and $h \in \mathcal{H}_e \mapsto U((c, h))$ both represent the preference relation $\succeq_e$. Therefore, there exists a continuous function W, which is increasing in its second argument, such that

$$\forall h \in \mathcal{H}_e, U((c, h)) = W(c, U(h)).$$

For any h in $\mathcal{H}_e$ and any $s \in S$ on can define a conditional act $h^s \in \mathcal{H}$ similarly to equation (17). Consider now two acts h and h' in $\mathcal{H}_e$ such that $h^s \succeq_e h'^s$ for all s. By definition of $\succeq_e$, we have $(c_0, h^s) \succeq (c_0, h'^s)$ for all s. Axiom A.3 implies then that $h \succeq_e h'$. The set $\mathcal{H}_e$ being isomorph to $\mathcal{H}^S$, we obtain that for any $h \in \mathcal{H}_e$, $U(h) = I(U \circ h)$ where $U \circ h : S \to Im(U) = [0, 1]$ is defined by $(U \circ h)(s) = U(h^s)$ and $I : B_0(\Sigma) \to \mathbb{R}_+$ is a continuous, strictly increasing function. Since any $h = (h_0, h_1, \ldots) \in \mathcal{H}$ can be viewed as $(h_0, (h_1, \ldots))$ where $(h_1, \ldots) \in \mathcal{H}_e$, we obtain:

$$U(h_0, \ldots) = W(h_0, I(U \circ h^1)).$$

It remains to show that I fulfill $I(x) = x$ for all $x \in Im(U)$. For this, consider any act such that $U(h^{s,1}) = x$ for all s. We have $U(h^s) = W(h_0, x)$ which is independent of s. With Axiom A.3, this implies that $U(h^s) = U(h)$ and therefore $I(x) = x$. ■

To end up proving Proposition 4, it remains to show that Axioms 6 and A.7 hold if and only if one can use a time aggregator W and a certainty equivalent that fulfill the same kind of restrictions as those derived in the risk setting. The sufficiency part of the proof is exactly the same as the one provided for the risk setting. Indeed, in Section B, the number $m \in \mathbb{N}_+$ and the probabilities $(\pi_1, \ldots, \pi_m) \in (0, 1)^m$ were considered to be fixed. Thus, the whole reasoning that was done considering the simple lottery $(\pi_1, x_1; \ldots; \pi_m, x_m) \in M([0, 1])$

can be reproduced here without any change by imposing that $m = card(S)$ and viewing the $x_1, \ldots, x_m$ as state-contingent realizations.

Necessity of Axiom 6 is obvious, as in the risk setting. For the necessity of Axiom A.7, which is a stronger monotonicity requirement that the one imposed in the risk setting, let us define the notion of time t conditional act as follows. For any $t \geq 1$, any $(\sigma_1, \ldots, \sigma_t) \in S^t$, and any act $h \in \mathcal{H}$ we set:

$$h^{\sigma_1,\ldots,\sigma_t} \in \mathcal{H} : (s_1, s_2, \ldots) \in \Omega \to h^{\sigma_1,\ldots,\sigma_t}(s_1, s_2, \ldots) = h(\sigma_1, \ldots, \sigma_t, s_{t+1}, s_{t+2}, \ldots).$$

This is a direct generalization to the notion of conditional act defined in equation (17).

Now consider any act h such that any time t conditional act is a constant act (i.e., the act h only depends on the information revealed during the first t periods). We can show that $U(h)$ is given by the terminal point of the backward recursion:

$$\begin{cases} V_t(\sigma_1, \ldots, \sigma_t, h) & = V(h^{\sigma_1,\ldots,\sigma_t}) \text{ for } \tau \geq t, \\ V_\tau(\sigma_1, \ldots, \sigma_\tau, h) & = I_{\tau+1}\left(s \mapsto V_{\tau+1}(\sigma_1, \ldots, \sigma_\tau, s, h)\right) \text{ for all } \tau \geq 0, \end{cases} \tag{54}$$

where V is the ex-post lifetime utility function, as defined in Section 5.1 and the I_τ are given by $I_t(\mu) = \beta^t I\left(\frac{1}{\beta^t}\mu\right)$. The property stated in Axiom A.7 is then found to hold for any pair of acts h and h' that only depends on information revealed in a finite number of periods. The extension to all pairs of acts is obtained by continuity.